Date of Submission: 6/6/2019
Word Count: 49 975

Computational and neural mechanisms of human aversive learning

Ondřej Zíka
St. Anne's College

DPhil in Clinical Neurosciences
University of Oxford

Acknowledgments

Firstly, I would like to thank the bodies that supported my education during and leading towards my PhD, namely, the Kellner Family Foundation, the Medical Research Council and the Engineering and Physical Sciences Research Council, and authors of tools that were essential for completing the work (particularly Linux/GNU, SPM, FSL, PSPM and jags).

I am eternally grateful to individuals who supported, encouraged and inspired me on this ten year long journey, particularly to those that went beyond their line of duty, who opened doors for me and who, for some inconceivable reason, believed in me. I wish to particularly thank Lucien Holmes, I-Chant Chiang, Guillaume Thierry, John O'Doherty, Andrea Reinecke and Nico Schuck. I am also extremely grateful to Katja Wiech and Rafal Bogacz who willingly (mostly) underwent four years of dealing with me and who were a great source of confidence and inspiration even at desperate times, and to Judith Appel and Lorika Skhreli, who aided me with data collection. A special big thank you to my sister Adela who is always on my side and who supported me throughout the years.

And most of all, I thank Laiza Poletti, Nini, who was the sole source of life force and without whom this would really not have much point.

I dedicate this work to the memory of Prof. Václav Nýdl, who inspired and catalyzed my hunger for discovery and especially the love for mathematics (which I obliterated by leaving to study psychology only to re-discovery it a few years later), and of whom I briefly think every time I apply the product rule.

Abstract

Aversive learning is characterised by rapid learning which is highly resistant to extinction. This has likely developed as an evolutionary mechanism but it often becomes maladaptive. In this thesis I investigate computational mechanisms governing the learning of aversive beliefs, neural structures that are involved in the process and test whether pharmacological intervention can be used to amend learning.

In a series of three experiments in healthy volunteers ranging in trait anxiety I employ a probabilistic aversive learning paradigm in behavioural (study 2 and 3) and neuroimaging setting (study 1). I first characterise the difference between acquisition and extinction, identifying that while acquisition learning is fast and reaches the target level, during extinction aversive expectations are consistently higher than true reinforcement rates. This extinction-specific overprediction increases over time but only in the low anxious group. High trait anxiety was associated with increased dissociability between acquisition and extinction which was driven by the activity in the dorsal anterior cingulate cortex. Computational modelling revealed that high anxiety is associated with the tendency to internally represent the learning environment as distinct states and switch between them. This finding was later supported by behavioural and computational evidence in study 2. The neuroimaging analysis suggests that state learning is processed by the dorsal anterior cingulate cortex, ventro-medial prefrontal cortex and the inferior parietal lobule. This finding highlights increased context-dependence of anxious individuals which bears importance to clinical practice. In study 3, I test the effect of Angiotensin-II receptor antagonist losartan on aversive learning showing that losartan specifically decreases learning rates without influencing perception or updating.

Abbreviations

BADS	Bayesian Adaptive Direct Search
MCMC	Markov Chain Monte Carlo
VB	Variational Bayes
BMS	Bayesian Model Selection
LME	Log Model Evidence
SEM	Standard Error of the Mean
GSR	Galvanic Skin Response
SCR	Skin Conductance Response
MNI	Montreal Neurological Institute
ROI	Region of Interest
AAL	Automatic Anatomical Labelling
FMRI	Function Magnetic Resonance Imaging
ERP	Event-Related Potentials
FSL	FMRIB Software Library
SPM	Statistical Parametric Mapping
dACC	Dorsal Anterior Cingulate Cortex
vmPFC	Ventro-Medial Prefrontal Cortex
dlPFC	Dorso-Lateral Preforntal Cortex
IPL	Inferior Parietal Lobule
RW	Rescorla-Wagner
PH	Pearce Hall

Table of Contents

Chapter 1

Literature Background and Introduction

Introduction

“For months after the attack, I couldn't close my eyes without envisioning the face of my attacker. I suffered horrific flashbacks and nightmares. For four years after the attack I was unable to sleep alone in my house. I obsessively checked windows, doors, and locks. By age 17, I'd suffered my first panic attack. Soon I became unable to leave my apartment for weeks at a time, ending my modelling career abruptly. This just became a way of life. Years passed when I had few or no symptoms at all, and I led what I thought was a fairly normal life, just thinking I had a "panic problem."

Then another traumatic event re-triggered the PTSD. It was as if the past had evaporated, and I was back in the place of my attack, only now I had uncontrollable thoughts of someone entering my house and harming my daughter. I saw violent images every time I closed my eyes. I lost all ability to concentrate or even complete simple tasks. Normally social, I stopped trying to make friends or get involved in my community. I often felt disoriented, forgetting where, or who, I was. I would panic on the freeway and became unable to

drive, again ending a career. I felt as if I had completely lost my mind. For a time, I managed to keep it together on the outside, but then I became unable to leave my house again."

*P.K. Philips, victim of mental and sexual abuse **

*https://adaa.org/living-with-anxiety/personal-stories/my-story-survival-battling-ptsd

The above testimony describes the severe consequence of a highly aversive experience, known as Post-Traumatic Syndrome Disorder (PTSD), a prevalent affective disorder estimated to affect 3 - 8% of the general population (Burri and Maercker 2014). PTSD is characterised by rapid learning that is prone to generalization to other contexts and by the near impossibility to overcome the associated fear, anxiety and catastrophising about future aversive events. While PTSD is arguably one of the most severe cases of aversive learning, it shares key features with common events that happen to every one of us on a daily basis. When we are stung by a wasp or bitten by a snake in a forest we often associate the fear and pain with the context where it happened. In some cases, this leads to the development of a phobia, and we then stay away from all forests. However, even if the fear is not extreme, we might be cautions for the rest of our lives whenever we are in or near a forest. In other words, it takes one unpleasantexperience to acquire the fear but many for it to go away.

Pain, fear, losses or emotional distress can be learned practically in an instance but it can take long time to forget them. This asymmetry appears to be a fundamental aspect of aversive learning (Bouton 2002). It has been reported in different forms and under different names in anxiety disorders (Arch and Craske 2009), fear learning (Mystkowski et al. 2006), pain (Rachman and Arntz 1991) or financial decision-making (Thaler et al. 1997). The resistance to the extinction of aversive associations has likely

been developed as a part of our evolutionary makeup to ensure survival (LeDoux 2000). A shrew that nearly got killed by an eagle on field will avoid that very field in the future even though it is buzzing with tasty beetles and grasshoppers. Yet, from an information processing perspective, this tendency constitutes an irrational bias. Under optimal accounts of learning, the rate of learning should be equal for gains and losses, fear and happiness, harm and relief (von Neumann and Morgenstern, 1947). That is, the shrew should revisit the field and update its expectation to encounter an eagle as quickly as it associated the bird with this field in the first place. Yet, that is not what happens - and for a good reason, as the consequences of underpredicting aversive events are more severe than those of overprediction.

The goal of my work is to understand what constitutes the difference between aversive learning and forgetting at the level of human behaviour, peripheral physiology and neural information processing. This is not a novel endeavour. Experiments in human and animal learning have tried to understand aversive learning throughout most of the 20th century. In contrast to most previous studies I employ computational modelling. This approach is relatively new to psychology and neuroscience, yet, it has already achieved remarkable results. Computational models allow us to explicitly model cognitive, neural and physiological processes which can't otherwise be observed. This methodology allows us to study the human brain as a machine, in a systematic and algorithmic way, which among other things carries the promise of improved and personalized treatment.

The goals of this thesis are to i) develop an algorithmic understanding of how humans learn to predict aversive events; ii) identify brain areas that are involved in such learning; iii) characterise the behavioural and physiological difference between aversive acquisition and extinction; iv) investigate the role of trait anxiety in aversive learning; v) test how different environment statistics influence acquisition and extinction and vi) explore whether both processes can be altered pharmacologically.

In Chapter 1, I will present a review of the existing literature on learning theory, the influence of environmental and individual factors on aversive learning, discuss the role of anxiety, the neural basis of aversive learning, optimal aversive learning, and a set of relevant computational theories and approaches that can help us understand the learning process in more detail. Chapter 2 summarizes the methods used in this investigation, including the experimental paradigm, physiological measures, functional MRI and computational modelling methods.

Chapters 3 to 5 each cover one experimental study, investigating the behavioural and neural differences between aversive acquisition and extinction (Chapter 3); the influence of environmental contingencies on learning strategies (Chapter 4); and the effect of angiotensin-II receptor antagonist losartan on aversive learning (Chapter 5). All three studies were designed to answer specific questions, but also address the three main goals of the thesis:

(1) What is the difference between acquisition and extinction?

(2) What role does trait anxiety play in acquisition and extinction?

(3) What computational mechanisms characterise aversive learning?

The thesis will be completed by an Overall Discussion section in which my experimental findings will be discussed in relation to the existing literature.

Learning theory

This thesis concerns itself with studying the difference of how aversive associations are acquired and extinguished. The anecdotal evidence presented in the introduction is supported by a large body of both animal and human studies. Aversive acquisition is consistently reported to be fast, while extinction is slow, incomplete and prone to relapse (Mineka and Zinbarg 2006; Bouton 2002; Craske et al. 2014a). Understanding the difference between the two processes has been a focal topic of

learning research since Watson and Rayner (1920) reported the famous little Albert case. An eleven month old boy was conditioned to associate loud sound with rabbits to subsequently exhibit prolonged avoidance behaviour towards animate and inanimate furry objects, indicating generalised fear. Aversive learning and extinction (or lack thereof) in particular has been the focus of extensive research due to its involvement in the development and maintenance of anxiety and related disorders, as well as their treatment (Hofmann 2008). Despite the improved understanding, the fundamental processes involved in aversive learning, particularly in extinction, are still debated. In addition to basic learning mechanisms, cognitive processes, such as attention or expectation, anxiety sensitivity or catastrophising have been shown to modulate extinction learning (Mineka and Zinbarg, 2006; Hofmann, 2008) The following section summarizes the underlying classical conditioning theory, theoretical propositions and experimental evidence for the differences between acquisition and extinction, as well as the role of environmental and individual factors in aversive learning.

Classical Conditioning

The concepts of acquisition and extinction are rooted in the theory of classical (i. e. Pavlovian) conditioning. In 1927, Pavlov reported the original set of conditioning experiments where dogs were observed to increase their salivation when a cue (tone / light) was repeatedly presented prior to the delivery of food. He referred to this type of learning as ‘associative’ because a previously neutral stimulus is being associated with the delivery of food and over time becomes predictive of future reward, i. e. the reward is now directly associated with the light.

Pavlovian learning occurs without an action or intention of the animal. It stands in contrast to instrumental learning where an action is required from the agent to achieve the reward / prevent a punishment. Although conditioning can occur without

conscious awareness (Squire 1992), it has been shown to play an important role (Lovibond and Shanks 2002). Pavlovian learning is a fundamental process of how brains learn, store, retrieve and manipulate information. Identifying cues which predict future rewards and warn against potential future harm are critical for an organism's survival. It has likely developed as an evolutionary mechanism that allows organisms to predict future events with a high level of accuracy at a low level of energy consumption. A number of mechanism have been found to influence learning including the predictability of future events (Mackintosh 1975), attention (Pearce and Hall 1980), prediction reliability (Kakade and Dayan 2002), conditioning schedule (Pearce and Bouton 2001) number of competing predictive cues (Wagner, 1981), conditioning context (Bouton, 2002), degree of evolutionary preparedness (Seligman 2016), properties of CS and perception of control (both Mineka and Zinbarg, 2006). Aspects relevant to this thesis will be discussed later in this chapter.

The terminology used in classical conditioning is based on the concepts of a stimulus and a response. If an initially neutral stimulus (conditioned stimulus, CS) becomes associated with a motivationally relevant stimulus (e.g., food or pain; unconditioned stimulus, US), the CS begins to predict the occurrence of the US and elicits a response (conditioned response, CR) that resembles the response to the US (e.g., withdrawal in case of pain). Acquisition refers to the learning process where a neutral stimulus becomes associated with the US. The acquired conditioned response occurs when the CS is now presented in the absence of the US. In contrast, extinction refers to the learning phase in which the previously conditioned stimulus is no longer followed by the US.

The extent to which acquisition or extinction occurs is assessed by monitoring the transfer of response from US to CS. The conditioned response measured in human experimental studies can vary depending on the study design. Perhaps the most common measure of autonomous response is the skin conductance (SC, also SCR for Skin Conductance Response). SCR measures the changes in sweat gland activity

reflecting the level of arousal. Another common methods include electromyogram (EMG), blink rate or startle response. In addition to autonomic measures, expectancy ratings have been used as a measure of subjective belief about the likelihood of aversive events / rewards (Jepma et al. 2018). (Dawson, Schell, and Banis 1986) demonstrated that explicit ratings can follow conditioned response, although this finding was not replicated by number of studies (Haaker et al. 2015). Collecting both physiological data and expectancy ratings can help researchers dissociate autonomous and cognitive components of learning.

In an experimental setting in humans, aversive learning is investigated using computerized tasks in which a cue (sound or visual cue, a colored shape or an abstract image) is repeatedly presented on the screen and either paired with aversive unconditioned stimulus or not. Unconditioned stimuli used in experimental studies involve electrical, heat or laser pain stimuli, a puff of air or a burst of white noise. In most conditioning studies the CR is measured, however, as mentioned earlier it is also common to obtain subjective shock expectancy ratings from the participant. Majority of studies uses high reinforcement contingency levels in acquisition and no reinforcement in extinction. An exception are fMRI studies where usually only non-reinforced trials are used for analysis, in which a conditioned stimulus can be reinforced on as little as 25% of trials.

Theories of Acquisition and Extinction

Early theories of learning describe and operationalise extinction as a process that mirrors acquisition (Pavlov 2010) Rescorla and Wagner, 1972). When a CS (e.g., tone, cue) is no longer followed by the US (e.g., shock), the agent gradually updates their expectation of the aversive event. Such theories explain the effect of slower extinction in terms of slower learning from a non-aversive stimuli, attention

(Mackintosh, 1975, Kakade and Dayan, 2002) or evolutionary optimal overprediction of aversive events (Rachman and Arntz, 1991). Learning about aversive contingencies is believed to override the previous associations. This view has been challenged by evidence from laboratory experiments (Bolles and Bouton, 1979) and clinical practice (Craske et al. 2014) in which a seemingly extinguished memory returned after a passage of time, return of context or presentation of unconditioned stimulus on its own.

Laboratory experiments have identified several distinct cases of relapse suggesting that the aversive association learned in acquisition is not forgotten. Particularly, the effects of reinstatement (Rescorla and Heth 1975), renewal (Bouton and Bolles 1979), spontaneous recovery (Brooks and Bouton 1993) and re-acquisition (Napier, Macrae, and Kehoe 1992) cast doubt on the original theory of extinction. Reinstatement occurs when following successful extinction a previously used US is presented on its own. Rescorla and Hath (1975) reported full recovery of the conditioned response upon presentation of the extinguished CS. The renewal phenomenon perhaps maps most closely onto clinical observation: when acquisition takes place in context A ('the world') and extinction in context B ('the therapist's office'), the conditioned response re-occurs once the agent finds itself back in context A. In addition to the reinstatement and renewal, arguably context-driven mechanisms, passage of time has been shown to increase return of aversive expectations (Pavlov, 1927; Robbins 1990), a phenomenon termed spontaneous recovery. Lastly, when an extinguished CS is presented together with the US again, rapid re-learning (i.e. re-acquisition) occurs (Napier et al. 1992). This evidence casts doubt on the view of extinction a process of mere 'unlearning'.

Bouton (2002) proposed that while acquisition is in fact one single learning

process, extinction involves at least two processes: first, the inhibition of previously acquired associations and second, learning about the new contingencies of the CS. This theory explains why previously conditioned responses tend to rapidly return: the original aversive association has never been overwritten. Instead, it was inhibited to be potentially re-used in the future. Bouton (2002) notes that extinction, and inhibition in general (Wagner, 1981), is a more context-dependent process than acquisition.

To challenge this view, Myers, Ressler, and Davis (2006) performed a set of experiments using reinstatement, renewal and spontaneous recovery paradigms and reported that extinction immediately after acquisition leads to a more lasting effect, and therefore to unlearning. The authors proposed that extinction is time-dependent and can therefore be described by a combination of the two original views. If it is performed immediately after acquisition, unlearning occurs. However, if it is performed 24 hours or longer after acquisition the conditioned associations have had time to consolidate and will behave similarly as described by Bouton (2002). This time-dependency effect would have a major impact on clinical applications. The studies using immediate extinction were repeated by number of research groups not replicating the results (Schiller et al. 2008a; Alvarez, Johnson, and Grillon 2007), casting doubt over the validity of the results. It is worth noting that in the original study by Myers et al. the immediate extinction was less effective than extinction 72 hrs after conditioning, but it appeared to be more permanent. In the end, both, immediate and delayed extinction, resulted in the same level of conditioned responding.

The above evidence suggest that acquisition and extinction are two separate processes. A more recent account of learning has highlighted that while they are often dissociable processes, they can also be a single learning process if

contingencies change gradually (Gershman et al. 2013). This account bears important clinical implications and it will be discussed in a later section. Prior to that, I will discuss cognitive processes influencing aversive learning and the role of anxiety.

Cognitive and Environmental Influences on Aversive Learning

While classical conditioning was originally treated as an unconscious autonomic mechanism, there is a large body of evidence for a significant role of cognitive processes in Pavlovian learning. In this part, the role of context, time, prior experience, stress, attention, and individual difference on aversive learning will be discussed. Anxiety will be purposefully omitted as it will be treated separately in a subsequent section.

Context

The context-specific effect on learning was first reported by Pavlov (1927) who observed the return of previously extinguished CR by presentation of US alone, a phenomenon known as reinstatement, described above. Bouton and Bolles (1979) reported that reinstatement only occurred when the US was presented in the same context (i.e. cage) as the original acquisition. In a similar study, contextual fear lead to reinstatement only when extinction comprised of only non-reinforced trials. When a partial reinforcement schedule was used, the chance of reinstatement decreased (Bouton and King 1986). Bouton (2002) suggested that cue that has been associated with both acquisition and extinction is more ambiguous and therefore more susceptible to contextual influences.

Renewal is a case of extinction relapse purely drive by context. When a US-CS association is acquired in context A and extinguished in context B, renewal of conditioned response is nearly instant and complete if context A is re-introduced in contrast to a case where context B remains (Bouton and Bolles, 1979). Interestingly,

the degree to which the conditioned response is renewed might be graded: when context C is introduced post-extinction renewal occurs as well, although not as strong as ABA renewal (Bouton and Brooks 1993). Furthermore, if context A is used in both acquisition and extinction and context is subsequently changed to context B, renewal occurs once again. This evidence suggests that a conditioned context might become a default prior to expectation of any new context.

Difference in context-specificity was proposed as one of the fundamental differences between acquisition and extinction (Bouton 1994). Acquisition was reported to generalize better than extinction (Bouton and Brooks 1993) which was in turn more context-specific (Rosas, Todd, and Bouton 2013). The change in context after conditioning without a decrease in the rate of reinforcement was reported to only lead to little change in conditioned responding (Bouton and King 1983; Harris et al. 2000). (Nelson 2002) specifically tested whether context-specificity in extinction is a generalization of context-specificity of inhibition. Conditioned inhibition was found to generalize across contexts, however, conditioning involving higher ambiguity, a form of uncertainty, was found to be more context-sensitive. Alternative explanation of higher context-specificity of extinction is that by definition acquisition always precedes extinction and the latter is therefore always an exception to a previously acquired rule, a form of primacy effect. Number of studies directly or indirectly support this idea (Swartzentruber and Bouton 1992; Nelson 2002), stressing the role of environmental ambiguity in context development. Harris et al. (2000) reported that a context-free acquisition became more context-dependent after extinction, especially in ambiguous environments. Perhaps learning comes to rely more on context as the uncertainty in observation increases. This idea will be discussed later from an information-processing perspective. It should be noted that the larger context-specificity of extinction has been challenged by a number of studies demonstrating context-specificity of acquisition (Starosta et al. 2016; Hall and Honey 1989).

An important question when discussing a context is what constitutes a context.

Bouton (2002) distinguishes between exteroceptive (room, stimuli, background, researcher) and interoceptive (bodily states, time passage, memory) contexts. In the area of learning, exteroceptive context could be the stableness / volatility of an environment (Behrens et al. 2007; Browning et al. 2015), time-dependent factors or a degree of partial reinforcement. Humans have been shown to create contexts and find patterns even in random sequence of events. How contexts may arise will be an important topic in Chapter 4. Interoceptive context can be based on the level of arousal, heart beat or memory. Importantly, pharmacological agents were shown to provide a conditioning context (MacArdy and Riccio 1991).

Time

Relatively little attention has been paid to multiple acquisition and extinction phases, i.e. scenarios where extinguished CR is re-conditioned and re-extinguished again. Yet, time-dependent changes might more closely resemble patient experience. In an experimental context, the analysis of multiple periods of acquisition and extinction often averages across multiple trials (Chang and Maren 2011) or entire sessions (e.g., Quirk 2002; Baum 1972). This is most likely due to high noise in single trial measurement of conditioned responses (most common measure is the GSR). Gottlieb and Rescorla (2010) specifically mention this challenge. A recent interest in computational modelling and single-trial analysis has lead to some remarkable publications (Bach, Tzovara, and Vunder 2018; Zhang et al. 2016; de Berker et al. 2016) but no study has specifically analysed time-dependent effect on extinction learning.

Past research investigating the impact of time on acquisition and extinction reported retainment of extinction memory past re-acquisition. This result supports the role of separate mechanisms (or contexts) of acquisition and extinction (Quirk 2002). Langton & Richardson (2009) found both extinction and re-extinction to be context-specific and susceptible to reinstatement. Baum (1972) repeated acquisition and

extinction on 5 consecutive days. Lack of extinction was apparent on day 1 and more prominent on day 2, however on days 3-5 extinction was nearly complete. This result suggests a possible meta-learning mechanism (i.e. learning to learn) over multiple sessions under which participants gradually learn the structure of the environment or task, a result similar to the one previously reported by (Harlow 1949).

Partial Reinforcement and Uncertainty

Partial Reinforcement

Most experiments discussed until this point use fully reinforced schedules. That is, during acquisition the CS is paired with an US on every trial (i.e., the reinforcement rate is 100%). While this approach may be valuable in the laboratory, natural environments are rarely fully deterministic. Learning about the likelihood of aversive events in a noisy environment is a crucial component of evolutionary adaptive behaviour. Experiments investigating the influence of partial reinforcement on aversive learning report the so called Partial Reinforcement Extinction Effect (PREE). When a partial reinforcement schedule is used during acquisition, the subsequent extinction is significantly slower (Li et al. 2016; Jenkins and Stanley 1950; Wagner, Siegel, and Fein 1967) compared to full reinforcement. It has been argued that full reinforcement provides a clearer context and the sudden lack of the US is interpreted as a new environment. Interestingly, a change of context accelerates extinction following partially reinforced acquisition (Bonardi, Honey, and Hall 1990; Boughner and Papini 2006).

Another phenomenon related to partial reinforcement is the so called occasional reinforced extinction. Under such a schedule, the CS is occasionally followed by the US during extinction (Bouton, 2004). This occasional reinforcement results in a more permanent extinction of the CR. This peculiarity of aversive learning has received a large amount of attention because it has a direct relevance for

exposure therapy (Craske et al. 2014). Three ideas explaining this phenomenon were put forward. The first hypothesis by Bouton et al. (2004) posits that this is due to the expectancy violation effect in which participants are less likely to expect the next trial to include a US since it has just occurred. The second explanation focuses on the attention paid to incoming the information. When an unexpected US occurs, the learning rate on the next trial increases (Pearce and Hall, 1980). This assumption is the hallmark of the Pearce-Hall model which will be explained in more detail later in this chapter. The third hypothesis focuses on the observed slower re-acquisition and the lower rate of reinstatement and spontaneous recovery following occasional reinforced extinction (Culver, Vervliet, and Craske 2015; Gershman et al. 2013). Gershman and colleagues reason that if extinction leads to the creation of a new learning process due to large errors (as posited by Bouton, 2002), a gradual, rather than rapid, change in contingency will lead to continuous updating of the original memory rather than to creation of a new one. This in turn means that extinction becomes a single learning process and that it will be more resistant to relapse. The experimental data of Gershman et al. (2013) as well as clinical reports (Craske et al., 2014) support this last hypothesis.

Uncertainty

Uncertainty has been shown to play a crucial role in learning (Dayan and Jyu 2003; Behrens et al. 2007). Recent advances in theoretical neuroscience have highlighted the probabilistic nature of information coding in the human brain (Knill and Pouget 2004), providing a neural and computational basis for representation of uncertainty (Schultz et al. 2008; Hsu et al. 2005)

Normative Bayesian accounts of learning often directly consider several types of uncertainty (Mathys et al. 2014a). The most general distinction is between reducible and irreducible uncertainty. While it is possible to learn, and thus minimise, reducible

uncertainty, irreducible uncertainty cannot be learned. For example, when parking a bicycle in the city center of Oxford one can learn which spots are more likely to be available in general, thus reducing the uncertainty of not finding a spot. On some days a random event, such as an accident, may alter the availability of bike parking spots. Since this event was unpredictable we refer to this type of uncertainty as irreducible. Number of studies have shown that the brain is able to track these two types of uncertainty separately. de Berker et al. (2016) was able to find such signals in galvanic skin response and pupil dilation data. Marshall et al. (2016) related different neurotransmitters to different types of uncertainty: noradrenaline to unexpected changes in the environment, acetylcholine attributes degree of uncertainty to events and dopamine links to representation of uncertainty. Payzan-LeNestour and Bossaerts (2011) linked unexpected uncertainty to neural activity in the brainstem and locus coeruleus, a primarily noradrenergic region.

Within the context of learning, acquisition and extinction, distinct types of uncertainty can be identified. I will briefly discuss perceptual, contingency and environmental uncertainty here, but in general, any computed variable has an inherent uncertainty component to it.

Perceptual uncertainty refers to the degree that a signal can be detected. Imagine standing in a heavy rain and having a bucket of water poured over your head. One might not even notice it. In the context of aversive stimuli this could be a pain in an already painful state, or simply the intensity of the pain which can change based on the degree of habituation and sensitisation. Contingency uncertainty refers to the likelihood of a certain event occurring, such as experiencing pain at a dentist. Lastly, environmental uncertainty then refers to the changeability of the environment: weather in the EU might be relatively stable while weather in the UK might change often. These three levels of uncertainty are often modelled in learning experiments, using Bayesian models such as the Hierarchical Gaussian Filter (Mathys et al. 2011 and 2014).

In the context of aversive learning uncertainty has often been described as a

feature that is aversive on its own: a certain pain was shown to be preferred to unstable chance of pain (Hiroshi and Yasuhiro 1982). In human fMRI experiment, uncertainty was found to be processed by the same neural structures as fear (Rosen and Donley 2006). Gorka et al. (2014) investigated the relationship of startle response and irreducible uncertainty in Panic Disorder. The authors reported that only participants with high intolerance to uncertainty showed increased startle response to safety cues. Browning et al. (2015) reported volatility encoding in pupil dilation. These findings suggest a key role of uncertainty in aversive learning, highlighting its link to physiological measures such as pupil dilation and startle response.

The Role of Anxiety in Aversive learning

A large body of research supports a key role of anxiety in associative learning (Lissek et al. 2005, Mineka and Zinbarg, 2006; Duits et al. 2015). Anxiety is often reported to play a disruptive role in aversive learning (Indovina et al. 2011; Browning et al., 2015; Edwards et al. 2017), particularly when at clinical level. By disruptive, here, I mean that it often has an undesirable influence, such as that it leads to excessive fear of more pronounced lack of extinction. In the contest of extinction, disruptive influence would be a cause that decreases the rate of extinction or that increases the rate of relapse.

Anxiety disorders and aversive learning

Clinical anxiety has been linked to increased fear generalization (Lissek et al. 2005), a lowered ability to distinguish between safe and threatening cues (Michael et al. 2007), lack of extinction (Sehlmeyer et al. 2011) and a higher chance of renewal and spontaneous recovery (Staples-Bradley, Treanor, and Craske 2018). Although the evidence for a disruptive role of clinical anxiety in aversive learning is supported by

many studies, it is actually by no means conclusive. A meta-analysis by Lissek et al. (2005) investigating the role of clinical anxiety in aversive acquisition and extinction is a frequently cited publication used to support the disruptive role of anxiety in learning. The results presented by the study are not nearly as strong as often implied. Across the 20 clinical studies which were included in the analysis, anxiety was found to accelerate acquisition and decelerate extinction. However, this effect was only present in "*simple, single-cue paradigms where only danger cues are presented and no inhibition of fear to safety cues is required*" (Lissek et al. 2005). Approximately one third of the included studies showed a negative effect size of disruptive role of anxiety meaning that anxiety lead to faster extinction or slower acquisition. This suggests that the exact role of anxiety might be more complex than commonly assumed. More specifically, anxiety might have positive, rather than negative, effect on learning under certain circumstances. A recent study in clinical anxiety patients (diagnosed with at least one of the following: generalized anxiety disorder, social anxiety disorder, specific phobia or panic disorder) by Marin et al. (2017) found decreased conditional responding to threat and safety compared to healthy controls. This large cohort study (N=61 patients and N=21 controls) bears particular importance as increased conditioned responding is often considered a biomarker of anxiety disorders (Kindt and Soeter 2014);

Trait Anxiety

A growing body of research aims at identifying precursors of clinical anxiety. Trait anxiety, in particular, has been proposed as a key risk factor (Mineka and Oehlberg 2008). While high trait anxiety is often associated with similar learning deficits as clinical anxiety, a number of studies reported no effect of trait anxiety on aversive learning or even a positive effect (Raes et al. 2009). The most frequently reported and cited finding is that high trait anxious individuals are unable to dissociate

between safe and harmful stimuli (Gazendam, Kamphuis, and Kindt 2013; Haddad et al. 2012). This is often operationalised as the difference between conditioned response (galvanic skin response or startle response) during extinction. High trait anxiety has been associated with a lack of inhibition of conditioned responding to extinguished stimuli (Gazendam et al. 2013; Indovina et al. 2011). Other studies did not replicate this effect (Torrents-Rodas et al. 2013; Pineles, Vogt, and Orr 2009; Otto et al. 2007). Torrents-Rodas et al. (2013) discussed the potential reason for this discrepancy. Most studies that found no effect of trait anxiety used high reinforcement rates (~70%+) while studies that used lower reinforcement rates found impaired inhibition in high trait anxious individuals. This suggests that the influence of trait anxiety depends on the degree of uncertainty. Extinction following a 33% reinforced acquisition is associated with higher uncertainty because the change is less apparent to participants. In contrast, extinction following an acquisition phase with 80% reinforcement rate would leave the participant more certain that the environmental contingencies have changed. This hypothesis is corroborated by other findings relating physiological measures to uncertainty and volatility (M. Zhang et al. 2016; Browning et al. 2015; Birrell et al. 2011). Li et al. (2011) showed that galvanic skin response measures correlate with the degree of contingency uncertainty, that is the highest SCR response was observed in periods just after a change in contingency. De Berker et al. (2016) demonstrated that irreducible uncertainty, i.e. volatility, uniquely predicts the level of experienced and reported stress. In line with this finding, Browning et al. (2015) reported that high trait anxious participants failed to estimate the environmental volatility which was reflected by a lack of volatility signal in pupil dilation measure in high trait anxious participants. An unpublished study by Wise et al. (in submission) found threat-related attention to be driven by aversive value rather than uncertainty, although no link to trait anxiety was found. Together, these studies show that apart from the effect of conditioning the SCR response is also related to uncertainty.

Trait anxiety has also been shown to influence several other aspects of

aversive learning, including an increased attention to threat (MacLeod and Mathews 1988), the inability to ignore distraction (Ansari, Derakshan, and Richards 2008; Bishop 2009; Moser, Becker, and Moran 2012), slower task switching (Ansari, Derakshan, and Richards 2008), contingency unawareness (Baas and Heitland 2015) and suboptimal learning adjustment to environmental volatility (Browning et al. 2015). While most studies support the disruptive role of trait anxiety in aversive learning, high trait anxiety has been reported to improve performance in extinction, but only under high cognitive load (Raes, et al. 2009). Likewise, it has been shown to improve perceptual sensitivity and lead to faster reaction times, but only under the threat of shock (Sussman et al. 2016). Lastly, it has been linked to better inhibitory effectiveness, but only under the condition of low mental effort (Edwards, Edwards, and Lyvers 2017). An unpublished study by Wise et al. (in submission) found faster learning from shock omission in high state anxious individuals, linking anxiety to improved safety learning.

Together, these reports suggest a complex role of trait anxiety in aversive learning. While some studies clearly show a negative effect on learning, others report no influence or even a positive effect. One feature that stands out is that the positive effects mostly occur under a specific condition such as high cognitive load or threat, so the improved performance might be due to increased vigilance or ability to cope with stress.

Mogg and colleagues proposed that the seemingly improved inhibition in fact reflects avoidance and a lack of cognitive processing of fear which in turn increases the chance of developing an anxiety disorder (Mogg, Mathews, and Weinman 1987). The authors argue that increased inhibition leads to the suppression of healthy processing of fearful memories which causes behavioural avoidance and persistence of fear. (Weinberg and Hajcak 2011) provide experimental ERP evidence that GAD patients fail to engage in such elaborate processing. In addition to providing an explanation for increased relapse after therapy, this theory would

suggest a divide between physiological response and reported measures of aversive learning.

Studies using explicit expectancy ratings often show a remarkable difference between physiological and explicit measures. A large study (n=377) by (Haaker et al. 2015) found a weak (r=0.13 - 0.19, depending on condition/phase) but robust association between fear ratings and trait anxiety but no relationship between anxiety and conditioned responses in either acquisition or extinction. Similarly, Kindt and Soeter (2014) found a rather modest effect of trait anxiety on explicit ratings but none on the startle response, which raises doubts regarding the role of trait anxiety as a biomarker for the development of clinical anxiety. A recent unpublished study by Atlas et al. (in submission) on instructed learning found that high trait anxious participants report a lower shock probability during extinction compared to the low trait anxious group. However, the same high anxious cohort showed uninhibited anticipatory GSR response during the same trials.

Taken together, although a large body of research suggests a disruptive role of anxiety in aversive learning, there is a remarkable body of opposing evidence. A detailed review of the clinical and trait anxiety literature suggests that a) anxiety does play a role in aversive learning; and b) this role can be both disruptive or beneficial, depending on context, the paradigm used, levels of uncertainty involved and the degree of motivation/incentive. A better characterisation of the precise role of anxiety in learning is crucial for our understanding the development of pain, anxiety and other affective disorders. It is one of the goals of this thesis to develop a computational account of the influence of anxiety on aversive learning.

Neural processing of aversive learning

Research into the neural basis of aversive and particularly fear learning has received a fair amount attention in animal (LeDoux 2012; LeDoux 2000) and human studies (Sevenster, Visser, and D'Hooge 2018; Sehlmeyer et al. 2009). The key structures identified in aversive acquisition are the amygdala (Dunsmoor et al. 2015; LaBar et al. 1998; J. E. LeDoux 2000; Phelps et al. 2004a), hippocampus (Knight et al. 2004), anterior cingulate cortex (ACC; Sehlmeyer et al. 2011) and the insula (Sehlmeyer et al. 2009), while in the extinction it is the amygdala (Phelps et al. 2004), hippocampus, the ACC (Sevenster et al. 2018) and the vmPFC (Schiller, et al. 2008b). Apart from these 'traditional' regions, a meta-analysis of fMRI studies highlighted the role of dlPFC in both acquisition (Fullana et al. 2016) and extinction (Fullana et al. 2018). Same regions were implicated in general aversive processing with the addition of the striatum (Jensen et al. 2003; Delgado et al. 2008; Seymour et al. 2005) and the periaqueductal gray (Roy et al. 2014)

In this section I will first discuss the neural model of aversive learning, followed by the proposed function of the key regions involved in aversive acquisition and extinction.

The Neural Fear Circuit

The prevalent neural model of fear conditioning has been presented by McNally et al. (2011). According to this model the aversive information (shock) is received by the dorsal horn which already holds an estimate of prediction, so it is here where first prediction error is generated. The information is then conveyed to the PAG which has been shown to be involved in fight-or-flight response. The PAG further passes the information to the thalamus which relays it further to the striatum, the amygdala and to

the PFC. The PFC provides higher order processing and communicates it back to the amygdala. At each level of this process an expectation is generated and conveyed back to preceding regions. In support of this model a study by (Mobbs et al. 2007) showed propagation of threat anticipation of signal from the vmPFC into the PAG in situation of threat proximity (McNally, Johansen, and Blair 2011)

Acquisition and Extinction

A review of fMRI studies on aversive acquisition and extinction found similar but dissociable networks involved in the two processes (Sehlemeyer et al. 2009). While acquisition was consistently found to be associated with the activity in anterior cingulate cortex, the amygdala and the insula, extinction additionally involved engagement of prefrontal cortex (although not in all reviewed studies). Further evidence suggests that rather than individual regions, the connectivity between multiple regions is crucial for successful extinction. For example, connectivity between the amygdala and vmPFC predicted the rate of extinction (Feng, Zheng, and Feng 2016). Milad and Quirk (2002) reported context-dependent modulation of fear learning via collaboration of vmPFC, hippocampus and the amygdala. I follow by initially discussing the proposed separate roles of the amygdala, the hippocampus and the ACC/vmPFC.

Amygdala and Dorsal ACC

In number fMRI studies, the amygdala has been linked to both acquisition and extinction (Knight et al. 2004; Hartley and Phelps 2010). Knight and colleagues as well as subsequent studies (Schiller et al. 2008) reported that amygdala seems to be particularly involved in learning just after a reversal, i.e. when attentional resources are

required. This is corroborated by studies suggesting that while the striatum processes Rescorla-Wagner-like learning, the amygdala is involved in tracking the current associability (Li et al. 2011; Roesch et al. 2012). The time-dependence of extinction was recently highlighted by (Morriss, Hoare, and van Reekum 2018). Within the amygdala, the basolateral nucleus (BL) has been identified as a source of fear-related plasticity, i.e. storing of the US-CS associations, while the central nucleus (CE) is thought to be involved in the expression of fear (Dunsmoor et al. 2015).

While physiological studies in animals and most early fMRI studies in humans (Knight et al. 2004; Phelps et al. 2004; LaBar et al. 1998) support the role of the amygdala in both processes, two recent meta-analyses (Fullana et al. 2016 and 2018) failed to support a decisive role of the region in aversive conditioning while other regions, such as the insula and the dorsal ACC (dACC) were found to consistently activate during both acquisition and extinction. The role of dACC is of particular interest as it is highly connected with the amygdala, vmPFC as well as the hippocampus. Additionally, a study investigating the use of MPVA in single-trial analysis only found a trial-to-trial learning signature in the superior frontal gyrus, including the dlPFC.

The role dACC had been consistently shown to play a role in both phases of aversive learning (Sevenster et al. 2018). Increase in dACC activity has been linked to extinction recall (Marin et al. 2017), however, in trait anxiety, Sehlmeyer et al (2011) observed a decrease in dACC activity during extinction. Number of studies have suggested that rather that the activity within the amygdala, it is the connectivity between the two regions that drives aversive learning (Sevenster et al. 2018). Fang et al. (2016) reported increased functional connectivity between the regions prior and post-conditioning. Additionally, the dACC-amygdala network was altered in PTSD patients (Brown et al. 2014).

Both, the amygdala and the dACC seem to play a central role in aversive learning. While the exact role of each of the regions is not entirely clear, animal studies in conjunction with fMRI studies suggests that they both track the aversive value

(Phelps et al. 2004; Delgado et al. 2008) and extinction success seems related to their connectivity.

Hippocampus

The hippocampus has previously been linked to both acquisition and extinction (Knight et al 2004; Hermann, Keck, and Stark 2014). FMRI studies examining the context-dependent recall of extinction in humans point to an important function for the hippocampus (Kalisch et al, 2006; Milad et al, 2007). In such studies, context is manipulated through changes in the background color or the visual scene in which the CS is presented. Extinction learning takes place only in one of the contexts, allowing activation correlated with context-dependent retrieval to be examined. Increased hippocampal activation was observed during extinction retrieval (Milad et al. 2007; Kalisch et al. 2006). Inactivation of the hippocampus before extinction learning impairs extinction recall on the subsequent day (Corcoran et al. 2005), suggesting that the hippocampus regulates fear expression both outside and within the extinction context.

Furthermore, activation in the hippocampus was positively correlated with vmPFC activation, providing further support for the notion that the hippocampus may mediate context-dependent extinction recall through connections with the vmPFC. Finally, confirming the function of the hippocampus in contextual reinstatement, a human lesion study found that contextual reinstatement of the CR was impaired in individuals with hippocampal lesions (LaBar and Phelps 2005), similar to the findings observed in rodents (Wilson et al, 1995).

Recent study in structural connectivity reported association between increased cingulate-hippocampal connectivity and GSR-indexed rate of reinstatement (Hermann et al. 2017). This suggest that people with more context-dependent extinction will suffer from higher rates of extinction relapse.

Lastly, trait anxiety has been associated with decreased connectivity between

hippocampus and vPFC (Indovina et al. 2011) suggesting a potential mechanism for influence of trait anxiety on lack of extinction.

vmPFC

The vmPFC has been specifically implicated as the primary region driving extinction by inhibition of the amygdala. Schiller et al. (2008) performed a reversal learning study in which they directly compared CS- with an CS+ after reversal to safety (i.e. CS-+). The authors found that the vmPFC was activated during the reversal period specifically in response to a previously harmful stimulus. This finding corroborates previous results of animals studies in which the infralimbic region (IL), animal counterpart to vmPFC/ACC, was found as the site of extinction consolidation (Quirk et al. 2000). Despite a wealth of evidence supporting the role of vmPFC in extinction, the exact nature of the signal is still debated. A frequent view is that the vmPFC might provide an inhibitory signal while new CS-US pairings are being learned by the amygdala. In the theory of extinction provided by Bouton (2002) the amygdala learns new US-CS pairings while the vmPFC inhibits and/or selects previous associations. Alternatively, this extinction-specific signal could also be a safety signal (or reward) and not be directly involved in inhibition. A recent fMRI study reported that the vmPFC signal in extinction was associated with enhancement of dynamic learning when shock omission was replaced by a neutral stimulus (Dunsmoor et al. 2019).

It might me important to consider the functional role of the vmPFC outside of the extinction literature. Attempts have been made to generalize the role of the vmPFC and the adjacent orbital-frontal cortex. A particularly interesting avenue is the proposed role of the vmPFC in representing the expected value of the current state (Chib et al. 2009; Levy and Glimcher 2012). Under this view, the vmPFC represents the brain's internal currency associated with the current expectation, integrating over range of contexts, actions and other influences to inform decision-making. The medial orbital gyrus has recently been shown to uniquely represent the the current state itself

(Schuck et al. 2016).

Linking aversive learning and involvement of the region in extinction can be crucial to a more general understanding of its role in the process. If the computational function of the OFC/vmPFC is to represent the current state then it should signal different states in acquisition and extinction. This processing should also only occur in people that in fact consider multiple states in the environment, participants tracking the value of one state should exhibit no difference between acquisition and extinction in the region. Furthermore, under this general functional view of the region, it is possible that the CS-US inhibition or safety signal found in previous studies is in fact a state signal.

Anxiety

The effect of trait anxiety on neural processing of aversive conditioning has been reported by a number of studies (Indovina et al. 2011; Bishop et al. 2007; Sehlmeyer et al. 2011; Marin et al., 2017). In particular, high trait anxiety has been linked to increased reactivity to fear cues in the amygdala and decreased inhibitory signal in the medial PFC (Indovina et al. 2011). Similarly, Barrett and Armony (2009) found that the amygdala response to CS+>CS- contrast scaled with trait anxiety. Sehlmeyer et al. (2011) reported sustained amygdala activation during conditioning and reduced dorsal ACC activity during extinction. These reports indicate neural regions associated with trait anxiety in aversive learning. Fear extinction was associated with increased BOLD activity in the vmPFC and MCC (Belleau et al. 2018).

In summary, multiple regions have been shown to be involved in the acquisition and extinction learning. In particular, the amygdala seems to play a role in both processes, although its contribution might be most dominant only post-reversal. The vmPFC is traditionally associated with either inhibitory control during extinction or

safety signalling. As alternative, the vmPFC, together with OFC, might equally likely be responsible for the representation of the current state and expected value. Activity in both regions is modulated by trait anxiety. The hippocampus has been consistently associated with providing contextual information. However, it seems to be involved in more than just that, it has also been shown to be active during initial learning. Recent work has highlighted the role of dorsal ACC in aversive learning. The region can similarly to amygdala, dissociate between CS+ and CS- but the signal appears to be more permanent. Additionally, the PAG is involved in early threat processing and may play an important role in probabilistic aversive learning.

Information Processing Approach to Learning

Computational modelling

Traditionally, models in psychology and neuroscience are specified using a verbal and graphical depiction of how a system or a given psychological phenomenon operates. A classical psychological model would for example be the Levels of Processing theory of memory recall by Craik and Lockhart (1972). The model proposes that the *depth* of encoding predicts the subsequent success of recall. For example, the model predicts that words which were presented to participants without a particular instruction will be remembered less than words participants were asked to read or memorise. This model makes predictions about recall based on qualitative descriptions of categories and criteria. Those can be interpreted differently by different researchers (E.g., what constitutes *depth* of processing), giving rise to ambiguity. As a consequence, such theories may be difficult to to prove correct or refute. An alternative to this approach is to define psychological models as a set of mathematical relationships. This method is familiar to those coming from sciences such as physics or engineering. By using explicit mathematical formulation, a model can be made to

predict unobserved data and can thus either be refuted or validated, without the ambiguity inherent to verbal or symbolic description. That being said, computational models suffer from their own ‘ambiguity’ that is related to the way they are used: the fitting method, the internal model dynamics or number of degrees of freedom can prioritize models that in fact aren’t accurate. This will be discussed in a later chapter.

Over the past 20 years, research in cognitive neuroscience has gradually started to adopt mathematical modelling. While previously, a typical experiment would average data by condition, process models allow researchers to understand the more fine-grained spatio-temporal dynamics of the neural and electrophysiological signals recorded. Additionally, multiple hypotheses formalised as models can be compared in terms of how well they explain the experimental data. This, in turn, allows us to make inferences about the cognitive, physiological or pharmacological processes that generated the data measured. Critically, the individual is seen as one biological system that is probed by presenting a stimulus and produces a given response. While previously, we would look at how the output changes as a function of the input, considering everything in between a black box (at least to some degree), computational models allow us to test hypotheses related to the course the information took between input and output. In an ideal setting, we would model the photon leaving a computer screen, the retina translating the message into the neural code, visual cortex reconstructs the image, associative cortex relating the input to our memories, amygdala calculating the aversive value of the stimulus, causing an increased demand for blood supply which we can indirectly measure by realigning hydrogen molecules and measuring their relaxation time, forming a k-space image, denoising the data via ten pre-processing steps, and using a general linear model to find out which of the brain signals is linked to the original stimulus. Additionally, we could also collect electrophysiological data with their own biophysical models, and incorporate them into our model of the process. To say that we have mastered any of these steps would be an outright arrogance. However, we are able to approximate this

process to a degree where we can dissociate between what information is most likely encoded by the amygdala. The methodology employed in this thesis will be discussed in detail in the Methods chapter. In short, the models used and compared here only formalize small part of the process described above, namely by which *process* the human brain learn about aversive associations. For the rest, (e.g., the estimation of how hydrogen atom relaxation relates to blood oxygenation), I will rely on methods commonly used in cognitive neuroscience.

Here, I present a short summary and motivation for models used throughout this thesis.

Rescorla-Wagner model

Rescorla and Wagner (1972, later RW) were the first to provide a computational account of associative learning that has received considerable attention and has been used widely since. The model has been applied to the study of reward learning (O'Doherty et al. 2004), social processing (Lockwood et al. 2018), pain (Roy et al. 2014) and others, including aversive learning (Schiller, Levy, et al. 2008; Delgado et al. 2008). The model assumes that an agent (i.e. the participant) keeps track of the US-CS association, that is, the likelihood that CS will be followed by the US, and updates it on an event-by-event basis based on the feedback received. When this model is fitted to the data, each participant is assumed to incorporate the new information at a different pace which is reflected in the individual learning rate (α).

Mathematically, the model tracks the value, Q, updating it as new outcomes, O become available, weighting the error by the learning rate parameter α, across all trials t.

$$(Eq.I) \qquad Q_{t+1} = Q_t + \alpha(O_t - Q_t)$$

Pearce-Hall model

While the RW model provides a useful theoretical tool it fails to provide an account of many experimental phenomena such as spontaneous recovery or context renewal. Perhaps the most successful extension of the RW model is the Pearce-Hall model (Pearce and Hall, 1980; later PH) which incorporates attentional mechanisms into the learning process. In the RW model, the learning rate parameter is assumed to be constant across the experiment. However, this seems unlikely as learning can, for instance, become slower over time. The PH model assumes that on each given trial the rate of learning depends on the current level of surprise. For instance, if the US is consistently omitted following the presentation of the CS, the prediction grows nearer to the received outcome (i.e., low US expectancy) to minimise the level of surprise and decelerate learning. The Pearce-Hall model predicts highest learning rates just after a change in US-CS contingencies which makes it a suitable candidate to model adaptive behaviour.

Building on Eq. I, the learning rate α now changes on each trial as a function of the prediction error. The parameter η controls the level of adaptability and the term α_t corresponds to current level of associability. Participants with high value of η will be more sensitive to changes in environmental contingencies, low values of η, such as when $\eta = 0$ the PH model will be equivalent to RW.

$$(Eq.II) \qquad \alpha_{t+1} = (1-\eta)\alpha_t + \eta|O_t - Q_t|$$

Theoretical models can be amended or combined based on a particular research question. For example, Li et al. (2011) used a version of the Pearce-Hall model that accounts for the degree of adaptability by weighing RW-like and PH-like learning. Similarly, Piray et al. (2018) combined PH and RW models with additional condition-specific weights (fear / happiness) to identify a condition-specific effect of

social anxiety on adaptive behaviour. In my context, an interesting model would be a RW model with different learning rate in acquisition and extinction.

Bayesian models of learning under uncertainty

While the above models provide a good description of the observed behaviour they do not consider the degree of uncertainty in the system. An alternative approach to modelling of the learning process stems from the Bayesian Decision Theory (Berger 1985). Behrens et al. (2007) developed a model of learning under uncertainty that tries to optimise reward predictions based on the level of current uncertainty and environmental volatility. The authors show that a region in the anterior cingulate cortex correlates with the expected reward predicted by the model, providing a biological basis for their model. Inspired by the Behrens model, (Mathys et al. 2011) have re-formulated learning under uncertainty as a series of Gaussian filters. Their model keeps track of perceptual uncertainty, contingency uncertainty and environmental uncertainty (volatility). It has been successfully applied to understanding of stress (de Berker et al. 2016) or autism (Lawson, Mathys, and Rees 2017). Importantly, despite coming from different theoretical backgrounds, the Bayesian models produce similar learning rate estimates as the Pearce-Hall model (Gershman et al. 2015).

State learning

While both Bayesian models and Pearce-Hall discussed above assume trial-by-trial updating of beliefs, neither fully explains the phenomena of reinstatement and context renewal discussed above. Both increase the update rate once the environment (i.e. shock probability) changes. However, in a number of animal and human experiments these changes often occur faster than predicted by the model, sometimes as quickly as after 1-2 trials (Bouton, 2004), suggesting that the learning process is no longer gradual but rather very abrupt. An intuitive explanation for the context-renewal

phenomenon may be a good lead in understanding what may have happened in the agent's mind. After a period of conditioning ("acquisition 1") the CS is no longer paired with US ("extinction 1"). As discussed above, Bouton (2002) suggested that extinction involves novel learning as well as inhibition of the previous memories. This account is supported by recent neuroimaging studies (for example Schiller et al. 2008b). In other words, the agent remembers that there is a state in which the same cue is predictive of aversive outcome. Context is the most apparent representation of such state: when context changes back to A in ABA renewal, the agent is cued to believe that a cue becomes predictive of aversion. In this example the state is explicit, i.e. directly observable. Importantly, a state can also be implicit and hidden (or partially hidden) to external observers. Such state could be defined by the likelihood of shock, environmental uncertainty or feeling of hunger for the matter. Coming back to our example, even when context is not explicitly changed and suddenly several USs occur in the absence of CS this might be interpreted by the agent as return of the original state. The idea of partially hidden states is an intuitive and potentially powerful theory to explain phenomena such as context renewal and reinstatement.

Redish et al. (2007) formulated a computational model in which the agent initially holds one state of the world and learns about cues in it using a reinforcement learning algorithm. Additionally, on each trial the agent decides whether there is enough evidence that it now finds itself in a new state and if so, a new internal state is created. The Redish et al. model uses a sophisticated state discovery mechanism without a biological interpretation but has been reformulated and simplified under Bayesian Decision Theory by Gershman, Blei, and Niv (2010) to provide a normative account of state switching behaviour.

Importantly, an agent that uses a state switching algorithm does not undergo extinction of fear memories. Similarly as proposed by Bouton (2002), extinction itself constitutes a different context. This has crucial implications for clinical practice. If a patient leaves the therapist's office believing that extinction is only specific to a

particular state they will be more likely to relapse (Craske, 2014). A multi-state representation of the learning environment can therefore be indicative of increased likelihood of fear renewal. This is supported by a number of experiments in which shock contingency was decreased gradually with occasional reinforcement that lead to improved extinction (Bouton and King 1986; Gershman et al. 2013); Craske et al. 2014). Under the state representation theory, this would mean that an agent has never created a second state and has rather overridden the associative value of the original state. This is an example of where algorithmic understanding of the learning process is crucial in designing optimal treatment strategies.

Use of computational models in aversive learning

While the majority of research involving computational models focuses on the domain of reward, the domain of aversive learning has also seen increased interest in the use of the method. Schiller et al. (2008) used a version of the Pearce-Hall model with decaying learning rates to understand how learning differs between conditions of safety, fear and post-fear safety. The authors reported faster learning in acquisition than after reversal (extinction), as indexed by fitted learning rates. A remaining question is whether the lack of extinction is merely a consequence of slower learning or whether slower learning is a consequence of another, a more fundamental, process. A recent study by Jepma et al. (2018) employed both reinforcement learning and Bayesian modelling to propose a self-reinforcing mechanism behind resistance to change in beliefs. The authors show that a bias in learning is reinforcing a bias in perception and vice versa which leads to a self-reinforcing effect. This work is a good example of how models can be used to discriminate between perceptual and learning components of information processing during learning. Unpublished work by Wise et al. (2019) used a so-called leaky beta model to investigate the effect of visual attention on aversive learning. The employed model is not based on reinforcement learning nor

does it have a proposed neural mechanisms. Nevertheless, it seems to explain the observed behaviour exceptionally well with very little free parameters. The authors build on the parametrization of the beta distribution which is defined by two parameters α and β. When either of the parameters is much higher than the other this will result in a skewed distribution. In the pain learning context this translates to having a high prediction of shock when many more shocks (i.e. high α) than no-shocks (β) were observed.

Links between information content and physiological processes have been drawn. Li et al. (2011) reported a correlation between a conditioned response (measured by GSR) and associability-dependent learning rate. The degree of associability is highest whenever highest errors are generated on the previous trial. This corresponds to the degree of uncertainty. (Tzovara, Korn, and Bach 2018) reported the encoding of both uncertainty in the outcome expectation in the anticipatory GSR. In the same paper, pupil dilation has been linked to outcome expectation while a number of other studies reported association between pupil dilation and uncertainty (Nassar et al. 2012; Koenig, Uengoer, and Lachnit 2018). Browning et al. (2015) reported association between pupil dilation and higher order uncertainty in a subsample of low anxious participants.

Computational modelling has been successfully applied to the study of aversive information processing in the brain. For example, Seymour et al. (2007) used a temporal-difference learning model in conjunction with fMRI to identify distinct regions in the ventral striatum involved in processing of gains and losses. Delgado et al. (2008) compared prediction error processing elicited by a primary (i.e. electrical shock) versus a secondary (monetary loss) punishers identifying the ventral striatum as a region processing both.

The methodology has successfully been applied to the study of pharmacological action (Pulcu et al. 2019; Bach et al. 2018). Pulcu and colleagues

investigated the role of losartan, a hypotensive agent, in appetitive and aversive instrumental learning. The authors compared appetitive and aversive learning rates on and off the drug concluding that it reduces aversive but not appetitive learning.

The above are merely a few examples of the use of computational modelling in the study of aversive learning. The relevant models will be discussed more thoroughly in Chapter 2: Methods.

Normative Aversive Learning

Evolutionary versus Normative Optimality

In the introduction I briefly touched on the topic of what constitutes optimal aversive learning. Here I will expand on the topic discussing the two main lines of reasoning.

It could be argued that an agent learns optimally if its predictions tightly reflect previous observations. In a changing environment where some seasons are marked by high predatory presence while others are relatively safe, learning the current probability of threat is crucial for survival. When the chance of a predator present is low, the animal can use the opportunity to gather resources that will ensure its survival over the winter. On the other hand, it is crucial to avoid being killed and so even environments which sometimes include a high chance of predator should be avoided. Instead of keeping track of seasonal changes in predatory frequency the energy should instead be preserved and perhaps other environments should be explored. Those two approaches provide competing predictions to what constitutes an optimal aversive behaviour. The former, termed here prediction optimality, dictates to accurately track the environment and be smart as to when to collect food whereas the latter, termed

here survival optimality, mandates to simply mark the current environment as threatening and avoid investing further energy on deriving accurate predictions.

An animal following the prediction optimality will likely get more food but one single inaccurate prediction would be fatal. Its neighbour following survival optimality might not get eaten by a predator and find a predator-free environment instead, however, they might also starve to death. Which strategy is chosen depends on the tradeoff between maximising reward and minimising punishment as well as effort. A choice of strategy might be determined by a particular environmental dynamics or individual differences. For example, some sets of environments might simply be too hard to navigate, so a species might develop in a way that they hibernate for months or even years at a time. Under such scenario the energy expenditure and threat are low but so is the reward. Cicadas of the the Magicicada genus have developed a strategy that takes such scenario into an extreme. Its nymphs will lay dormant for 13 or 17 years before all hatching in one season. In another environment predatory presence is highly predictable, for example due to a migratory pattern, and so it is optimal to monitor changes and detect the point when the predators have left. This would require higher energy expenditure and higher chance of being wrong (i.e. coming out too early or staying out too long) but it might yield a higher reward.

Moving to humans, in a classical scenario of individuals suffering with chronic pain, patients will often initially track the current likelihood of pain but over time they generalize pain to all situations and contexts which in turns leads to social isolation, decreased mobility and mental health issues. Under this scenario, a gradual shift in strategy has occurred where the initial exploratory behaviour shifted to a habitual one. In this case, it can be argued that such adaptation leads to overall negative consequences. The benefit of being active and meeting friends bears less weight than the fear of return of pain. An important component here is the temporal change in prediction: while patients start as prediction optimisers they gradually become overpredictors, i.e. survival optimisers. Arntz (1991) argued that this is a consequence

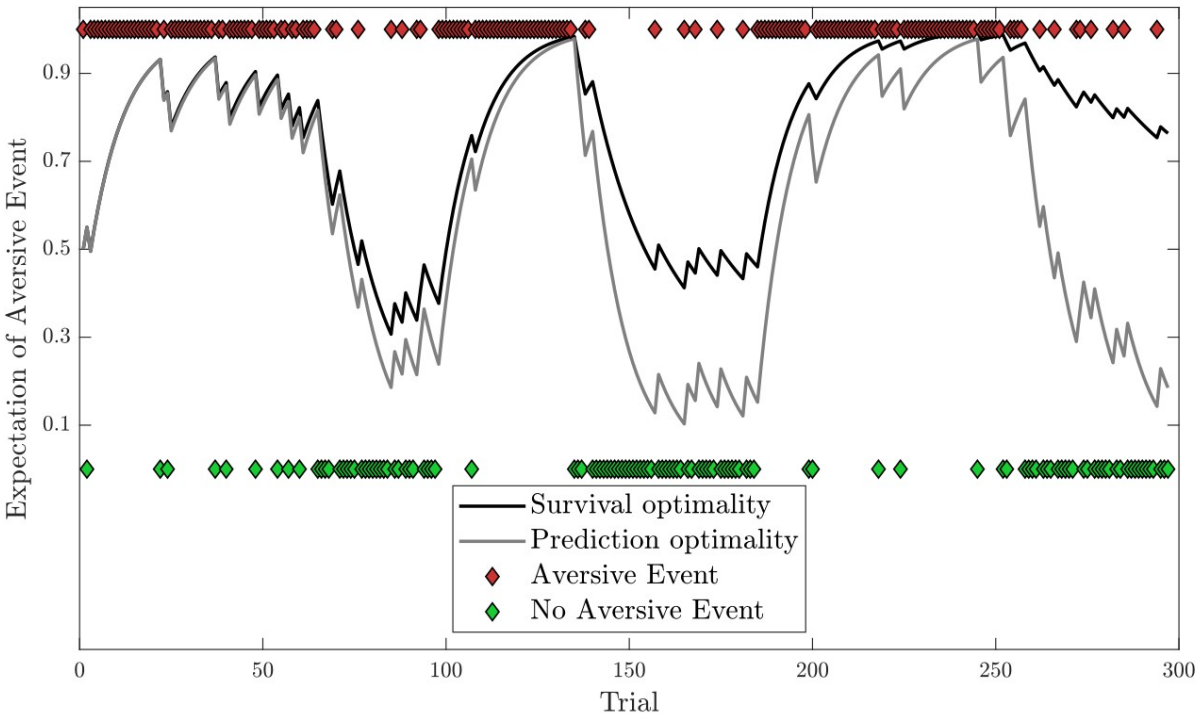
of maladaptation. The author suggests that overpredicting pain is evolutionarily optimal because while threat overprediction leads to mild aversive outcomes, threat underprediction can be fatal. However, in the context of chronic pain this transfer is maladaptive because aversion overprediction will lead to social isolation and can impact mental health which has been linked to worsening of chronic pain symptoms.

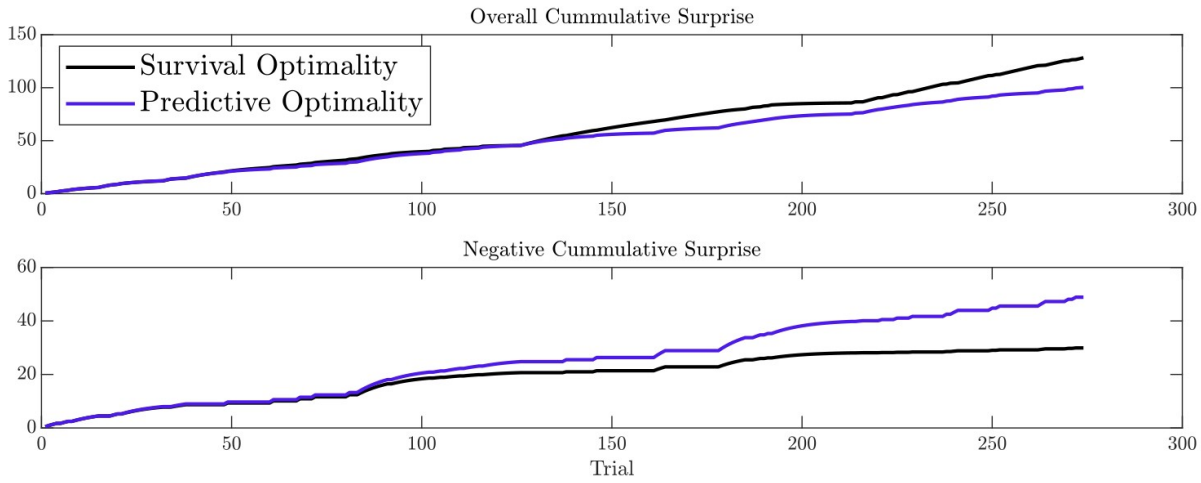
Similar evidence was presented by LeDoux (2012). The fear learning module is dominated by subcortical and limbic structures, making it 'quick and dirty', and more prone to false negatives rather than false positives. This asymmetry between learning from aversive and omission of aversive event is crucial for understanding mechanisms of acquisition and extinction. The evolutionary view would suggest that lack of extinction is in fact an optimal strategy because most threats will be predicted. This view can be reformulated using the concept of surprise from Shannon Information Theory (Shannon 1948) as minimisation of negative surprise. If minimising future negative surprise is the goal of an optimal agent then aversion should nearly always be predicted. An orthodox evolutionary optimiser would never come from shelter and therefore die. A more likely scenario of evolutionary optimisation is a gradual transfer from prediction optimiser to evolutionary optimiser where over longer periods of time the lack of extinction should become progressively more pronounced. This is similar to the case of chronic pain patients who gradually begin to overpredict the probability of pain as described earlier. If an agent isn't susceptible to such adaptation and follows prediction optimality, i.e. the goal is to minimise an overall error in prediction, then the optimal agent should always undergo full extinction. Under this prediction optimality, lack of extinction is a sub-optimal feature that should be minimised.

Shannon surprise is calculated using prediction P and an outcome O. In case of binary outcomes where "aversive event" is coded as 1 and "no event" is coded as 0, Shannon surprise s on a trial t is given by Eq. III. Note that this measure of surprise will correlate with unsigned prediction error. The main difference is that large prediction errors will be enlarged more than small or moderate ones.

$$(Eq.III) \qquad s_t = -log|O_t - (1 - P_t)|$$

Figure 1 *shows expectancy of aversive event as a function of trial. Predictions of two computational models are presented. Both start as prediction optimisers but one learns to follow the survival optimality (black) while the other continues to follow the prediction optimality (grey). In other words, the earlier model tries to minimise the negative surprise while the latter one minimises the overall surprise. The two panels below show how much cumulative surprise has been experienced by either agent. While survival model (blue) experienced more overall surprise (top panel) it experienced significantly less negative surprise overall (bottom panel).*





Under this view, lack of extinction can either be a consequence of suboptimal predictive optimisation learning or optimal survival optimality learning.

States and Contexts

Another issue to consider when discussing learning optimality is the concept of state learning discussed in the previous section. In order to increase the accuracy of prediction an agent might associate events, internal and external states with contexts. This model-based learning strategy leads to accurate predictions but state space might become intractable and computationally expensive. Additionally, as discussed previously, state-like representation of learning might lead to higher rates of relapse. Reconciling these ideas with the two optimalities discussed here, state vs gradual learning should be considered an additional factor where both can be either survival or prediction optimal. A survival optimal state learner would assign higher probability to a state in which aversive event is more likely, eventually becoming resistant to switch to a low-aversion state while predictive optimiser would learn to switch between the two states symmetrically.

Having discussed what might be considered optimal aversive learning, in most

clinical situations extinction is still a desirable process and therefore normative learning is preferred. It is, however, important to consider the eventuality that patients might engage in survival optimality behaviour.

Thesis Goals

The presented research raises number of fundamental questions about the nature of aversive learning. Despite large volume of studies, a complete theory of computational and neural mechanisms that underlie acquisition and extinction is lacking. While most evidence suggests that extinction involves some additional components, the nature of these is debated. Inhibition of previous US-CS pairings in extinction has been highlighted as the main difference between the phases. This mechanism likely operates in addition to general tendency to over-predict the likelihood of aversive events, making extinction slower and incomplete. Another contributing factor is the uncertainty caused by partial reinforcement which has also been shown to decrease the rate of extinction. Furthermore, temporal factors have been shown to play an important role. In untreated patients, lack of extinction increases over time, however, in experimental studies participants learned to dissociate between the two phases over time and rapidly changed their anticipation when an environmental change has occurred. This suggests that in at least some settings people engage in structure learning while elsewhere, especially when facing large aversive prospects, they settle on overprediction and stop undergoing extinction. One potential explanation could be that in some environments it is beneficial to follow the current contingencies while in others it is better to save energy and settle on overprediction. Additionally, it is possible that there is a differential tendency of some individuals to learn the structure of the environment and represent it as different states. For example, in number of behavioural studies high anxious individuals were shown to suffer from increased rates

of post-extinction relapse. It has been proposed that this is due to increased context sensitivity. The idea of multi-state representation is an interesting conjunction between fundamental computational problem optimisation (knowing the structure saves energy) and behavioural effects reported in extinction research (relapse, context-specificity). While the idea has been discussed (Gershman, Blei, and Niv 2010) it hasn't been directly tested.

In this thesis I will use probabilistic learning to first characterise the differences between acquisition and extinction at a behavioural, neural and computational level. I will then investigate how the difference changes over time and as a function of reinforcement contingencies. Next, I will test the idea that gradual learning leads to structure learning and investigate whether there is an association with anxiety. I will also test whether it is possible to use drug enhancers to increase the rate and permanence of extinction.

With the previous research in mind and realistic goals in vision, as a part of my doctoral thesis I set to perform a series of three studies aimed at answering the following seven questions:

1) What is the behavioural difference between acquisition and extinction in probabilistic aversive learning and how do these change over time?
2) What role does trait anxiety play in probabilistic aversive learning?
3) Is there an association between individual differences such as trait anxiety and increased tendency for state rather than gradual learning?
4) What algorithms best explain the experimental learning data?
5) What brain regions are involved in aversive acquisition and extinction learning? And where in the brain does anxiety influence the process?
6) What is the influence of environmental statistics on learning? Do clearer contingencies improve extinction?
7) How does the angiotensin-II antagonist losartan influence aversive probabilistic

learning and extinction in particular?

I will now briefly discuss motivations, a priori hypotheses and considerations related to each of the questions:

1) How do acquisition and extinction in probabilistic aversive learning differ with respect to behaviour and peripheral-physiological parameters?

The fundamental goal of this thesis is to improve our understanding of how the lack of extinction arises as a consequence of a Pavlovian learning process. In order to answer this question I will first characterise the difference considering behavioural and physiological measures, for example by calculating learning rates, switch points, mean reported probabilities or skin conductance response. Extinction is hypothesised to be slower and take longer than acquisition. A time-dependent factor is predicted to influence the rate of extinction, due to convergence to evolutionary optimality. Additionally, lack of extinction is hypothesised to be present at the conscious and autonomic level.

2) What role does trait anxiety play in probabilistic aversive learning?

Trait anxiety has been proposed as a risk factor in development of clinical anxiety. Building on findings from Q1, I will study what role does trait anxiety play in aversive learning. I hypothesise that anxiety will play a disruptive role, particularly, that it will lead to longer and slower extinction. I expect anxiety to influence the subjective ratings and galvanic skin response levels during extinction.

3) Is there an association between individual differences such as trait anxiety and increased tendency for state, rather than gradual, learning?

Using computational models I will test whether some individuals learn in a gradual manner while others tend to learn the structure of the environment. I expect some individuals to represent the environment as multiple states. If state learning does occur, I will use the state prediction error (SPE) and state identity data to test where in the brain are states represented. A priori, based on the previous studies I would expect this to occur in the vmPFC or the OFC. However, SPE's were also linked to lateral PFC and inferior parietal lobule (Glaescher et al. 2010).

4) What algorithms best explain the experimental learning data?

Following from questions 1 – 3, I will use computational models to investigate how the differences arise as a consequence of learning. For example, is lack of extinction solely a consequence of faster learning from shocks than from shock omissions or is there an additional process at play? Computational models will be fitted to behavioural and physiological data to investigate any potential discrepancies between information processing at the different levels. A question of particular interest is whether people tend to increase their learning rate after a change in contingency, or whether they instead maintain an alternative state in memory and therefore switch to a new state abruptly.

5) What brain regions are involved in aversive acquisition and extinction learning? And where in the brain does anxiety influence the process?

I will use functional MRI to identify regions specific to acquisition and extinction learning. Based on the literature review, both processes are hypothesised to engage the amygdala, hippocampus, dACC and the insula. Additionally, inhibitory signal in extinction is hypothesised to involve ventro-medial prefontal cortex (vmPFC). Level of vmPFC engagement is hypothesised to be modulated by level of individual trait anxiety. Following recent work on aversive learning, the role of periaqueductal gray in acquisition, extinction and prediction error processing will be investigated.

6) What is the influence of environmental statistics on learning? Do more distinct contingency changes improve extinction?

Following from previous work on gradual extinction, I will test whether large versus small changes in pain probability encourage particular learning strategy (state switching versus gradual learning). It is hypothesised that large jumps in contingencies will encourage participants to form a multi-state (i.e. multi context) representation of the environment while environments with smaller contingency changes will encourage gradual Rescorla-Wagner-like learning. Consequently, if gradual extinction does occur, smaller changes in contingencies will lead to complete and more permanent extinction.

7) How does the angiotensin-II antagonist losartan influence aversive probabilistic learning?

Pharmacological agents were previously used as enhancers of Cognitive Behavioural Therapy (CBT). Using the Classical conditioning paradigm from from previous sections I will investigate the role of the angiotensin-II receptor antagonist losartan on aversive acquisition and extinction. Several recent studies found an effect on aversive learning, particularly aversive prediction errors. Here I will use computational modelling to identify which part of the learning process is affected by the drug and discuss how this can be used to enhance CBT.

Plan Overview

To address the seven research questions outlined in the previous section, three experimental studies were designed. Study 1, presented in Chapter 3, will use fMRI and GSR to characterise behavioural, physiological and neural processes involved in acquisition and extinction. It is followed by Study 2, presented in Chapter 4, which uses a behavioural setting to investigate the role of environmental contingencies on learning

strategies, and Study 3, presented in Chapter 5, which will use a double blind placebo design to examine the influence of losartan on aversive learning. Since there is a non-marginal overlap in methodology across the three studies, the methods used throughout the projects will be presented in Chapter 2. Study-specific methods will be discussed in each relevant chapter.

Chapter 2

Methods

This chapter outlines the methodology used for data collection and analysis in three experimental studies included in this thesis. All three projects used the same paradigm and to some degree behavioural and physiological measures. The modelling approach differs slightly between studies depending on the particular research focus, but the general methodology is largely similar. Study-specific methods will be presented in each relevant chapter. This section also contains information about the studied sample. Table 1 presents an overview of the methods used in the three studies.

	Study 1: Learning Asymmetry	**Study 2: State Learning**	**Study 3: Losartan and Learning**
fMRI	**YES**	**NO**	**NO**
GSR	**YES**	**YES**	**NO**
Pupillometry	**NO**	**NO**	**YES**
Probabilistic Learning Task	**YES**	**YES**	**YES**

Table 1: *Methods used in the different studies.*

Behavioural Paradigm

The goal of the paradigm is to investigate how participants learn to predict the likelihood of an aversive event and how they update their expectations of the aversive event to occur on a trial-to-trial basis. To this end, I used a Pavlovian probabilistic learning task in which participants learned to associate three visual cues (abstract images) with the delivery or omission of a painful electrical stimulus (shock). On each trial, participants were presented with one of the cues which could be followed by the electrical stimulation. Throughout the experiment, one of the cues was followed by a shock in 75% of trials while no stimulus was applied in the remaining 25% of trials (‘harmful cue’). For the second cue, contingencies were reversed, i.e., the electrical stimulus was applied in 25% of trials and omitted in the remaining 75% (‘safe cue’). For

the third cue, the probability of stimulus delivery switched - unbeknownst to the participant - between 75% and 25% with switches occurring every 30 (+/- 3) trials in study 1 (‘reversal cue’). Participants are instructed to pay attention to all three cues as the probability that they may be followed by a shock may change at any point. The reversal cue is presented more frequently than the two stable cues (see Table II) Outcome measures were (changes in) shock probability ratings which are provided by the participant for each trial following cue presentation, galvanic skin response recorded throughout the experiment 1 and 2, pupil dilation recorded in experiment 3, and functional MRI data recorded in experiment 1.

Task

Trial Structure

Before the start of the task (in study 1 just once in the beginning, on study 2 for each session and for study three on each visit) three abstract fractal cues were randomly selected from a pool of 20 and randomly assigned to the three conditions. Where task was demonstrated to subjects or where there was a training sessions different set of cues were used for the actual task. On each trial one of the three visual cues is presented on a screen (either in the scanner or in a behavioural lab) to which participants were asked to indicate how likely this cue will be followed by a shock on the present trial. Underneath the cue there was a slider with value starting at a random position. Shock expectancy ratings were provided using a keyboard to move a slider on a 0% to 100% scale (0% = “no chance of pain”; 100% = “certain pain”) and participants submitted their answer by pressing a confirmation button with the index finger of their right hand. Once the participant has submitted their answer, the painful stimulation is either delivered or omitted after an inter-stimulus interval, 2-4 s in study 1

(different for each study). 2 seconds after stimulus delivery or omission, the visual cue is replaced by a fixation cross that is presented in the centre of the screen. Figure 1 presents a schematic of a typical trial from the fMRI study 1.

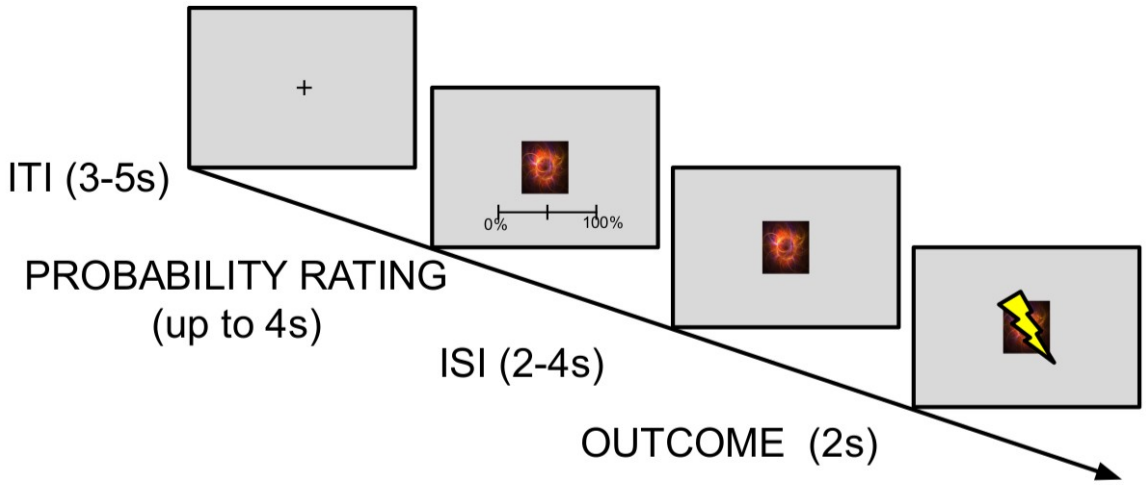


Figure 1: *Trial Structure*

The overall experiment is split into four sessions to allow for brief breaks. Parameters for studies 2 and 3 differ from those of Study 1 with respect to the overall number of trials, reversal frequency, number of sessions (block of trials between breaks), number of phases (acquisition/extinction), number of trials per cue, shock probabilities, length of inter-trial-interval and inter-stimulus-interval to accommodate the respective requirements of the study (for details see Chapters 3-5). An overview of the study-specific parameters is provided in Table II.

Task Structure

The task consists of several hundred trials (variable between studies, see Table II) during which the three cues are presented in a random order. Schedules were randomized across participants. In each schedule, it was ensured that all three trial types (i.e., harmful, safe and reversal) and outcomes (shock/no-shock) are spread

evenly across the experiment. This was done by first generating a 'schedule map' saying what phase types and in which order will follow. Then each phase was generated separately and checked for correct frequencies before adding it to the current candidate schedule. The reversal cue was presented on 40% of trials, the safe and harmful were both shown on 30% of trials. This ensured that within each sequence of trials during which no contingency switch occurred, participants on average completed twelve reversal trials, nine harmful trials and nine safe trials. To ensure that reinforced and non-reinforced trials did not cluster, the following criteria were considered during generation of appropriate task schedules: 1) For a given cue the actual reinforcement rate should not deviate by more than 5% from the underlying probability in each phase, that is, for example the harmful cue (shock probability: 75%) had to be reinforced on 70 - 80% of trials within each phase; 2) The overall reinforcement rate of should not deviate by more than 5% from the underlying probability. For the stable cues, this meant the same as in 1 except across all of the phases, for the reversal cue this meant that the overall rate of reinforcement lied between 45 and 55%. 3) For the reversal cue, at least three of the first five trials following reversal had to be of the new common outcome, i.e., after a switch from high to low shock probability, at least three no-shock events occured in the first five trials.

A total of 10 000 candidate schedules was generated. A Rescorla-Wagner model with $\alpha = 0.15$ was then used to assess the schedule: model predictions were generated by the model using the proposed set of outcomes and a mean of the predicted values was taken within each phase. If it lied within +/-5% of the underlying probability it was accepted. Finally, out of the schedules matching the above criteria, eight were selected at random, four starting with high shock probability and four starting with low shock probability. For each participant, one of those eight schedules was chosen at random.

Importantly, the probability of contingency change was assigned so that in 10-15% cases a switch did not occur. This was done to ensure participants pay attention

and with the prospect of analysing switch anticipation behaviour.

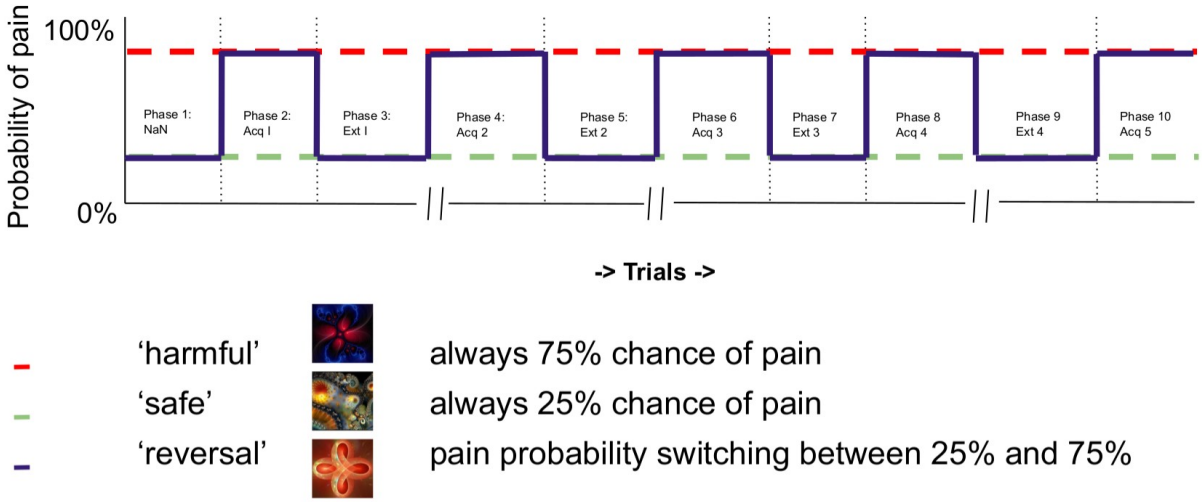


Figure 2 *Task structure*

Table 2: *Overview of Parameters for Studies 1 - 3*

	Study 1: Learning Asymmetry	Study 2: State Learning y	Study 3: Losartan Learning
Visits	2	1	3
Sessions per visit	1	3	1
Trials per session/visit	Training session: 60-90; Testing session: ~250	210 per session (~630 per visit)	Visit 1: 150, Visit 2: 310, Visit 3 150
Reversal frequency (trials per phase)	30 +/- 2	35 +/- 10	30 +/- 5
Number of phases	8-9	6 per session, 18 in total	Visit 1: 5-6; Visit 2: 10-11; Visit 3: 5-6
Cues	3	3/session	3/visit
Cue frequency (harmful, safe, reversal)	30-30-40	20-20-60	25-25-50
Shock probabilities	25-75%	Session A: 40-60%, Session B: 25-75%; Session C: 10-90%	25-75%
Inter-trial interval (ITI)	3-5s	1-2s	2s
Inter-stimulus interval (ISI)	2-4s	1-2s	1s

In experiments 1 and 3 the task was split into 3-4 sessions. This was done due to time limit of one scanning session and to re-calibrate the electrical stimulation.

Participants were told that these breaks occur at random times for practical reasons and that they should treat the entire experiment as one long continuous experience.

While the presented paradigm follows the classical Pavlovian conditioning task used to investigate aversive learning, it bears a few differences from previous studies exploring acquisition and extinction.

Firstly, most previous studies split the experiment into blocks and report data averaged per block or at the maximum across several trials (e.g., early vs late trials). This is mostly done due to low signal-to-noise ratio in galvanic skin responses which are often used as indicators for aversive learning (see below), but is also found in studies using shock probability ratings like in the present study. In my paradigm I hope to utilize trial-by-trial information in all measures. While extraction of single trial GSR signal may be challenging, recent advances in skin conductance modelling allow for this possibility (see for example Tzovara et al. 2018).

Secondly, in the proposed experiments the reinforcement rate in acquisition and extinction is never fully deterministic. Early classical conditioning studies often use 100% reinforcement during acquisition and 0% during extinction. While this is useful for the study of fundamental processes involved in the two phases, there are two reasons for using a probabilistic paradigm 1) Uncertainty has been shown play an important role in aversive learning; 2) reinforcement is scarcely deterministic in real world scenarios (both points were discussed in Chapter 1).

Thirdly, the vast majority of studies only have a single change from acquisition to extinction while the present paradigm includes several switches between the two. Similarly to point two, a single change may be useful to identify the most distinct differences and features. However, it does not capture a key feature of real-life acquisition and extinction, namely the period switch between them. As described earlier, both anxiety and pain disorders are characterised by periods of heightened anticipation of fear and pain and relative lack thereof. In the paradigm proposed here,

the participant experiences at least three acquisition and three extinction phases to allow for the investigation of more time-dependent processes and the evolution of learning strategies over the course of the experiment.

Electrical Stimulation

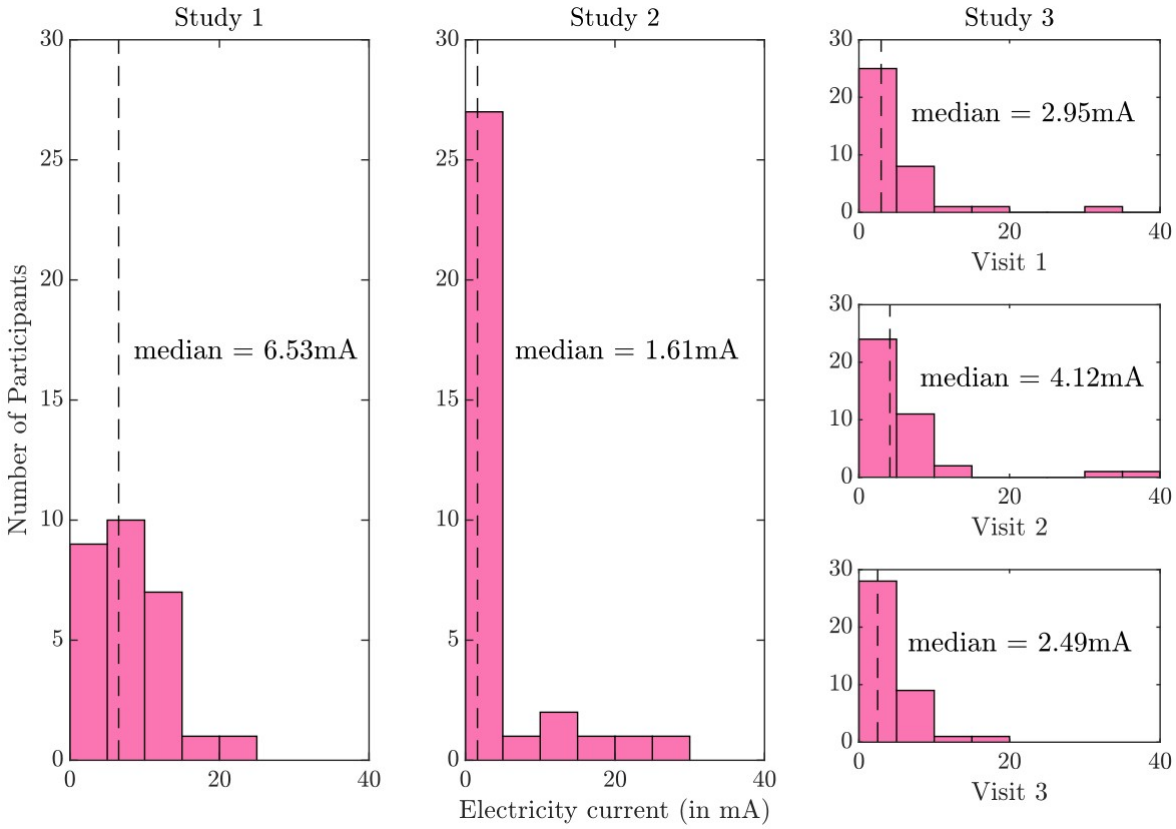
Electrical stimuli were applied using a commercial electric stimulation device (Constant Current Stimulator, model DS7A; Digitimer, Hertfordshire, UK), delivering a 2 monopolar square waveform pulse via a concentric silver chloride electrode attached to the back of the left hand.

Electrical stimuli used were calibrated individually before each of the four sessions to match an intensity of 8 on a scale ranging from 0 (= 'not painful') to 10 (= 'unbearably painful'). The 8/10 pain level was defined as a sensation that is painful but acceptable for a given amount of trials (study-specific number corresponding to 50% of trials). Three qualitative anchor points were defined to help standardize the calibration across participants: 1/10 which was defined as the intensity at which the sensation becomes painful (pain threshold); 8/10 is a sensation that is painful but acceptable; and 10/10 which would be level of pain that is too strong to be tolerated. The calibration procedure used followed the Method of Limits. Upon each stimulus delivery participants were asked to report how painful the sensation was on a Numerical Rating Scale ranging from 1 to 10. The current of delivered stimuli was drawn from a normal distribution with an increasing mean. When a pain level higher than 8 was reported by the participant the next stimulus was always lower than the previous one. If participants reported an intensity of 8/10 on 3 out of 5 stimuli, the mean stimulation intensity used in these 5 trials was used for the upcoming session.

The electrical stimulation differed slightly between experiments. In Study 1 which involved brain imaging the delivered shock was a sequence of stimuli while in

Studies 2 and 3 a single pulse was applied. The adaption in Study 1 was necessary because the MR environment required stimulus calibration to be done automatically. Stimulus intensity was therefore varied by changing the number of very brief stimuli applied in one sequence instead of increasing or decreasing the current of a single stimulus. The median current (calculated by multiplying the base stimulation by number of microstimulations) for Study 1 is therefore higher than for the other two studies.

Figure 3: *Distribution of mean shock intensities used in Studies 1 - 3*



Behavioural Data

Shock Probability Ratings

Shock probability rating were collected upon the cue presentation on each trial.

Ratings were entered using a slider that was presented on the computer screen in front of the participant which they could move to the left and right until it reached the desired position and submitted their answer by pressing the button under their index finger. Note that participants could not see the numerical equivalent to the position of the slider on the scale but only the endpoint labels of the scale (0% and 100%). This was done to eliminate visual number processing as a value signal confound in the MRI and physiological data. Participants had 4 seconds to submit their rating, if they failed a message flashed on the screen saying 'MISSED TRIAL', the failure was recorded and the trial was restarted.

Ratings provided a principle dependent measure which as used to describe average tendencies such as mean probability in acquisition vs extinction, post reversal learning, single trial model-free learning rate and for switch point detection. Additionally, the ratings were used to fit the behavioural models and calculate parametric modulators in the MRI analysis (see below).

Mean Probability Measure

To identify general differences between trial types (i.e., harmful, safe and reversal cue trials) in the data sets, expectancy ratings were averaged in a standard per-condition manner. When differences between acquisition and extinction were sought, the initial four post-reversal learning trials were cut to remove the effect of initial learning. Four was determined by averaging ratings from all reversal trials and including trials where the cohort had completed learning, i.e., reached at least 60% in acquisition and at least 40% in extinction.

Model-free Learning Rates

Learning rates are commonly estimated as a single learning rate across all trials of the task or condition. However, as I intend to characterise trial-by-trial learning,

we used the Rescorla-Wagner learning rule to calculate the learning rate α for each trial.

Eq. I $$P_{t+1} = P_t + \alpha_t(O_{i,t} - P_t)$$

Eq. II $$\alpha_t = \frac{P_{t+1} - P_t}{O_t - P_t}$$

where $$0 \leq \alpha_t \leq 1$$

Since this is not a commonly adopted metric, I formally compared our model-based learning rates to the more commonly used model-free learning rates (full analysis in Appendix I). Model-free learning rates across all cues and visits correlated moderately with model-based learning rates (r=.51). This correlation was stronger in the reversal cue (r=.65). A mixed ANOVA with the factors group and visit performed on the model-free learning rates resulted in the same set of statistically significant results as an identical analysis performed on model-based learning rates in study 3. Importantly, both methods have their strengths and drawbacks. Model-free learning rates capture learning including additional effects of reporting noise, attention, cognitive biases and other factors, so they can often be negative, however, they do provide a trial-by-trial measure of learning. Model-based learning rates lack single trial sensitivity as they are estimated per condition. The model can only capture the information processing it is explicitly designed for, so additional processes will not be fitted unless the model is modified.

Change Point and Change Steepness

A key hypothesis of this thesis is that extinction is delayed relative to acquisition when the same amount of evidence is presented. In order to quantify the onset of both processes and estimate the subjective change point (i.e., when participants begin to

increase their probability ratings during acquisition and to decrease their ratings during extinction) we employed the CUSUM method (Page, 1954) in which a change point is estimated by first computing a cumulative sum over a segment of data (here: shock probability ratings) and subsequently finding the maximum/minimum.

First, a segment of 10 trials before and 10 trials after the reversal was extracted. By reversal I mean the objective change in contingencies, either from high to low or from low to high. The mean probability rating across the segment was subtracted from each rating and the cumulative sum was calculated (See Figure 4). In the case of acquisition, the CUSUM minimum was identified as true switch point. In the case of extinction this was the maximum of the computed series. In addition to the change point, the switch steepness was estimated. To this end, I extracted 5 trials pre and 5 trials post the previously identified true switch point to which a sigmoidal function was fitted using Eq III.

Eq. III $$f(x, a, c) = \frac{1}{1 + e^{-a(x-c)}}$$

where *x* represents the data time series, *a* stands for the fitted sigmoid steepness and *c* is the midpoint. Keeping the value of *c* fixed at the previously estimated switch point, we used the Bayesian Adaptive Direct Search algorithm to fit the function to the data, obtaining the estimate of steepness (see Figure 4).

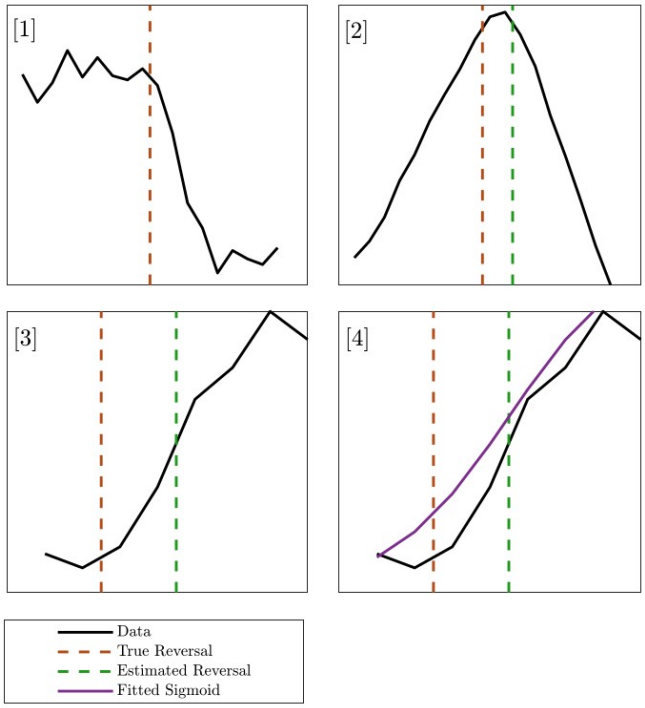


Figure 4: *The CUSUM procedure was used to find the steepest post-reversal learning point. [1] shows the raw data and the true reversal point (i.e., beginning of the extinction phase as defined by the reinforcement schedule), [2] shows the cumulative sum of the data and the subjective reversal point (i.e. change point) in green, [3] shows the data (inverted) aligned with the newly identified switch point, and [4] shows a sigmoidal function fitted to the data segment from [3].*

Questionnaires

A range of psychometric questionnaires were administered to assess individual differences in trait anxiety, anxiety sensitivity, depression, pain catastrophizing, pain vigilance, intelligence and attentional control. Participants filled in the questionnaires prior to their first visit using an on-line interface. All questionnaires were analysed according to their respective manual.

Study 1 and 2

- State-Trait Anxiety Inventory (STAI; Spielberger, 1983);

- Fear of Pain Questionnaire (FPQ-III; McNeil & Rainwater, 1998);
- Pain Vigilance and Awareness Questionnaire (PVAQ; McCracken, 1997);
- McWilliams & Asmundson, 2001);
- Pain Catastrophizing Scale (PCS; Sullivan, Bishop & Pivik, 1995);
- Centre for Epidemiological Studies Depression Symptoms Index (CES-D; Radloff (1977).

Study 3

- State-Trait Anxiety Inventory (STAI; Spielberger, 1983);
- Anxiety Sensitivity Index Revised (ASI-R) (Taylor and Cox 1998),
- Neuroticism subscale of the Eysenck Personality Inventory (EPI) (Eysenck and Eysenck 1975);
- Behavioural Inhibition Scale (BIS) (Carver and White 1994);
- Beck Depression Inventory (BDI) (Beck, Steer, and Brown 1996);
- Attentional Control Scale (ACS) (Derryberry and Reed 2002);
- National Adult Reading Test (NART) (Derryberry and Reed 2002; H. E. Nelson 1982; “National Adult Reading Test,” n.d.)).

Galvanic Skin Response

Data Acquisition

While participants performed the task, their galvanic skin response (GSR) was recorded in Studies 1 and 2 using the Biopac (BIOPAC Systems UK Inc.) system which is a non-invasive method commonly used to record electrophysiological data. The recording involved attaching a set of adhesive disposable electrodes (Biopac EL509) to the palm of the participants’ left hand. For better recording quality a conducting gel (Biopac GEL101) was applied to the electrodes. Data were recorded at a sampling frequency of 500 Hz in Study 1 and of 1000 Hz in Study 2. Because the electrodes were attached to the same hand where the electrical stimuli were applied, early deflections in the GSR signal were caused by the electrical stimulus. This short burst of electricity that in reality lasted 2 ms was detectable in most participants. In the GSR waveform, such signal, as seen in the grand mean waveform, lasted between 0

to 500 ms before returning to the baseline. This short latency makes it separable from the peripheral and central nervous system response which usually onsets 1-2 seconds after the stimulus and settles after 5-6 seconds. See Figure 5 for an example of SCR.

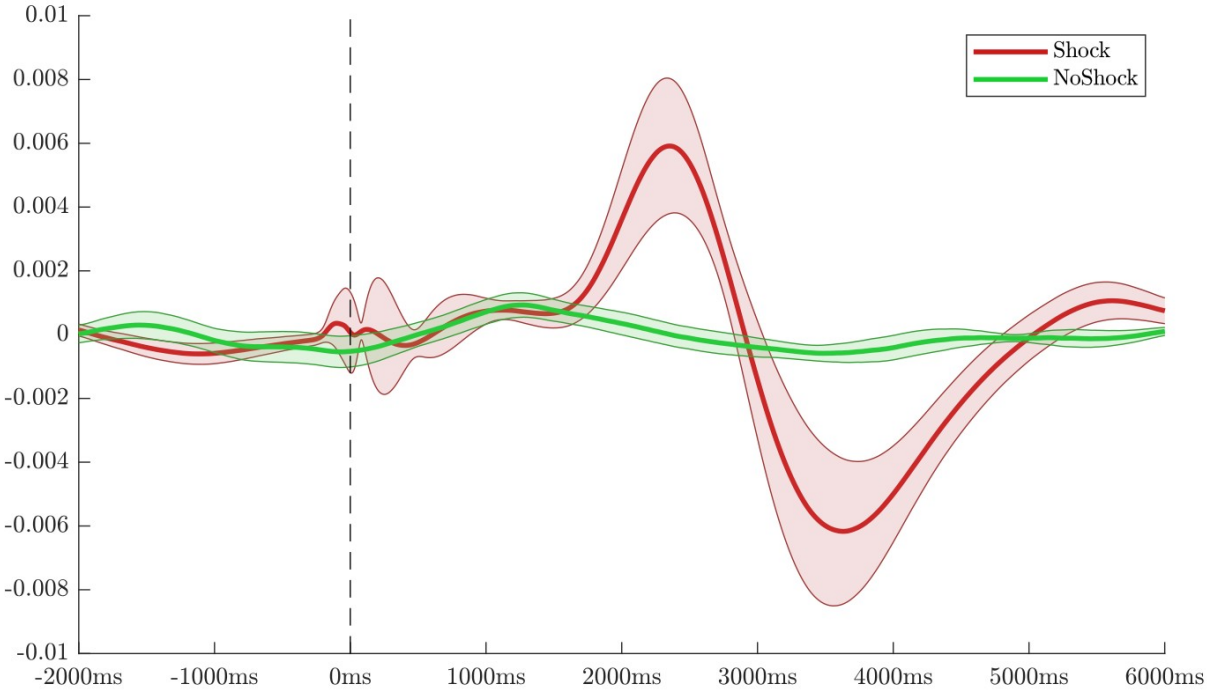


Figure 5 *Mean SCR response to shock and shock omission (shaded area +/- 1 SEM). The initial ~[0 to 500ms] post-stimulus signal increase is a direct result of the electrical stimulation while the large component at ~[1500 - 6000 ms] is a result of change in sudomotor nerve activity.*

Signal Processing

Raw data were saved in the *.acq format and imported to Matlab using PSPM (Dominik R. Bach and Friston 2013). The data were filtered using a bandpass filter between 0.0159 and 5 Hz. The bounds were selected according to recommendation for model-based SCR analysis by (Staib, Castegnetti, and Bach 2015). No further preparation of data was performed.

GSR Analysis

Skin conductance data were analysed using two methods: 1) general linear model (Bach et al. 2009); and 2) dynamic causal model for anticipatory SCR (Bach, Daunizeau, et al. 2010).

Model-based analysis of GSR signal has been proposed (Bach et al. 2009); (Bach and Friston 2013; Bach, Flandin, et al. 2010). The model-based analysis uses the known shape and latency window of the GSR to generate a time series of the predicted response and to regress it against the recorded data. The general linear model (GLM) is used to estimate the degree of observed signal associated with a given stimulus, measured by regression coefficients β for each regressor of interest. This methodology explicitly models the expected response allowing for signal detection even in temporally overlapping events. Similarly to MRI, the outcome measure of the analysis is a per-condition beta.

The second method, dynamic causal modelling (DCM) by (Bach, Daunizeau, et al. 2010) provides an alternative modelling method to the GLM. While the GLM provides a description of the data using estimated response function, the DCM takes a more elaborate approach. It uses a generative model of anticipatory, evoked and spontaneous fluctuations that link sudomotor nerve activity with the recorded skin conductance. Variational Bayes is used to invert the model to estimate the most likely original signal. Bach, Daunizeau, et al. (2010) showed that this method can explain a significantly larger portion of the variance than methods 1 and 2. The method has successfully been used to estimate single trial signal and as an error function for model fitting (Tzovara, Korn, and Bach 2018). As an outcome measure, the DCM provides an estimate of anticipatory activity on a per-trial basis.

The GLM served as the main analysis to identify difference in GSR among different conditions while the DCM was used to estimate single trial amplitude and peak latency. Dispersion of the signal was capped at 0.3 s.d., based on sympathetic

nervous system data recorded by (Gerster et al. 2018).

Challenges in the analysis of GSR data

Cross-trial interference

In studies 1 and 2 the priority was to collect as many trials as possible. Therefore, the elicited GSR on a given trial overlapped with the response to cue onset of the subsequent trial. In Study 1 where the inter-trial interval (ITI) was 3 - 5 s, each trial contained approximately 500 ms of the previous trial's period. In Study 2, this overlap was more substantial (ITI = 1 - 2 s). While in a model-free analysis this would be problematic (in fact the peak-and-trough results demonstrate strong effect of previous outcome), the GLM approach models all events and therefore accounts for the effect of previous outcome. In the GLM and DCM results in study 1, no effect of previous outcome is present, in study 2, in GLM, the effect of previous trial was present.

Trigger-signal offset

Due to the setup used, there was a slight offset between the recorded timings in Matlab and the GSR signal. This occurred despite more precise functions from the PsychToolbox 3 were used, most likely due to length of the cables involved. In Study 2, we recorded Matlab timestamps as well as triggers sent to the Biopac system and 50ms lag in signal. Overall, the offset poses no significant issue. For the GLM, the response function is adjusted to the data and only takes data acquired until 500 ms after cue onset into account. In the DCM analysis which estimates parameters in anticipation of outcome I initially detected a weak effect of the upcoming outcome when 500ms pre-outcome was removed.

Pupillometry

Data acquisition

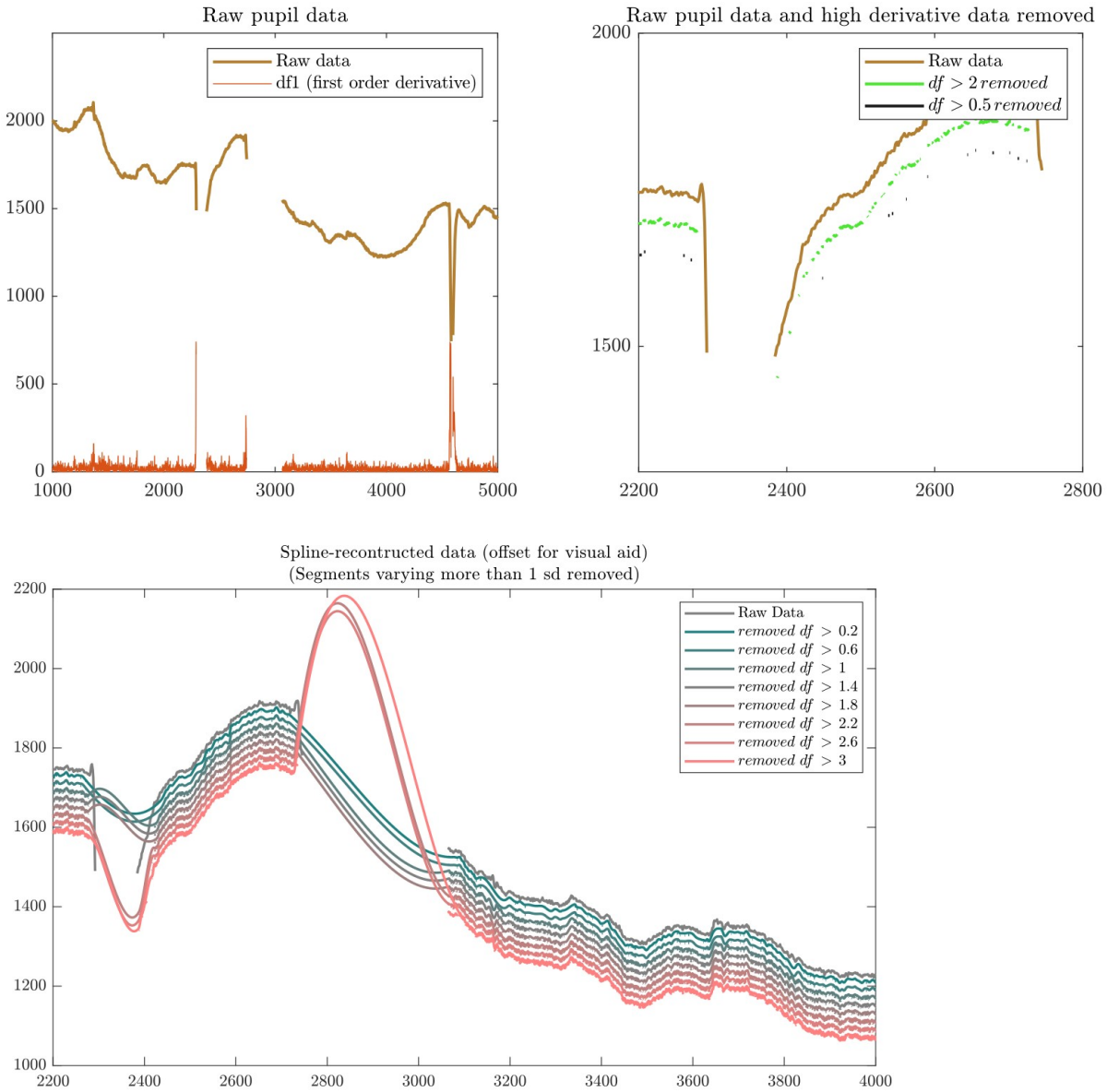
Pupil dilation data were recorded in study 3 (only visit 2). Tower mount setup with a screen distance of 70 cm was used. The EyeLink recording software was used. Prior to the start of each session the eye-tracker was calibrated using a nine-point calibration. The lighting conditions were kept constant for every participant

Data Preprocessing

Data were processed using custom MATLAB scripts following the appropriate guidelines (Kret and Sjak-Shie 2019) and parameters chosen in similar studies (Nassar et al., 2012; Browning et al., 2015; Pulcu and Browning, 2017). The European Data Format (EDF) data were converted to ascii format. Segments of blinks and missing data were identified by calculating the first order derivative (df) and excluding data points with $df > 1.6$. Data segments that deviated more than one standard deviation from the mean of the entire session were excluded from the data set. Where data gaps were shorter than 2s, Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) was used to interpolate the missing data. Trials with more than 50% interpolated data points were excluded. Figure 5 shows details of the interpolation procedure. The resulting data were next filtered using a bandpass filter between 0.1 and 4 Hz. Subsequently epochs locked to cue onset were extracted and corrected using a pre-event baseline window of [-0.1 to -0.5 s].

Figure 6: *Interpolation steps of pupil signal. The top left panel shows raw pupil data together with the first derivative (df1). In order to reconstruct missing data, a suitable df1 cutoff value had to be selected to remove sharp changes in the data that were caused by recording errors or blinks. The top right panel shows the signal following removal of data points exceeding df = 2 (green) and df = 0.5 (black). Any remaining*

data in the middle of the gap (s.d. > 1) were removed to prevent interpolation from including them. The bottom panel shows data that was removed and spline reconstructed for derivatives ranging between 0.2 and 3. The derivative cutoff of 1.6 was found to best preserve the shape of the original signal without including blinks and other artifacts.



Data Analysis

To analyze pupil dilation data we employed the same approach as described by (Pulcu and Browning 2017b). Data for each event (i.e., cue onset, subject's response and outcome) were analysed separately. Event-locked pupil size grand

means were calculated and plotted. To increase statistical power, the post-event data were binned into 0.5 seconds wide categories. For cue onset and outcome, a 1.5 s post event window was analysed. For the response event the window was shorter (i.e., 1 s). A shorter window was selected to prevent the inclusion of outcome-related signal. Binned data were analysed using a mixed ANOVA that included time as a repeated measure variable and, in Study 3, drug manipulation as a between subject variable. For the outcome phase, event type (shock or no shock) was added as another within-subject factor.

Additionally, the general linear model approach was used to identify the effect of expectation, outcome and prediction error on pupil dilation. A sliding window of 50 ms was used as dependent variable in a linear regression incorporating current expectation, outcome and unsigned prediction error as predictor variables (see Eq. IV). The resulting beta values were subsequently averaged across participants and analysed by binning to categories 0.5s wide. Two seconds post each event were included, resulting in a repeated measures variable of four bins.

Eq. IV $$X \sim \beta_0 + \beta_1 * P_i + \beta_2 * O_i + \beta_3 * |PE|$$

Functional Magnetic Resonance Imaging

Data Acquisition

Functional magnetic resonance imaging (fMRI) data were only recorded in Study 1. For each participant we acquired one high resolution structural scan (T1-weighted MP-RAGE, 1x1x1 mm, TR=1900 ms, TE=3.97 ms, flip angle: 8 degrees), four runs of functional scans (multi-band with acceleration factor 3 and GRAPPA in-plane acceleration factor 2, TR=1570 ms; TE=30 ms, FoV = 216 mm, voxel size: 2x2x2 mm, flip angle 70 degrees, 30 degree tilt, pre-scan normalization was off), a fieldmap for each functional scan and a diffusion tensor imaging (DTI) scan (multi-band

acceleration factor of 4, TR=2951 ms, TE = 83.8 ms, FoV = 204 mm, voxel size: 1.5m isotropic). This particular echo-planar imaging (EPI) sequence was selected for its higher field of view and significantly better signal-to-noise ratio in subcortical areas and medial orbitofrontal cortex compared to a multi-band 4 sequence commonly used in EPI studies in the department. During acquisition, data were visually inspected for quality and signal dropout.

Data Preprocessing and Analysis

The data were preprocessed using a combination of the SPM12 (Penny et al. 2011) and FSL5 (Jenkinson et al. 2012) toolboxes. The MRI data preprocessing including the following steps:

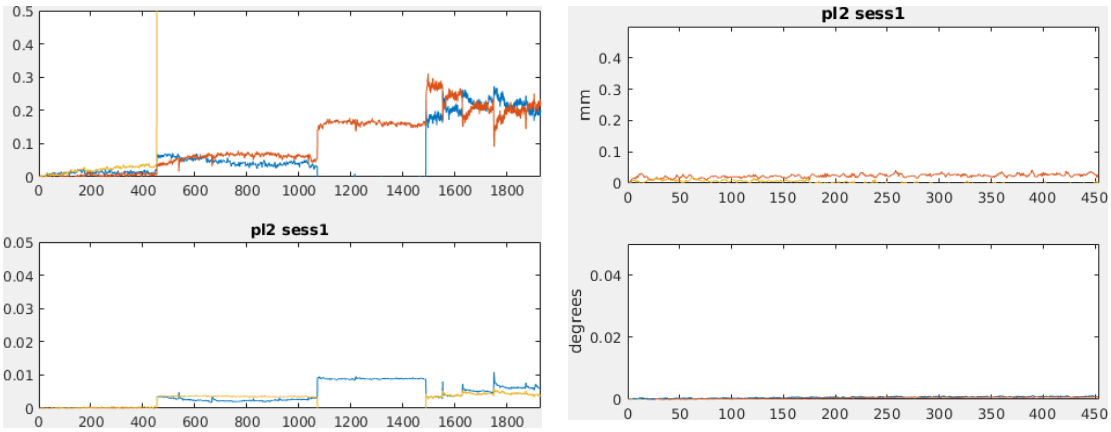
1) Brain extraction of EPI data using FSL's Bet2 tool;
2) Bias-field correction on EPI data using SPM12's bias field calculation and spm_math to apply it;
3) Pre-smoothing of EPI data (4mm kernel) for AROMA, using SPM12's smoothing function.
4) AROMA motion correction - an ICA-based method for motion correction from FSL (Pruim et al. 2015);
5) Temporal re-slicing (SPM12);
6) Combining phase and magnitude maps into a Fieldmap (SPM12);
7) Realignment and unwarping using fieldmaps (SPM12);
8) EPI coregistration with the structural scan (SPM12);
9) Structural scan segmentation (SPM12);
10) Normalization to MNI space (SPM12);
11) Smoothing with 8mm kernel (SPM12).

Instead of including motion parameters in the general linear model to correct for motion correction, we opted for an ICA-based method from the AROMA toolbox which is part of FSL 5.0. This technique has previously been shown to successfully remove

motion noise (Pruim et al. 2015). This is particularly important in studies involving strong stimuli with a sudden onset (e.g., painful electrical stimuli) as the delivery of the stimulus can trigger a withdrawal response. An example of motion parameters before and after the application of AROMA are presented in Figure 6.

The preprocessed data were analysed using a GLM approach. Each model was convolved with the canonical haemodynamic response function from SPM and regressed against the data. Signal from a priori defined regions of interest was extracted using the MarsBar toolbox (Brett et al. 2002). The specific models and analyses used are described in Chapter 3. As a data check I constructed a model with separate regressor for cue outcome, response and the two possible outcomes (shock / no-shock) and tested the CS+>CS- contrast in a whole brain analysis.

Figure 7: *Motion parameters estimated for datasets without (left 2 panels) and with AROMA applied (right 2 panels). Upper panels show displacement parameters and lower panels show rotation parameters.*



Computational Modelling

Most computational models employed in this thesis are based on reinforcement learning. Different forms of the Rescorla-Wagner (RW) model were used to model a particular process or dissociate various sub-processes. In this section, I will

first describe the methodology used for model specification, model fitting, parameter recovery and model comparison. I will then follow by presenting the models used throughout the thesis. The models were either selected from existing literature, or where no available model existed written by the author.

Methods in Modelling

Parameter Fitting

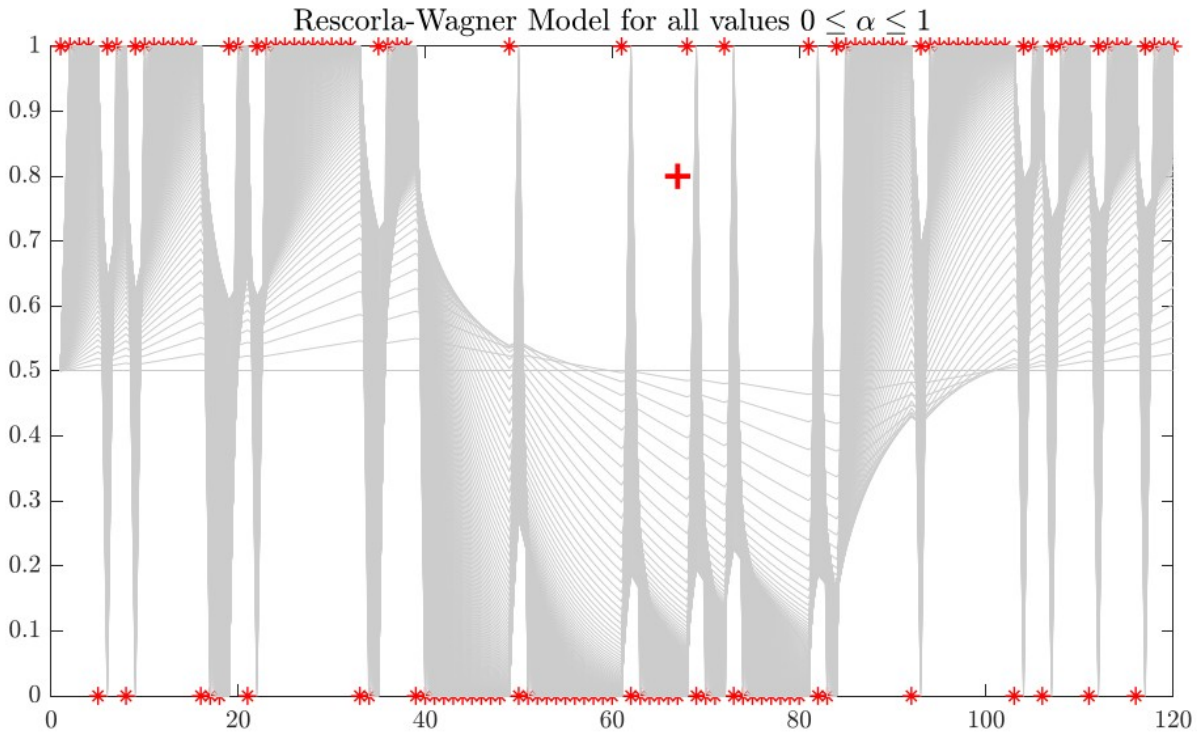
A computational model is a set of plausible scenarios under a given hypothesis. The parameters of a model characterise which patterns of data the model can capture. For example, as mentioned in Appendix I, the Rescorla Wagner model has one free parameter, α, which is defined for values $0 \leq \alpha \leq 1$.

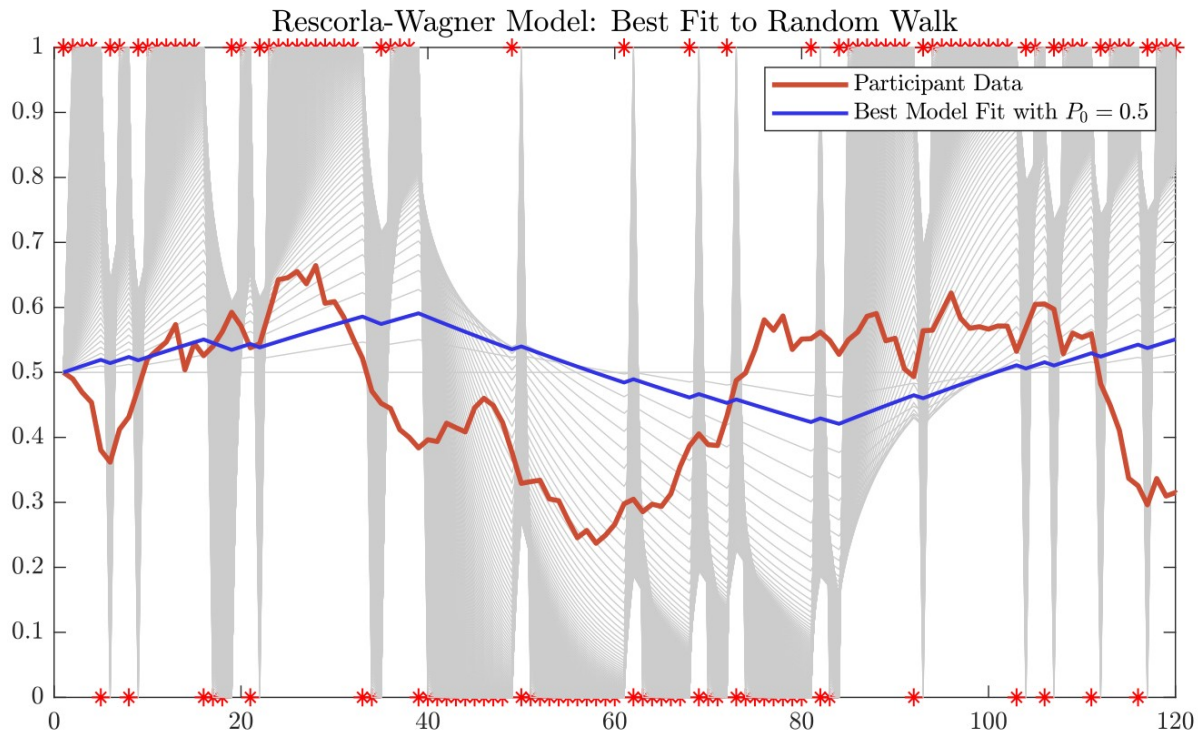
Eq. V $$P_{t+1} = P_t + \alpha_t(O_t - P_t)$$

If a participant uses a higher or lower learning rate (e.g. increases their shock expectation following shock omission), the model will not be able to capture their behaviour and the fit will fail - in these cases the optimisation usually outputs one of the extreme values. Figure 8, upper, demonstrates all possible trajectories a RW model is able to fit (gray). Red asterix represent either shock or no-shock. Those are plausible under the hypothesis that the data were generated using a mechanism similar to the process defined by the model, given the set of outcomes and the starting value of 0.5. Note that the model would not be a good fit, for instance, for a participant who reported a shock probability of 0.8 on trial 67 (marked with red cross) as under no combination of free parameters would the model capture that point. In Figure 8, lower panel, I plot data generated by a “random” learner that ignores the outcomes (red) and a best fitting Rescorla-Wagner (blue). This would be a participant who is insensitive to the outcome

but remembers the rating they provided previously. When the Rescorla-Wagner model is fitted to this participant, the algorithm performs a procedure during which it attempts to find such values of parameters that minimise the error between model estimate and the data. If we manually went through all the grey lines and calculated a cumulative error from the data for each of them, the blue line would produce the smallest error. In this case, the value of the single parameter α that minimised this error is $\alpha = 0.0098$. The data were generated by random but the model still tried to fit them within its range. This dummy example roughly illustrates the model-fitting procedure.

Figure 8: *Gray lines on both panels show range of possible trajectories that the RW model can capture for different values of α. Lower panel shows a time series that ignore the shock (red dots on top) and noshocks (red dots on the bottom), and how such participant will not be fitted well by the model because they don't share a common fundamental generative process.*





Formally, if P is a set of probabilities reported, R is set of values predicted by the model under a given value α, and n is the number of trials, then the optimisation algorithm tries to minimize the least squares error

Eq. VI $$\hat{P}_{LS}(\alpha) = argmin_{\hat{P}(\alpha)}[E\{[P - R(\alpha)]^2|\alpha\}]$$

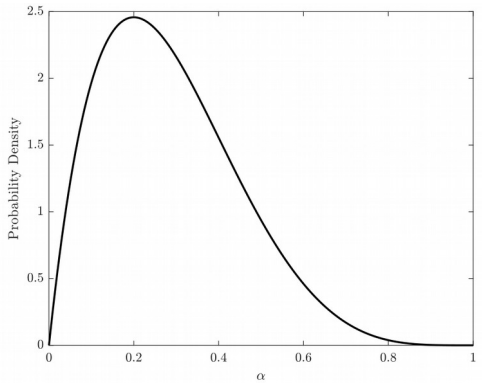
Note that the squared error only considers two points: the model estimate and the data. While this is sufficient in many cases, it sometimes is important to consider the expected distribution of the data, especially when the residuals are not normally distributed. An alternative error function would be the maximum likelihood. Unlike least squares, the likelihood makes an assumption about the shape of the distribution of all variables. For example, consider a case where the model predicts the value of 0.7 on a given trial while the participant reported 0.4. If we now assume that the standard deviation of the model prediction is 0.2 then we can ask what is the probability of observing the value 0.4 under the known expected distribution $\mathcal{N}(\mu = 0.7, \sigma = 0.2)$. (Note: the standard deviation does not have to be assumed, it can be estimated from

the data). Evaluating the Normal probability density function we obtain the value of 0.6476. The likelihood error function adds an additional layer of information: the uncertainty in the parameters and therefore in the model. While this approach is not often necessary as normally distributed residual sum of squares (RSS) will lead to the same minimal parameter values as least squares (Takezawa 2012), it becomes important if we are interested in Bayesian parameter estimation.

To capitalise on the amount of information available, we can use Bayesian inference to obtain a posterior distribution of parameter values. In order to do that we need one more piece of information: the prior probability density on parameter values. We can make assumptions about the most likely values of certain parameters based on our prior beliefs or literature search (i.e. empirical priors). For example, high values of α are not very likely given that most previous experiments using the Rescorla-Wagner model reported mean values between 0.1 and 0.35. Using this knowledge we can construct a prior guess as to how our parameter α will be distributed, such as shown in Figure 9.

Figure 9: *Example of a beta distribution with* $\alpha = 2, \beta = 5$.

$$\alpha \sim Beta(2,5)$$



We can then use Bayes Theorem to calculate our posterior parameter distribution of α:

Eq VII $$P(\alpha|D, M) \propto P(D|M, \alpha) \, . \, P(\alpha|M)$$

where M is the model and D is our data, $P(\alpha|M)$ represents our prior guess and $P(D|M, \alpha)$ the data likelihood.

This approach can be very beneficial, especially if we have few data and high number of parameters, or in hierarchical models. The Bayesian posterior can be approximated using the variational Bayes (VB) method (Beal, 2003) or it can be sampled using sequential sampling algorithms such as Markov Chain Monte Carlo. The VB method is an approximation to minimum negative log likelihood and model evidence (Daunizeau, Adam, and Rigoux 2014; Acerbi et al 2018). While this optimisation method is commonly used, MCMC is considered more robust. MCMC uses sequential sampling (e.g. Metropolis-Hastings) algorithms to map the landscape of posterior probability density (Brooks et al. 2011; for review see Forstmann, Ratcliff, and Wagenmakers 2016). The advantage of the VB approach is undoubtedly its speed, variational estimation can take under a minute where MCMC might take 20 or 30 times as long. However, compared to MCMC methods it is more likely to converge to a local minimum.

In this thesis, most models are fitted by minimising Eq VI. Where Bayesian models are used (usually hierarchical models) MCMC is given preference since computation/time is not a limitation. However, due to the limitations in the JAGS language (in which Bayesian graphical models are specified) certain operations are not possible to program. In those cases, VB approximation was used. The non-Bayesian optimisation was performed using Bayesian Adaptive Direct Search (BADS; Acerbi and Ma 2017). The MCMC model specification and sampling was performed using JAGS (Lee and Wagenmakers 2014).

Observation Functions

Observation function refers to the mapping of agents' internal states to observable variables such as subjective ratings in my case. Until now we have assumed that there is an exact correspondence between the agent's internal belief and the reported probability (i.e. value). However, evidence from value-based decision making research suggests that there is a non-linear distortion in reported probabilities at both ends of the probability spectrum (Kahneman and Tversky 2000). This non-linearity can be modelled using a sigmoidal function such that:

$$P_t = \frac{1}{1 + e^{-a(Q_t - b)}} \quad \text{Eq VII}$$

where P_t represents the reported probability, Q_t the internal belief, a the steepness of the function (the steeper the more binary the responses), and b stands for the mid-point which can be interpreted as a general bias towards one or the other end of the spectrum. While this function often dramatically improves model fits, upon closer inspection the values predicted by the model are often implausible due to overfitting.

It also needs to be pointed out that there is no need to map the internal belief only onto reported probabilities. The GSR signal was previously used to fit learning models (Tzovara et al. 2018; Li et al. 2011). In this thesis I will use both subjective value and GSR amplitude signal as a target function. Furthermore, the sigmoidal response function will be explored but it will not be used as a standard part of any model. The default mapping will therefore be:

$$P_t = Q_t \quad \text{Eq VIII}$$

Model Comparison

Standard Models

To dissociate between competing hypotheses, models need to be compared.

This can be done in a number of ways. One common way is to calculate the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) for each model and compare the obtained values. In doing so, it is important to penalize for model complexity - models with a higher number of free parameters will have more degrees of freedom and are therefore more likely to fit better. Both AIC and BIC penalize more complex models by adding a penalization constant which is multiplied by the number of parameters for the final score (in both cases a lower score means better fit, so addition means penalization). In this thesis I mostly used the AIC score, although BIC is also calculated.

Formally, the Akaike Information Criterion is defined as:

Eq IX $$AIC = -2 * ln(\mathcal{L}) + 2k$$

where $\mathcal{L}$ is the likelihood (seen above) and k corresponds to the number of free parameters. Alternatively, AIC can also be calculated using residual sum of squares using:

Eq X $$AIC = N * ln(RSS/N) + 2k$$

where residual sum of squares (RSS) corresponds to the sum of the squared errors and N to the number of data points. The RSS/N term therefore corresponds to the sample variance.

Similarly, we can calculate the Bayesian Information Criterion:

Eq XI $$BIC = ln(N) * k - 2ln(\mathcal{L})$$

and

Eq XII $$BIC = N * ln(RSS/N) + k * ln(N)$$

While these approaches are commonly used they are still partially susceptible to overfitting. We want each model to capture true, and therefore generalizable, processes in the data. Models with more free parameters will have a higher tendency to explain all the variance in the data but this might simply be due to overfitted. One way to mitigate the effects of overfitting is to penalize for number of free parameters. Additionally, to check whether overfitting has taken place a cross-validation procedure can be used to see how well the fitted model predicts new data. A model that predicts data better is more generalizable. To perform this procedure, one can randomly remove (i.e. "hide") segments of data and fit the model to the remaining data set, to subsequently use the fitted values to predict the "hidden" portion of the data set.

In this thesis I performed this procedure by using 70 - 80% of the data for model fitting and the remaining 20-30% to assess model predictions.

Bayesian Models

Bayesian models are compared using a quantity known as model evidence (a.k.a. marginal likelihood). Model evidence is the normalization term $p(D|M_m)$ from the Bayes' Theorem:

Eq XIII $$p(\theta|D, M_m) = \frac{p(D|\theta, M_m)\, p(\theta|M_m)}{p(D|M_m)}$$

where θ is a set of parameters of a given model M_m and D is the observed data.

When calculating the posterior probability, both MCMC and VB method ignore the model evidence term because it cannot be calculate directly and the estimation/sampling can be performed without the normalization term by calculating the posterior proportional to the prior multiplied by the likelihood. Model evidence is an extremely useful quantity when comparing models. It tells us how good a given model

is on average at predicting the observed data, taking model complexity (i.e. across parameters) into account. Note that model complexity is not only determined by the number of parameters as assumed by AIC and BIC. Lee and Wagenmakers (2014, page 103) demonstrated that even two models with the same amount of parameters can have different a complexity which is simply a measure of the space of scenarios it can capture. The beauty of using marginal likelihood for model comparison is that due to processes inherent to the Bayes' Theorem the model will automatically apply Ockham's razor - the simplest and most predictive model will have the highest model evidence.

While in most cases it is not possible to directly calculate the model evidence, several methods have been proposed to approximate it. In addition to approximating maximum likelihood, the VB procedure also estimates model evidence. In MCMC, a measure known as Deviance Information Criterion (DIC) has been used to evaluate model performance (Spiegelhalter et al. 2002). A number of additional methods have been proposed, such as Savage-Dickey density ratio (Dickey 1971) or the Product Space method (Lodewyckx et al. 2011). In this thesis I will use the DIC measure to compare Bayesian models as it is automatically estimated by JAGS.

Models

In this section I introduce the main models used throughout the thesis. The models were selected due to their common use in the learning and decision-making literature, some were specified to capture a particular learning strategy. Some study-specific versions of these models are omitted here and presented in the respective chapter. A summary of models is presented in Table 1. For each model a parameter recovery procedure was run. This involved generating sets of all plausible parameter combinations that lead to value predictions between 0 and 1, and randomly selecting 200 of those combinations. The selected parameter combinations were then used to generate artificial data, these were then fitted using the BADS algorithm three times (to

ensure that local minimum was not selected) with different initial parameter values. The best fitting option was selected and the estimated parameters were compared to the original set of parameters. Recovery rates are Pearson's ρ correlation coefficients.

Table 3: Model Overview

Name	Parameters		Recovery Rate (Pearson's ρ)
Rescorla-Wagner (RW1)			
	$\alpha \in [0, 1]$	Learning rate	1.00***
Outcome-specific Rescorla-Wagner (RW2)			
	$\alpha_- \in [0, 1]$	Shock learning rate	1.00**
	$\alpha_+ \in [0, 1]$	No-shock learning rate	1.00**
Dynamic Decaying or Growing Learning Rate (RW4_DecGrowth)			
	$\alpha_0^+ \in [0, 1]$	Starting shock learning rate	0.3**
	$\alpha_0^- \in [0, 1]$	Starting no-shock learning rate	0.35***
	$\lambda_+ \in [0, 2]$	Shock learning rate decay	0.68***
	$\lambda_- \in [0, 2]$	No-shock learning rate decay	0.70***
Pearce-Hall (PH5)			
	$\alpha_0 \in [0, 1]$	Starting learning rate	0.97***
	$\eta_0 \in [0, 1]$	Starting Associability	0.67***
	$\pi_+ \in [0, 1]$	Shock Associability Weight	1.00***
	$\pi_- \in [0, 1]$	No-shock Associability Weight	0.91***
	$\kappa \in [0, 1]$	Scaling of Associability	0.96***
Pearce-Hall with Decay (PH6)			
	$\alpha_0^+ \in [0, 1]$	Starting shock learning rate	0.89***
	$\alpha_0^- \in [0, 1]$	Starting no-shock learning rate	0.91***
	$\eta_+ \in [0, 1]$	Shock Associability	0.84***
	$\eta_- \in [0, 1]$	No-shock Associability	0.79***
	$\lambda_+ \in [0, 1]$	Shock Decay	0.69***
	$\lambda_- \in [0, 1]$	No-shock Decay	0.67***
Surprise-Minimizing Model			
	$\alpha_+ \in [0, 1]$	Shock learning rate	0.23*
	$\alpha_- \in [0, 1]$	No-shock learning rate	0.18 n.s.
	$\theta^+ \in [0, 1]$	Positive surprise weight	0.8***
	$\theta^- \in [0, 1]$	Negative surprise weight	0.78***
State-Switching Model			
	$\alpha_+ \in [0, 1]$	Shock learning rate	0.97***
	$\alpha_- \in [0, 1]$	No-shock learning rate	0.80***

	$P_0^{S_1} \in [0, 1]$	Starting value of state 1	0.33**
	$P_0^{S_2} \in [0, 1]$	Starting value of state 2	0.33**
	$\eta \in [0, 20]$	Trials considered	0.62***
*p<0.05 **p<0.01 ***p<0.0001 n.s. = not signifiacnt			

Rescorla Wagner

As explained before, the Rescorla-Wagner learning rule is defined by the equation

$$\text{Eq XIV} \quad P_{t+1} = P_t + \alpha(O_t - P_t)$$

where P_t is the current value, or probability in our case, at the trial t, while O_t is the current outcome, coded as $O = 1$ for shock and $O = 0$ for shock omission. This model has 1 free parameter, $\alpha \in [0, 1]$. The model will be referred to as RW1.

To account for differential learning from shock and no-shock outcomes, this model can be extended to have separate learning rates for shock, $\alpha_+ \in [0, 1]$ and for shock omission $\alpha_- \in [0, 1]$. This model will be referred to as RW2.

Dynamic Learning Rate Models

Both above versions of the RW model have fixed learning rates, i.e. the learning rate does not change over trials. It has, however, been shown that this assumption is not necessarily valid and learning rates tend to change at least in two ways.

Decaying or Growing Learning Rate

A number of studies have reported a gradual decrease or growth in the speed of learning (ref). I therefore specified a LR Decay/Growth model (termed RW4_DecGrowth) to account for this pattern of learning rate evolution. The model is

build on Eq XIV (RW2) but the learning rate α now changes on a trial-to-trial basis such that

Eq XV $$\alpha_t^i = \alpha_0^i * \lambda_i^t$$

where i stands for outcome, t for ordinal occurence of a given outcome, $\lambda \in [0,\ 2]$ is a decay/growth parameter which will take values $\lambda < 1$ for decay in learning rate, $\lambda = 1$ for stationary learning rates and $\lambda > 1$ for growing learning rates.

Pearce-Hall Model

Several models have been proposed to account for attention paid to the stimuli based on their uncertainty (Mackintosh, 1975; Pearce and Hall, 1980). One version of the Pearce-Hall learning rule was presented by Pearce et al. (1982) and has been used in a number of studies (Piray et al. 2019; J. Li et al. 2011; Tzovara, Korn, and Bach 2018; Morriss, Saldarini, and van Reekum 2019) (Piray et al. 2019). This model keeps track of a current level of associability η, that is the degree to which attention is paid to new stimuli driven by recent surprise, and uses a weighing parameter π to update associability as follows:

Eq XVI $$\eta_t = (1 - \pi_i)\eta_{(t-1)} + \pi_i|O_t - P_t|$$

The current associability level is the scaled by κ to form the current learning rate:

Eq XVII $$\alpha_t = \kappa * \eta_t$$

The parameter π can also be as a weighing parameter between constant and dynamic learning.This model is referred to as PH-RW5.

Pearce-Hall with decay

To combine both previously seen dynamic features of learning, attention and decay, I designed a model that includes both. It is important to consider learning rate decay as my behavioural data suggested decrease in learning rates over time:

Eq XVIII $$\alpha_t = \alpha_{(t-1)} + \eta_i|O_i - P_i| - \lambda_i\alpha_{(t-1)}.$$

Here, η is an outcome-specific volatility parameter while λ controls the rate of learning rate decay. This model is referred to as PH6.

Surprise-Minimising Model

To test how participants' learning is driven by a combination of positive and negative expected surprise I designed the following model building on the concept of absolute Shannon surprise (Shannon 1948) which is defined as:

Eq XIX $$s_t = -log|O_t - (1 - P_t)|$$

Here, I assume a two step learning process where value Q is initially learned by an standard outcome-specific Rescorla-Wagner mechanism and subsequently updated to P by a combination of expected positive and negative surprise.

Eq XX $$Q_{(t+1)} = P_t + \alpha_i(O_t - P_t)$$

Eq XXI $$P_{(t+1)} = Q_{(t+1)} - \theta^- log|1 - (1 - Q_{(t+1)})| + \theta^+ log|0 - (1 - Q_{(t+1)})|$$

where $-log|1 - (1 - Q_{(t+1)})|$ represents the expected absolute negative surprise, i.e.

if a shock occurred, and $-log|0-(1-Q_{(t+1)})|$ represents the absolute positive surprise, i.e. if shock did not occur. The parameters θ^+ and θ^- represent sensitivity to positive and negative surprise. For example, if the agent wants to increase their prediction to minimise the negative prediction error, the parameter θ^- will be high. Absolute negative expected surprise needs to be added, but it has negative log in front, so to add the quantity we need to subtract the term, and vice versa in the case of positive surprise.

State-Switching Model

To test whether some participants represent the environment as two states I developed a state-switching model based on Rescorla-Wagner updating. The idea of state switching has previously been presented by Redish et al. (2007) and Gershman (2010). The principle idea is that the agent associates a certain level of reinforcement with a particular state and on each trial decides whether the most recent outcome means that the state on next trial will be the same or whether it has likely changed. The fundamental problem in state learning is state discovery. If we assume that an agent starts learning in one state in a gradual RW-like manner when and how is a new state created if observations don't match the existing state. Here, I don't aim to answer this question. I simply present a model which assumes two states in hope to identify participants that represent the learning environment in a similar fashion.

A state S_s is characterised by a mean value, i.e. identity:

Eq. XXII $$S = \{\mu_s\}$$

defined as sum of previous values associated with the state:

Eq. XXIII $$\mu^{S_s}_{(t+1)} = \sum_{n=t}^{i=1} P^{S_s}_{i\ldots t}$$

Two states are created and their initial values are estimated from data as free parameters. On each trial a state-specific error ϵ is calculated as the difference

between mean of η past outcomes (i.e. the recent objective value) and the mean value of the state being considered, μ_{S_i}. η controls the amount of past outcomes considered, it is an estimated free parameter.

$$\epsilon_t^{S_s} = |\mu_{S_s} - \sum_{n=\eta}^{i=1} O_{t-i}| \quad \text{Eq. XXIV}$$

Errors for both states are compared and the state with the lower error is chosen. Here, if state 2 has a lower error the following updating is performed:

$$if\ \epsilon_t^{S_1} > \epsilon_t^{S_2} \quad \text{Eq. XXV}$$

$$P_{(t+1)}^{S_2} = P_t^{S_2} + \alpha_i(O_t - P_t^{S_2}) \quad \text{Eq. XXVI}$$

$$R_{t+1} = P_{(t+1)}^{S_2} \quad \text{Eq. XXVII}$$

$$P_{(t+1)}^{S_1} = P_t^{S_1} \quad \text{Eq. XXVIII}$$

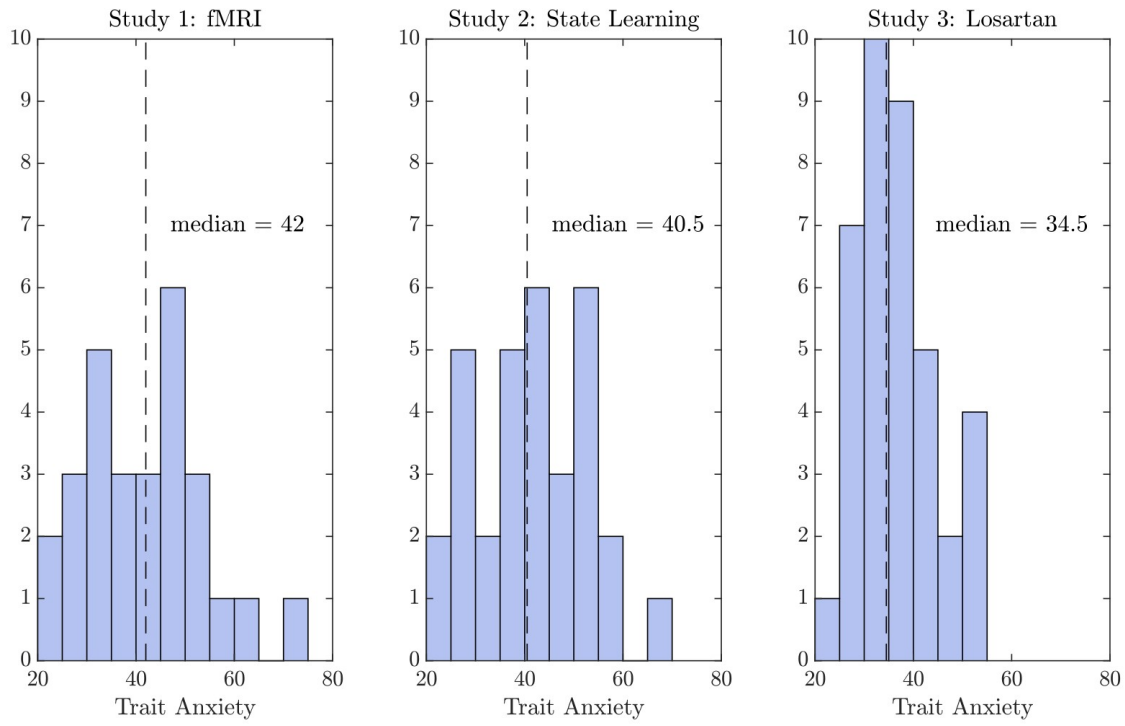
The current value of the winning state is updated using the state's previous value and outcome-specific learning rate α_i. The final value R_{t+1} is assigned as the value of the winning state. The value of the losing state remains the same. The mean value of the winning state is updated.

In addition to the error used to classify states described above, two other errors are generated on each trial: the prediction error, which is the the difference between the prediction of the model P_{S_t} and outcome; and state prediction error which is the difference between current state identity μ_{Ss} and outcome. While these will often be correlated they reflect a fundamentally different quantity, prediction error only considers one trial, while state prediction error is a running mean of all trials classified under that state. It is the state identity that constitutes the model's approximation to subject's hidden state.

Population Sample

The demographics of each study is described in each relevant chapter. In total, 119 participants completed the tasks (study 1: 34, study 2: 40, study 3: 45) from which 101 were included in the final analysis (study 1: 28, study 2: 33, study 3: 40), 55 female. No participant took part in more than one study. All participants were recruited using local advertisement and an in-house recruitment system provided by the Department of Experimental Psychology. Since trait anxiety is a critical variable in this thesis, I present the distribution of trait anxiety scores per study. Medians were used in studies 1 and 2 to split the sample

Figure 10: *Distribution of trait anxiety scores per study*



Chapter 3:

Neural and computational mechanisms of aversive acquisition and extinction

Introduction

The following chapter reports an investigation into behavioural, physiological and neural mechanism of aversive acquisition and extinction. Much of the relevant research was already discussed in Chapter 1. Here, I will focus on the main motivations and relevant points and proceed with the methods, results and discussion.

A lack of extinction is a widely studied phenomenon due to its prevalence in many situations involving negative prospects. Aversive experiences were shown to lead to rapid learning but are highly resistant to extinction (Lissek et al. 2005). This mechanism has likely developed as a survival tool to anticipate danger, but it becomes maladaptive in many everyday situations.

Acquisition is a term referred to the process of associating as neutral stimulus with an unconditioned stimulus (US) which produces a conditioned response (CR). Extinction refers to the reverse process of *dissociating* the now CS from US thus diminishing the CR. While the nature of acquisition is relatively uncontroversial, the processes governing extinction are less clear, particularly due to its increased contextual dependence, slowness relative to acquisition and phenomena like reinstatement and context renewal where the conditioned response is rapidly re-acquired (Bouton and Bolles, 1979). This suggest that extinction is likely more than mere ‘unlearning’ and the original memory is likely retained.

Theories on aversive learning disagree about the exact nature of the difference between acquisition and extinction. One approach argues that the observed lack of extinction is a consequence of aversive event over-prediction (Artnz, 1991), a faster learning from aversive than appetitive/neutral events, and that extinction is in fact a form of unlearning. This implies forgetting of previously US-CS pairings. Others, on the

other hand, have suggested that there is a separable learning process involved in each phase and memory for previous US-CS pairing is retained post extinction (Bouton, 2002). The evidence for the earlier view comes from studies in which extinction is performed immediately after acquisition, and the authors show that such extinction leads to decreased levels of fear reinstatement (Myers et al 2006). Further evidence comes from studies on gradual extinction in which extinction is not abrupt but contingencies change in a gradual manner (Gershman et al 2013). While this is often not made explicit, slower extinction may be a consequence of mere differential learning from shocks versus no-shocks, although this phenomenon on its own could not explain renewal, reinstatement and spontaneous recovery. In fact, a recent study demonstrated that replacing noshocks with neutral events improved extinction learning (Dunsmoor et al. 2019). The alternative views that acquisition and extinction are fundamentally different and dissociable processes because of their inherent mutual temporal dependence. Bouton (2002) proposed that extinction is slower because on top of learning the new contingency, the old contingency has to be inhibited and retained in memory. This view has received much support in animal and fMRI studies which identified a the engagement of the ventro-medial prefrontal cortex (vmPFC) in humans and the infralimbic (IL) cortex in animals to be active specifically in extinction. One issue with this views is that it assumes that the agent knows *when* to start learning new associations and start inhibiting the old ones, implying that a change in contingency is automatically understood as *extinction*. This may be the case in scenarios of full reinforcement but in probabilistic environments this is often not the case.

Reconciling these two views, Redish et al (2007) suggested that an agent performs a gradual step-by-step learning initially before identifying different contexts. It has been argued that the main difference between the phases is that they provide a context difference. This is shown in experiments involving relapse of fear memories following passage of time, return of context or presentation of US on its own (e.g.

Bouton and Bolles, 1979). While most research uses externally induced contexts to study these phenomena, it may be argued that there is an inherent contextual component to acquisition and extinction that can be thought of as ‘period of danger’ and ‘period of safety’. Relapse phenomena then may be consequence of perceived context shift in which Arntz’s (1991) argument ‘better be be safe than sorry’ certainly plays a role. Gershman et al. (2013) demonstrated that occasional reinforcement during extinction decreased the chance of relapse, suggesting that if contingency switches less abrupt, the agent will perceive the learning as one context and unlearning will therefore mean a ‘true’ extinction that overrides previous learning.

Therefore, when exploring the lack of extinction, the problem can be broken down to two main components: biased learning (differential learning from shocks and no-shocks) and contextual influence (state learning). In this chapter, I will use a probabilistic learning paradigm to investigate the learning process considering the effect of outcome (i.e. shock/no-shock) as well as phase (acquisition/extinction), delineating these two influences on learning. I will also explicitly test whether the environment is represented as different states and discuss the implications.

Research on the difference of acquisition and extinction has found a number of regions (amygdala, hippocampus, dorsal anterior cingulate cortex, dorso-lateral prefrontal cortex and the insula) to be involved in both phases of learning (Sehlmeyer et al. 2009). While the amygdala is considered a key structure in aversive conditioning (Phelps et al. 2004b), a recent meta-analysis of neuroimaging studies in humans failed to find consistent amygdala engagement in acquisition and extinction and instead highlighted a role of the dorsal anterior cingulate cortex (dACC) in both phases of learning (Fullana et al. 2018). A recent literature review highlighted the connectivity between the dACC and the amygdala as the critical aspect of fear circuit governing

fear learning (Sevenster, Visser, and D'Hooge 2018). Another neural structure consistently involved in acquisition and extinction is the hippocampus (Knight et al. 2004). FMRI studies examining the context-dependent recall of extinction in humans point to an important context-related function for the hippocampus (Kalisch et al, 2006; Milad et al, 2007). A recent study by Marek et al. (2018) demonstrated how contextual information from the hippocampus leads to relapse of fear.

Although many brain regions are engaged in both acquisition and extinction, the ventro-medial prefrontal cortex (vmPFC) has specifically been implicated in extinction (Phelps et al. 2004b; Schiller, et al. 2008). The role attributed to the vmPFC (or mPFC in some studies) is the inhibition of existing fear associations in the amygdala (Sevenster, Visser, and D'Hooge 2018). This notion has been supported by number of fMRI studies, but vmPFC engagement has also been interpreted as a safety or context signal rather than an inhibitory signal. Reviewing wider research into the role of vmPFC casts doubt on the idea that this region is involved in the inhibition of fear memories. Instead, vmPFC was linked to expected value representation (Chib et al. 2009; McNamee et al. 2013) and more recently to the representation of the current state (Schuck et al. 2016). The latter one can be of particular relevance because a switch from acquisition to extinction constitutes the switch of a context that can be internally represented as a different state. If the vmPFC is in fact involved in state representation it should process error information that only agent *aware* of a state can. To test whether a given region is involved in state learning, a computational model of state learning must be defined and single-trial predictions from such model extracted. I will operationalize this by fitting the state switching model described in Chapter 2 to the data and extracting state prediction errors (SPEs) for each trial.

I will use the fMRI imaging and computational models to investigate the nature of information processing in regions previously implied in aversive acquisition and extinction. Furthermore, I will test the idea that the problem of acquisition/extinction is fundamentally a problem of state (i.e. structure) learning.

High trait anxiety has been highlighted as a key vulnerability factor in the development of anxiety disorders. It is often associated with an inability to dissociate between danger and safety (Kindt and Soeter, 2014) due to lack of fear inhibition, sub-optimality of tracking environmental volatility (Browning et al. 2015) or increased reactivity to fear cues (Indovina et al. 2011). Particularly, the study of Browning et al. (2015) suggests that highly anxious individuals focus on recent outcomes, failing to learn the underlying structure of the task. In the context of extinction, anxiety has been associated with decreased extinction learning and higher chance of relapse. Neural studies into the role of anxiety in fear conditioning and extinction identified a positive correlation between amygdala response to CS+>CS- and trait anxiety (Indovina et al. 2011; Sehlmeyer et al. 2011). This finding is often associated, or directly based on, the difference between acquisition-extinction GSR response. Most early studies index conditioning as GSR only, and subjects were excluded if they showed no distinction. This may be one of the reasons why a recent meta-analysis failed to find a significant association between the amygdala and conditioning.

Previous studies have tied the influence of anxiety on differential learning to activity in the dACC, albeit with varying findings. While some studies found a negative association between phase difference and trait anxiety (Sehlmeyer et al. 2011), arguing that the ACC plays an inhibitory role in extinction, others have found a positive correlation (Barrett and Armony 2009), assigning the ACC similar role as to the amygdala.

The main goal of the first study is to investigate and characterise the difference between aversive acquisition and extinction at the behavioural, peripheral-physiological and neural level, and to explore the role of trait anxiety in aversive learning. To this end, I will employ a probabilistic learning paradigm while recording GSR and fMRI data and use computational modelling to understand the information processing occurring

during aversive learning. Based on evidence from previous research I hypothesise that extinction will show significant differences from acquisition in all employed measures, in particular, that extinction will be incomplete, slower and that it will start later. This phase difference is hypothesised to be reflected prefrontal regions previously implicated in extinction processing (dlPFC, VMPFC). Furthermore, I predict that while differential treatment of shock and no-shock outcomes plays an important role, there is a fundamental state-dependent difference between the two phases driven by different internal contexts which will be reflected in the behavioural and modelling data. Regarding the influence of trait anxiety, I hypothesise extinction in high trait anxious individuals will be slower and less complete than in low trait anxious participants due to high trait anxious individuals' inability to learn the higher-order structure of the task. Neural correlates of this groups differences will be explored.

Methods

Participants

Participants were recruited using poster advertising and local recruitment site of the Department of Experimental Psychology, University of Oxford. All participants took part having given full informed consent, and the study was approved by the local research ethics committee. Participants were right-handed, displayed normal pain thresholds at the site of stimulus application and had no history of neurological or psychiatric disease or chronic pain.

Thirty three participants completed the study. One participants was excluded due to technical failure of equipment and one due to an incidental finding in the MRI scan. Data from four additional participants were discarded because they failed to meet our criterion for differential learning which was defined as significantly (as tested

by paired t-test) higher mean probability ratings for the harmful than the safe cue. In total, 28 participants were included in the final behavioural analysis. Further 5 MRI data sets had to be excluded due to poor data quality (3 due to incomplete MRI data acquisition, 1 due to problems with normalization and 1 due an incidental finding). The final MRI data set consisted of 23 participants (15 female, mean age: 25.4 years).

Experimental Design

This study used the basic version of the probabilistic aversive learning task described in Chapter 2. The two contingency levels were set to 75% and 25%. A 'safe' cue was always reinforced at 25%, a 'harmful' cue was reinforced on 75% of trials and contingencies of the reversal cue switched between the two levels every 30 trials (+/- 2). Over four scanning sessions within the same visit, each participant completed 8-9 phases to ensure that they completed 3 acquisition and 3 extinction phases at minimum. If the first phase was an extinction phase (i.e., the reversal cue was reinforced in 25% of trials), it was excluded from the analysis as we were only interested in responses to this reinforcement schedule following an acquisition phase (i.e., when the reversal had been reinforced in 75% of trials). Each participant had at least one 4th acquisition or extinction, some had both. This was due to the probabilistic nature of the contingency switch for the reversal cue which occurred at 75% probability. When data are split by half in the analysis section, they are split by the second scanning break. The scanning breaks coincided with the contingency changes of the reversal cue, every scanning session included 2-3 phases (randomly assigned). As described in Chapter 2, participants provided shock probability ratings on each trial. To obtain a second measure of cue-specific value, we also assessed cue preference in 5% of trials. On these trials, participants were presented with two of the three cues and were prompted to indicate which cue they prefer by pressing one of two buttons, without a time limit. These preference trials were excluded from the main analysis and

they were analyzed separately.

The duration of each trial was slightly longer than in my other studies (see Chapters 4 and 5) due to MRI timing requirements. The inter-trial-interval was jittered between 3 to 5 seconds and the inter-stimulus interval between 2 and 4 seconds. An additional jitter was introduced because participants' shock probability ratings were self-paced and could take between 0 and 4 seconds.

Experimental Procedures

All eligible individuals who had expressed an interest in taking part in the study were invited to a short introductory session during which all procedures including the task and electrical stimulation were explained to them and questions were answered before they signed the consent form. During this session, participants also completed a practice run of the task with 60 trials to test if differential learning could be established successfully. Participants were instructed as follows: "On each trial, you will see one of three abstract pictures in the centre of the screen that may be followed by an electrical stimulus applied to the back of your left hand. Whether an image is followed by an electrical stimulus can vary between pictures: some may predict a stimulus whereas others may signal that no stimulus will be delivered. Finally, pictures can change from being predictive of the electrical stimulus to predicting that no stimulus is delivered and vice versa. Please pay attention to the picture cues and whether you can tell from seeing these pictures if a stimulus will be delivered. On each trial, you will be prompted to indicate by using a slider on the screen how likely you think it is that the picture will be followed by electrical stimulus."

As in the actual experiment, they had to provide shock probability ratings on each trial. They were only invited to take part in the actual experiment, if their mean ratings for harmful cue was significantly higher than their rating for the safe cue.

The actual experiment that was performed 1 - 3 days after the practice session, involved performing the full task in the MRI scanner during simultaneous recording of

skin conductance responses and pupil dilation as well as brain responses using fMRI. Upon the participants' arrival, a second consent form was obtained. All participants were screened for MRI safety by a trained radiographer at the Wellcome Trust Center for Integrative Neuroimaging. Participants were then asked to complete the STAI-TRAIT questionnaire before they were positioned in the scanner and reminded of the task instructions. No information was given about the number of cues, the contingencies or changes in contingencies. Importantly, it was stressed that while the scanner has to be stopped at random points to re-initiate the scanning protocol and to re-calibrate the intensity of the stimulation, they should treat the entire task as one continuous experience, as during the practice session. Upon completion of the task, a structural and diffusion tensor imaging (DTI) scan were acquired. After leaving the scanner room participants had to rate each cue for mean shock probability, likeability and visual appeal. Participants were also asked whether any of the cues changed their probability of being followed by shock at any point. Lastly, participants were informed about the full nature of the task and debriefed.

Analysis of Behavioural Data

Behavioural data were analyzed using custom scripts in Matlab and R. Prior to all analyses participants were assessed on whether they could learn the difference between safe and harmful cue (using a t-test). In this study, the following measures were used: mean reported shock probability per condition (cue type, phase, half), model-free learning rates, switch point, switch steepness and surprise. All measures except for the Shannon surprise were described in Chapter 2.

Surprise was assessed by calculating Shannon surprise for each trial (equation below). As discussed earlier, this measure correlates with prediction error. The main difference is that Shannon surprise is exponentially larger as prediction errors get larger.

Eq. I $$s_t = -log|O_t - (1 - P_t)|$$

Data were analyzed by investigating the difference between cues and between acquisition and extinction. To test for any changes over time, split by half was introduced. In measures occuring at the outcome time, split by outcome type (shock/noshock) was appropriate. Where continuous measures such as trait anxiety were obtained either a median split was performed to binarize the group, or, where appropriate Pearson's correlation was used.

GSR Analysis

GSR data were analyzed using the GLM and DCM methods described in Chapter 2. GLM provides a per-condition beta estimate and can only be used for evoked responses, not for anticipatory signal, while DCM provides a single-trial measure of anticipation. The PsPM 4.02 toolbox (Bach 2014) was used for both analyses in this study.

General Linear Model (GLM)

I used the GLM method described by (Bach and Friston 2013). The GLM used in this study included 76 regressors. At cue onset, I modelled cue type (3 levels: harmful, safe, reversal), phase (2 levels: acquisition, extinction), half (2 levels: first and second half of the experiment), current outcome (2 levels: shock or no shock) and previous outcome (2 levels: shock or no shock). At the time-point of outcome, I modelled cue type (3 levels: harmful, safe, reversal), phase (2 levels: acquisition, extinction), half (2 levels: first and second half of the experiment) and current outcome type (2 levels: shock or no shock).The mean activation in each session was included as additional regressors. Shock events were modelled at the time of shock delivery.

No-shock event were modelled at the time of cue offset, that is, 2s after the shock would have occurred if this were a shock trial.

The design matrix was convolved with a canonical SCR function which is part of the PsPM 4.02 toolbox. Estimated betas were then analyzed in the same manner as behavioural data, averaging per condition, using parametric statistics.

Dynamic Causal Modelling (DCM)

The DCM method used has been described by (Bach, Daunizeau, et al. 2010). The DCM non-linear model was used to estimate the amplitude of the anticipatory GSR response during the inter-trial-interval, i.e., after the participant had provided their shock probability rating until 0.5s prior to outcome delivery. The model uses variational Bayes optimisation to estimate the mean amplitude during each trial. This method therefore provides a single-trial estimate that can be analyzed by averaging per condition or used as a cost function in model fitting.

fMRI Analysis

The MRI data acquisition and analysis pipeline is described in Chapter 2. The preprocessed data were analyzed using the SPM12 package. MarsBar toolbox was used for region of interest analysis.

The data were analysed using two types of models: i) a basic model including all events and event types without any parametric modulation; and ii) a model with a parametric modulator at the time of outcome (see below for details).

Basic Model

The basic model had a separate regressor for a combination of each levels of the following variables: session (sessions 1-4), cue type (harmful, safe, reversal), half (first and second half) and phase (1-2; acquisition/extinction) at the cue onset time; and

session (sessions 1-4), cue type (harmful, safe, reversal), half (first and second half), outcome (shock, no shock) and phase (acquisition/extinction) at the time of outcome, plus 4 regressors modelling the mean for each session, resulting in a total of 148 regressors. Cue onset was modelled as a box function lasting from cue onset to response. Outcome was as modelled using a stick function.

Once first level models were estimated, contrasts of interest were specified and calculated. For ROI analyses, the model was re-estimated for each region of interest (see below). The resulting betas were then baseline-corrected by subtracting the mean beta value for a given session. The resulting data were analysed by averaging betas in their respective conditions and calculating parametric statistics to test for within group differences.

Exploratory whole brain analyses was also performed. To this end, contrast images were entered into one-way t-test for the second (group) level analysis. The resulting t-scores, and their corresponding z-scores, were used to report group level results.

Parametric Models

Five models involving parametric modulators were estimated. All included the full set of regressors of the Basic Model but outcome time was modulated by a different parametric modulator (PM). The five following variables were used in this type of analysis: model-free learning rate, shock probability rating, Shannon surprise, Pearce-Hall prediction error (PHPE) and state prediction error (SPE). Model-free learning rates and Shannon surprise were introduced earlier (see methods here and Chapter 2), shock expectation refers to the probability submitted by the participant. The remaining two quantities are predicted by the two competing behavioural models, the Pearce-Hall model and the state-switching model, both introduced in Chapter 2. In the case of PHPE, the quantity is the difference between model prediction P_t and outcome O_t. State prediction error is computed as the difference between the outcome and current

state identity (here computed as a mean state probability) and the outcome $O_t - \mu_{(s_t,t)}$ (see Chapter 2).

Regions of Interest (ROIs)

The following regions of interest for this study were included based on *a priori* hypotheses about their involvement in aversive learning: amygdala (AMY), hippocampus (HIPP), ventro-medial prefrontal cortex (vmPFC), anterior insula (AI), dorsal anterior cingulate cortex (dACC) and dorso-lateral prefrontal cortex (dlPFC). The mask for AI was derived from the AAL atlas (Tzourio-Mazoyer et al. 2002), for hippocampus and amygdala from the SPM Anatomical Toolbox and for vmPFC, dACC and dlPFC from Neurosynth.org, association maps thresholded at $z > 3$. All ROI analyses were performed by overlaying region of interest in the MNI (Montreal Neurological Institute) space with individual subject's gray matter map that was extracted and normalized during segmentation.

Reporting fMRI results

Results of ROI analyses are considered significant at p<0.05. Whole-brain fMRI results are reported at p<.001, uncorrected. All coordinates are reported in millimeters in relation to the origin in the standard MNI space.

fMRI Model Comparison

A recently developed technique by (Soch and Allefeld 2018) uses Bayesian Model Selection (BMS) to perform a voxel-wise analysis of different model likelihoods. This approach is motivated by a well known short-coming of the model-based fMRI approach: estimating multiple correlated parametric modulators at once leads to either estimation failures or unreliable estimates. To directly compare the nature of

information processed by a given voxel, two GLMs can first be estimated separately using the the same set of regressors with the difference only in the parametric modulator. Subsequently, the likelihood for each model can be computed and the models compared. The authors propose several different ways this can be done. In this chapter I used the cross validated log model evidence at the single subject voxel level:

$$cvLME(m) = \sum_{i=1}^{S} log(y_i \mid \theta, m)p(\theta \mid \cup_{j \neq i}, y_j, m)d\theta$$

Eq. II

where y stands for observed data on sample i (one TR) and θ are the model m parameters. The cross-validated version of log likelihood was used to eliminate the effect of the prior. At the group level, the log model likelihood is calculated across the group

$$log\, p(Y_i) = \sum_{i=1}^{N} log\, p(y_i)$$

Eq. III

Where $Y = \{y_1 \ldots y_N\}$ stand for the group data and N for number of subjects. Finally, the fixed-effects Bayesian model selection is applied to calculate the probability for each model given the data:

$$p(m_i \mid Y) = \frac{p(Y \mid m_i)p(m_i)}{\sum_{j=1}^{M} p(Y \mid m_j)p(m_j)}$$

Eq. IV

This results in a probabilistic map of relative model likelihoods called selected model maps (SMM) which are shown in the results section.

This methodology provides a useful tool when comparing the type of information that is processed in a given brain region without estimating all variables as parametric modulators in the same model. In my case, I employed this technique to

compare regions that predominantly process prediction error (Pearce-Hall) and those that mainly process state prediction error.

Modelling Methods

To explore the type of learning participants show in this study, six learning models were included in the model comparison. Each model was specified using custom MATLAB scripts and optimized 30 times using the Bayesian Adaptive Direct Search (Acerbi and Ma 2017). The best fitting version of the 30 runs according to AIC score was included in the model comparison. This procedure ensured that a local minimum was not chosen, increasing the likelihood of true minimum being selected to represent the model. Model comparison was based on mean AIC scores.

The six models included are: 1) Rescorla-Wagner; 2) Outcome-sensitive Rescorla-Wagner; 3) Dynamic model with decaying or growing learning rate; 4) Pearce-Hall with a separate decay term; 5) State-Switching model based on RW; and 6) Surprise-minimising model. All models are described in Chapter 2.

Results

Behavioural Results

Starting Probability

Motivated by identified confound in Chapter 5, I tested whether the starting probability of the reversal cue (i.e., 75% or 25% at the beginning of the task) had an influence on the reported shock probability or learning rates in any of the three cues. Six mixed ANOVAs were performed with starting probability (75% or 25%) as a

between subject factor, and phase (acquisition or extinction) as within subject factors, three (one for each cue) for model-free learning rates and three for reported probability. There was no main effect or interaction of starting probability in any of the tests, although the main effect of starting probability in the safe cue was near significance (p=0.078).

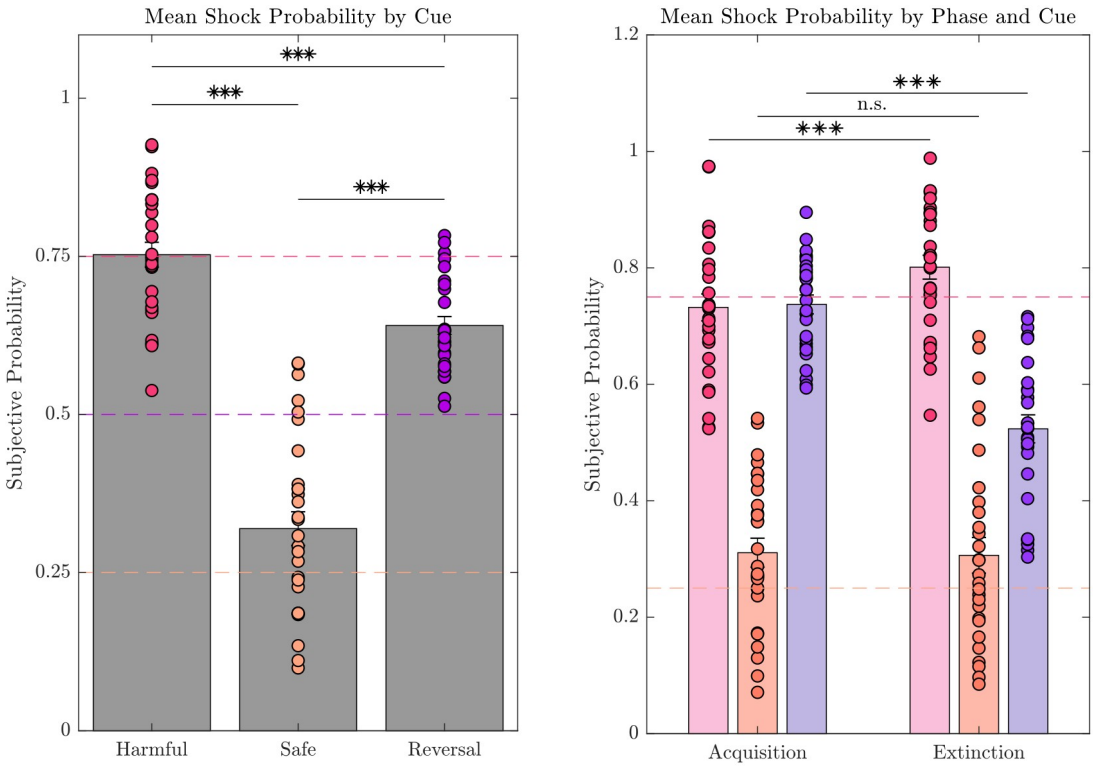
Mean Shock Probability Ratings

To explore general differences between cues, cue-specific subjective probability ratings were compared across phases. A one-way ANOVA found a significant main effect of cue, $F(1, 27) = 100.15, p = 0.000$ with highest ratings for the harmful cue, followed by the ambiguous reversal cue and the safe cue. Bonferroni-corrected post-hoc t-tests found all three cues to be significantly different from each other, $t_{harm>safe}(27) = 11.61, p = 0.000$, $t_{harm>rev}(27) = 4.66, p = 0.000$, $t_{rev>safe}(27) = 9.97, p = 0.000$. Furthermore, using a one-way t-test I tested whether the mean probabilities for the two stable cues differed significantly from the true levels of reinforcement (i.e., 75% for harmful and 25% for safe). Ratings for both cues did not differ significantly from their respective objective probabilities, indicating differential learning from the stable cues. Results are presented in Figure 1a.

I next replicated the above analysis including phase as a within-subject factor. In addition to previous results the test found a main effect of phase, $F(1, 27) = 11.87, p = 0.002$ with higher ratings during acquisition, and a phase*cue interaction, $F(1, 27) = 28.71, p = 0.000$. A series of post-hoc t-tests found the harmful and the reversal cues to differ between phases. The harmful cue was significantly higher in extinction than in acquisition, $t(27) = -4.71, p = 0.000$ whereas the reversal cue was significantly lower in extinction, $t(27) = 5.81, p = 0.000$. No difference was found in the safe cue. For the reversal cue, I next used a one-way t-test to test whether the mean probability differed from the

underling rates of reinforcement. Mean probabilities in the reversal cue did not differ from the true reinforcement in acquisition. However, during extinction, the mean reported probabilities were significantly higher than the true probability, $t(27) = 10.94, p0.000$ indicating that participants overestimated the likelihood of receiving a shock during extinction Results are presented in Figure 1b.

Figure 1 *left panel presents mean probabilities by cue, and right panel mean probabilities by cue and phase. The overlaid scatter plots show individual subject scores. Dashed lines show true reinforcement rates at 25%, 50% and 75%.*



Model-free Learning Rates (MFLR)

Next I analyzed the model-free learning rates. Full set of results in presented in Tables 1 -3. In the harmful cue there was a main effect of outcome, shock learning rate ($\bar{x} = 0.41, \sigma = 0.18$) was higher than no-shock learning rate ($\bar{x} = 0.15, \sigma = 0.10$).

Further, there was a phase by outcome interaction. Post-hoc tests found this to be driven by significantly higher no-shock learning in acquisition ($\bar{x} = 0.21, \sigma = 0.14$) than in extinction ($\bar{x} = 0.15, \sigma = 0.08$), $t(27) = 2.76, p = 0.010$. In the safe cue there was a significant main effect of outcome, no-shock learning ($\bar{x} = 0.34, \sigma = 0.14$) was faster than shock learning ($\bar{x} = 0.23, \sigma = 0.12$). In the reversal cue, there was a significant main effect of outcome, shock learning ($\bar{x} = 0.33, \sigma = 0.11$) was significantly faster than noshock ($\bar{x} = 0.20, \sigma = 0.08$). Furthermore, there was a significant interaction between phase and outcome. No-shock learning was faster in extinction ($\bar{x} = 0.23, \sigma = 0.10$) than in acquisition ($\bar{x} = 0.17, \sigma = 0.08$), $t(27) = -4.82, p = 0.000$ and shock learning was faster in acquisition ($\bar{x} = 0.35, \sigma = 0.13$) was faster than in extinction ($\bar{x} = 0.31, \sigma = 0.12$), $t(27) = 2.29, p = 0.030$ which however didn't survive the correction.

These results show that in harmful and safe cues learning is faster from the more frequent and confirmatory outcome. In the reversal cue, where frequency of both outcomes is matched, there was a bias towards faster learning from shocks which is further modulated by phase, noshock learning was faster in extinction when it's more frequent and informative.

Table 1: *Learning rates in the harmful cue*

	SS	d.f.	MS	F	p
(Intercept)	9.034	1	9.034	240.819	0.000
Error	0.975	26	0.038		
(Intercept):Phase	0.003	1	0.003	0.439	0.514
Error(Phase)	0.164	26	0.006		
(Intercept):Outcome	1.287	1	1.287	60.919	**0.000**
Error(Outcome)	0.549	26	0.021		
(Intercept):Phase:Outcome	0.076	1	0.076	11.685	**0.002**
Error(Phase:Outcome)	0.169	26	0.007		

Table 2: *Learning rates in the safe cue*					
	SS	d.f.	MS	F	p
(Intercept)	9.448	1	9.448	196.055	0.000
Error	1.301	27	0.048		
(Intercept):Phase	0.000	1	0.000	0.009	0.926
Error(Phase)	0.081	27	0.003		
(Intercept):Outcome	0.381	1	0.381	18.118	**0.000**
Error(Outcome)	0.568	27	0.021		
(Intercept):Phase:Outcome	0.003	1	0.003	0.458	0.504
Error(Phase:Outcome)	0.168	27	0.006		
Table 3: *Learning rates in the reversal cue*					
	SS	d.f.	MS	F	p
(Intercept)	8.110	1	8.110	230.461	0.000
Error	0.950	27	0.035		
(Intercept):Phase	0.003	1	0.003	0.921	0.346
Error(Phase)	0.077	27	0.003		
(Intercept):Outcome	0.470	1	0.470	89.154	**0.000**
Error(Outcome)	0.142	27	0.005		
(Intercept):Phase:Outcome	0.073	1	0.073	18.452	**0.000**
Error(Phase:Outcome)	0.107	27	0.004		

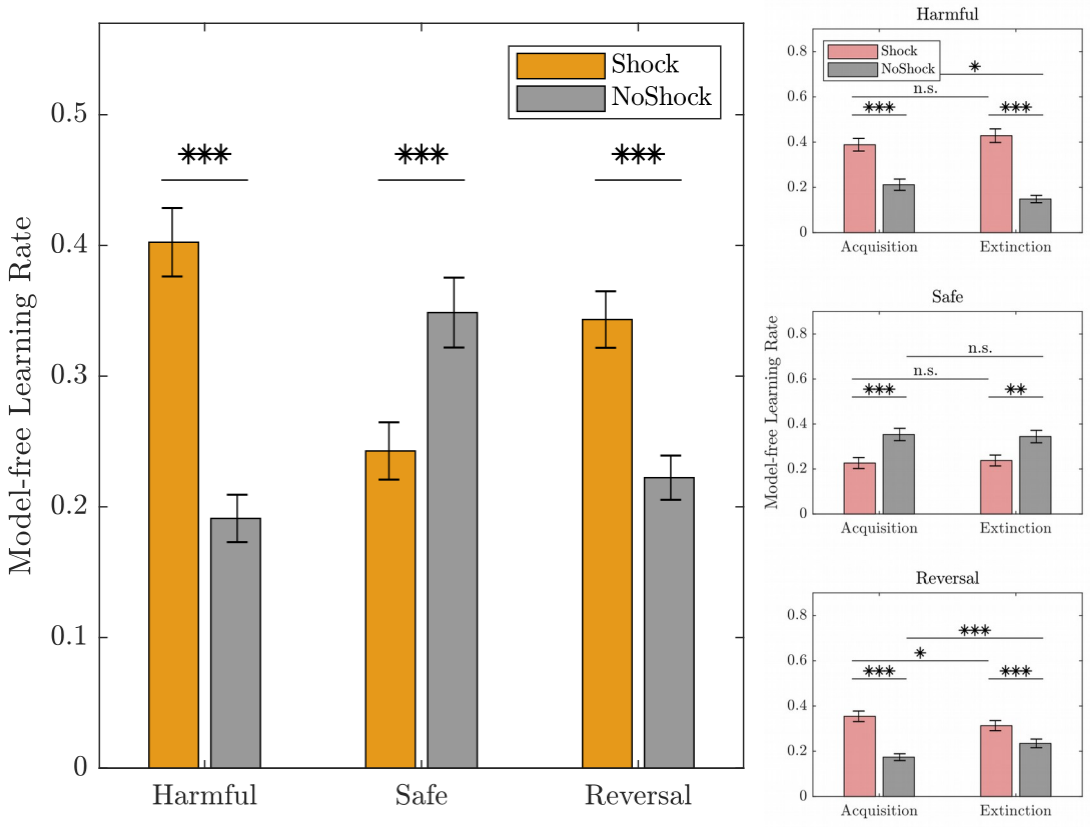
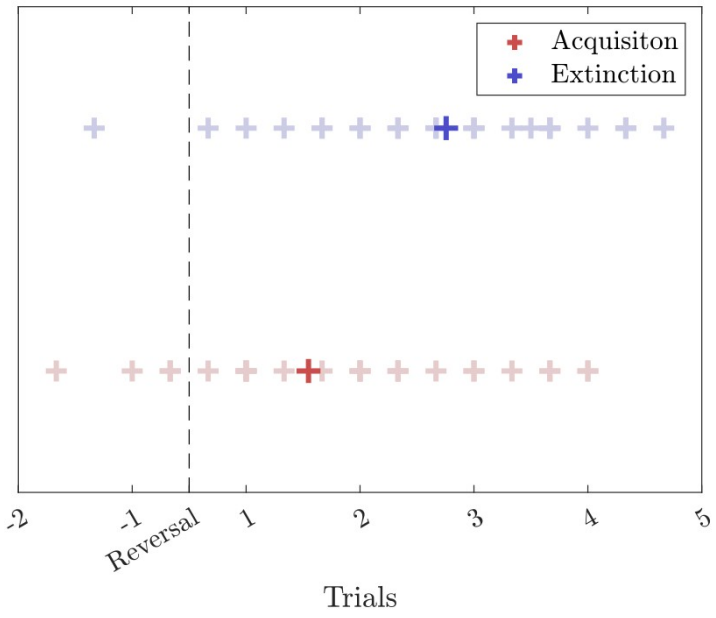


Figure 2 *shows model-free learning rates by cue. Learning in the stable cue was faster for the relevant outcome (shock-harmful and no-shock for safe). Learning from shock was also faster in the reversal cue (left). The right panel shows MFLRs by phase, outcome and cue. Learning rates in the reversal cue are sensitive to the phase.*

Switch Point

Only the reversal cue was included in the switch point analysis. The per-subject averaged switch points revealed a significantly later switch in extinction compared to acquisition, $t(27) = -3.575, p = 0.001$. The data are presented in Figure 3. No difference was found in the steepness of the switch.

Figure 3 *shows the mean and individual post reversal switch points split by extinction (blue) and acquisition (red).*



Changes in Behavioural Measures Over Time

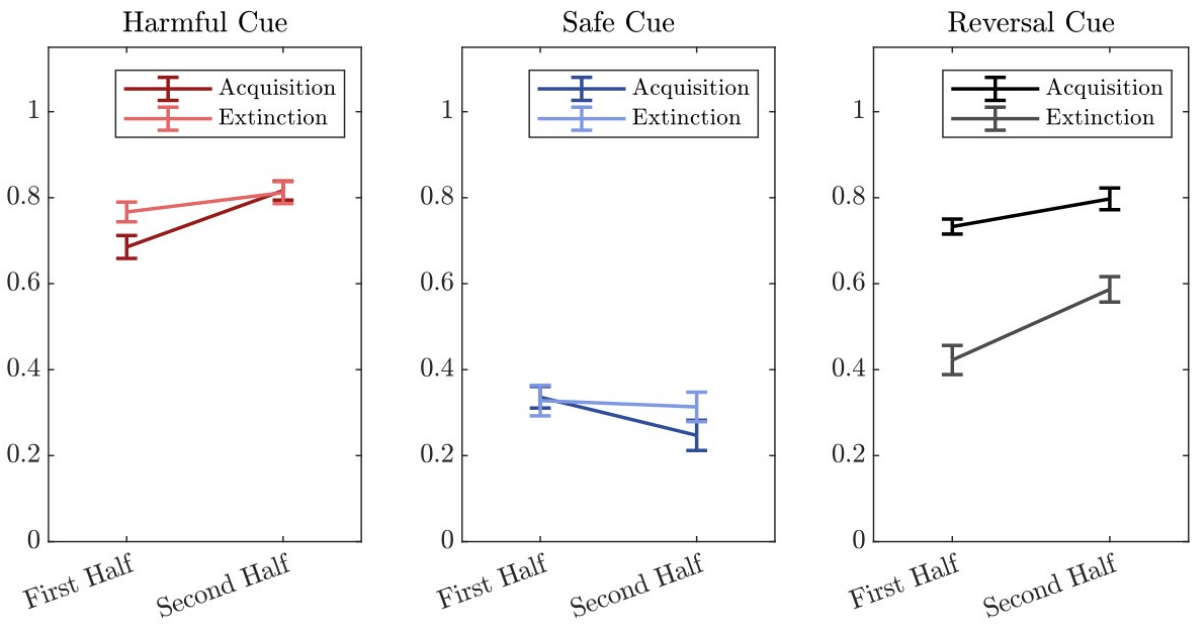
Mean Shock Probability Ratings

To investigate any changes in shock probability ratings occurring over time, the probabilities were split into those acquired during the first half of the experiment and those obtained in the second half and a new variable “half” was introduced to the

per-cue ANOVAs in addition to the factor phase (acquisition, extinction). For the harmful cue, we found main effects of half, $F(1, 27) = 25.17, p = 0.000$ with higher ratings in the second half ($\bar{x} = 0.81, \sigma = 0.14$) compared to the first half ($\bar{x} = 0.73, \sigma = 0.12$) and phase, $F(1, 27) = 9.34, p <= 0.000$ with higher ratings during extinction ($\bar{x} = 0.79, \sigma = 0.12\bar{x} = 0.75, \sigma = 0.10$) than acquisition ($\bar{x} = 0.75, \sigma = 0.10\bar{x} = 0.79, \sigma = 0.12$) , as well as a significant interaction, $F(1, 27) = 9.60, p = 0.000$. Post hoc test revealed that this was driven by increase in acquisition probabilities over time $t(27) = -5.69, p = 0.000$. There was also an increase in extinction did not survive the correction. In the safe cue, there was a main effect of half, $F(1, 27) = 7.84, p = 0.009$ with higher ratings in the first half ($\bar{x} = 0.33, \sigma = 0.14$) than the second half, ($\bar{x} = 0.28, \sigma = 0.16$). Neither the main effect of phase nor the interaction with half reached significance. For the reversal cue, the main effect of half, $F(1, 27) = 15.60, p = 0.000$, higher ratings for the second ($\bar{x} = 0.69, \sigma = 0.12$) than the first half ($\bar{x} = 0.57, \sigma = 0.10$) and phase, $F(1, 27) = 62.26, p = 0.000$, higher ratings in acquisition ($\bar{x} = 0.76, \sigma = 0.09$) than for extinction, ($\bar{x} = 0.50, \sigma = 0.10$), as well as their interaction, $F(1, 27) = 8.43, p = 0.007$ were significant. The interaction was driven by a significant increase in extinction probability between first ($\bar{x} = 0.42, \sigma = 0.13$) and second half ($\bar{x} = 0.57, \sigma = 0.14$), $t(27) = -4.27, p = 0.000$. The increase in ratings for acquisition trials did not reach significance (p=0.03). These results are summarized in Figure 4.

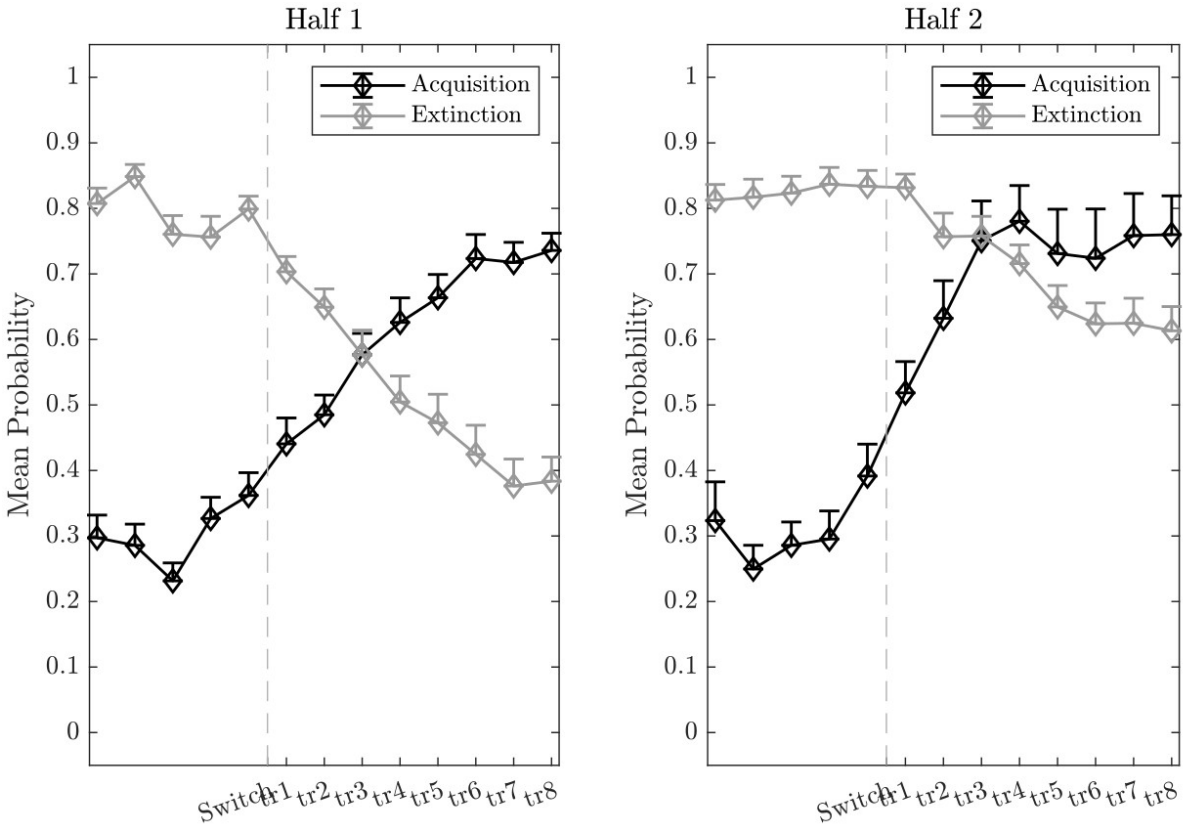
Figure 4 *presents mean shock probability ratings split by cue and half. The mean probability changed over time in all three cues. While participants' ratings of the stable cues became more extreme towards their respective boundaries, ratings for the reversal cue showed a the lack of extinction that was already present in the first half*

but increased in the second half of the task.



To expand on the previous analysis, I next extracted trails for reversal cue and plotted single-trial probabilities locked to the time of the contingency reversal. As depicted in Figure 5, participants show considerably less extinction in the second half of the experiment than in the first half. Furthermore, participants slightly increased their shock probability ratings 1-2 trials before the actual reversal which suggests a certain degree of anticipation in the acquisition phase.

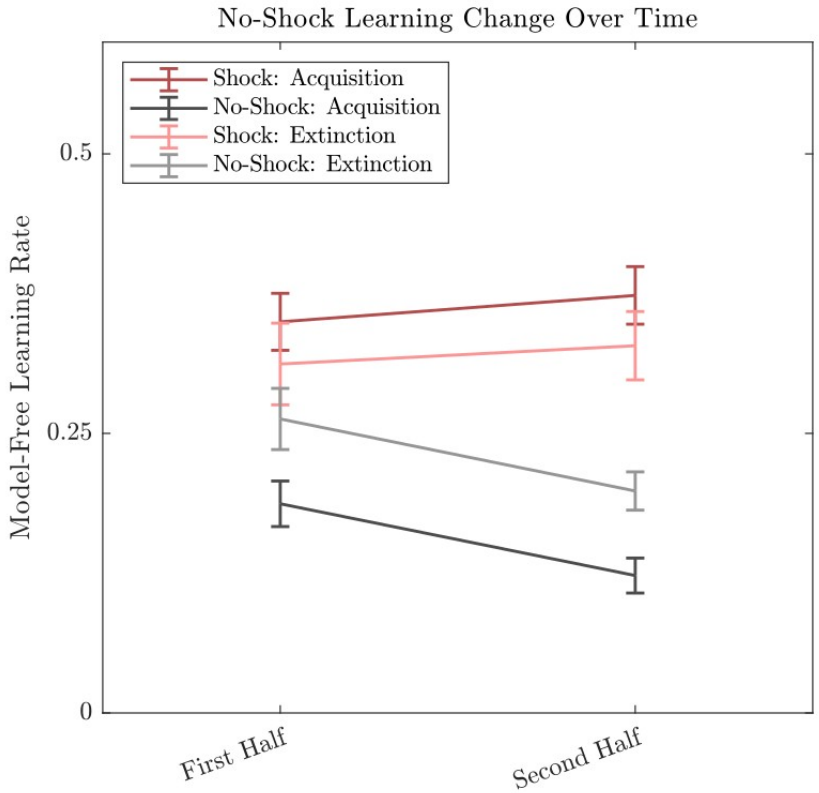
Figure 5 *single trial probabilities averaged across participants locked to time of reversal. While extinction learning was slightly impaired in the first half (left plot) the lack of it became progressively more pronounced in the second half (right plot).*



Learning Rates

To investigate whether the time-dependent lack of extinction is driven by changes in learning rates, I next analyzed model-free learning rates as a function of time, phase and outcome. This ANOVA found a significant main effect of outcome type, $F(1, 27) = 153.04, p = 0.000$ with faster learning from shock than from no shock, and significant interactions between half and outcome, $F(1, 27) = 7.32, p = 0.011$ and phase and outcome, $F(1, 27) = 17.32, p = 0.000$. Post hoc tests found the first interaction to be driven by a decrease in no-shock learning over time, $t(27) = 3.24, p = 0.0031$ and the second interaction by faster no-shock learning in extinction compared to acquisition, $t(27) = -5.41, p = 0.000$. Results are presented in Figure 6.

Figure 6 *shows the interaction effect of phase and outcome type on model-free learning rates. No-shock learning is significantly faster in extinction.*



Surprise

Next, I investigated the amount of Shannon surprise as a function of valence (positive/negative) and time. In the harmful cue, there was a significant main effect of valence, $F(1, 27) = 70.90, p = 0.000$ more positive than negative surprise, and half $F(1, 27) = 4.22, p = 0.049$ (with stronger surprise in the second than the first half), and a half-valence interaction, $F(1, 27) = 12.00, p = 0.001$. Post-hoc t-test found that this interaction was driven by increase in positive surprise over time, $t(27) = -2.87, p = 0.008$ and decrease in negative surprise over time, $t(27) = 5.03, p = 0.000$. In the safe cue, there was a main effect of valence,

$F(1, 27) = 33.09, p = 0.000$, and an interaction between half and valence, $F(1, 27) = 4.27, p = 0.0048$. Post−hoc tests revealed that this interaction was driven in decrease of positive surprise over time, $t(27) = 3.52, p = 0.002$. In the reversal cue, the main effect of valence $F(1, 27) = 74.61, p = 0.000$, half $F(1, 27) = 7.88, p = 0.009$ and their interaction, $F(1, 27) = 25.69, p = 0.000$ were all significant. Post−hoc tests found a significantly higher positive than negative surprise in both halves, $t_{first}(27) = -5.50, p = 0.000; t_{second}(27) = -8.86, p = 0.000$. Positive surprised increased significantly from first to second half, $t(27) = -4.55, p = 0.000$ while negative surprised decreased over the same interval, $t(27) = -4.42, p = 0.000$.

This results suggests that there is a change in experienced surprise in all three cues over time. Importantly, in the reversal cue the there is an increase in positive and decrease in negative surprise over time.

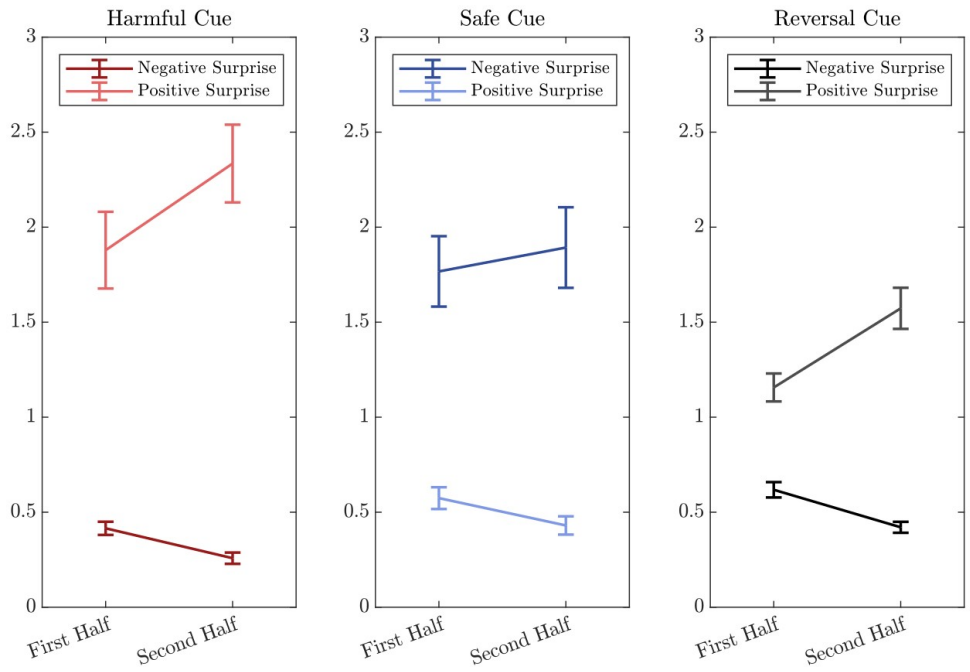


Figure 7 *depicts the experienced level of surprise by cue and half.*

Behavioural Measures: Influence of Trait Anxiety

Mean Shock Probability Raings

I first tested whether trait anxiety influenced mean probability ratings by adding trait anxiety as a between-subject factor to the ANOVAs testing for the effect of half and phase separately for each cue. There was no main effect or interaction of anxiety in any of the stable cues. In the reversal cue, there was a main effect of anxiety, $F(1, 26) = 8.27, p = 0.011$ with higher ratings in low than high anxious individual, and an interaction between anxiety and half, $F(1, 26) = 5.31, p = 0.028$. Post hoc test revealed that there was a significant difference between high and low anxious individuals in the second half, $t_{second}(26) = 4.06, p = 0.003$ while no difference was found in the first half. The data are presented in Figure 8. Additionally, I also correlated trait anxiety scores with the mean shock probability ratings separately for each half and phase, which revealed a significant negative correlation between trait anxiety and mean probability for extinction in the second half, $r(27) = -.47, p = 0.011$. Altogether, the data show that high trait anxious individuals report significantly lower shock probabilities during extinction in the second half extinction, thus approximating the true reinforcement rate more closely.

Figure 8 *mean probabilities split by trait anxiety, half and phase. High trait anxious individuals report significantly lower probability during second half extinction.*

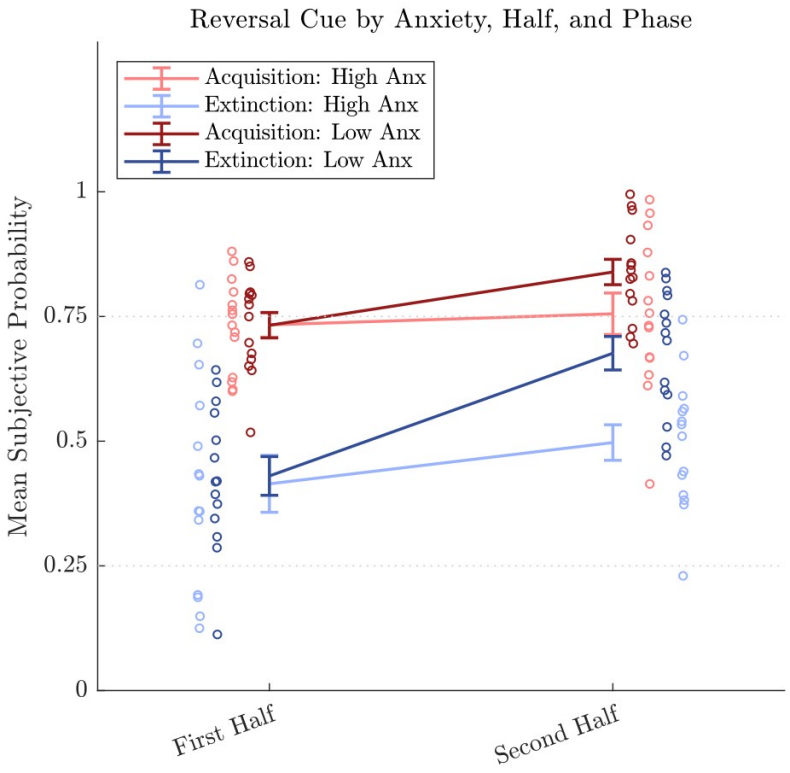
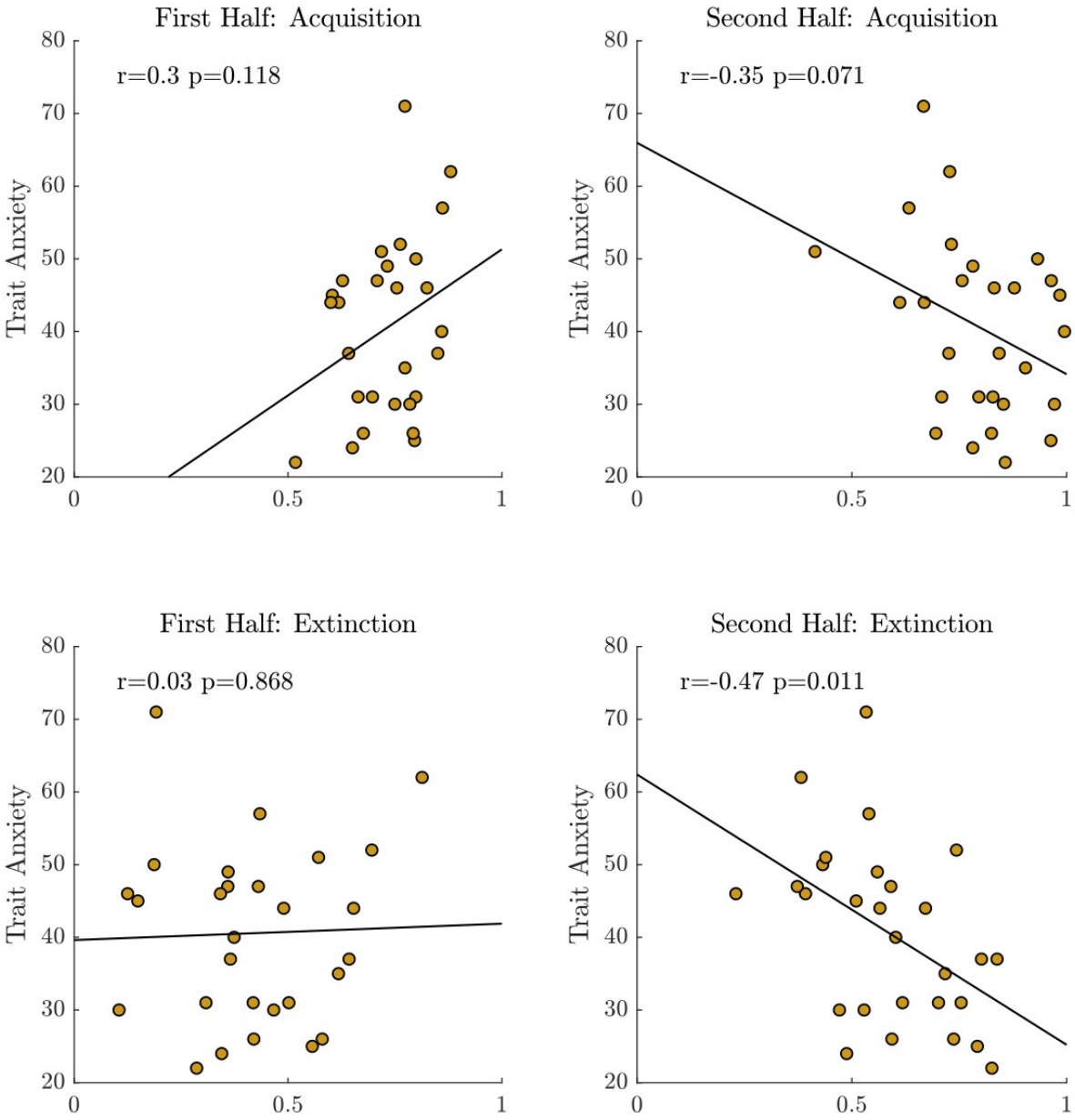


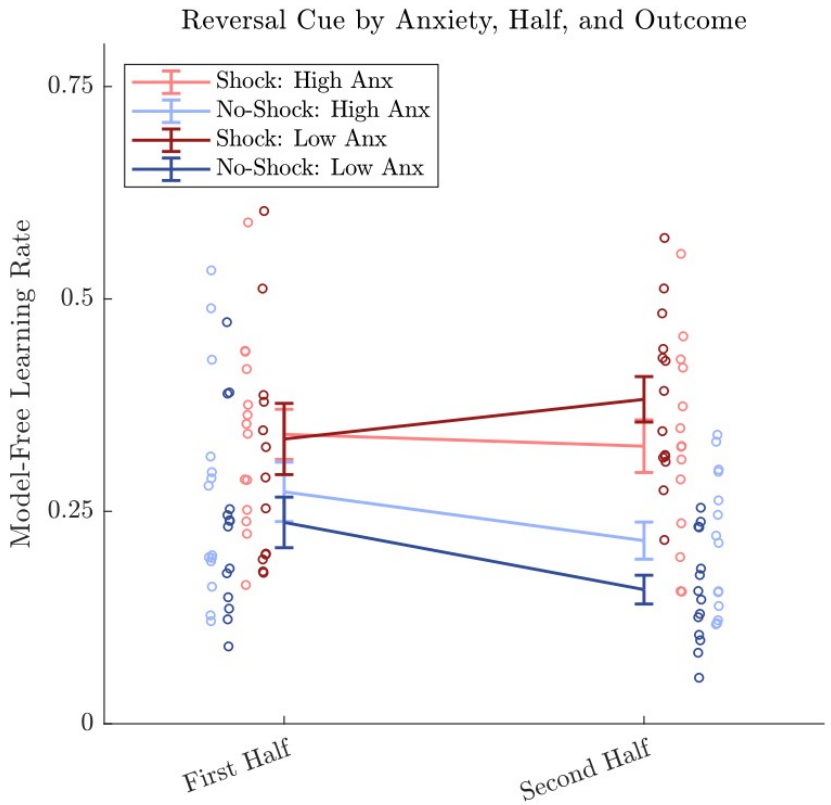
Figure 9, (below) *shows correlations between trait anxiety and mean probability split by phase and half. During extinction trials in the second half (lower right panel), trait anxiety correlated negatively with shock probability ratings, suggesting better extinction in the second half in high trait anxious individuals.*



Learning Rates

Testing whether model-free learning rates are influenced by anxiety, I extended the earlier model by incorporating trait anxiety as a between-subject factor in addition to the within-subject factors half and outcome. The test showed a significant interaction between anxiety and outcome, $F(1, 27) = 8.13, p = 0.008$. A post-hoc test revealed that this was driven by a significantly bigger difference between shock and no-shock in the low compared to the high anxious group, $t(26) = 2.85, p = 0.008$.

Figure 10 *model-free learning rates split by outcome, half and anxiety. The data show a significant interaction between anxiety and outcome.*



Switch Point

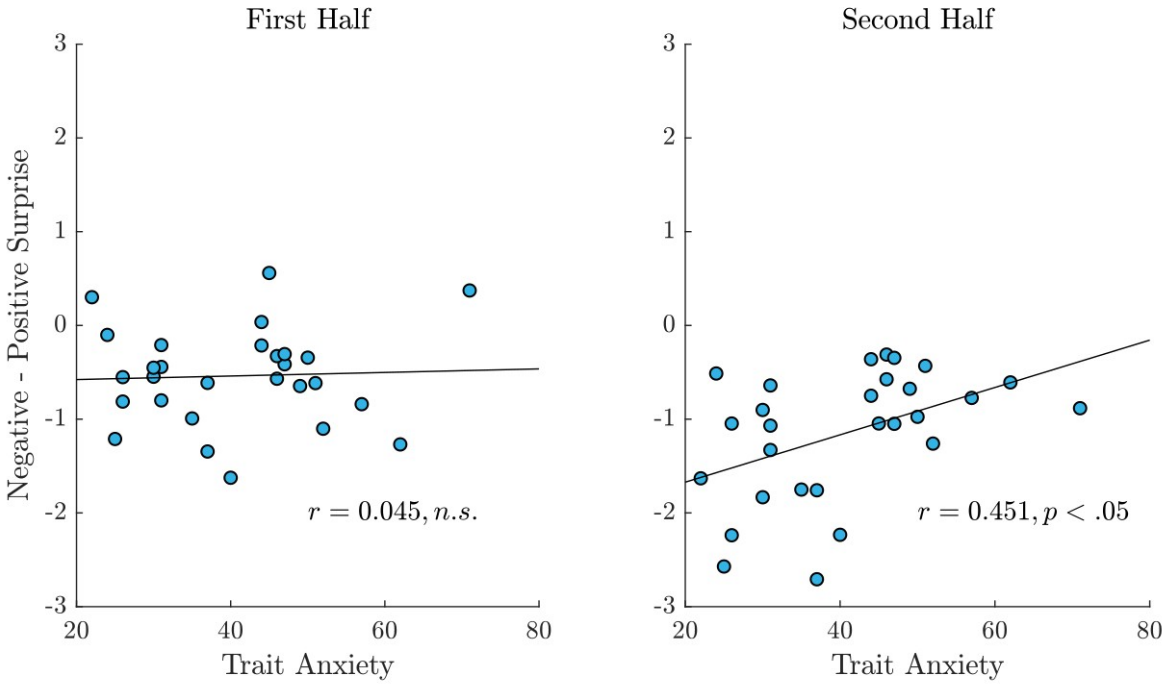
Analysis of switch point and switch steepness found no relationship between anxiety and either of the measures.

Surprise

To test whether there is a relationship between trait anxiety and experienced surprise over time in the reversal cue, the mean positive surprise was subtracted from the mean negative surprise separately for the first and the second half. While there was no relationship between surprise valence difference in the first half, the difference between positive and negative surprise was significantly positively correlated with trait

anxiety in the second half, $r(27) = .451, p < .025$, suggesting that high trait anxious individuals experience more symmetric amounts of positive and negative surprise in the second half while low trait anxious experience more positive surprise than negative surprise (Figure 11).

Figure 11 *The difference between positive and negative surprise split by half as a function of trait anxiety.*



Galvanic Skin Response

The GSR data were analyzed using the more standard GLM approach as well as using the dynamic causal modelling (DCM) approach described in Chapter 2. The two analyses focused on different aspects. While the GLM was used to assess skin responses to cue onset and outcome, the DCM was set up to estimate an anticipatory signal based on the GSR amplitude between response and outcome.

GLM Results

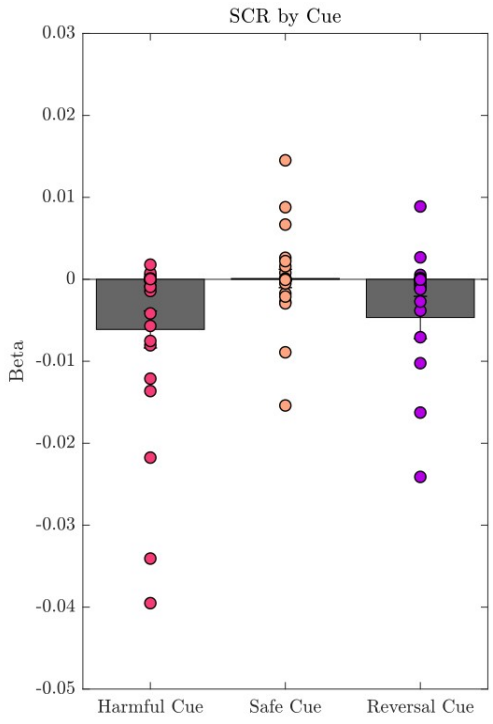
The GLM results are all presented in estimated betas. It is important to remember that negative betas have no direct physiological interpretation and are likely a consequence of latency variation of the SCR evoked by delivery of the shock. They can therefore only be interpreted relative to other conditions, not as absolute measures.

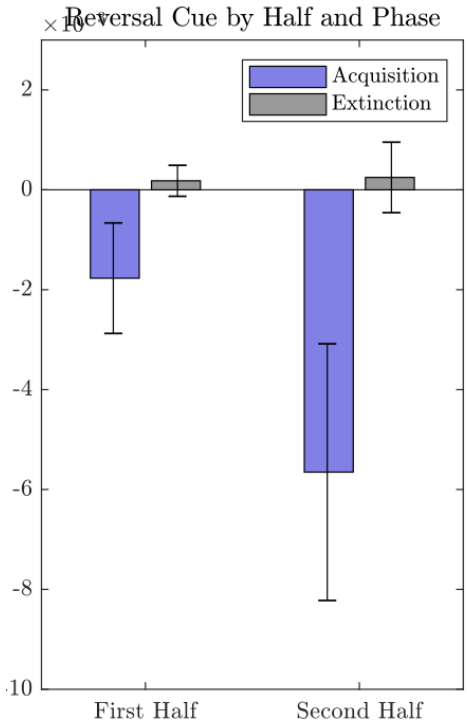
First, I checked whether there was an influence of the previous outcome type on the current trial at the time of cue onset. A t-test found no difference between previous shock and no-shock trials. $t(23) = 0.11, n.s.$. Consequently, all trials were included in the subsequent analyses.

Next, I compared the SCR response to the three different cues. A one-way ANOVA showed a significant main effect of cue, $F(1, 22) = 7.25, p < .01$. Post-hoc tests found a significant difference between harmful and safe cue, $t(23) = -3.20, p < .01$, and between safe and reversal cue, $t(23) = -2.09, p < .05$.

I tested whether there was an effect of phase and half on the reversal cue. A two-way ANOVA found a significant main effect of phase, $F(1, 23) = 9.28, p < .05$ with stronger responses during extinction than acquisition. The Interaction with half was not significant. Plots for this section are presented in Figure 12.

Figure 12 *The left panel shows mean SCR betas by cue; the right panel shows the mean beta SCR response for the reversal cue, split by half and phase.*





DCM Results

The DCM analysis resulted in a single-trial measurement of anticipatory amplitude. I tested whether there was a main effect of previous outcome. This was not significant ($p>0.2$).

I next examined the relationship of GSR amplitude with the reported probability, model-free learning rate, optimal Bayesian value and volatility. The Bayesian single-trial values were predicted using a canonical optimal Bayesian learner developed by Behrens et al. (2007). Overall, there was a very weak relationship between single-trial amplitude estimates and any of the examined variables. The strongest correspondence was found for the Bayesian value where 10 out of 69 measurements correlated significantly with GSR amplitude, 7 positively.

Table 4: *Number of participants with a significant correlation between GSR amplitude and value, learning rate, Bayesian value and Bayesian volatility. [out of 23 total]*

	Harmful Cue	Safe Cue	Reversal Cue

Subjective Probability	0	0	2 (2 pos)
Model-free Learning Rates	5 (4 pos, 1 neg)	0	2 (2 pos)
Optimal Bayesian Value	2 (2 pos)	3 (1 pos, 2 neg)	5 (4 pos, 1 neg)
Optimal Bayesian Volatility	2 (1 pos, 1 neg)	0	1 (neg)

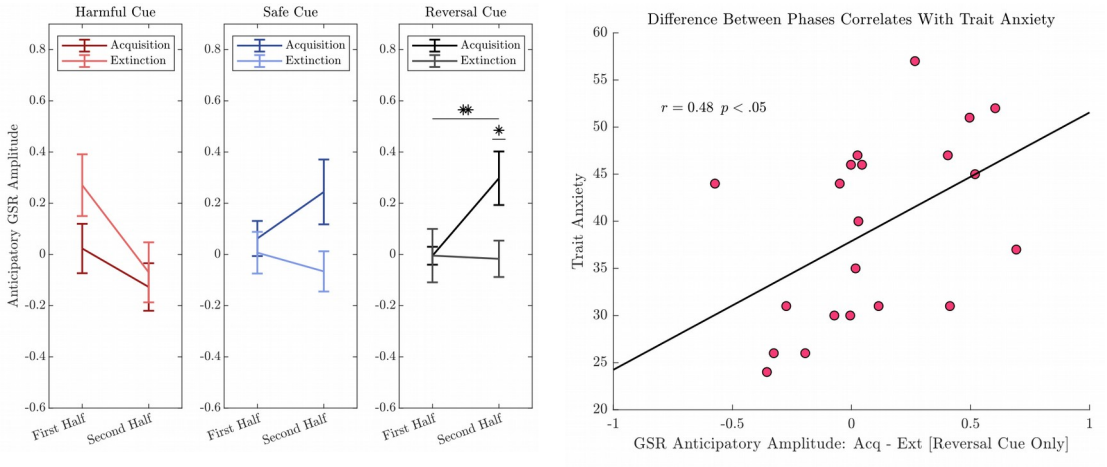
Next, I averaged the data by half and phase separately by each cue (Figure 13). In the full model (cue, half, phase), there was a significant interaction between half and phase, $F(1, 20) = 5.83, p < .05$, which was driven by a significant difference between acquisition and extinction in the second half. Furthermore, there was a significant interaction between half and cue, $F(1, 20) = 4.01, p < .05$, which was driven by a decrease in anticipation of harmful cue from first to second half, $t(20) = 2.31, p < .05$, and an interaction between phase and cue, $F(1, 20) = 4.20, p < .05$, driven by the difference between harmful and safe cues in acquisition, $t(20) = -2.50, p < .05$. I next split the analysis by cue. In the harmful cue, there was a main effect of half, $F(1, 20) = 5.31, p < .05$, anticipatory amplitude was significantly lower in the second half. In the safe cue, there was no main effect or interaction. In the reversal cue, there was a significant interaction between half and phase, $F(1, 20) = 4.40, p < .05$. A post-hoc test identified that this was driven by increase in anticipatory amplitude in the second half acquisition, $t_{sec>first}(20) = 2.811, p < .05$, $t_{2^{nd}acq>2^{nd}ext}(20) = -3.11, p < .01$.

The same model was extended by adding anxiety. There was a main effect of anxiety, $F(1, 20) = 6.27, p < .05$, the low trait anxious individuals had overall higher anticipatory GSR amplitude than low anxious individuals.

I next subtracted the mean GSR signal in extinction from acquisition and correlated the resulting quantities with trait anxiety. Similarly to the results for shock probability ratings, trait anxiety was found to be positively correlated with the difference

in anticipatory GSR between acquisition and extinction, $r(20) = .48, p < .05$.

Figure 13: *The left panel shows a significant increase in anticipatory GSR in the second half acquisition phase of the reversal cue. The right panel shows a significant correlation between mean phase difference and trait anxiety.*



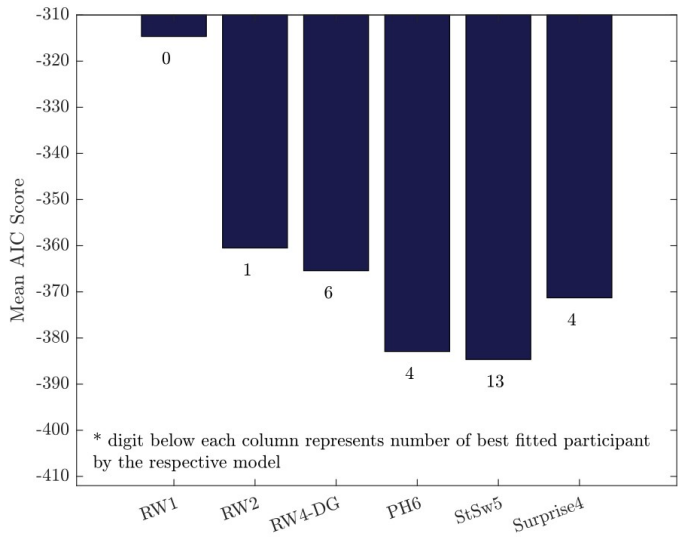
Modelling Results

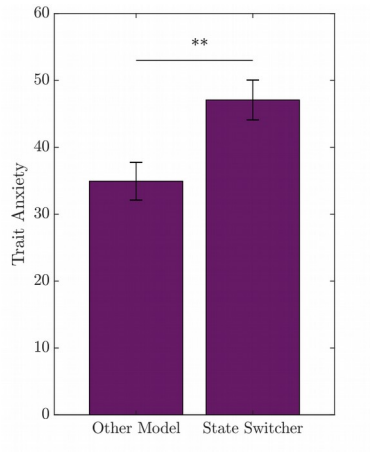
The model comparison results are presented in Figures 14 (left) and 15. Models were ranked from 1 to 6 within participant based on their respective AIC scores with the best fitting model being ranked as 1. The overall rank is based on the mean AIC scores.

The state-switching (StSw5) model had the lowest mean AIC score of -384.6. It was also the best fitting model in most subjects (13/28). The second best model was the Pearce-Hall learner with independent decay term (PH6) which had a mean AIC score of -382.9. All other models significantly trailed behind the two first models. The

state switcher either fitted best or it was 3rd/4th (i.e. it rarely was the second best fit) while the Pearce-Hall model ranked very highly in most participants (Figure 15). This suggests that the state switcher captures a particular pattern of behaviour that can clearly be found in some but not in other individuals. Since a number of behavioural measures had been shown to differ depending on trait anxiety, I compared anxiety scores in subjects best fitted by the state switcher and in subjects best fitted by another model. Somewhat surprisingly, there was a significant difference, $t(26) = -2.96, p < .01$ which indicates that the learning in high trait anxious was best fitted by the state switcher had significantly higher trait anxiety (see Figure 14 right).

Figure 14: *The left panel shows the mean AIC scores and number of participants best fitted by each model. The right panel shows mean trait anxiety for participants best fitted by the state switcher versus other models.*





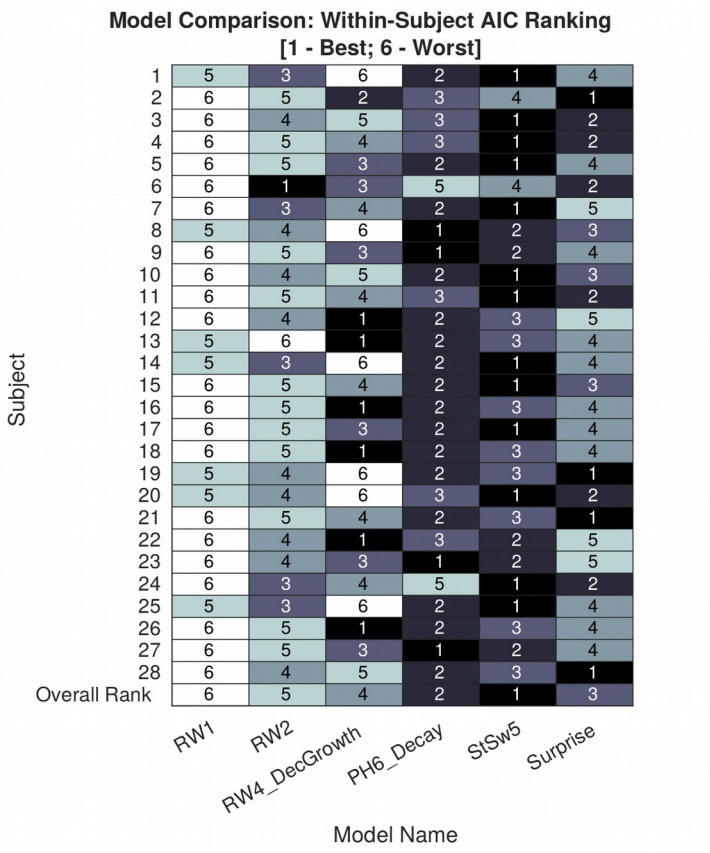


Figure 15 (left): *Model ranking based on individual AIC scores.*

fMRI Results

Data Quality and Sanity Checks

All images were checked for visual artifacts such as wraparound, ghosting and

signal dropout. Statistical checks for fMRi data outliers and biases were performed using the CanLab quality control feature (https://github.com/canlab). Additionally, motion parameters of preprocessed images were estimated and visualised to ensure that motion artifacts were removed by the AROMA ICA tool.

Quality Control

The main four conditions (factorial design by outcome type and half) at the time of outcome were analyzed. In particular the distributions and means of effect sizes of the first level contrast images were compared to i) compare the differences between conditions that might drive the findings rather than a region-specific cognitive process; ii) to identify general condition-invariant tendencies in the data and outliers.

There was an overall tendency for negative effects as shown in Figure 14b. Breakdown by individual condition found that this was driven by mean decrease in the second half. While mean COPE values in the gray matter in the first half were $$\bar{x}_shock = -0.44, SEM = 0.27$$ and $$\bar{x}_noshock = -0.52, SEM = 0.38$$, there was a significant drop in the second half, $$\bar{x}_shock = -1.25, SEM = 0.31$$ and $$\bar{x}_noshock = -1.14, SEM = 0.40$$. Crucially, the difference in effect sizes for shock and no-shock events were not different either in the first or second half. This suggests that our main results of differential time-dependent treatment of shocks and no-socks is not a whole-brain artifact.

Four contrasts were identified as outliers due to large negative effect sizes. These corresponded to subjects 19 and 21 in First Half: Shock condition and to subjects 6 and 8 in First Half:NoShock condition. The data points were retained as no difference between the conditions was observed.

Figure 15a: The figure shows distribution of COPE effect sizes per condition

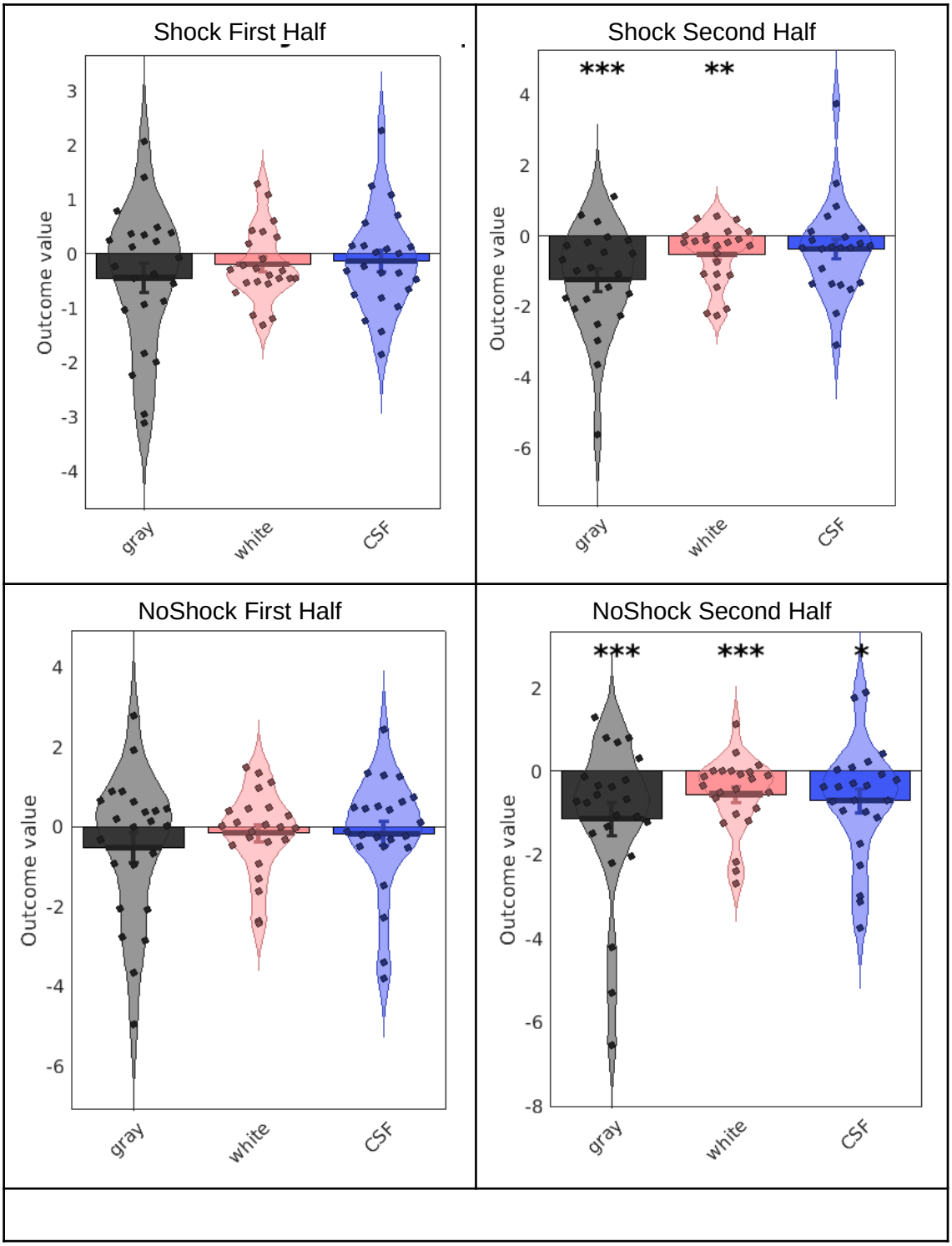
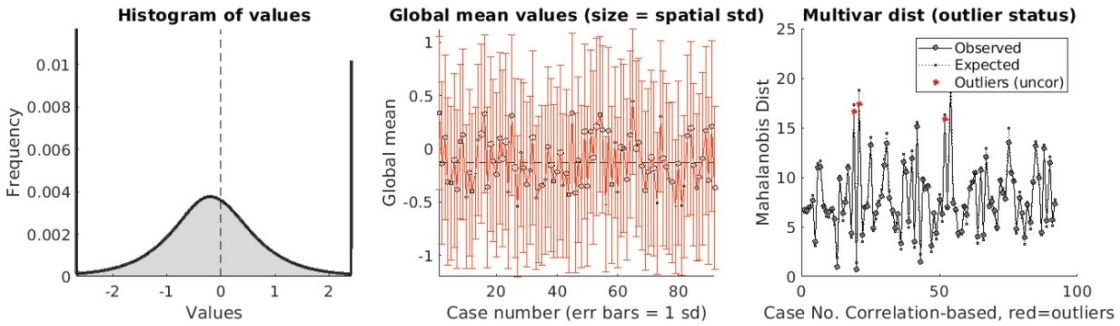


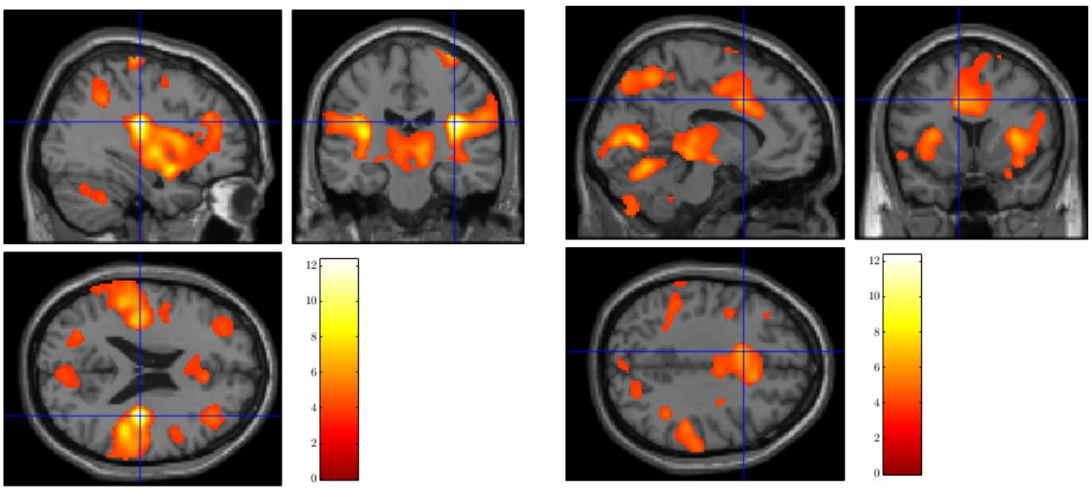
Figure 15a: Distribution of COPE effect sizes across all four conditions (left) and individual subjects (center and right).



Neural Response to Noxious Stimulation

To further check the quality of the data, correct timing, and processing, I tested whether the expected regions are sensitive to shock delivery. To formally define the regions relevant for pain processing, I adopted the Neurologic Signature of Pain (Wager et al. 2013) as a binary mask. There was a significantly higher response in the NPS following shock compared to shock omission, $t(22) = 4.99, p_{uncorr} < .001$. I followed up by running a whole brain analysis on the same contrast. The hypothesised areas for this contrast involved the insula, thalamus, PAG and the somatosensory cortex. All hypothesised regions were found significant at p<.001, uncorrected. The analyses found two large clusters ($N_{k_1} = 28724,\ N_{k_2} = 3590$). First with peak voxel activation at $x = 36, y = -20, z = 22$, $t(22) = 12.29, p < .05,\ FWE$, encompassing the insula, somatosensory cortex, thalamus, PAG and lateral prefrontal cortex. The second cluster with peak activation at x=-10, y=12, z=36, $t(22) = 6.57, p < .05,\ FWE$, included the medial cingulate cortex (MCC).

Figure 16: *Both panels show group level results for shock>noshock contrast [left panel: peak voxel 1; right panel: peak voxel 2], thresholded at p<.001, uncorrected.*



Conditioning Effect

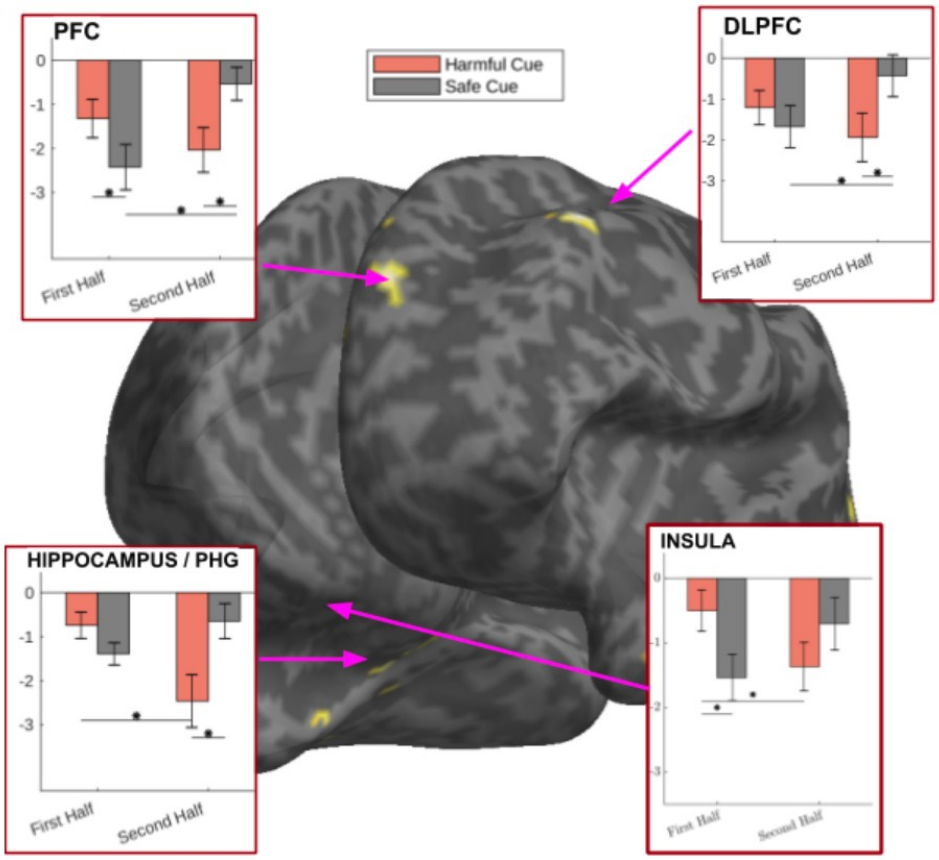
In previous fMRI studies on classical conditioning the amygdala was found to show differential activation in the CS+>CS- contrast (Phelps et al. 2004b; Indovina et al. 2011). To identify regions involved in conditioning, I extended the contrast to account for an *increase* in the (conditioned) differential response to the CS+ (harmful cue) and CS- (safe cue) over time as recommended by (Morriss, Hoare, and van Reekum 2018). Table 5 presents the full set of statistical results for a whole brain analysis for the contrast ($\text{harm}_{\text{first}} - \text{safe}_{\text{first}}$) < ($\text{harm}_{\text{second}} - \text{safe}_{\text{second}}$), at the time of cue onset, thresholded at $p < .0005$ and a cut at minimum al cluster extent of size 10 contiguous voxels.

Table 5: Whole Brain results for Conditioning Sensitivity Map

Region	T-score	Z-score	Cluster Size	P-value	x	y	z
Middle Frontal Gyrus (L)	5.66	4.4	181	0.00001	-40	22	52
Superior Frontal Gyrus (L)	5.09	4.09	187	0.00002	-18	60	30
Fusiform Gyrus (R)	4.54	3.77	99	0.00008	40	-38	-16
Medial Temporal Gyrus (R)	4.51	3.76	66	0.00009	60	0	-32
Brainstem	4.3	3.62	14	0.00014	-6	-36	-22
Occipitotemporal Transitional Region (L)	4.29	3.62	27	0.00015	-50	-62	14
Anterior Insula (R)	4.12	3.51	22	0.00022	34	14	-14
Precuneus (L)	4.1	3.5	16	0.00023	-24	-50	0
Cuneus (R)	4.09	3.49	39	0.00024	8	-80	36
Inferior Temporal Gyrus (L)	4.06	3.47	56	0.00026	-42	-24	-20
Parahippocampal Gyrus (R)	4.03	3.45	17	0.00028	30	-32	-12
Somatosensory Cortex (L)	3.99	3.43	22	0.00031	-38	-4	66
Temporal Pole (R)	3.87	3.34	15	0.00041	42	-16	-24

I inspected the pattern of ROI activity in a priori defined regions. overlaping with a priory hypothesis While the pattern of the dlPFC, PFC and hippocampus showed an increased dissociation of harmful and safe cue over time, the insula showed the opposite pattern: the difference was large initially and decreased in the second half.

Figure 17 *Whole-brain analysis showing an increased dissociation between safe and harmful cues over time (contrast* $harm_{first} - safe_{first} < harm_{second} - safe_{second}$*).: DLPFC (-40, 22, 52), PFC (-18, 60, 30), insula (34, 14, -14) and hippocampus/parahippocampal gyrus (30 -32 -12), Regions were selected from all results to illustrate the pattern of activity.*



Main fMRI Analysis

The main fMRI analysis follows up on the observed behavioural and modelling results with three main questions:

1) Which regions track the change in no-shock learning over time ?
2) Which regions reflect the increasing difference between acquisition and extinction in high trait anxious individuals?
3) Which regions are involved in state prediction error processing (as observed in high trait anxious individuals), as opposed to Pearce-Hall prediction error processing?

1) Time-Dependent Change

Our behavioural analysis of shock probability ratings had revealed a decrease in the difference between acquisition and extinction from the first to the second half of the experiment. To follow up on this observation, I explored which of our ROIs showed a change in differential activation over time. These analyses revealed a significant *increase* in phase dissociation in the right amygdala ($t(22) = -2.30, p = 0.032$), left hippocampus $t(22) = -2.19, p = 0.040$, and bilateral dACC (left: $t(22) = -2.23, p = 0.037$, right: $t(22) = -2.20, p = 0.038$) at the time of cue onset. No region appeared to map onto the behavioural finding in which value difference *decreased* over the two halves.

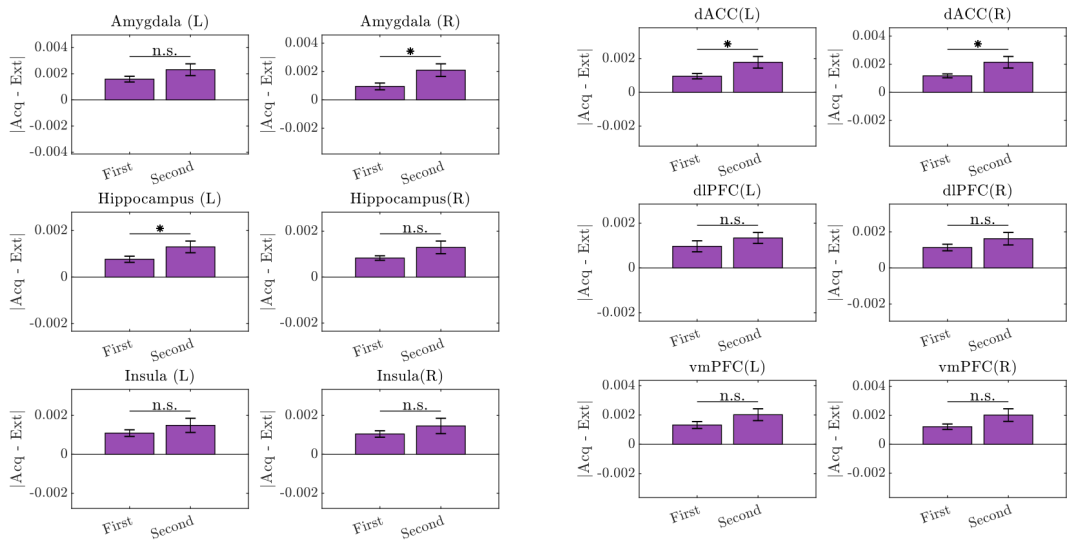


Figure 18: *ROI results on changes in the difference between acquisition and extinction over the course of the experiment.*

My behavioural results showed that the decrease in difference between acquisition and extinction was due to increasing overprediction of shock likelihood in extinction trials which in turn was driven by a decrease in no-shock learning during extinction trials in the second half. To follow up on this behavioural finding, I investigated if any of the pre-defined ROIs showed a change in response to shock omission in the second compared to the first half. Of the twelve investigated ROIs, only the right dlPFC showed a significant increase in response to no-shock outcomes $t(22) = -3.01, p = 0.006$.

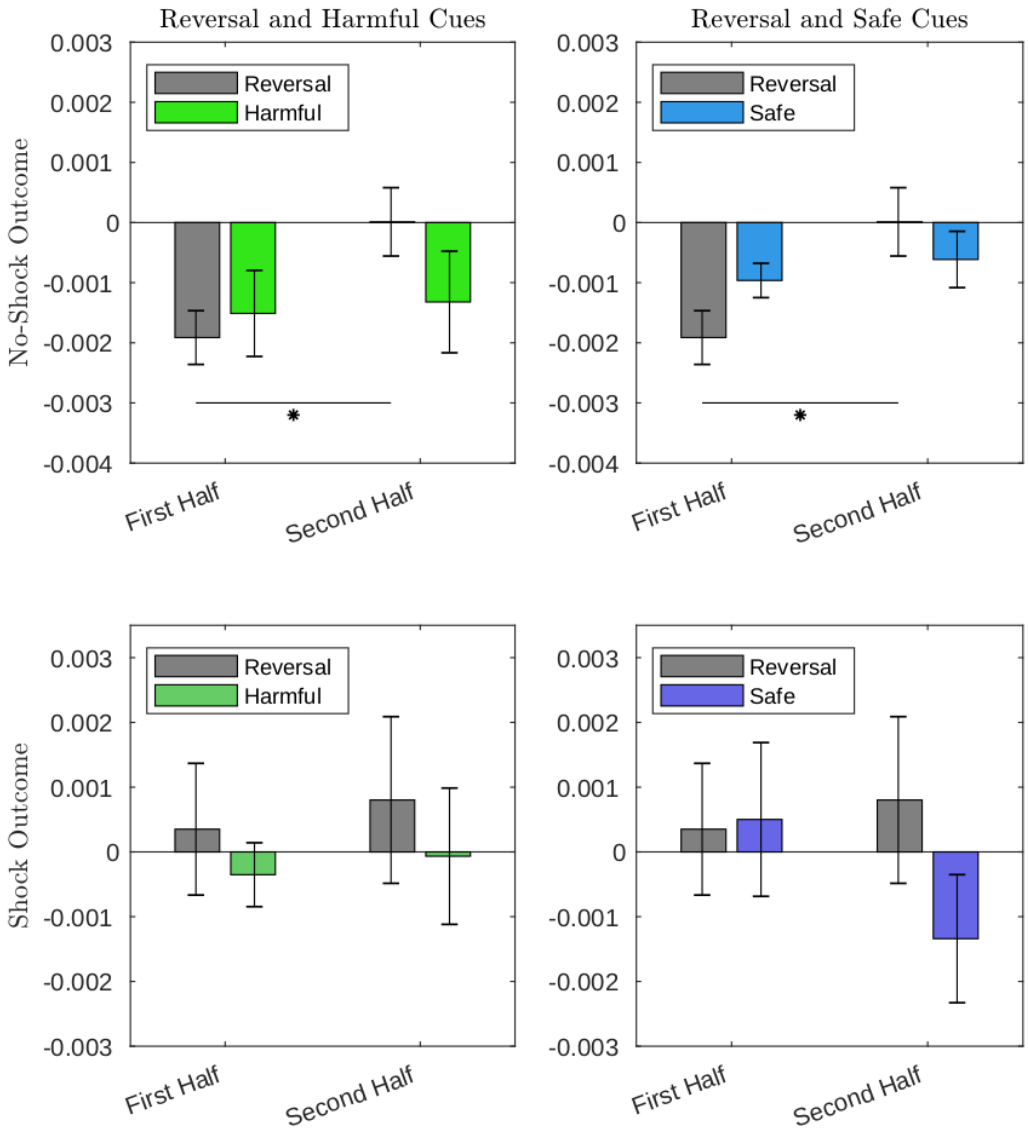
Follow-up analyses

First, I explored whether this pattern of change in the dlPFC was specific to the reversal cue. To this end, I compared the activation pattern during extinction trials in the reversal cue condition to activation changes in the harmful and safe cue conditions. As shown in Fig. 19 (upper panel), the no-shock specific increase in dlPFC activity over time was not found in either of the stable cues,

$t_{harm}(22) = -0.19, p = 0.851; t_{safe}(22) = -0.69, p = 0.496$.

Second, I investigated whether a similar change in dlPFC activity occurred when shocks were delivered. As shown in Fig. 19 (lower panel), this was not the case, indicating that the increase in dlPFC activity was indeed specific to shock omission in extinction trials.

Figure 19 *shows the significant increase in dlPFC activity to shock omission from the first to the second half of the experiment. In contrast, neither of the stable cues showed this change over time.*



For completion, I also ran whole-brain analyses testing for regions that either *increased* or *decreased* their responding to no-shock in the second half. There was no suprathreshold cluster for decreased activation. However, several regions increased their activity in the second half (contrast first<second). Apart from the known cluster in the right dlPFC, a significant increase was found in the middle frontal gyrus, cerebellum and posterior cingulate cortex. The whole brain results thresholded at $p < 0.001$, cluster size $n > 10$, are presented in Table 6.

Table 6 *Whole brain results for*

Region	T-score	Z-score	Cluster Size	P-value	x	y	z
Middle Frontal Gyrus (R) - orbital part	4.24	3.58	20	0.00017	46	52	-4
Middle Frontal Gyrus (R) - dorsal part (part of the right dlPFC ROI)	3.85	3.33	15	0.00044	42	22	40
Cerebellum (L)	3.84	3.32	11	0.00045	-6	-58	-50
Middle Frontal Gyrus (R) - medial part	3.77	3.28	11	0.00053	20	38	28

Model-based fMRI Analysis of the Key Regions

To investigate the nature of processing in the right dlPFC I considered the following variables previously associated with behavioural aspects of the study: model-free learning rate, Shannon surprise, aversive anticipation (subjective probability of shock) and state prediction error based on the state-switching model. For *each* of the four variables I estimated a separate model including the key variable as a parametric variable of outcome. Here, I report mean betas in the region. None of the four tested variables were found to be encoded by the right dlPFC, suggesting that the change of activity in right dlPFC is not related to pain anticipation, surprise, learning rate or state learning.

In order to explore the role of our 12 ROIs, I tested (i) whether their

engagement was related to any of our model-based parameters or the subjective shock probability and ii) whether they showed a change in processing over time. Significant processing was assessed by comparing the distribution of betas for the given parametric modulator against 0 (one-way t-test).

Activity in the right hippocampus scaled with the MFLR $t(22) = -3.55, p = 0.002$, ($\bar{x} = 3.9 * 10^{-3}, \sigma = 5.3 * 10^{-3}$). A more fine-grained analysis showed that this was driven by significant MFLRs encoding in the second, $t(22) = -3.1, p = 0.005$ but not in the first half, $t(22) = -0.77, p = 0.449$ but the difference between halves was not significant. Furthermore, there was a significant overall MFLR encoding in the left vmPFC, $t(22) = -2.11, p = -0.047$ which was stronger in the second compared to the first half in both, left, $t_{right,first>second}(22) = 2.32, p = 0.030$, and right, $t_{left,first>second}(22) = 2.45, p = 0.023$, vmPFC. None of the other three variables was found to be encoded significantly in any of the ROIs.

2) High Anxiety in Extinction

In our behavioural analysis, high trait anxiety had been associated with increased dissociation of acquisition and extinction, both overall (in behavioural and GSR data) and increasingly over time (in behavioural data). To test which neural structure reflects this difference between high and low trait anxious individuals, I correlated a) the difference between acquisition and extinction with trait anxiety; and b) anxiety with the *increase* in phase difference over time (i.e. acq-ext first half < acq-ext second half). From all tested regions only the dorsal ACC showed a significant correlation between trait anxiety and absolute phase difference in both, left, $r(22) = 0.475, p = 0.022$, and right, $r(22) = 0.595, p = 0.003$ ROI. Importantly, activity in the same region reflected the anxiety-dependent increase in phase

dissociability over time (left, $r(22) = 0.419, p = 0.047$, right, $r(22) = 0.499, p = 0.015$)

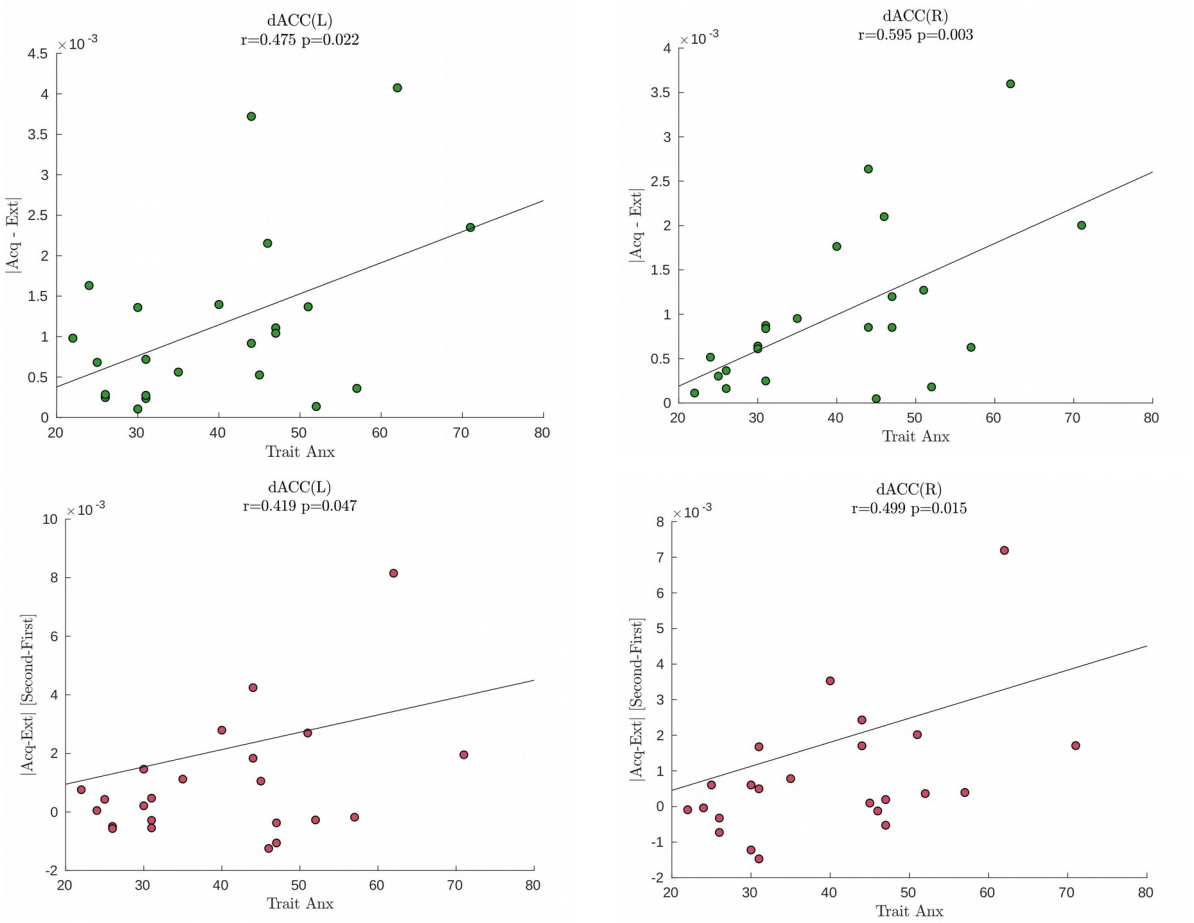


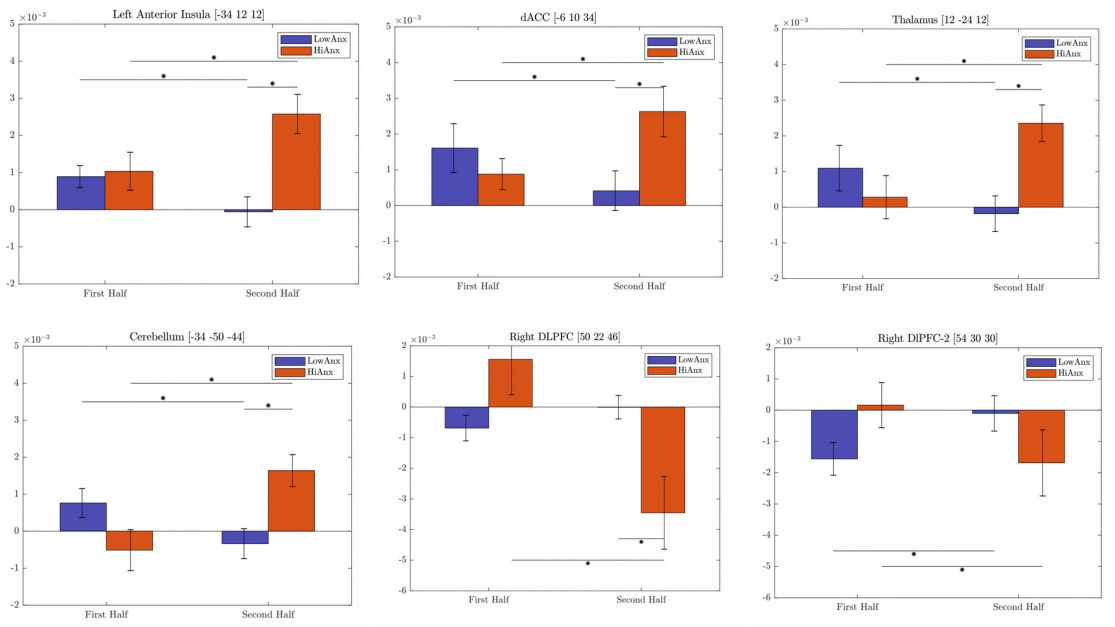
Figure 20 *trait anxiety positively correlates the absolute difference between acquisition and extinction activity in the dorsal ACC (top two panels) and with the increase in acquisition - extinction difference over time (bottom two panels)*

Next, I extended the analysis to the whole brain. Results are presented in Table 7. To explore the pattern of interaction in relevant regions, I next plotted the peak voxel activity (Figure 21).

Table 7: Whole brain results for regions responding more in the second than first half in extinction trials

High > Low Anxiety second > first

Region	T-score	Z-score	Cluster Size	P-value	x	y	z
Cerebellum	5.53	4.3	152	0.00001	-34	-50	-44
Anterior Insula / White Matter (R)	4.68	3.83	21	0.00006	28	26	10
Brainstem	4.52	3.73	23	0.00009	-6	-42	-40
Thalamus	4.44	3.69	21	0.00011	12	-24	12
Inferior Temporal Gyrus (R)	4.43	3.68	17	0.00012	46	-46	-10
Cerebellum	4.36	3.64	13	0.00014	-40	-62	-28
Dorsal Anterior Cingulate Cortex (L)	3.99	3.4	28	0.00033	-6	10	34
Medial Cingulate Cortex (L)	3.96	3.38	20	0.00036	-10	-4	40
Anterior Insula (L)	3.92	3.35	13	0.00040	-34	12	12
Cerebellum	3.86	3.32	14	0.00045	16	-42	-36
High < Low Anxiety second > first							
Middle Frontal Gyrus (R)	4.88	3.94	58	0.00004	50	22	46
Middle Frontal Gyrus (R)	4.37	3.65	33	0.00013	54	30	30
Angular Gyrus (R)	3.97	3.39	21	0.00035	42	-60	52



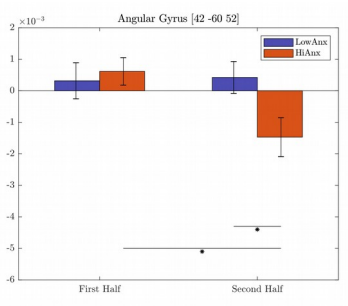


Figure 21 *Patterns of peak voxel activations for the contrast investigating the change in response to the reversal cue in extinction trials as a function of anxiety group. Blue bars represent the low anxious sample while red bars represent high anxious subjects. These were performed to understand what is drives the whole brain contrast.*

Results in this section suggest that primarily the bilateral dorsal ACC is associated with the increased dissociability between acquisition and extinction, and the further increase in the phase difference over time. Furthermore, several other regions showed a differential pattern for both phases which changed over time. While a set of regions showed increased responding in the second half in those with higher trait anxiety (left anterior insula, left dorsal anterior cingulate, thalamus and cerebellum), another set of regions showed the opposite pattern: their responses were decreased the in second half in the more trait anxious group (two clusters in right dlPFC and angular gyrus).

3) State Learning

As revealed by our model comparison, state learning was the predominant strategy used by the participants and particularly by those who scored high on trait anxiety. To follow up on these results, I aimed to identify brain regions that are involved in state learning.

To check whether any of the pre-defined ROIs encode state prediction error, as opposed to gradual aversive prediction error as defined by the Pearce-Hall model, I compared the betas for each of the two parametric modulators for each ROI. Based on

a recent study by (Kuchibhotla et al. 2019) latent beliefs (or states) can only be decode from non-reinforced trials, this analysis was restricted to trials in which no shock was delivered.

The results of this analysis show that the left anterior insula ROI significantly encoded the state prediction error, $t(22) = -2.52, p = 0.020$. To extend my analysis from the pre-defined ROIs, I performed a whole brain analysis to identify regions encoding the state prediction error on no-shock trials in the second half. The reason for only including trials from the second half is twofold: Firstly, state learning occurs later in the task when structure has been discovered and secondly, the main difference between groups was found in the second, not in the first, half. The whole brain analysis revealed that only the left anterior insula encoded a state prediction error at $p_{uncorr.} < 0.001$ level, $t(22) = 4.30, p = 0.00015$, cluster size $n = 19$, $x = -30, y = 22, z = -6$. This result indicates that the anterior insula is involved in processing state error.

fMRI Bayesian Model Comparison

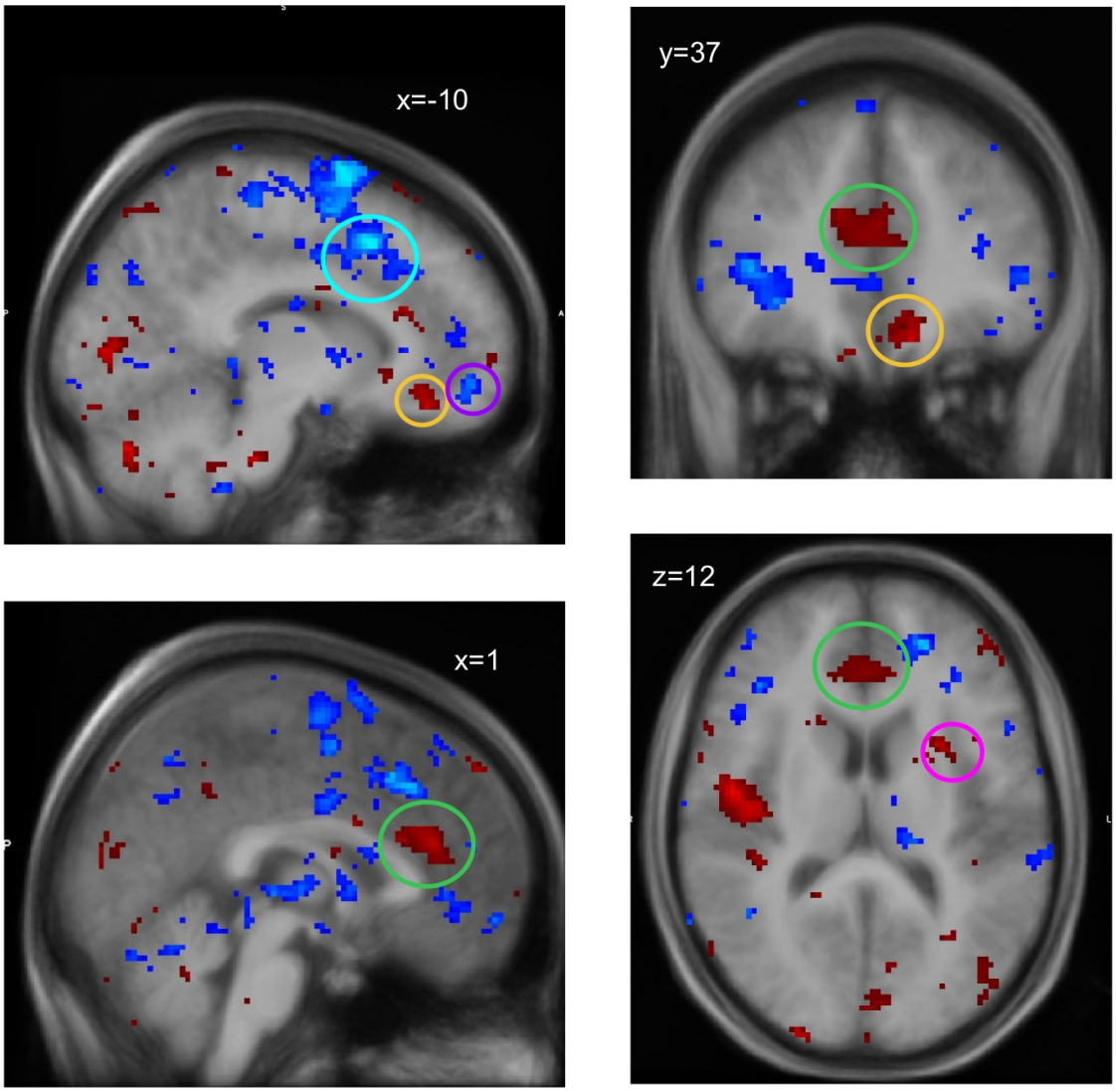
To directly test which brain regions are more involved in state learning than Pearce-Hall learning and vice versa, I employed a Bayesian model comparison technique developed by (Soch and Allefeld 2018), see full description in methods. The approach uses a Bayesian fixed-effects model selection routine that results in normalized Bayes factors. Since in my case two models are considered, a model with probability > 0.5 is more likely in a given voxel. To eliminate conflict regions with high uncertainty I report results when one model is at least 1.5 times more likely than the other, that is with a normalized probability of 0.6 and more.

The analysis aims to follow up on two a priori hypotheses and two experimental results from this study. Firstly, the OFC/vmPFC has previously been implicated in state representation (add references). I therefore test whether SPE is more likely in this region than PHPE. Secondly, a study by Glaescher et al (2010) identified SPE signal in

the lateral PFC as well as the IPL/angular gyrus. I will investigate SPE processing in those regions. Thirdly, I found that the dACC is associated with higher phase dissociability in high anxious individuals. As it can be argued that an increased dissociability might correspond to increased state awareness, I tested whether the dACC encoded SPEs more than PHPEs. To summarize, the following regions were investigated: OFC/vmpFC, lateral PFC, angular gyrus, dACC (bilateral) and left anterior insula.

The analysis revealed clusters of high SPE probability in the vmPFC (posterior portion, shown in orange), dACC (shown in green) and inferior parietal lobule (IPL). The SPE cluster in the dACC largely overlapped with our dACC mask (see Figure 23). I also found a significant cluster in the left anterior insula (pink) mask but a closer inspection found that this did not overlap with the cluster reported in our previous analysis. High PHPE probability clusters were also found in the dACC/MCC (light blue) located posterior to our dACC mask and SPE activation. Lastly there was a PHPE signal in the vmPFC (anterior portion, shown purple).

Figure 22: *Probabilistic model likelihood maps thresholded at p>0.6. Red clusters correspond to regions in which state model representation was more likely and blue clusters correspond to regions where the Pearce-Hall prediction error was more likely to be represented.*



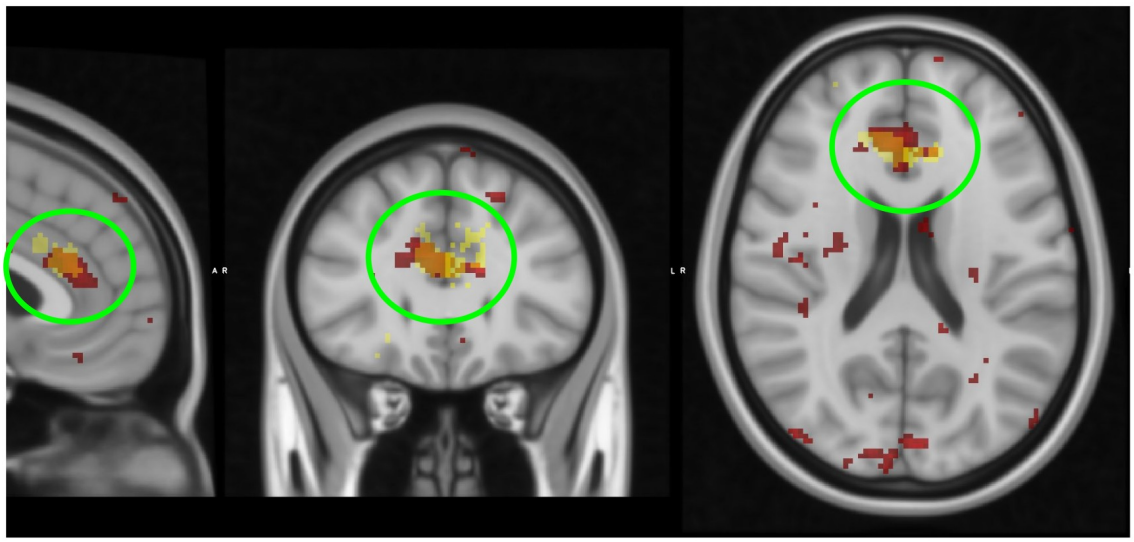


Figure 23: *Overlap between state prediction error cluster (red) from BMC analysis and a priori bilateral dACC mask (yellow). View coordinates: x=4, y=29, z=20.*

The Bayesian model comparison identified several clusters where SPE was at least 1.5 times more likely than PHPE. Some of those regions, particularly the dACC and vmPFC, directly corresponded to our a priori defined masks. SPE model was also more likely in the IPL which has previously been reported to process state prediction errors (Gläscher et al. 2010) .

Discussion

In the present chapter I explored the difference between acquisition and extinction in aversive learning using computational modelling and fMRI. The behavioural results show that post-reversal updating of shock expectancy rating started significantly sooner in acquisition than extinction trials and that while participants converge to the objective shock probability in acquisition trials, they tend to overpredict and never reach the true reinforcement rate of 25% in extinction. Furthermore, I show that reduced extinction increases over time and is driven by a

decrease in learning from shock omission. This finding is linked to the activity in the dlPFC which increases in no-shock trials during the second half of the experiment. Investigating the effect of trait anxiety on acquisition and extinction, I found lower shock expectancy during extinction in the second half of the experiment in high trait anxious individuals. The increased dissociability between acquisition and extinction correlated with the recorded GSR signal as well as with activity in bilateral dACC. Furthermore, high trait anxious population was best fitted by the state-switching model, which suggests a multi-state representation of the environment.. As revealed by Bayesian model section, brain activity specific to state prediction errors was found in the same portion of the dACC that encoded increased dissociability between the two phases in high anxious individuals.

Differences between acquisition and extinction learning

As reported previously (Li et al. 2016), participants were able to accurately increase their shock expectancy when the objective shock probability increased (acquisition) whereas they showed a consistent bias to overpredict shock occurrence when the objective shock probability decreased. I next focused on investigating whether this lack of extinction is a consequence of mere slower no-shock learning or whether there is a contextual component, mirroring the two main competing theories of extinction. Analysis of behavioural data revealed that updating towards new contingency occurs later in extinction than in acquisition. This result supports the theory of Bouton (2002) that there is an additional component to extinction learning. If extinction was solely a result of decreased no-shock learning, the switch point would occur at both phases at the same time according the the Rescorla-Wagner learning rule. Instead, I found that shock learning was significant faster than no-shock learning. Additionally, learning rates are modulated by the phase. During acquisition,

participants learned faster from shock delivery than shock omission, whereas the opposite was true during extinction trials. This suggests that participants were able to learn that our reinforcement schedule created different contexts and that they represent the recent history of outcomes as a particular context or state that is used to guide their predictions.

Changes in extinction over time

To explore how acquisition and extinction change over time, data from the first half of the experiment were compared to those from the second half.. This analysis revealed that the lack of extinction is particularly prominent in the later stage of the experiment while acquisition showed no time-dependent differences (Fig. 5). This finding raises the question what drives this increasing lack of extinction. In the previous paragraph I demonstrated that participants seems to be able to learn the contingencies, however, this only seems to be true for the initial stages of the task. Analysis of model-free learning rates revealed that shock learning remains unchanged learning from shock omision decreases over time. This suggests that over time, participants begin to disregard new evidence that suggests a lower likelihood of shock and rather continue to expect outcome delivery. Similar findings have been reported before, albeit for positive outcome. In a study by Park and colleagues in which participants engaged in a gambling task, alcohol dependent individuals kept choosing the option that used to be rewarded in the past but was no longer followed by the positive outcome (Park et al. 2010). There are number of potential explanations for this phenomenon. Firstly, participants may anticipate the worst outcome to be prepared in case the noxious stimulus is delivered. This would support the theory put forward by Arntz (1991) in which the author argues that aversion overprediction is a crucial survival component. Anticipating a shock on every trial ensures that the organism is always well prepared when the shock actually occurs. As discussed in the

introduction, this behaviour may be seen as evolutionary optimal considering that it prevents overall harm. To explore this hypothesis further, I used the Shannon surprise to quantify positive and negative surprise. While positive surprise increased significantly over time, negative surprise decreased (Fig. 7). To explicitly test whether participants actively sought to decrease negative surprise, I included the Surprise-Minimising model in the model comparison. Although four subjects (out of twenty eight) were best fitted by this model, this strategy was not dominant across the group, suggesting that minimisation of negative surprise is a byproduct rather than the main motivation. Secondly, it could be argued that participants disengaged from the task in the second half and report to expect shock delivery at a high probability but in fact not anticipate it.

To dissociate between the two explanations, anticipatory GSR activity and fMRI data were examined. Differences in GSR response to acquisition and extinction cues only reached significance in the second half of the experiment, suggesting that people were engaged in the task and learned to dissociate periods of danger from those of safety. This observation is supported by evidence
from MRI data where several regions, particularly the bilateral dACC and the left hippocampus, showed increased dissociability between acquisition and extinction in the second compared to the first half.

But is there any evidence for an active suppression or at least discarding of new evidence that should trigger updating of expectations to reflect the lower objective probability of negative outcome in extinction trials? A shown in Fig. 19 we found an increase in dlPFC activity over time in trials that should have triggered extinction, namely, when shocks were omitted in most trials. Importantly, this increase in dlPFC activity was not seen in trials where the shock was delivered or following presentation of the stable cues. Furthermore, dlPFC activity did not scale with MFLR, anticipation, surprise or state prediction error which suggests that rather than directly processing

learning rates, the dlPFC is involved in a form adaptive control. In line with this interpretation, the dlPFC has recently been linked to managing adaptive flexibility (Park et al. 2010; Beylergil et al. 2017), down-regulation of fear (Ironside et al. 2018) and strategic control (Venkatraman and Huettel 2012). It might therefore serve to optimise the management of resources and, instead of tracking trial-by-trial changes, settle on a general adaptive rule which results on overall overprediction of pain.

Anxiety and extinction

Based on a review of the trait anxiety literature I hypothesised that high trait anxiety will lead to *less* extinction. Contrary to the hypothesis the behavioural data showed that high trait anxious participants reported lower probabilities in extinction (Fig. 8). This was apparent in the mean probability data as well from significant negative correlation between anxiety and mean probability during extinction trials of the second half. In contrast to the tendency of the overall sample which showed higher mean probability ratings in the second compared to the first half, the high trait anxious group did not significantly increase their ratings in extinction trials. This stands in sharp contrast to the low anxiety group which reported nearly the same probability in acquisition as in extinction in the second half. This group difference is perhaps best illustrated by Figure 8 which shows the mean probabilities by half. This result directly contradicts much of what we know about trait anxiety. High trait anxious individuals were reported to have decreased ability to dissociate between acquisition and extinction (Kindt and Soeter 2014), to be unable to inhibit conditioned response (Haaker et al. 2015) and adapt learning rate to environmental volatility (Browning et al. 2015). In my sample, high trait anxious group was *better* in at least two of these measures. I would argue that what has been called an inability to inhibit conditioned response (Indovina et al. 2011) could as well be improved ability to differentiate between contexts. Data presented in this chapter suggest that high trait anxiety is

linked to better dissociation between acquisition and extinction at the behavioural, physiological and neural level. As a consequence, the group is “better“ at extinction. One potential explanation for this is that highly anxious individuals preferentially engage in state rather than gradual learning and when presented with sufficient evidence for a change in contingency, they abruptly switch to another belief state. In my task belief, ‘states’ would most likely correspond to acquisition and extinction as there is a period of high shock probability and a period of low shock probability.
To explicitly test this idea, I designed a state switching model based on the Rescorla-Wagner model and compared it to five other earning strategies including the Pearce-Hall model which has been widely successful at describing adaptive behaviour. The state switcher explained best the overall behaviour of the sample, suggesting that participants do engage in state switching behaviour at least to some extent. To test whether adopting a state switching strategy is related to trait anxiety, I investigated which participants are best fitted by the model. Subjects best fitted by the state switcher were significantly higher in trait anxiety than those best fitted by other models. To further verify this notion, I next tested where in the brain trait anxiety influenced aversive learning i and where, if anywhere, state prediction errors were processed. As shown in Fig. 23, the bilateral dorsal ACC was involved in both processes. The difference between acquisition and extinction in the region significantly correlated with trait anxiety as did the increase in differential activation between acquisition and extinction from the first to the second half. I next used Bayesian Model selection for fMRI to test which regions were more likely involved in state learning compared to PH learning. Again, the dorsal ACC (together with the vmPFC) was the region most likely involved in state learning. Together, these results strongly support the notion that high anxiety is associated with increased state learning reflected in the dorsal ACC. A similar tendency for “improved” extinction in highly anxious individuals was recently observed in two other studies (Wise et al. and Atlas et al. both under review).

Our findings are of direct relevance for clinical manifestations of anxiety. High trait anxiety has previously been associated with an increased tendency for fear reinstatement (Lissek et al. 2005). This is particularly undesirable in the clinical context where extinction training is one of the main behavioural methods of exposure therapy. An increased tendency to represent environment as multiple states means increased context-dependent learning. As a consequence, anxious individuals will be more likely to relapse following a seemingly successful therapy session. Paradoxically, what initially may have seemed as better extinction, may therefore in fact be a disadvantage in clinical treatment.

There are a number of limitations in this study and general comments that I would like to address. Firstly, the location of my SPE signal in the IPL resembles a previously published study reporting state prediction errors. (Gläscher et al. 2010) used a two-step decision making task to dissociate between model-based and model-free instrumental learning. Their results found the model-free prediction error, which in my case corresponds to gradual learning, in the ventral striatum, and the model-based prediction error, which corresponds to state prediction error, in the IPL/IPS and lateral PFC. While their task is fundamentally different due to the use of instrumental, as opposed to Pavlovian learning, my state prediction errors in fact constitute a form of Pavlovian model-based prediction error as recently defined by (Dayan and Berridge 2014). Interestingly, the model selection analysis also revealed a high probability SPE cluster in the IPL. While this finding requires further investigation, it is an interesting observation, especially since my study design differed considerably from the one adopted by Glaescher and colleagues.

Secondly, I cannot rule out that my new model captured hypervigilance behaviour instead of state switching. High anxious individuals might be more prone to follow instruction and focus on the task more closely. It might be argued that if that was the case and they were in fact not using a state switching strategy. However, since the

state switching model was best in explaining participants' learning, state learning might indeed have been the dominant strategy, albeit due to hypervigilance.

Conclusion

In conclusion, I performed a model-based fMRI study to investigate the computational and neural dynamics of aversive learning. I found acquisition learning to be faster and to start sooner than extinction learning. Additionally, I found less extinction over the course of time, leading to increasing pain overprediction which was associated with decreased no-shock learning in the second half of the experiment. The dlPFC was the only region reflecting these behavioural findings. dlPFC engagement was increased in response to no-shock outcomes in the second half. As this region was not involved in processing of learning rate, expectation, surprise or state prediction error processing, I suggest that the dlPFC might instead play a role in strategic control of behaviour. Contrary to our expectation, high trait anxiety was, associated with decreased shock anticipation during extinction. This was, however, driven by the tendency of high trait anxious individuals to represent the two phases as distinct states which has been linked to an increased risk of relapse of clinical forms of anxiety . The dorsal ACC was associated with increased dissociability between acquisition and extinction as well as with higher probability of state prediction error processing, providing neural evidence for a link between anxiety and state learning.

Chapter 4:

The effect of environmental contingencies and trait anxiety on aversive learning strategies

Introduction

Accurate learning in a noisy environment provides an evolutionary benefit. At the same time, minimising energy expenditure in the form of attentional and computational resources is crucial for survival, preventing the cognitive system from being overhauled. This tradeoff between computational efficiency and accuracy has been explained by a number of theories such as dual-system learning theory, (Daw, Niv, and Dayan 2005) or hierarchical reinforcement learning (Botvinick, Niv, and Barto 2009).

In model-free reinforcement learning (RL), updating occurs on a trial-by-trial basis with prediction error being weighted by a learning rate. Optimal outcome prediction is inherently dependent on i) the noise in the environment; and ii) any underlying environmental structure, such as how often does environment change and between what levels of reinforcement. Let us consider the two most extreme cases of noise. In a completely noise-free deterministic environments the optimal strategy would be to to make a prediction that fully copy the most recent outcome. In the Rescorla-Wagner model this would translate to having a learning rate of 1. In the opposite case where outcomes are entirely random the best strategy would be to maintain predictions at some stable level (e.g. always predict reward) which translates to learning rate of 0. In most environments, however, there is some underlying rate of reinforcement which will produce reward/punishment in a probabilistic manner. To best predict the next outcome the optimal agent will gradually update their expected value by weighing the new information. This weight is commonly referred to as learning rate. This is the fundamental mechanism of reinforcement learning: expectations are updated gradually, with the degree of updating being contingent on the magnitude of the prediction error as well as individual learning rate. If we now consider an environment in which there is an

underlying structure, e.g. 20 reinforced and 20 non-reinforced trials keep alternating, the reinforcement learning agent will take several new outcomes to adjust its prediction accurately. As an alternative, it may learn that there is a higher order structure in the environment that explains changes in outcome. In our example, the agent would assume that there are two states - one in which the outcome is always delivered and one in which the outcome is always withheld. In the state in which outcome is delivered, outcome omission signals a switch to the state in which outcome is omitted and vice versa. Investigating the difference between gradual and structure learning is the topic of this chapter.

Gradual learning, as described above, ensures that when faced with a clear expectancy violation, an agent does not change its belief drastically. Additionally, learning can be accelerated when large prediction errors are generated (Botvinick, Niv, and Barto 2009; John M. Pearce and Hall 1980). This makes RL a suitable strategy when exploring a new environment. However, over longer periods of time gradual updating may become ineffective and computationally wasteful, so state learning becomes a more efficient strategy (Gershman, Blei, and Niv 2010). The amount of trials since the last change need to be monitored and when a change occurs predictions can change abruptly to minimise future prediction error. This difference between a gradual and structure learner is presented in Figure 1.

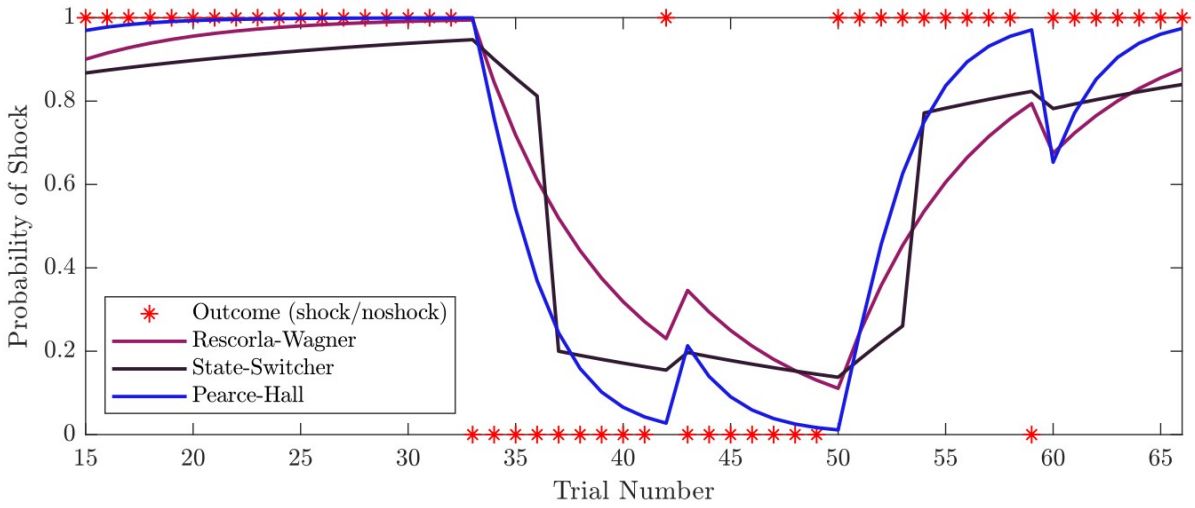


Figure 1: *Different predictions made by a Rescorla-Wagner, Pearce-Hall and a state-switching models.*

A crucial component of structure learning is the correct detection of an underlying pattern. In the example above, the agent has to decide whether each surprising outcome is a rare event that can be discarded or whether it signals a change in the underlying structure of the environment. This is particularly true in probabilistic environments that we have to navigate on a daily basis. As an example, (Behrens et al. 2007) showed that participants adjust their estimate of volatility based on the true volatility of the environment, i.e they use their volatility estimate to adjust learning rates. Interestingly, this search for an underlying structure can lead to maladaptive learning. Browning et al. (2015) used the same task to show that high trait anxious individuals fail to adjust their learning appropriately to environmental volatility. A good structure learner will learn meaningful features that can be generalised and used for accurate predictions. In contrast, a learner which infers wrong patterns or who considers new data as overly informative will make inaccurate predictions. Identification of a causal relationship between two or more environmental variables leads to the creation of a internal generative model that is subsequently used to make predictions. The internal model combines external information, such as rewards or contexts, with internal information, such as the perceived current

experimental phase. While the external causes are known, the internal ones have to be inferred. This is a difficult problem as internal states or contexts can be task relevant (i.e. experimental block) or task irrelevant (i.e. homeostasis, weather).

A large body of research has focused on understanding the different mechanisms governing aversive learning and identified considerable differences between aversive acquisition and extinction. While acquisition is characterised by rapid initial learning, extinction is slow and often incomplete. As seen in Chapter 1, extinction is believed to be a separable learning process. Additionally, as the cases of reinstatement and renewal show, fear relapse is strongly context-dependent (Bouton 2002).This posits a major problem in clinical treatment which relies on prolonged effectiveness and generalization of extinction training. If acquisition and extinction are seen as different contexts then experience in one might be seen as separate from experience in the other, this making updating in extinction separate from updating in acquisition. One way to think about context is to consider it an internally represented state. For example, a particular context or a period of relatively stable reinforcement rate can internally be considered as a state.

In a typical fear conditioning experiment, acquisition involves repeated pairing of US with CS. This can follow a partial or a full reinforcement schedule. A participant can therefore categorize the environment into at least two different states: a state of high shock probability and a state of low shock probability. For now, let us assume that these correspond to acquisition and extinction. Acquisition is characterised by a high frequency of shocks while extinction is a phase with relatively low shock frequency. It is assumed that for the two phases to be represented as two distinct states, a certain level of distinctiveness must be in place. It is one of the goals of this study to identify the necessary level of reinforcement rate between acquisition and extinction that would lead to gradual learning, as opposed to state switching. It must be noted here that while these may correspond to the acquisition and extinction

phases, the internal belief state is not directly known. A participants' acquisition may differ from what a research paper might report as acquisition.

(Redish,et al. 2007) proposed a model that combines the reinforcement learning approach with structure learning. An agent initially learns using RL until a meaningful structure is discovered. If the reinforcement pattern changes abruptly, large prediction errors are generated and as a consequence a new state is created provided the new pattern is consistent and sufficiently different from those already established (Gershman et al. 2013). However, if changes in environmental probabilities are gradual, a updating of the of assumptions about the existing state is more likely than creation of a new one . Understanding when participants represent environment as one versus multiple states has important implications for exposure therapy. If extinction is represented as a separate state the relapse back to the 'acquisition' state is more likely. If a gradual extinction occurred and the agent has overiden the memory of a high shock probability state, they are less likely to relapse. The method of gradual extinction is known as occasionally reinforced extinction in the clinical literature and it is a well documented method that improves therapy success rate (Craske et al. 2014b).

High trait anxious individuals are known to have a higher relapse risk than low anxious individuals (Lissek et al. 2005). Trait anxiety has previously been linked to deviations from normative aversive learning behaviour (Browning et al. 2015). In my previous data set (Chapter 3) I found behaviour of high trait anxious participants to be preferentially matched by a multi-state learning rather than single state (i.e. gradual) learning model, indicating that high trait anxious individuals are more likely to assume different states rather than update their existing representation of contingencies. As a consequence, the original representation remains unaltered and can be reactivated at any time. Such mechanism may account for higher rates of relapse in high anxious individuals reported by clinical studies. A question that remains to be answered is whether structure as opposed to gradual learning is preferred with increasing

distinctiveness of states.

Here, we will use computational models that represent competing theories of learning together with behavioural and physiological measures to test i) which reinforcement rates encourage gradual learning versus state switching strategy; and ii) whether trait anxiety relates to preferential reliance on state-learning rather than gradual learning. To encourage different learning strategies we employ a task where shock contingency changes are either small, medium or large. Low reinforcement rate differences between acquisition and extinction are thought to encourage more gradual learning while large differences is likely to lead to a state-switching behaviour. At medium reinforcement rates, the participant is under most pressure as to which strategy to employ. It is here that we expect the biggest difference between low and high anxious participants.

We hypothesise that an increase in the difference in contingencies between acquisition and extinction will promote state-switching learning. As a second hypothesis, we predict that high trait anxiety will lead to an increased likelihood of state switching behaviour.

Methods

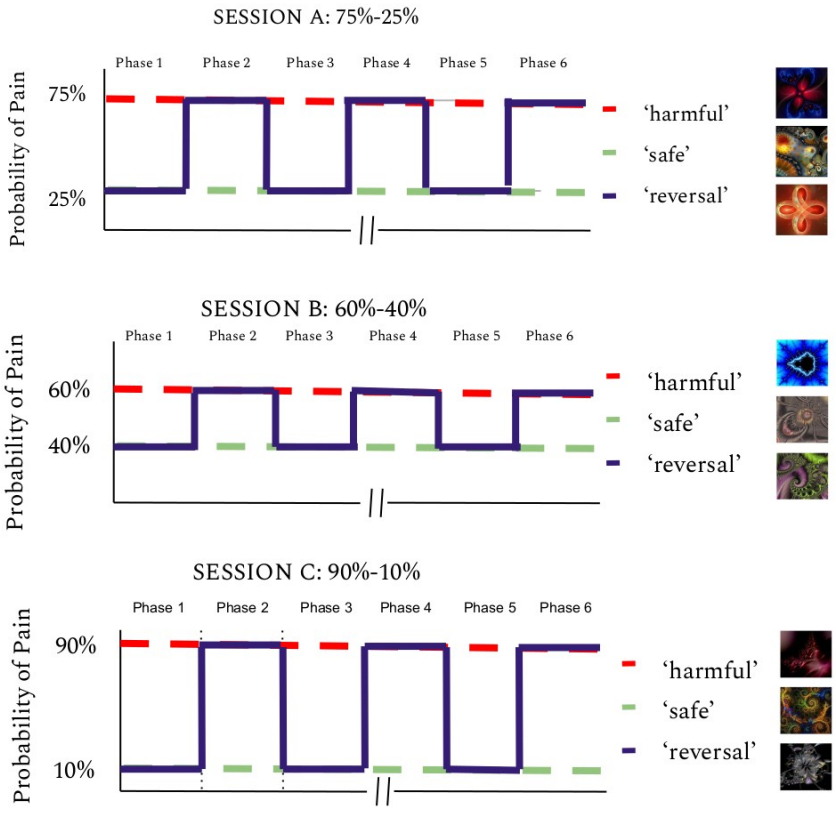
The majority of the methods employed in this study is described in Chapter 2. Here, I only add information on methods that have not been introduced yet or specifications that deviate from those given in the Methods chapter.

Task

The task comprised three sessions of the probabilistic aversive learning task described in Chapter 2. In order to investigate the effect of shock contingency on

aversive learning, each session used different probabilities of shock. The reversal cue always switched between two probability levels and the stable cues shadowed those without the switches. The three contingency levels were 60/40%, 75/25% and 90/10% chance of shock. Figure 1 shows a schematic of the design. The amount of trials per phase was normally distributed around the mean of 35 with a minimum and maximum of 20 and 50 trials, respectively (+/- 15 trials). The three sessions were presented in a randomized order that was counterbalanced prior to the experiment. Shock intensity was re-calibrated before each session. Upon completion of the experiment, participants completed a set of computerized questionnaires. In addition to the psychometric questionnaires common to Study 1 and Study 2, visual appeal [-5 to 5], cue liking [-5 5] and overall shock likelihoods [typed in] associated with each image data were rated by the participant.

Figure 2: *Schematic of the experimental design*



Participants

40 healthy participants completed all three sessions of the task (22 female, mean age $\bar{x} = 25.8, \sigma = 5.05$). Five participants were excluded due to failure to distinguish between the safe and the harmful cue. This was assessed by comparing reported probabilities for the two cues using a repeated measures t-test and averaging the p-value over the 75/25 and 90/10 conditions. Subjects with an average p-value of lower than 0.05 were excluded. One subject misunderstood the scale and switched their responding halfway through the experiment. One subject reported not receiving any stimulation for the last 30 trials of session 3. Data of the remaining 33 participants were included in the final analysis.

Models

The main purpose of this study was to dissociate between gradual learning and a state switching strategy. To capture gradual learning, I employed the most commonly used version of the Pearce-Hall (Tzovara, Korn, and Bach 2018; Li et al. 2011). To capture state-switching behaviour, I previously used the state-switching model based on reinforcement learning described in Chapter 2. The model, however, has several drawbacks. In particular, it assumes 2 states and therefore cannot dissociate between one and multi-state learning. It also does not propose a mechanism of how new states are discovered. To address those limitations, I developed a new state-switcher model which can account for new state creation as well as state switching. This is an important feature because the new model can explain gradual (one-state) as well as state-switching (multi-state) behaviours but it can not account for Pearce-Hall-like learning.

Beta State Switcher

This model is based on the leaky beta model recently used by Wise et al. (in submission). The original version of this model implicitly assumes one state. Here, I extend the model to account for multi-state representation of the environment.
The beta distribution is suitable for probabilistic learning as it is regularized for values between 0 and 1. At the same time, its parametrization can have a naturalistic interpretation for binary outcomes. It is parametrized by two parameters α and β such that when $\alpha = \beta = 1$, the probability density is a uniform distribution between 0 and 1, i.e. all states are equally likely. Generally, in any scenario where $\alpha = \beta$ the distribution will be symmetrical. If we now increase the value of α by 4, the beta probability density becomes large near $P(P_{(shock)} = 1)$. In our case, this would correspond to the participant receiving 4 shocks in a row. Figure 3 presents the evolution of beta PDF as a function of increasing number of shocks (corresponding to the α parameter) and shock omissions (corresponding to the β parameter).

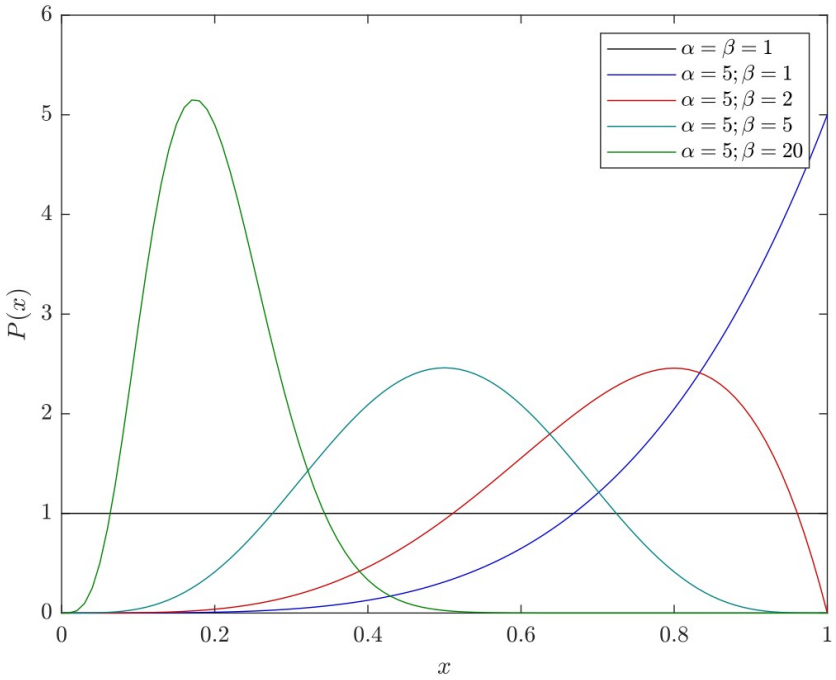


Figure 3: *Behaviour of the beta distribution as its input parameters change*

Eq. I $$\mu_{Beta(\alpha,\beta)} = \frac{\alpha}{\alpha+\beta}$$

Eq. II $$\sigma^2_{Beta(\alpha,\beta)} = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

The value predicted by the agent on each trial is the mean of the beta distribution given the parameters of the current state (using Eq 1). The sum of α and β is inversely proportional to the standard deviation of the distribution (Eq. 2). The more outcomes the agent has experienced, the more certain it is about the probability of an outcome on the next trial. In a changing and noisy environment, this model with no free parameters converges to the mean. To make the model adjustable to recent experience, forgetting has to be introduced. On each trial, both α and β decay by the amount set by the free parameter $\lambda \in [0,\ 1]$. On each trial, this model adds 1 to either α if the outcome was shock or to β if the outcome was shock omission. To account for differential learning from shock and no-shock outcomes, an additional constant is added for each outcome type. I define this constant as $\tau^+ \in [-1,\ 1]$ for shock and $\tau^- \in [-1,\ 1]$ for no-shock. These can be loosely interpreted as attentional weights and they provide the model with the ability to dissociate between shock and no-shock updating. Positive values of τ corresponds to extra weight on the given outcome while negative values correspond to downweighting of that outcome. On each trial, the values of α and β are updated as follows:

Eq. III $$if\ O_t = 0: \quad \beta_{t+1} = \lambda(\beta_t + 1 + \tau^-);\ \alpha_{t+1} = \lambda\alpha_t;$$

Eq. IV $$if\ O_t = 1: \quad \beta_{t+1} = \lambda\beta_t;\ \alpha_{t+1} = \lambda(\alpha_t + 1 + \tau^+);$$

where O_t stands for the outcome, coded 1 for shock and 0 for no-shock and t

represents the trial number. λ represents the degree of memory leak common to both parameters.

The version of the model described up to here has been shown to describe gradual learning well (Wise et al. in submission). Here, I extend the model to include multi-state learning. In line with the Pearce-Hall quantification of associability, I assume that the agent tracks the current level of surprise I defined as:

Eq. V $$I_t = I_{t-1} + \alpha_{surp}(O_t - P_t) - \lambda_{surp} I_{t-1}$$

where α_{surp} corresponds to surprise learning rate and λ_{surp} to surprise decay. Note that the sign of I matters here, large negative values of I correspond to sudden omission of shock and large positive values correspond to large positive surprise.

The learning process starts by assuming one state. On each trial, following the outcome, the current level of surprise is compared to the current state uncertainty (Eq VI). If the surprise level exceeds a certain threshold, state is either switched or a new state is created. Creation of too many states is discouraged by making each subsequent state harder to create. A state s is defined by its mean and uncertainty that are calculated using the state-specific α_s and β_s and Equations 1 and 2. The decision on each trial is:

Eq VI if: $I_t > \sigma_{t,s} * \eta * max(S)$: switch or create

else: stay in the current state

where $max(S)$ represents the number of states already created. The parameter η controls the subjective threshold for state switch / new state creation. A more state-switchy individual will have a low level of η while a one-state learning will have a higher value of η.

If the level of surprise exceeded the expected deviation in surprise, an expected value estimate is calculated by adding surprise to the old state value $P_{new} = p_{old} + I_t$ and testing whether any other existing state has that value within its expected range using the same decision rule as described above. If exactly one state matches, the agent switches to it. If two states match, the likelihood of each candidate state is calculated, and the state with the highest likelihood is switched to. If no state matches a new state is created as follows:

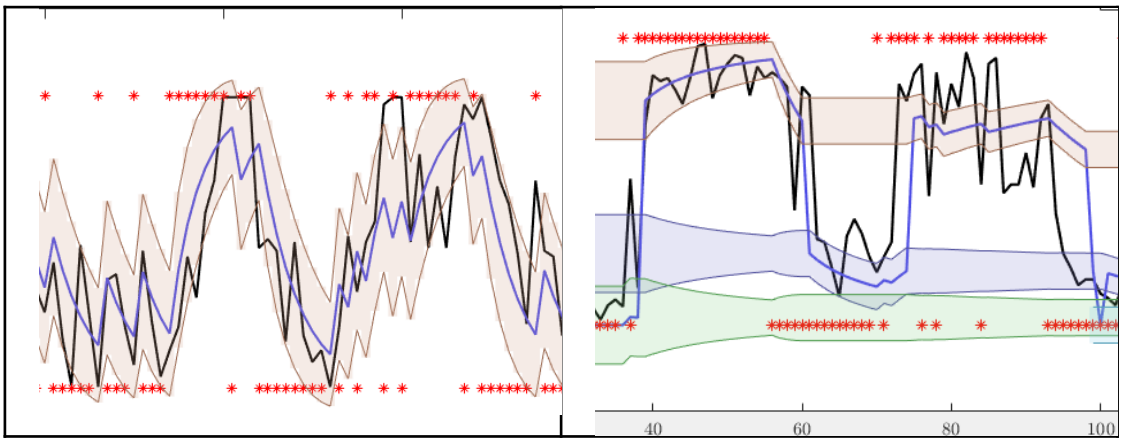
Eq. VII $P_{s_{new}} = P_{s_{old}} + I_t$

Eq. VIII $\sigma^2_{s_{new}} = \sigma^2_{Beta(1,1)}$

The value of the new state is the previous value plus the surprise. The uncertainty of the new state is the variance of $Beta(1, 1)$ distribution.

Fitting the model to a typical ‘gradual learner’ and ‘state-switcher’ suggests that the model performs well at capturing both strategies (see Figure 4).

Figure 4: *Example of the beta state switcher fit to gradual and state-switching participant. Each shaded area represents one state.*



Results

Investigation of potentially confounding factors

To check whether our data were influenced by the i) within-session starting probability (either high or low) or ii) the order in which sessions were presented we performed a series of mixed ANOVAs with mean probability as the dependent variable. Additionally, since most analyses assume independence between cues, the co-dependence of cues on each other was examined.

Starting Probability

A mixed ANOVA for each session was run. The following independent variables were included: i) cue type (safe, harmful, reversal) and session (60/40, 75/25 and 90/10) as within-subject; and ii) starting probability of the reversal cue (high/low) and trait anxiety as between-subject. None of the analyses found a main effect or interaction of starting probability.

In the reversal cue condition, an additional set of mixed ANOVAs was run to check for any effect of starting probability. Phase (acquisition/extinction) was added as a within-subject variable. As above, no main effect or interaction of starting probability was found.

Order Effects

For each session, a mixed ANOVA was run with cue as a within-subject variable, and ordinal block number (e.g. session 60/40 occurred 1st / 2nd / 3rd) and trait anxiety as a between-subject factors. None of the tests showed a significant main effect or interaction of order.

In the reversal cue, phase as added to the analysis. No main effect or interaction between order and phase was found.

Learning Leakage

The three cues presented on each session occur in a pseudo-random order. There is a possibility that outcomes for one cue influence the probability adjustments for the other cues. To check whether the three cues tend to influence each I considered two learning models, one assuming cue independence and other modelling cue co-dependence.

Independent vs Co-dependent Updating

To check for any mutual influence of the three cues on each other, a model based on the Rescorla-Wagner learning rule was built which uses a weighted combination of the current value of the three cues as the expectation term, as defined by the equation:

$$P_{A_{t+1}} = P_A + \alpha(O_t - (\omega_A P_A + \omega_B P_B + \omega_C P_C))$$

where P_A, P_B and P_C represent the current values of the three cues and ω_A, ω_B and ω_C represent the weights. If only information related to cue A is used for updating the resulting value of ω_A will be 1 while the two others will be 0. We compared this model to a version of Rescorla-Wagner in which the three cues are updated independently. The AIC scores for each block are presented in Figure 5. AIC scores for the independent updater were lower in all three blocks, suggesting co-dependent updating as less likely to occur than independent updating.

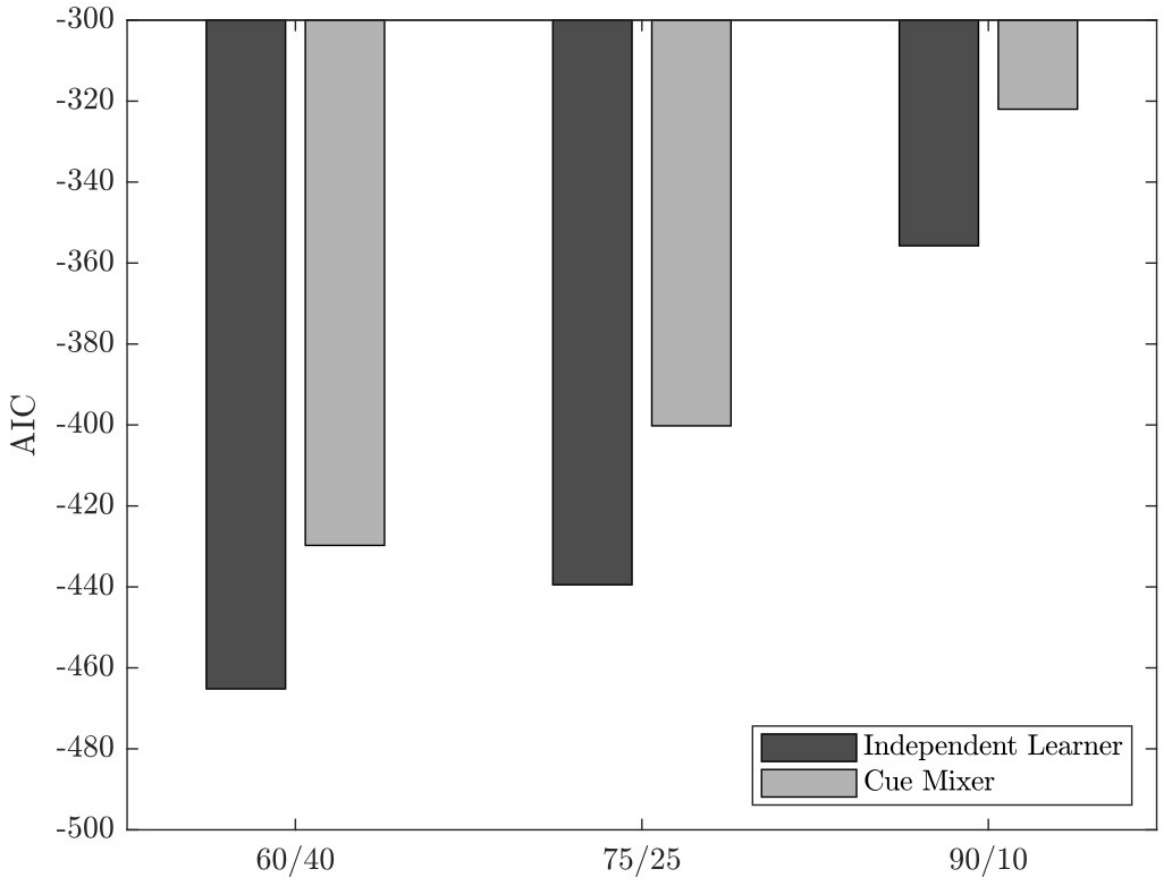


Figure 5: *Mean AIC scores for cue independent model and cue-mixing model.*

Behavioural Results

Post-Task Questionnaires

The following data were collected at the end of the experiment using computerised questionnaires. Three questions were asked for each cue: what was the averaged reinforcement rate, how much did the participant like the cue considering the outcomes and how did it visually appeal to them. The three variables were tested using a repeated measures ANOVA with session and cue as factors. Results are presented in Table 1.

Table 1: *Post-task reported mean probabilities*

	SS	*d.f.*	*MS*	*F*	*p*
Intercept	926,328.658	1	926,328.658	1,450.335	0.000
Error	23,631.898	37	638.700		
Session	473.175	2	236.588	1.144	0.324
Error(Session)	15,308.602	74	206.873		
Cue	109,786.544	2	54,893.272	88.502	**0.000**
Error(Cue)	45,898.567	74	620.251		
Session*Cue	2,545.333	4	636.333	1.239	0.297
Error(Session*Cue)	75,988.222	148	513.434		

In post-task reported mean probabilities there was a main effect of cue (see Table 1). Post-hoc t-tests found the differences between all three cues to be significant, ratings were significantly higher for the harmful ($\bar{x} = 70.0, \sigma = 15.0$) than safe , ($\bar{x} = 27.32, \sigma = 12.2$) $t(37)_{h>s} = 11.90, p < .001$, and harmful than reversal cue, ($\bar{x} = 57.77, \sigma = 16.3$), $t(37)_{h>r} = 4.15, p < .001$, and significantly higher for the reversal versus the safe cue $t(37)_{r>s} = 8.88, p < .001$.

Table 2: *Post-task visual appeal ratings*

	SS	*d.f.*	*MS*	*F*	*p*
Intercept	0.046	1	0.046	0.003	0.960
Error	679.288	38	17.876		
Session	4.108	2	2.054	0.434	0.650
Error(Session)	359.892	76	4.735		
Cue	46.228	2	23.114	3.750	**0.028**
Error(Cue)	468.439	76	6.164		
Session*Cue	6.746	4	1.687	0.371	0.829
Error(Session*Cue)	691.254	152	4.548		

In post-task reported visual appeal there was a main effect of cue (see Table 2). The cue means and standard deviations were: harmful, $\bar{x} = -0.45, \sigma = 1.90$, safe, $\bar{x} = 0.44, \sigma = 1.57$, reversal $\bar{x} = -0.017, \sigma = 2.01$. The main effect was driven by a significant difference between harmful and safe cue, $t(38) = -3.11, p = 0.004$. No other difference was significant.

Table 3: *Post-task cue liking ratings*

	SS	*d.f.*	*MS*	*F*	*p*
Intercept	137.892	1	137.892	7.138	0.011
Error	734.108	38	19.319		
Session	25.510	2	12.755	3.018	0.055
Error(Session)	321.157	76	4.226		
Cue	432.484	2	216.242	22.905	**0.000**
Error(Cue)	717.516	76	9.441		
Session*Cue	9.858	4	2.464	0.491	0.743
Error(Session*Cue)	763.476	152	5.023		

In post-task reported cue liking there was a main effect of cue (see Table 3). Post-hoc t-tests found the differences between all three cues to be significant, ratings were significantly higher for the harmful ($\bar{x} = -1.86, \sigma = 2.18$) than safe, ($\bar{x} = 0,83, \sigma = 1.78$) $t(38)_{h>s} = -6.49, p = 0.000$, and harmful than reversal cue, ($\bar{x} = -0.84, \sigma = 2.19$), $t(38)_{h>r} = -3.17, p = 0.003$, and significantly higher for the reversal versus the safe cue $t(37)_{r>s} = -3.66, p = 0.000$.

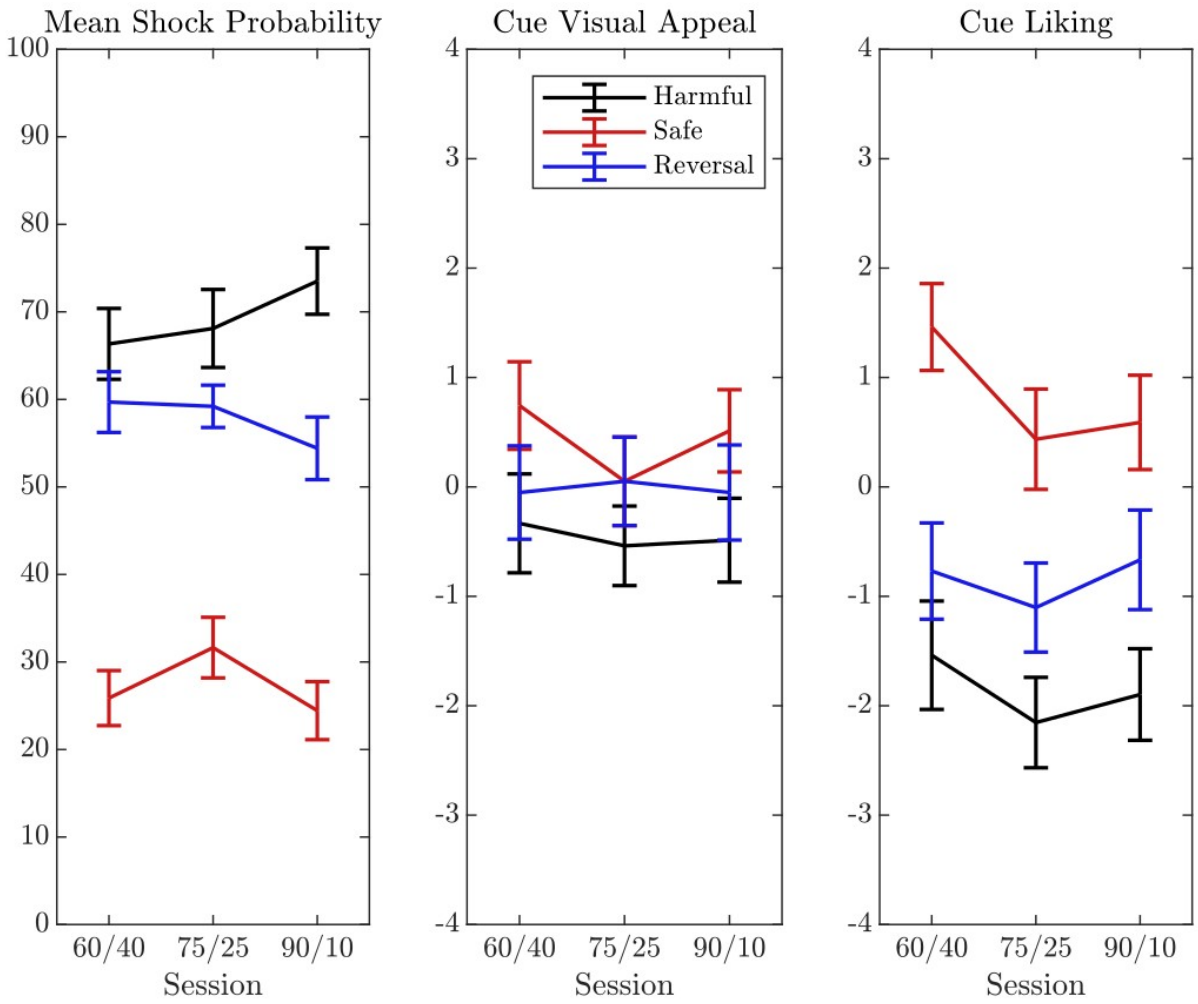


Figure 6: *Post-task ratings of mean probability, cue visual appeal and cue liking.*

The post-task questionnaire revealed that participants were able to dissociate between the cues and correctly follow the reinforcement levels which projected into how much they liked the cues. Visual appeal data suggest that the conditioning also influenced how visually appealing were different fractals found as there was a significant difference between the safe and the harmful cue.

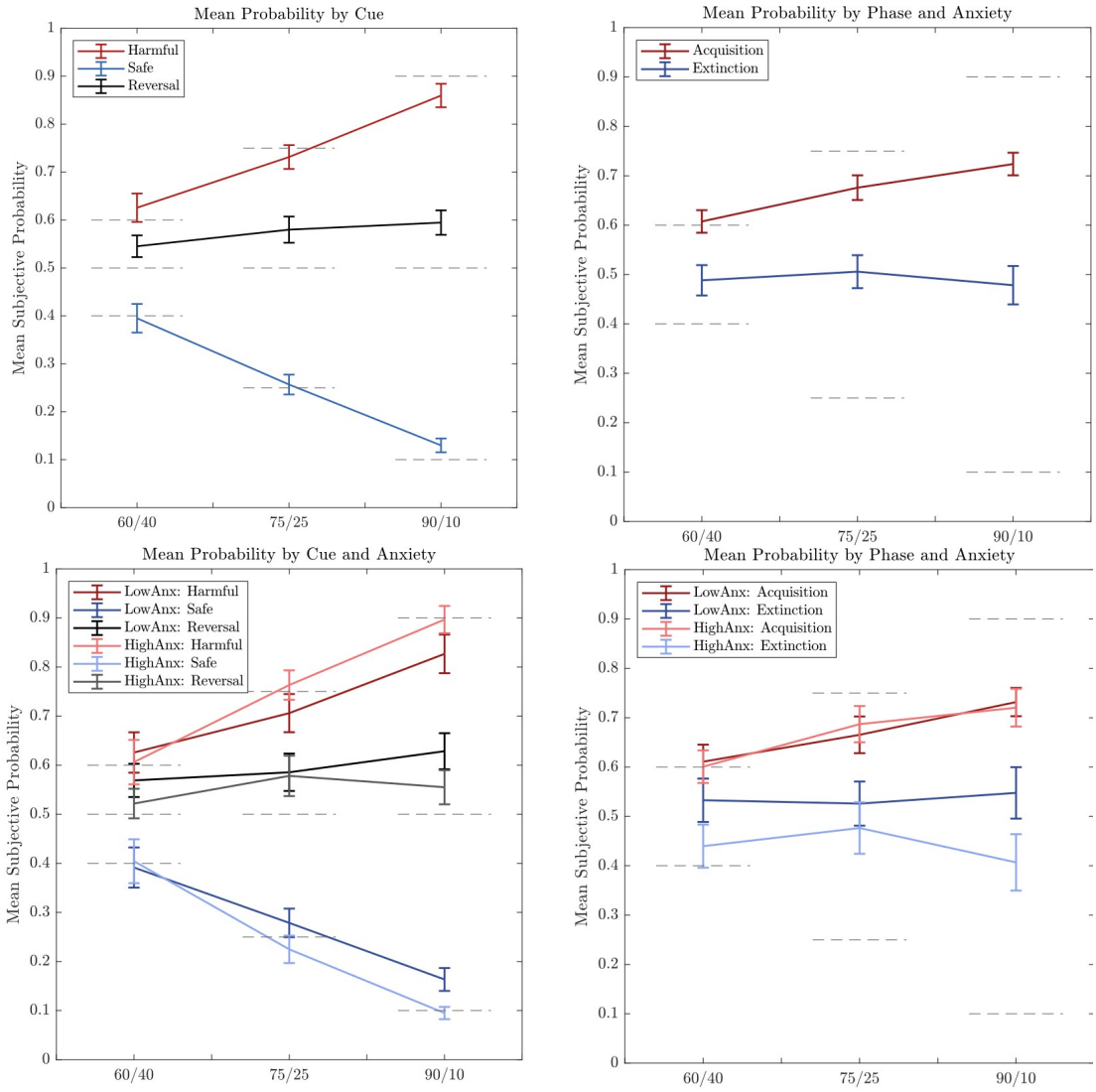
Mean Probability

Reported trial-by-trial subjective probabilities were averaged per cue and session. A mixed ANOVA including cue and session as within subject factors and median-split trait anxiety as a between-subject variable was performed. Results are

summarized in Table 1, graphical depiction of the data is shown in Figure 6. The results show a significant main effect of cue, $F(2, 60) = 187.15, p = 0.000$. The cue means and standard deviations were: harmful, $\bar{(x)} = 0.73, \sigma = 0.12$, safe, $\bar{(x)} = 0.26, \sigma = 0.07$, reversal $\bar{(x)} = 0.57, \sigma = 0.12$. Post-hoc t-tests revealed a significant difference between all combination of cues, harmful cue was significantly higher than safe cue $t(32)_{h>s} = 19.19, p = 0.000$, harmful cue was significantly higher than safe cue $t(32)_{h>r} = 6.63, p = 0.000$, and reversal cue was significantly higher than the safe cue$t(32)_{r>s} = 13.00, p = 0.000$. Furthermore, there was a significant cue*session interaction. In the **harmful** cue the session means were: 60/40 $\bar{x} = 0.61, \sigma = 0.17$; 75/25: $\bar{x} = 0.73, \sigma = 0.14$; 90/10: $\bar{x} = 0.86, \sigma = 0.13$, the reported probabilities were higher in session 75/25 than 60/40, $t(32) = -3.75, p = 0.0001$, higher in session 90/10 than 75/25, $t(32) = -5.73, p = 0.000$, and higher in session 90/10 than in 60/40, $t(32) = -8.73, p = 0.000$. In the **reversal** cue, the means were: 60/40 $\bar{x} = 0.55, \sigma = 0.13$; 75/25: $\bar{x} = 0.58, \sigma = 0.15$; 90/10: $\bar{x} = 0.59, \sigma = 0.15$. 90/10 session was significantly higher than 60/40 session, $t(32) = -1.6, p = 0.048$ which, however didn't survive correction. No other difference in the reversal cue existed. In the **safe** cue, the means were: 60/40 $\bar{x} = 0.38, \sigma = 0.16$; 75/25: $\bar{x} = 0.26, \sigma = 0.11$; 90/10: $\bar{x} = 0.14, \sigma = 0.07$, the reported probabilities were lower in session 75/25 than 60/40, $t(32) = 3.78, p = 0.001$, lower in session 90/10 than 75/25, $t(32) = 5.97, p = 0.000$, and lower in session 90/10 than in 60/40, $t(32) = 8.34, p = 0.000$.

Table 4: Mean probability by cue, session and anxiety					
	SS	d.f.	MS	F	p
(Intercept)	78.882	1	78.882	1,836.289	**0.000**
Anxiety	0.015	1	0.015	0.349	0.559
Error (Anxiety)	1.289	30	0.043		
Cue	11.319	2	5.659	187.148	**0.000**
Cue*Anxiety	0.091	2	0.046	1.512	0.229
Error (Cue)	1.814	60	0.030		
Session	0.003	2	0.001	0.140	0.870
Session*Anxiety	0.007	2	0.003	0.311	0.734
Error (Session)	0.639	60	0.011		
Cue*Session	2.160	4	0.540	40.663	**0.000**
Cue*Session*Anxiety	0.079	4	0.020	1.482	0.212
Error (Cue*Session)	1.594	120	0.013		

Figure 7: *Top to plots show mean probability by cue (left) and reversal cue split by phase (acquisition/extinction; right). Lower two plots show the same data median-split by trait anxiety.*



Motivated by the analysis in Chapter 3, I split the reversal cue into its respective phases (acquisition and extinction). A mixed ANOVA with phase, session and anxiety as factors found a significant main effect of phase, $F(1, 30) = 50.61, p = 0.001$ with higher ratings for acquisition ($\bar{x} = 0.70, = 0.13$) than extinction ($\bar{x} = 0.45, = 0.19$) and a significant interaction between phase and session, $F(2, 60) = 9.73, p = 0.001$. Means for this interaction are presented in Table 5. Post hoc test found that this

interaction was driven by increase in acquisition between 60/40 to 75/25, $t(32) = -3.46, p = 0.002$ and higher probability in acquisition in 90/10 than 75/25, $t(32) = -2.42, p = 0.021$ and higher probability in acquisition in 90/10 than 60/40, $t(32) = -5.2, p = 0.000$, where the second comparison did not survive correction. There was no significant difference between any of the comparisons in extinction. The interaction between binary anxiety and phase neared significance (p=0.073). The full set of results for this analysis is presented in Table 6. To limit the effect of first few post-reversal trials, the same analysis was repeated with the first three post-reversal trials removed. The statistical results remained the same. These results suggest that while the ratings get higher with true reinforcement rate in acquisition, no extinction there is not difference between the sessions.

Table 5: *Means and standard deviations for session by phase interaction*

	60/40	75/25	90/10
Acquisition	$\bar{x} = 0.62\sigma = 0.14$	$\bar{x} = 0.70\sigma = 0.14$	$\bar{x} = 0.77\sigma = 0.14$
Extinction	$\bar{x} = 0.47\sigma = 0.19$	$\bar{x} = 0.47\sigma = 0.20$	$\bar{x} = 0.42\sigma = 0.25$

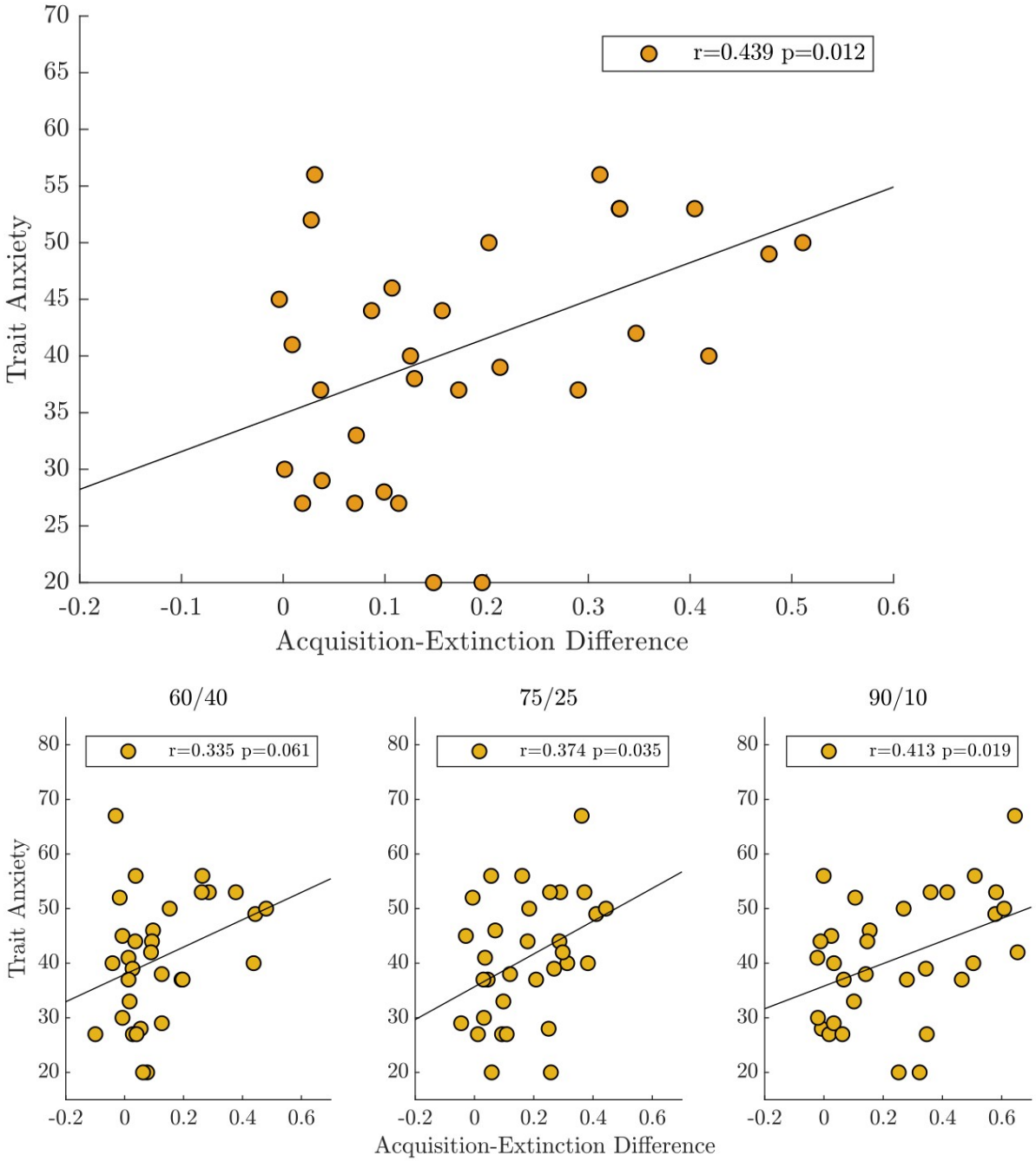
Table 6: *Reversal cue mean probability by phase, session and anxiety*

	SS	d.f.	MS	F	p
(Intercept)	64.316	1	64.316	747.393	**0.000**
Anxiety	0.107	1	0.107	1.243	0.274
Error (Anxiety)	2.582	30	0.086		
Phase	1.575	1	1.575	50.606	**0.000**
Phase*Anxiety	0.107	1	0.107	3.443	0.073
Error (Phase)	0.933	30	0.031		
Session	0.108	2	0.054	2.475	0.093
Session*Anxiety	0.031	2	0.016	0.717	0.492
Error (Session)	1.313	60	0.022		
Phase*Sesson	0.135	2	0.067	9.731	**0.000**
Phase*Session*Anxiety	0.008	2	0.004	0.552	0.579
Error (Phase*Sesson)	0.415	60	0.007		

As binarizing trait anxiety leads to a loss of information, I next performed a series of correlation analyses investigating the relationship between anxiety and the difference in probability between acquisition and extinction. Collapsing data over all

sessions, the difference between acquisition and extinction was significantly correlated with trait anxiety, $r(32) = 0.44, p = .012$ Splitting the data by individual sessions, the correlation remained significant in the 75/25 and 90/10 conditions, $r_{(75/25)} = 0.37, p = .035$, $r_{(90/10)} = 0.41, p = .019$, all per-session results were marginally significant with FDR corrected values at $p_{60/40} = 0.061, p_{75/25} = 0.0525, p_{90/19} = 0.0525$ Together, these data indicate that with higher trait anxiety the difference between acquisition and extinction gets bigger. The scatter plots are presented in Figure 8.

Figure 8: Correlations between phase difference and trait anxiety. The upper plot shows the acquisition-extinction difference collapsed across the three session, the lower panel shows the relationship with trait anxiety per session.

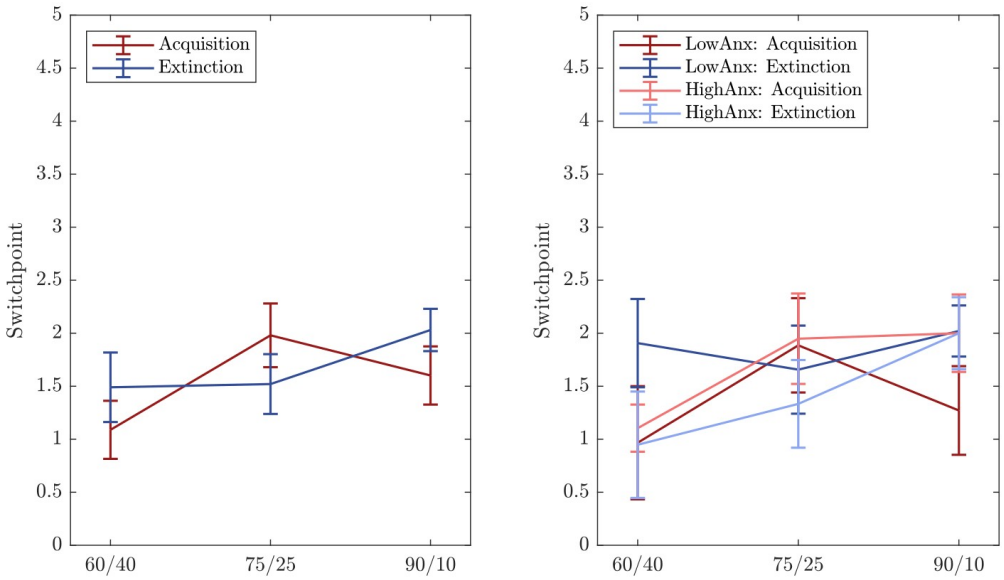


Switch point and steepness

The cumulative sum analysis was used to estimate the steepest point of learning after reversal. The analysis focused on two measures: switch point, measured in relation to the true reversal in trials, and switch steepness, measured as the α (steepness) parameter of the fitted sigmoid (see Chapter 2 for details).

For the switch point, a mixed ANOVA with phase and session as within-subject factors and anxiety as a between subject factor revealed a main effect of session, $F(2, 58) = 3.27, p = .045$. However, none of the post-hoc t-tests reached significance. Figure 9 presents the summary of switch point data.

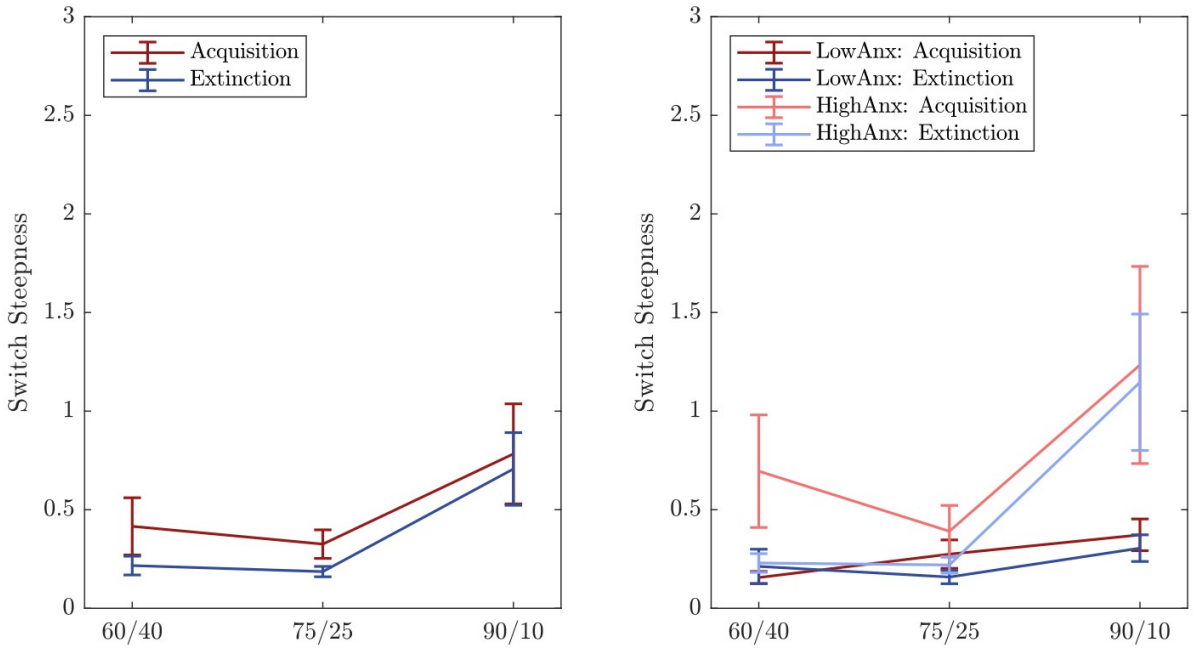
Figure 9: *Post-reversal switch point (indexed as number of trial post reversal) by phase (left) and phase and anxiety (right)*



Switch steepness data were analyzed using the same test as switch point. The main effects of session, $F(2, 60) = 8.39, p = .000$, and anxiety, $F(1, 30) = 5.62, p = .024$, as well as the interaction between session and anxiety, $F(2, 60) = 4.39, p < .017$ reached significance. Post hoc tests found a significantly higher switch steepness in the 90/10 session compared to 60/40,

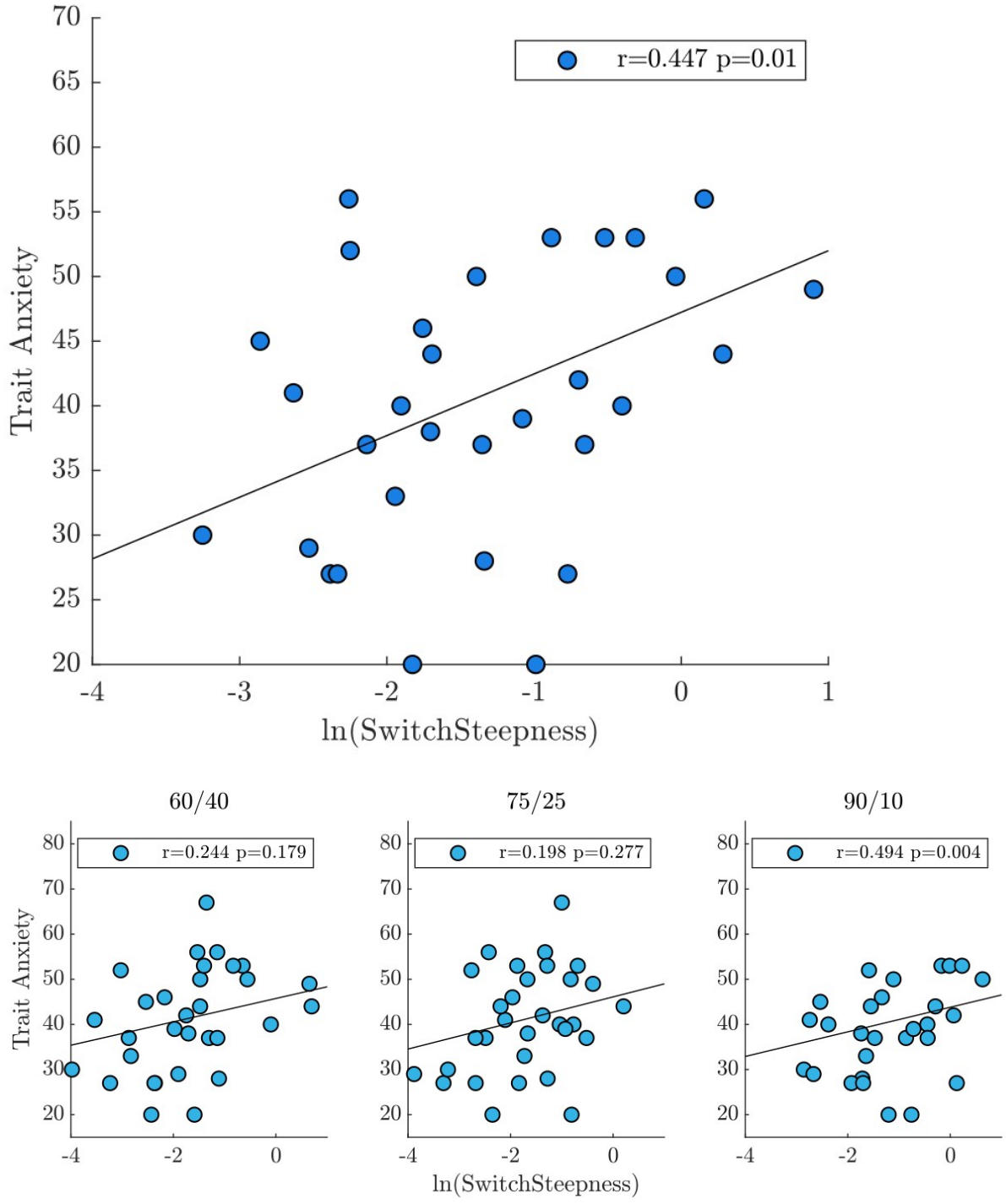
$t(32) = 2.74, p < .01$, and to 75/25, $t(32) = 2.85, p < .01$. No difference was found between 75/25 and 60/40.Following up on the main effect of anxiety, low anxious group ($\bar{x} = 0.25, \sigma = 0.18$) had significantly lower switch steepness than high anxious group ($\bar{x} = 0.65, \sigma = 0.66$). Post−hoc tests of the session*anxiety interaction found significantly higher switch steepness in the high anxious group in the 90/10 condition, $t(30) = 2.40, p < .05$, but this did not survive Bonferroni correction (p=0.023). Results of this analysis are summarized in Figure 10.

Figure 10: *Post-reversal switch steepness by phase (left) and phase and anxiety (right)*



To further explore the relationship between anxiety and switch steepness using the full anxiety scale a correlation between steepness and anxiety overall and per-block was run. Switch steepness significantly correlated with trait anxiety across the three sessions, $r = .447, p < .01$. Splitting the data by individual session, this correlation was only significant in the 90/10 block, $r = .49, p < .01$ which remained significant even after correction for multiple comparisons (Bonferroni, $p_{crit} = 0.0167$). Results for this analysis are presented in Figure 11.

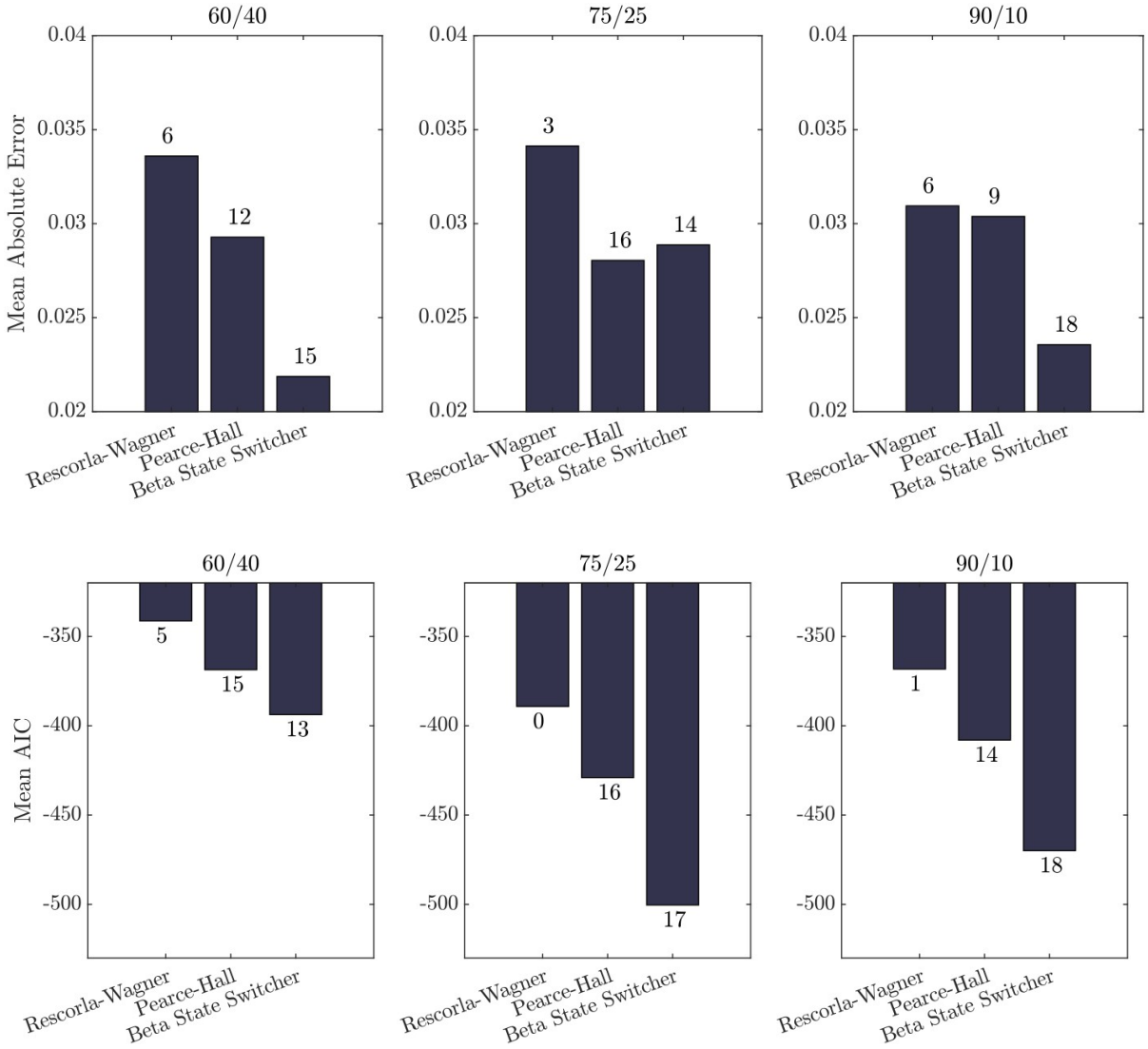
Figure 11: *Correlation between trait anxiety and switch steepness overall (top) and by session (bottom three).*



Modelling results

Three models were included in the model comparison: the outcome-sensitive Rescorla-Wagner model (a standard RW model with separate learning rate for shock and no-shock), the commonly used Pearce-Hall model (Li et al. 2011; Tzovara et al. 2018; Dunsmoor et al. 2019) and the novel beta state switcher introduced in this chapter. The model comparison was performed using cross validation as well as the standard AIC approach. To explore the former, the models were fitted to 70 - 80% of the data and used to predict the remaining data. The metric used for model comparison is the mean error between predicted and actual data. 30 fitting attempts were used for all models. In the 60/40 and the 90/10 session, the beta state switcher predicted data most accurately whereas in the 75/25 session the Pearce-Hall was most accurate closely followed by the beta state switcher. The results are summarised in Figure 12. For comparison, I include standard AIC scores but main conclusions are drawn from the cross-validation (CV) approach. The AIC and CV results differ: the mean AIC in 60/40 session was lowest for the beta state switcher but most participants were fitted best by the Pearce-Hall (as in the predictive metric).

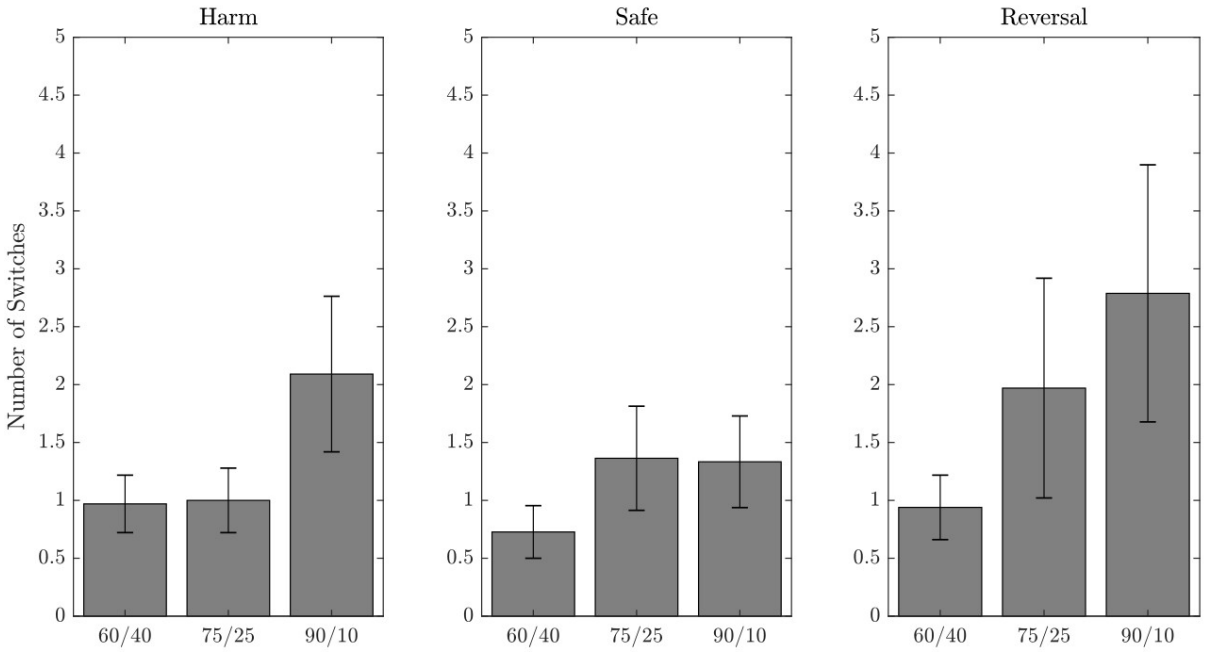
Figure 12: *Model comparison by session. Upper row shows the predictive power of models (lower is better), lower row shows mean AIC score (lower is better). The number above each bar tells the amount of participant best fitted by that model in that session.*



The model comparison results show that the state-switching strategy is most common in two of the three sessions (60/40 and 90/10). I next analyzed whether there is a relationship between anxiety and learning strategy by calculating mean anxiety per model. While mean anxiety for participants best fitted by the state switcher was higher than for the participants best fitted by PH, the difference was not significant in any of the three comparisons.

Since the beta state switcher achieved the lowest overall prediction error and AIC it is considered the overall winning model and its parameters were analysed. It is a new model, so several checks were performed to ensure that the model captures the characteristics it was designed to assess. Next, I also analysed the internal as well as free parameters of the model, such as the number of state switches estimated by the model. Participants switched states most frequently in in the 90/10 condition and, as expected, in the reversal cue condition (Figure 13). An ANOVA that included cue and session as within subject factors found a main effect of session, $F(2, 64) = 3.71, p < .05$. Mean number of switches per session were: in 60/40 $\bar{x} = 0.87, \sigma = 0.97$, in 75/25 $\bar{x} = 1.44, \sigma = 2.39$, in 90/10 $\bar{x} = 2.07, \sigma = 2.35$. The main effect was driven by significantly more switches in the 90/10 compared to the 60/40 condition, $t(32) = -2.71, p = 0.011$. Furthermore, the overall number of states estimated by the model correlated with the switch steepness, $r(32) = .47, p = 0.0061$. This confirms that the model captures at least some degree of state−switching behaviour.

Figure 13: *Mean number of switches predicted by the beta state-switcher model per session and cue.*



The parameter estimates of the beta state switcher, namely shock weight, no shock weight, leak, switch threshold, surprise learning rate and surprise decay (Figure 14) were analyzed using ANOVA with cue and session as within subject factors (Table 7). Subsequently, parameters per cue and session were correlated with trait anxiety (Table 8). Trait anxiety was found to correlate with cue-specific informative outcomes, that is positively, $r(32) = .44, p < .05$, with shock weight in the harmful cue and negatively, $\mathring{(}32) = -.43, p < .05$, with the no-shock weight in the safe cue in the 90/10 condition. The η parameter estimates how high is the threshold for a state switch. In the 90/10 condition, the η parameter was found to negatively correlate with trait anxiety, $\mathring{(}32) = -.40, p < .05$, indicating that participants had a lower threshold for state switching the higher they scored on the trait anxiety scale.

Table 7: *Statistical analysis of parameter estimates*

Shock weight τ_{sh}					
	SS	d.f.	MS	F	p
Intercept	11.380	1	11.380	17.780	**0.000**
Error	20.483	32	0.640		
Session	0.908	2	0.454	1.243	0.295
Error(Session)	23.374	64	0.365		
Cue	2.445	2	1.223	2.381	0.101
Error(Cue)	32.869	64	0.514		
Session*Cue	2.818	4	0.704	2.494	**0.046**
Error(Session*Cue)	36.159	128	0.282		
Noshock weight τ_{nosh}					
	SS	d.f.	MS	F	p
Intercept	5.936	1	5.936	7.513	**0.010**
Error	25.282	32	0.790		
Session	0.302	2	0.151	0.520	0.597
Error(Session)	18.605	64	0.291		
Cue	0.967	2	0.484	1.082	0.345
Error(Cue)	28.622	64	0.447		
Session*Cue	0.326	4	0.081	0.244	0.913
Error(Session*Cue)	42.701	128	0.334		
Leak λ					
	SS	d.f.	MS	F	p
Intercept	244.334	1	244.334	4,022.167	**0.000**
Error	1.944	32	0.061		
Session	0.525	2	0.263	7.444	**0.001**
Error(Session)	2.257	64	0.035		
Cue	0.031	2	0.016	0.673	0.514
Error(Cue)	1.480	64	0.023		
Session*Cue	0.117	4	0.029	1.166	0.329
Error(Session*Cue)	3.207	128	0.025		
Switch threshold η					
	SS	d.f.	MS	F	p
Intercept	8,766.158	1	8,766.158	862.296	**0.000**
Error	325.314	32	10.166		
Session	14.775	2	7.388	0.793	0.457
Error(Session)	596.072	64	9.314		
Cue	33.426	2	16.713	1.893	0.159
Error(Cue)	565.021	64	8.828		
Session*Cue	27.141	4	6.785	0.741	0.566
Error(Session*Cue)	1,171.891	128	9.155		
Surprise learning rate α_{surp}					

	SS	d.f.	MS	F	p
Intercept	32.834	1	32.834	358.488	**0.000**
Error	2.931	32	0.092		
Session	0.678	2	0.339	4.652	**0.013**
Error(Session)	4.665	64	0.073		
Cue	0.048	2	0.024	0.207	0.814
Error(Cue)	7.457	64	0.117		
Session*Cue	0.173	4	0.043	0.456	0.768
Error(Session*Cue)	12.158	128	0.095		
Surprise decay λ_{surp}					
	SS	d.f.	MS	F	p
Intercept	78.178	1	78.178	607.699	**0.000**
Error	4.117	32	0.129		
Session	0.634	2	0.317	3.296	**0.043**
Error(Session)	6.151	64	0.096		
Cue	0.084	2	0.042	0.483	0.619
Error(Cue)	5.554	64	0.087		
Session*Cue	0.571	4	0.143	1.081	0.369
Error(Session*Cue)	16.910	128	0.132		

In the **shock weight** parameter, there was a significant interaction between cue and session (see Table 3). Post hoc one-way ANOVAs were run for each cue and session. There was a main effect of cue in session 90, $F(2, 96) = 6.13, p = 0.003$ which was driven by significantly higher shock weight parameter in the reversal cue ($\bar{x} = 0.47, \sigma = 0.60$) than in the safe cue ($\bar{x} = 0.11, \sigma = 0.68$), $t(32) = -2.69, p = 0.011$, and than the harmful cue ($\bar{x} = -0.06, \sigma = 0.60$), $t(32) = -3.82, p = 0.001$.

In the **leak** parameter, there was a main effect of session. The means were: 60/40, $\bar{x} = 0.95, \sigma = 0.03$, 75/25, $\bar{x} = 0.92, \sigma = 0.13$, 90/10, $\bar{x} = 0.85, \sigma = 0.16$. There was a significant difference between the harmful and the reversal cue, $t(32) = 3.74, p = 0.001$, and between safe and reversal cue, $t(32) = 2.07, p = 0.047$ but only the earlier survived correction.

In the **surprise learning rate** parameter, there was a main effect of session.

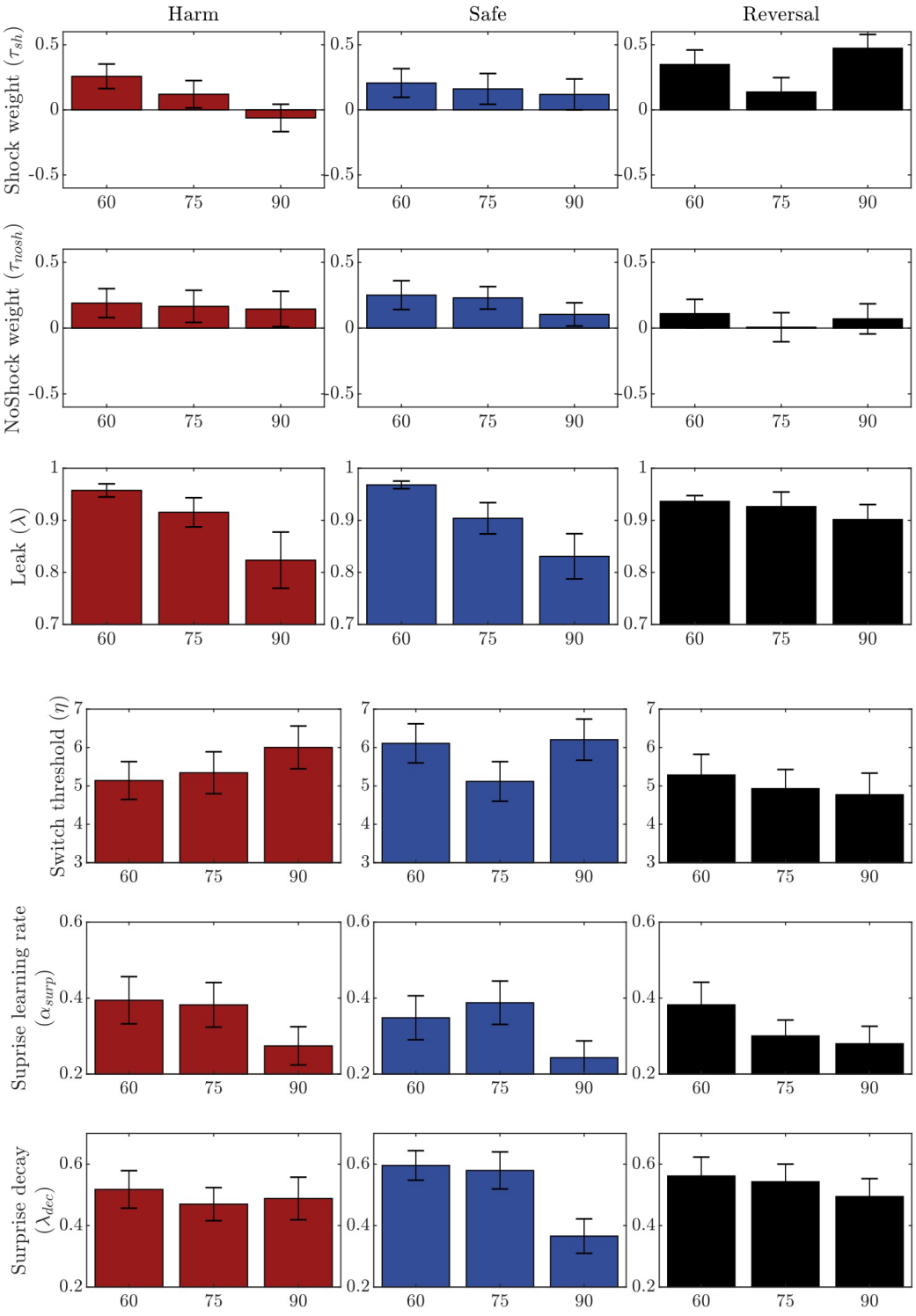
The means were: 60/40, $\bar{x} = 0.38, \sigma = 0.18$, 75/25, $\bar{x} = 0.36, \sigma = 0.16$, 90/10, $\bar{x} = 0.27, \sigma = 0.15$. There was a significant difference between the harmful and the reversal cue, $t(32) = 2.86, p = 0.007$, and between safe and reversal cue, $t(32) = 2.43, p = 0.021$ but only the earlier survived correction.

In the **surprise decay** parameter, there was a main effect of session. The means were: 60/40, $\bar{x} = 0.56, \sigma = 0.15$, 75/25, $\bar{x} = 0.53, \sigma = 0.21$, 90/10, $\bar{x} = 0.45, \sigma = 0.21$. There was a significant difference between the harmful and the reversal cue, $t(32) = 2.66, p = 0.012$, no other comparison was significant.

Table 8: *Correlations between parameter estimates and trait anxiety*

	60/40	75/25	90/10	60/40	75/25	90/10
	Shock weight τ_{sh}			Noshock weight τ_{nosh}		
Harmful Cue	0.058	0.134	**0.444**	-0.099	-0.047	-0.285
Safe Cue	-0.096	-0.340	**-0.434**	-0.078	0.100	0.172
Reversal Cue	-0.116	0.177	0.048	-0.015	0.114	-0.005
	Leak λ			Switch threshold η		
Harmful Cue	-0.004	0.222	-0.343	0.011	0.130	-0.033
Safe Cue	0.165	**-0.442**	-0.067	-0.048	-0.137	-0.112
Reversal Cue	-0.285	-0.235	0.056	0.097	0.308	**-0.402**
	Surprise learning rate α_{surp}			Surprise decay λ_{surp}		
Harmful Cue	-0.106	-0.098	-0.222	0.042	0.005	0.002
Safe Cue	-0.276	0.012	0.303	-0.045	-0.250	0.195
Reversal Cue	0.157	0.164	-0.246	-0.248	0.062	-0.230

Figure 14: *Mean parameter estimates of the beta state switcher model by session and cue*



Galvanic Skin response

GLM Results

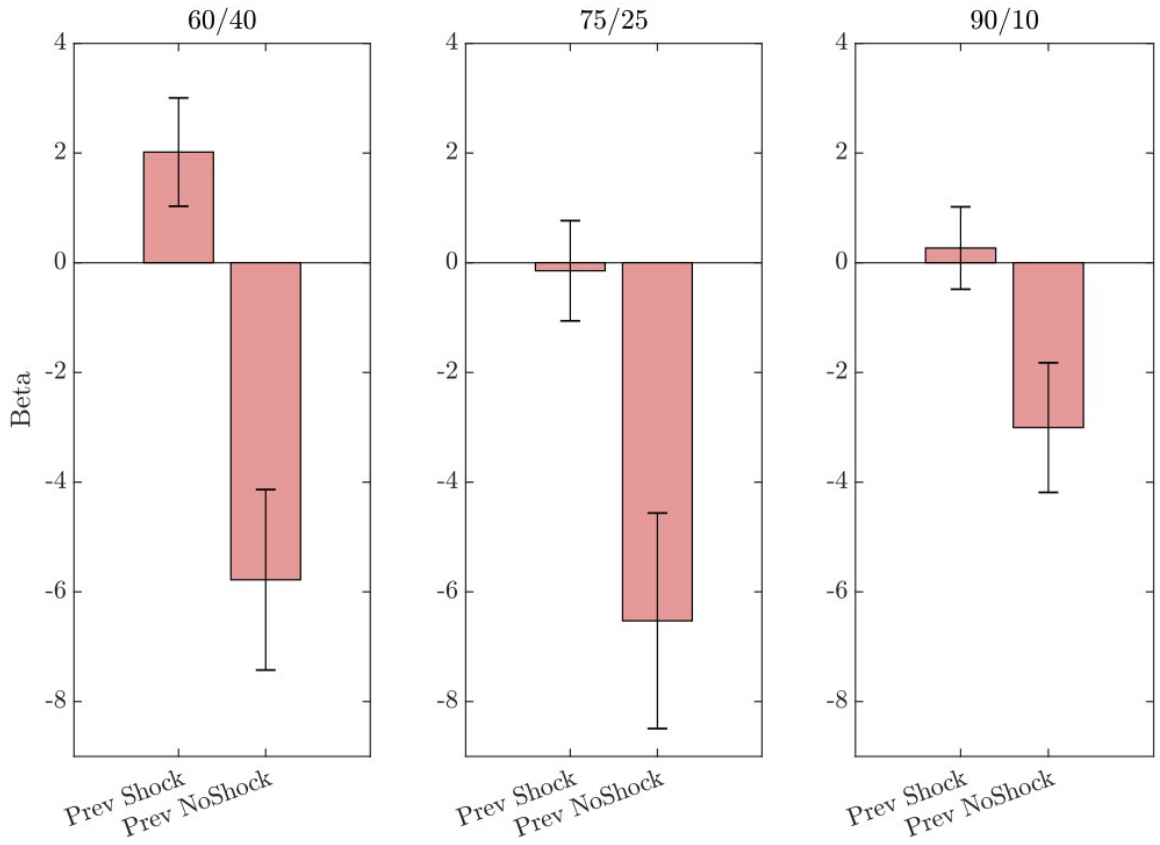
The data were analyzed using the GLM approach described in Chapter 2. Initial checks revealed residual influence of the previous outcome on the GSR amplitude in

the current trial: the SCR's generated in response to the shock stimuli lasted up to 6 seconds. Additionally, since the SCR latency following a shock is variable this can introduce a confounding effect in estimation of betas. Therefore, to ensure that the data do not include responses to the previous shock, only trials where no shock had been delivered in the previous trial were included.

Table 9: *GSR data, ANOVA*

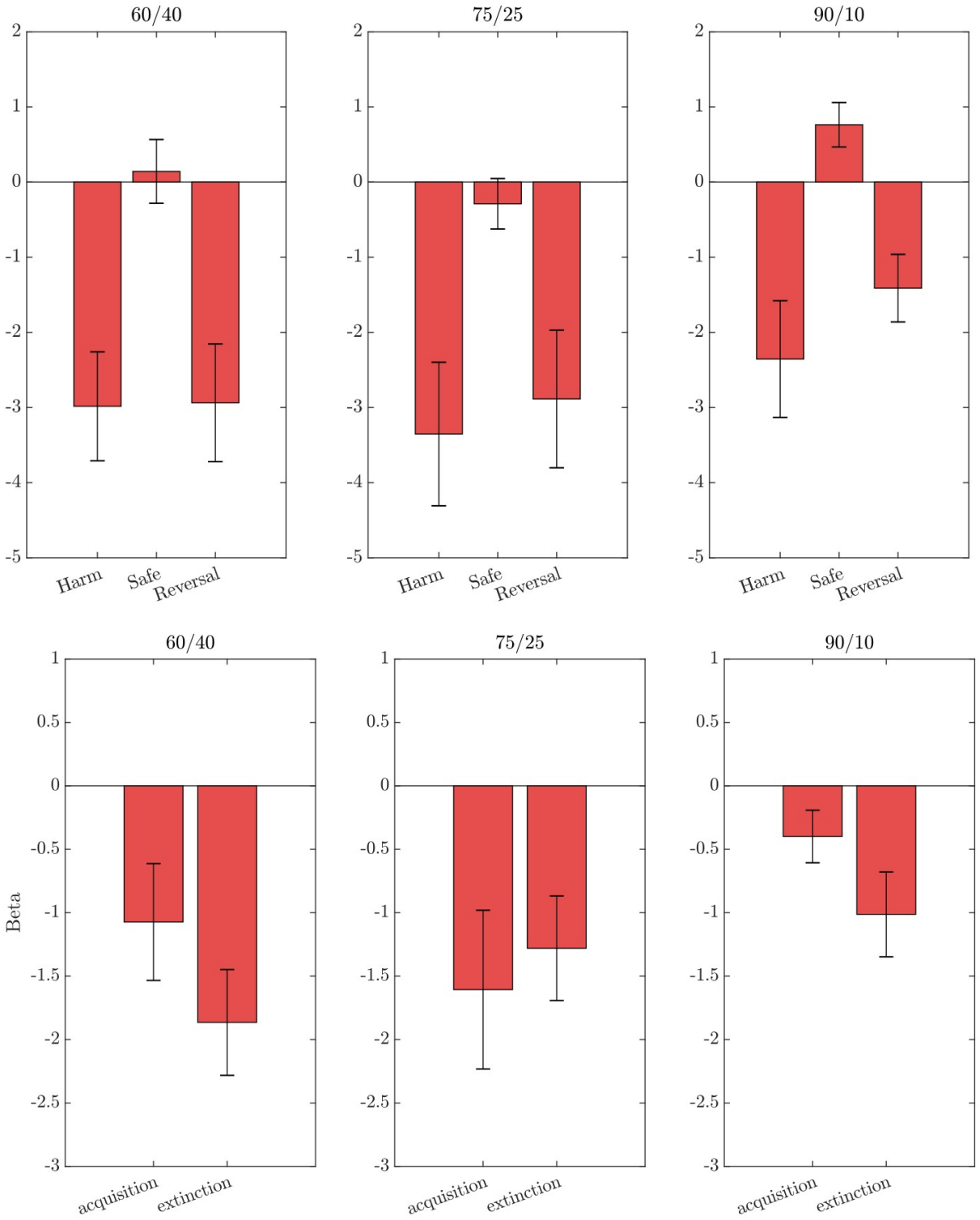
	SS	d.f.	MS	F	p
Intercept	896.575	1	896.575	19.780	0.000
Error	1,359.832	30	45.328		
Session	129.377	2	64.689	3.671	0.031
Error (Session)	1,057.178	60	17.620		
PrevOutcome	1,573.898	1	1,573.898	9.430	**0.005**
Error (PrevOutcome)	5,007.155	30	166.905		
Session*PrevOutcome	166.071	2	83.035	2.275	0.112
Error (Session*PrevOutcome)	2,190.165	60	36.503		

Figure 15: *Effect of previous outcome on estimated betas*



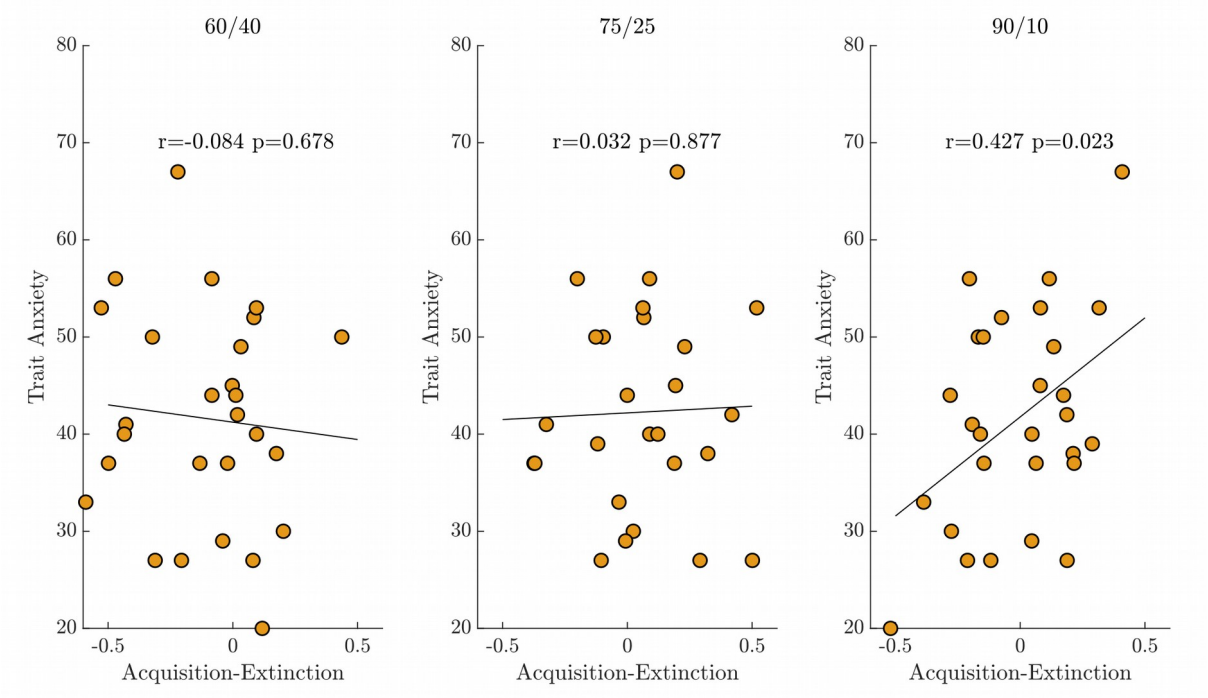
Comparing SCRs between sessions and cues, an ANOVA found a significant main effect of cue, $F(2, 60) = 22.23, p < .001$ which was driven by differences between safe and harmful, $t(30) = -5.22, p < .001$ and between safe and reversal cues, $t(30) = -5.40, p < .001$ (Figure 14, upper panel). The response to harmful was similar a s response to the reversal cue. Taking only the reversal cue, I next tested the difference between acquisition and extinction. Neither the main effect of phase or session nor the interaction between phase and session reached significant (Figure 14, lower panel).

Figure 16: *Mean betas for cue by session (upper) and for reversal cue by phase (lower panel).*



Similarly to Chapter 3, I also estimated single trial anticipatory responses (GSR amplitude) using a DCM model. Furthermore, I tested whether GSR differences between acquisition and extinction scaled with trait anxiety. While there was a significant correlation in the 90/10 condition, this did not survive FDR correction ($p_{corr} = 0.065$).

Figure 17: *Per-session correlations between trait anxiety and acquisition-extinction difference*



Discussion

In the present study I investigated how environmental statistics influences aversive acquisition and extinction and learning in relation to trait anxiety. I found that while higher contingencies were reported with increasing reinforcement level in acquisition, in extinction there was no difference even between the most extreme cases (10% vs 40% of shocks). Next, I tested whether higher difference between

acquisition and extinction promotes state-switching strategy. Switches in the 90/10 condition were significantly steeper than in 75/25 and 60/40. There was no difference between 75/25 and 60/40 in switch steepness, suggesting that there is a step change in strategy between 75/25 and 90/10 which I attribute to state switching. While the 90/10 condition was best fitted by the state switching model, the 75/25 seemed to favour gradual learning (best fitted by Pearce-Hall). Investigating the role of anxiety, the increased switch steepness was driven by the high trait anxious individuals, suggesting increased tendency for state switching. Furthermore, a critical parameter in the beta state switching model controlling the threshold for creating new states significantly negatively correlated with trait anxiety, providing further proof that high anxiety is associated with state, rather than gradual, learning.

The goals of this project were to understand 1) how different shock contingencies influence learning and learning strategy and 2) whether high trait anxious individuals tend to represent the environment as multiple states and switch between them.

Both, post-task reported probability and averaged trial-by trial shock probability ratings show that the subjects followed the experimental manipulation. This was also reflected in cue liking: the safe cue was liked the most while the harmful cue was liked the least. This was not due to visual appeal of the randomly allocated cue. In the reversal cue, we found a slight overprediction bias in the reversal cue on all three visits. Splitting the reversal cue data by acquisition and extinction, participants could distinguish between the phases. There was a degree of over-prediction in extinction and under-prediction in acquisition caused by data averaging that included even trials immediately post-reversal. This effect remained when the first three post-reversal trials were excluded suggesting a general tendency not to report extreme values, since this under/over-prediction is strongest in the 90/10 condition. Overall, the data suggest that participants were able to learn the manipulated contingencies.

To address the first question, I first compared the mean probabilities between sessions. For acquisition the mean probability ratings increased with underlying shocks rate. In contrast the mean probability was the same on all three visits in the extinction phase, indicating that the resistance to extinction does not depend on the actual reinforcement rate. It needs to be noted that shock probability ratings differed considerably during extinction which could not be explained by starting probability, order of blocks or cue co-dependence. Adding trait anxiety into the analysis, there was a near-significant difference between high and low trait anxious individuals. A parametric analysis found a significantly positive correlation between the level of trait anxiety and the difference in shock probability ratings between acquisition and extinction, suggesting that lack of extinction is at least partially driven by anxiety. This effect was significant in the 90/10 and 75/25 conditions, although neither survived the strict Bonferroni correction (see discussion below). Secondly, while post-reversal switch point did not differ between sessions, the switch steepness in the 90/10 condition was higher than in the two other conditions. Interestingly, the steepness did not differ between the 60/40 and 75/25 conditions. This might be a signature of different learning strategies employed, suggesting that i) the state-switching learning occurs at high contingency levels; and ii) that the critical difference between low and high reinforcement levels at which new states start to be created lies between 0.5 and 0.8. To test the assumption that switch steepness relates to the use of state switching strategy we correlated the number of states that the beta state switcher estimated with switch steepness. The significant correlation provides a degree of evidence in support of our assumption.

Lastly, to directly assess whether state switching is more likely to occur at high contingency difference levels I analysed the model comparison data and the fitted

parameters. The beta state-switcher had a better model fit (mean prediction error) in the 90/10 than in the 75/25 condition, suggesting that as the contingency difference between acquisition and extinction increases, participants were more likely to create multiple states. This is further supported by the observation of a linear increase in the number of state switches with an increase in phase contingency difference.
As revealed by the model comparison, more subjects were fitted by the beta state switcher in the 60/40 than the 75/25 condition. While this seems to counter our hypothesis it is important to note that the one-state version of the beta state switcher captures Rescorla-Wagner-like learning pretty well. It doesn't, however, have a mechanism to accelerate learning like PH which is why the main conflict between state-switching and PH learning strategies occurs in the 75/25 and 90/10 conditions.

The behavioural and modelling data support our hypothesis that increasing the difference between reinforcement levels of acquisition and extinction leads to increased reliance on multi-state representation of the environment.

Relationship between anxiety and learning strategies

The second goal of this study was to investigate whether high trait anxious participants tend to represent the environment as multiple, rather than one, states. The behavioural and modelling data converge to support this hypothesis. In our behavioural results, higher trait anxious participants showed better discrimination between acquisition and extinction and higher post-reversal switch steepness. In the modelling results, trait anxiety correlated negatively with the threshold for a new state creation in the 90/10 condition.

High anxiety has previously been associated with lack of extinction and higher renewal/reinstatement rates (Mineka and Zinbarg 2006). Here, I show that one of the

potential causes for an increased tendency to relapse is a multi-state representation of the environment, i.e. high trait anxiety is associated with more context-dependent learning. Previous clinical reports (Craske et al. 2014b) highlight the need for extinction generalization as a requirement for successful treatment of anxiety disorders. In exposure therapy, extinction training is performed in a clinical setting. High anxious individuals are known to associate extinction with the specific context (i.e. therapist's office) but fail to transfer it to everyday life, particularly to contexts where conditioning took place. Here we show that apart from context being driven by external stimuli (i.e. color of the room), it is also possible to create an internal context based on pattern in observations. Occasionally reinforced extinction is a method previously used to enhance permanence of exposure therapy. Under such treatment, for example in exposure therapy for phobia, the feared object is occasionally paired with the unconditioned stimulus during extinction. From the information processing perspective, such reinforcement schedule leads to medium prediction errors that mostly lie within the expected range for a given state. This is in contrast to a scenario with no occasional reinforcement where the learning rate weighted prediction errors become too large, which leads to the creation of a new state (i.e. internal context). Our results show that high trait anxious individuals have a lower threshold for the creation of a new state in highly variable environments (90/10 condition). In other words, less surprising outcomes will lead to the creation of a new internal context in more trait anxious individuals. This is the first time anxiety has been associated with increased tendency to represent environment as multiple states.

In our behavioural analysis, high anxiety was associated with increased dissociability between acquisition and extinction in the behavioural ratings, similarly to Chapter 3, but no discrimination in the GSR analysis. While the subjective ratings result contradicts a number of previous studies (Staples-Bradley, Treanor, and Craske 2018; Andreatta and Pauli 2017) the GSR result is in line with previous studies that found no difference in SCR response between acquisition and extinction (Torrents-

Rodas et al. 2013). Recent work by Atlas et al. (in submission) using an instructed learning task showed that high trait anxious participants performed sudden post-reversal jumps when instructed, resulting in an increased difference between mean acquisition and extinction ratings. This resembles my finding that high trait anxious individuals tend to create and switch between multiple states. In an fMRI study using instructions, instructed learning was associated with neural processing in the OFC and the striatum while gradual learning was associated with activation of the amygdala. This neural dissociation would support the difference in ratings and GSR results. The amygdala has previously been associated with the GSR signal (Li et al. 2011) while the striatum and the OFC are believed to support higher-order learning as well as state representation (Schuck et al 2015).

Limitations

The present study suffers from a number of limitations. Foremostly, while we show that switch steepness is associated with the number of states created by the beta state switcher model, it is not possible to definitely reject alternative hypotheses given the unobservable nature of the hidden subjective states. For example, both, the model and the switch steepness measure, could pick up on similar patterns in the learning data that do in fact not reflect state switching. To mitigate this, several steps were undertaken: the data fits (both model and sigmoid fits) were visually inspected, and cross validation (rather than fit) was used for model comparison which practically eliminates the danger of overfitting.

Our task was designed to maximise the amount of trials within a reasonable time frame. For that reason, the task was not optimised for GSR analysis. Signals from the previous trial affected the signal on the current trial. To mitigate this issue,

we used the GLM approach in which all events are explicitly modelled and convolved with the canonical SCR function. While this method helps to decouple overlying signal it assumes latency invariance of the SCRs. Future work should employ methods accounting for latency variability of aversive SCRs. While the dynamic causal modelling (DCM) approach solves this issue it is not yet implemented for overlapping signals. Future tasks should consider longer inter-stimuli and inter-trial intervals to prevent inter-trial signal overlap.

In Chapter 3 I reported that the mean anxiety of participants best fitted by the state-switching model is significantly higher than those fitted by the Pearce-Hall model. While we did not directly replicate those individual differences in the model fit, it must be noted that here a different model was used to capture state-switching behaviour. Most importantly, the beta state switcher has internal mechanisms that make it suitable to capture Rescorla-Wagner style learning as well as state switching. That in turn means that state-switching and some gradual learning participants will be well fitted by the model.

While our beta state switcher model fits and predicts that data with high accuracy it is a descriptive account of the data and has no direct biological interpretation. An alternative version of this model can be formulated in which the basic learning mechanism does not rely on the leaky beta model but rather on a biologically supported learner such as the Rescorla-Wagner model.

Conclusions

In summary, we investigated the influence of environmental contingencies on probabilistic aversive learning and differential learning strategies. Our data support the notion that environments with higher difference between acquisition and extinction contingencies encourage multi-state representation of the environment. Furthermore,

we tested whether trait anxiety increases the rate of state learning. Both our behavioural and modelling results support the hypothesis that high trait anxious individuals have a higher tendency to represent the environment as multiple states. This has potentially important implications for exposure therapy.

Chapter 5

Angiotensin II antagonist losartan decreases learning rate but does not alter perception or updating in aversive learning

Introduction

Anxiety disorders are very common and disabling conditions and although exposure-based cognitive-behavioural therapy (CBT) is widely deemed effective, its effects are often only moderate and difficult to maintain (Craske et al. 2014). There has recently been a growing interest in amplifying the effects of CBT through a combination of therapy and pharmacological interventions designed to improve neuroplasticity, the key cellular mechanism underlying learning (de Kleine, Rothbaum, and van Minnen 2013; Fitzgerald, Seemann, and Maren 2014). The drug investigated most commonly is cycloserine, which targets glutamatergic transmission. Clinical trials have shown that the drug significantly increases the number of treatment responders when administered before CBT (Ressler et al. 2004; Otto et al. 2010). However, a growing number of studies looking at the effects of cycloserine on CBT outcome have reported no or even detrimental effects on patients' well-being (Andersson et al. 2015; Mataix-Cols et al. 2014). Such effects occur because the drug drives the extinction of fear, which is thought to reduce anxiety, but also increases fear learning, which is a cause of anxiety. As a consequence, patients' anxiety symptoms can worsen if cycloserine is given before an unsuccessful session of CBT (Smits et al. 2013). These findings call for alternative drugs that can selectively enhance fear extinction and/or reduce excessive acquisition.

The angiotensin receptor antagonists (ARA) losartan has recently been highlighted as a potential enhancer of exposure therapy in anxiety disorders (Shekhar 2014). ARAs are routinely used to treat hypertension and cardiovascular diseases. Angiotensin receptors are found in the central nervous system (Allen, Zhuo, and Mendelsohn 1999), particularly in the hypothalamic-pituitary-adrenal axis (HPA; Krause et al., 2011), the central nucleus of the amygdala (Shekhar 2014) and the locus coeruleus (Gong et al. 2015). ARAs have been linked to a number of cognitive

processes, including stress, anxiety (Armando et al. 2007; Krause et al. 2011; Saavedra, Sánchez-Lemus, and Benicky 2011), memory and learning (Mechaeil et al. 2011); Marvar et al., 2014; (Gard 2004; Kułakowska et al. 1996).

In clinical studies, ARAs have been reported to improve symptoms of depression and anxiety in type 2 diabetes patients (Pavlatou et al. 2008). In a large cross-sectional study in army veterans long-term intake of losartan was significantly associated with decreased symptoms of post-traumatic stress disorder (Marvar et al. 2014; Khoury et al. 2012). Animals studies revealed that losartan can enhance the consolidation of extinction memories, while leaving fear acquisition, baseline anxiety and blood pressure unaffected (Marvar et al., 2014). A similar role of losartan in fear extinction was more recently reported by (Lazaroni et al. 2016), who demonstrated the importance of the Angiotensin (1-7)/Mas axis in fear memory extinction. Braszko (2002) reported an effect of losartan and angiotensin II on fear acquisition. The authors administered Angiotensin II which lead to increase in fear acquisition, and subsequently compared the effect of angiotensin type 1 (AG_1) and 2 (AG_2) receptor blockade on fear acquisition measured by conditioned response. Administration of losartan (AG_1 blocker) lead to reduced fear acquisition. While this finding suggests involvement of losartan in acquisition it must be noted that in animals that were not treated with free Angiotensin-II there as no difference between losartan and saline on fear acquisition, and those were both slower than the AG-treated group.

In contrast to these documented effects of losartan on learning and memory in animals, only little is known about the role of losartan in human aversive learning. The only available study so far found a reduction in the aversive but not in the appetitive learning rate (Pulcu et al. 2018). The authors administered a single dose (50mg) of losartan to participant performing a decision-making task incorporating wins and losses. Separate learning rates were then estimated for the two types of outcomes, identifying decrease in learning from losses to be specific to the losartan group.

In controlled experimental settings, aversive learning is commonly studied using a fear conditioning paradigm that involves an acquisition phase during which a cue is paired with an unconditioned stimulus, US (e.g., a painful electrical stimulus; i.e. acquisition) and an extinction phase during which the cue is no longer followed by the US (Craske et al., 2018). Extinction learning refers to the rate at which the agent decreases their expectation of aversive outcome. This can either be measured by autonomic responses, such as galvanic skin response (GSR), or a self-reported measure of US expectancy.

Here, I investigated the effect of losartan on aversive probabilistic learning in humans using shock expectancy ratings and pupil dilation as two measures of learning. I also employed Bayesian modelling together with pupil analysis to investigate which part of the learning process (learning rate, perception or updating) is impacted by the drug. Recent work in computational psychiatry has highlighted the use of mathematical learning models in understanding the brain mechanisms underlying learning in health and disease (Browning et al. 2015; Schlagenhauf et al. 2014; Lawson, Mathys, and Rees 2017). Reinforcement learning and Bayesian learning models allow us to infer hidden cognitive variables and components of learning that would otherwise be inaccessible. Most commonly estimated hidden parameter is the learning rate. However, the learning process has at least two other major components that can contribute to the observed behaviour: perception and updating (see O'Reilly et al. 2013; Mathys et al. 2014; Jepma et al. 2018). In the given context, perception is how intensely the participant perceived the outcome, i.e. some outcomes might be perceived more intensely in some subjects. Updating is a quantity by which a participants tends to over- or under- update systematically across all trials. If only the learning rate is modelled, a particular group or visit difference may seem to impact the learning rate while in fact it impacts the perception or updating. This is particularly important in drug studies as there is an increased chance that they will influence perception in particular. In the present study I will use a Bayesian model that

incorporates and estimates all three components at the same time, thus suggesting which factor is most likely impacted by the group.

For a similar reason I will employ pupillometry on visit 2 where the drug is likely to have the biggest effect. Using a regression analysis, I will examine differential encoding of shock expectancy, outcome processing and prediction error. This, together with the Bayesian model, will inform our analysis relinating learning, perception and updating. Pupil dilation is a suitable measure as it has been shown to reflect uncertainty and prediction error processing in learning and decision making paradigms (Colizoli et al. 2018; Lavín, Martín, and Jubal 2014; Koenig, Uengoer, and Lachnit 2018). Additionally, pupil size has been linked to the activity of the locus coeruleus (LC; (Joshi et al. 2016), a nucleus in the brainstem containing, among others, Angiotensin II receptors. While studies employing pupil dilation often focus on the role of norepinephrine in the LC (Aston-Jones and Cohen 2005), in the present study we considered the effect of losartan on pupil dilation due to the presence of angiotensin receptors in the LC as a mechanism of how the drug might impact learning.

The goal of the present study is to develop a procedural and neural understanding of the effect of losartan on aversive learning and aid the assessment of losartan as a potentially CBT-enhancing drug. To this end, I compared the effect of losartan and placebo on fear acquisition and extinction using a probabilistic learning paradigm in conjunction with computational modelling and pupil dilation analysis. I hypothesise that the administration of losartan will affect learning via decrease in learning rate in acquisition, preventing the reinstatement of fear.

Methods and Materials

Participants

Forty-five healthy volunteers (age 18–39) were recruited through local advertisements. Due to a lack of studies exploring the effect of losartan on learning outcomes, we estimated our sample size based on the only available study into the effects of losartan on cognitive function in healthy volunteers (Michaeil et al., 2011). Calculations suggested 20 participants per group to achieve effect sizes of *d*=0.9 and a statistical power of 80% (α-level 0.05). We aimed for an overall sample size of 45 to allow for the exclusion of participants who failed to meet the basic learning criteria described below. Participants were included in the study if they had no history of a DSM-IV Axis I disorder as assessed using the Structured Clinical Interview for DSM-IV (First et al. 2015). Participants also had to have been free from CNS-active medication for at least six weeks, have a body mass index between 18 and 30 kg/m^2, and have no first-degree family member with a history of a severe psychiatric disorder. The study was approved by the Oxford University Research Ethics Committee (R29583/RE001), and all participants gave written informed consent. Five participants had to be excluded from the study at the point of data analysis, two due to technical failure of the equipment, and three (all placebo group) because they failed to show learning of the difference between low and high probability cues (assessed by a repeated measures t-test), leaving 21 participants in the losartan group (7 female; $\bar{x} = 26.09, \sigma = 5.32$, range $19 - 39$) and 19 in the placebo group (10 female; $\bar{x} = 25.00, \sigma = 4.23$, range $19 - 35$). The exclusion assessment was made independently by two of the investigators, who were both blind to the participants' group assignment.

Materials and Study Design

The study involved three visits to the Department of Psychiatry at the University of Oxford. *Visit 1* (Screening visit) included medical and psychiatric screening followed by an introduction and completion of a short version of the fear learning task (described below). *Visit 2* (Main testing visit) included completing a battery of psychological questionnaires, drug administration, one hour of waiting time and a long version of the fear learning task. In addition to behavioural data, pupil dilation data were collected. *Visit 3* (Follow-up visit) took place one day after the main testing visit. Participants completed a questionnaire to assess the occurrence of side effects and a shorter version of the fear learning task.

In a double-blind design, participants were randomly allocated to one of two groups: one group that received a 50mg single oral dose of losartan (Cozaar, Merck Sharp & Dohme Ltd.) or a second group that received a placebo that was matched to the active drug in appearance (microcrystalline cellulose; Rayotabs, Rayonex GmbH). Dosing of losartan was guided by the intention to reduce fear without triggering hypotensive effects which had been achieved in a previous study using 50mg losartan. Testing started one hour after capsule intake, when drug peak plasma levels are reached (Lo et al. 1995, Ohtawa et al. 1993). Mood and physiological symptoms were assessed by the participants on a visual analogue scale (VAS) before administration of medication and prior to testing before their heart rate and blood pressure were measured using an Omron 705IT sphygmomanometer. At the end of the testing session both participant and experimenter were asked to indicate whether they thought the participant had received losartan or placebo.

Fear Learning Task

To investigate acquisition and extinction learning, the same probabilistic aversive learning paradigm as described in Chapter 2 was employed. Each of the three

sessions consisted of 150 (visits 1 and 3) or 300 (visit 2) trials. The length of one phase was always the same, what differed between the short and the long version was the amount of phases which were 5-6 and 10-11, respectively.

Outcome measures

Unless specified otherwise, reported data refer to the reversal cue.

Shock probability ratings

A shock expectancy rating was collected on each trial. To investigate the influence of losartan on acquisition and extinction, group and visit specific mean probability ratings were calculated separately for trials in which the objective shock probability was low (extinction phase) and trials in which the shock probability was high (acquisition phase). Probabilities across trials of interest were averaged to form a mean shock probability. For example, the mean shock probability during acquisition represents the averaged expectancy ratings over all trials in which the true shock probability for the reversal cue was low, i.e. extinction phase.

To analyze post-reversal learning patterns, probabilities reported on each trial after a reversal occurred were averaged across a specific phase (acquisition / extinction) and group.

Model-free Learning Rates

Learning rates are commonly estimated as a single learning rate across all trials of the task or condition. However, as we intend to characterise trial-by-trial learning, we used the Rescorla-Wagner (eq. I) learning rule to calculate the trial-specific learning rate α (eq. II).

Eq. I $$P_{t+1} = P_t + \alpha_t(O_{i,t} - P_t)$$

Eq. II $$\alpha_t = \frac{P_{t+1} - P_t}{O_t - P_t}$$

where $$0 \leq \alpha_t \leq 1$$

We performed a model-based/model-free learning rates comparison to ensure the validity of the latter (full analysis in Appendix I of Chapter 2). Model-free learning rates across all cues and visits correlated moderately with model-based learning rates (*r*=.51). This correlation was stronger in the reversal cue (*r*=.65).

Pupillometry

Pupillometry (diameter area) was recorded was during visit 2. The recording was performed using the tower-mount EyeLink II eye-tracking hardware and the appropriate recording software (SR Research Ltd., Ottawa, Canada). Binocular pupil diameter and x and y gaze coordinates were recorded at a 500 Hz sampling frequency. Prior to the start of the session the eye-tracker was calibrated using a nine-point calibration. The lighting conditions were kept constant for every participant and the experiment took place in a room with a sealed window.

Data were processed using custom MATLAB scripts following the appropriate guidelines (Kret & Sjak-Shie, 2018) and similar studies (Nassar et al., 2012; Browning et al., 2015; Pulcu and Browning, 2017). The EDF data were converted to ASCII format. Segments of blinks and missing data were identified by calculating the first-order derivative (*df*) and excluding data points with $df > 1.6$. Data segments that deviated by more than one standard deviation from the mean of the entire session were excluded from the data set. Where data gaps were shorter than 2s, Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) was used to interpolate the missing data. Trials with more than 50% interpolated data points were excluded. The resulting data were subsequently filtered using a bandpass filter between 0.1 and 4 Hz. Lastly,

each trial was baseline corrected using 0.5 prior to event, and z-transformed.

Statistical Analysis

For all analyses, the full statistical results are presented in tables and only significant or directly relevant insignificant results are mentioned in the text.

Questionnaire, Physiological and Sociodemographic Data

Sociodemographic and questionnaire data were summarized by mean and standard deviation and tested between groups using either a t-test or a χ^2 test. The effect of losartan on physiological and VAS measures between baseline and peak level was assessed using a mixed ANOVA with time (baseline/peak) as a within-subject and group as between-subject factor.

Starting Probability

Each session started randomly with either high or low probability of shock (i.e. either with 75% or 25% chance of shock). While the starting reinforcement rate of the reversal cue was randomized it is important to check whether it exerts any effect on learning. For example, participant with 75% of shocks in their first 30 trials might develop a different initial association with the reversal cue, considering it another harmful cue. This, in turn, might bias their learning towards the category of “high chance of shock”. To check whether the starting probability (SP) on each visit plays a role it was added into analyses where it was possible as a between-subject variable. Since a given participant didn’t always start with the same SP this between subject factor was different for each visit.

Probability Ratings Data

Behavioural Data Statistical analyses were carried out using Matlab (MathWorks), R and JASP using two-tailed tests at $\alpha = 0.05$. Potential drug-induced changes in shock probability ratings and model-free learning rates were assessed using the appropriate parametric tests (ANOVA, t-test). The basic mixed ANOVA model for all analyses consisted of the within-subject factor VISIT with three levels (visit 1, visit 2, visit 3) and the between-subject factor GROUP (placebo group and losartan group). Where appropriate, the factors PHASE (within-subject factor with the levels acquisition and extinction) and OUTCOME, (within-subject factor with the factors shock and no shock) were included. In order to explore the influence of trait anxiety on extinction learning, the sample was split into those scoring high on the STAI-T questionnaire and those scoring low (median split; between-subject factor with the levels low and high anxiety). In all analyses the starting true shock probability of the reversal cue for a given visit was included as a covariate. This between-subject variable had two levels (75%, 25%).

To investigate how pain expectation changes after reversal, data for each reversal were extracted and averaged per visit, phase, group and starting probability irrespective of the series of outcomes that occurred. Data were analyzed in the same way as the mean probabilities but additional variable TRIAL was introduced into the model. Post-reversal trials were averaged into groups of 3 (i.e. 1-3, 4-6, 7-9, 10-12). Importantly, only participants that were actively learning were included in this analysis. Only phases were the mean probability on five trials preceding reversal was below 50% for acquisition or above 50% for extinction were included. A participant was included only if at least one acquisition and/or one extinction met this criterion, otherwise the participant was excluded from that given visit. This meant that a given participant could be included on visit 1 but not in visit 2. This pruning resulted in a

subset of ‘active’ learners (n=30 on visit 1, n=33 on visit 2 and n=29 on visit 3).

Model Parameters

Parameters estimated by the HBM model included the learning rate, perceptual distortion and updating distortion. All parameters were estimated as part of the same model. Statistical analysis was performed on per-parameter basis. Importantly, Bayesian estimation provided information about uncertainty of each parameter estimate in each subject. To maximize on the information contained in the group analysis this lower-level uncertainty was incorporated into group-level statistics in the following manner: Instead of calculating the standard deviation using the usual formula it was calculated by pooling the variance of individual parameter estimates (i.e. single parameter Bayesian posteriors) using eq. IV

Eq. III $$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + ... + (n_k - 1)s_k^2}{n_1 + n_2 + ...n_k - k}}$$

where n is number of samples drawn from the posterior distribution; k is the number of individuals and s is the sample standard deviation.

Parameter estimates were analyzed by group (losartan/placebo; between) and visit (1-3; within) by performing either a within or between (as appropriate) subject t-test for all combinations of visit and group (15 comparisons in the HBM model, 66 in the HBM_sp). The resulting p-values were adjusted using Bonferroni correction. Multiple t-tests were used to again preserve the lower level uncertainty estimate which would be lost in a standard ANOVA.

Pupil Data

To analyze changes in pupil dilation we employed the same approach as

described by (Pulcu and Browning 2017a). Data for each event (cue onset/outcome) were analysed separately. Event-locked pupil size grand means were calculated across the group. To increase statistical power the post-event data were combined into 0.5 bins. For cue onset and outcome, a 1.5 s window was analysed. For the response event the window was set to 1 sec to avoid the inclusion of outcome-related signal. Binned data were analysed using a mixed ANOVA that included TIME as a repeated measure variable and GROUP as a between subject variable. For the outcome phase, event type (i.e., shock or no shock) was added as another within-subject factor.

Additionally, a general linear model approach was used to investigate the effect of expectation (indexed as shock probability rating on a time point in pupil dilation), outcome and prediction error on pupil dilation. Data from a sliding window of 50 ms was used as dependent variable in a linear regression analysis incorporating current expectation, outcome and unsigned prediction error as predictor variables (eq. III). The resulting beta values were subsequently averaged across participants and analysed by split into 0.5s bins. Two seconds post each event were included, resulting in a repeated measures variable of four bins.

Eq. IV $$X \sim \beta_0 + \beta_1 * P_i + \beta_2 * O_i + \beta_3 * |PE|$$

Computational Modelling

Parameters of Bayesian models were estimated by sampling the posterior probability of the model using the Markov Chain Monte Carlo method, as implemented in the JAGS (Just Another Gibbs Sampler) toolbox. In total, 80 000 samples were obtained and thinned down by a factor of 4, resulting in 20 000 posterior samples per model.

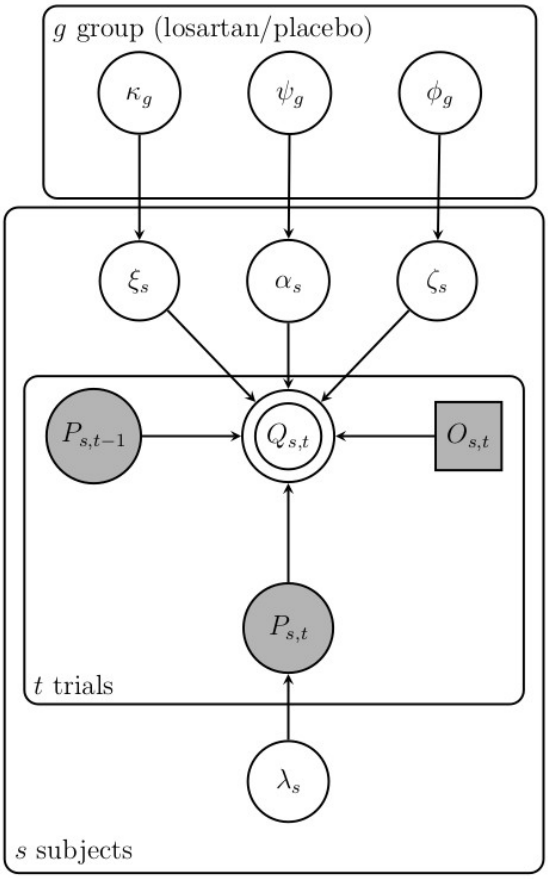
Hierarchical Bayesian Models (HBM)

To explicitly model the effect of learning, perceptual and updating bias I initially specified this model as a non-Bayesian reinforcement learning model with separate parameters for perception (ζ), learning (α) and updating (ξ).

Eq. V $$Q_{s,t} = P_{s,t-1} + \xi_s + \alpha_s(O_{s,t} + \zeta_s - P_{s,t-1})$$

where $\alpha \in [0,\ 1]$, $\xi \in [-1,\ 1]$ and $\zeta \in [-1,\ 1]$

However, the estimability of the ζ parameter of the model was low ($r = 0.32$, assessed by parameter recovery), the Bayesian approach was adopted to minimise the information loss inherent to the method: instead of using point values of parameter estimates, the uncertainty in estimation is incorporated into the model in a hierarchical manner. The HBM, In addition to modelling single-subject's learning rate (α), perceptual bias (ζ) and updating bias (ξ), estimates population hyperparameters separately for losartan and placebo groups and for all three parameters (ψ corresponding to learning rate hyperparameter, ϕ to perceptual bias hyperparameter and κ to updating bias hyperparameter). Comparing the distribution of hyperparameters allows for a direct comparison of the influence of losartan versus placebo. The HBM attempts to identify any group tendencies separately for the losartan and placebo groups. For example, $\alpha_{s=1}$ represents the learning rate for participant 1, and because this participant belongs to the losartan group α_s is assumed to come from a group level Normal distribution with mean $\psi_{losartan}$ and standard deviation λ_α. Therefore, the model has 6 hyper-parameters (3 parameters x 2 groups), with additional 3 parameters ($\lambda_\alpha, \lambda_\xi, \lambda_\zeta$) controlling the width of hyperparameter distribution, a set of subject-specific parameters (3 for each subject x 40 subjects) and an additional set of value uncertainty parameters λ_s (1 per subject, 40 in total). Figure 1 shows the full graphical model.



$$
\begin{aligned}
P_{s,t} &\sim \text{Normal}\big(Q_{s,t}, \lambda_s\big) \\
Q_{s,t} &= P_{s,t-1} + \xi_s + \alpha_s\big(O_{s,t} + \zeta_s - P_{s,t}\big) \\
\psi_g &\sim \text{Beta}\big(1, 1\big) \\
\phi_s &\sim \text{Uniform}\big(-1, 1\big) \\
\kappa_s &\sim \text{Uniform}\big(-1, 1\big) \\
\lambda_s, \lambda_\zeta, \lambda_\alpha, \lambda_\xi &\sim \text{Gamma}\big(0.001, 0.001\big)
\end{aligned}
$$

$$
\alpha_s \sim \begin{cases} \mathcal{N}(\psi_{losartan}, \lambda_\alpha) & \text{if } group = 1 \\ \mathcal{N}(\psi_{placebo}, \lambda_\alpha) & \text{if } group = 2 \end{cases} \tag{1}
$$

$$
\zeta_s \sim \begin{cases} \mathcal{N}(\phi_{losartan}, \lambda_\zeta) & \text{if } group = 1 \\ \mathcal{N}(\phi_{placebo}, \lambda_\zeta) & \text{if } group = 2 \end{cases} \tag{2}
$$

$$
\xi_s \sim \begin{cases} \mathcal{N}(\kappa_{losartan}, \lambda_\xi) & \text{if } group = 1 \\ \mathcal{N}(\kappa_{placebo}, \lambda_\xi) & \text{if } group = 2 \end{cases} \tag{3}
$$

Figure 1: *Graphical representation of the Bayesian Hierarchical Model*

To further account for the influence of starting probability on each visit and to test whether including this variable improves model evidence I extended the set of hyperparameters for each factor (perception/learning/updating) from 2 (placebo/losartan) to 4 (placebo-sp25/placebo-sp75/losartan-sp25/losartan-sp75). This model is referred to as HBM_sp.

Model Comparison

Two Bayesian models were compared using the Divergence Information Criterion (DIC) which was estimated as part of the MCMC sampling procedure. Lower DIC value represents more evidence for the given model. The main purpose of the model comparison was to identify on which visit the starting probability plays a role.

Results

Group Matching and Drug Side Effects

The two groups were well-matched on sociodemographic and clinical parameters (Table 1). There were no group differences in heart rate, blood pressure and VAS rating changes from baseline to drug peak level, all $F(1, 38) < 2.21$, all $p > .15$ (Table 2). Furthermore, neither the participants nor the experimenter were able to indicate above chance whether the participant had been allocated to the drug or the placebo group (experimenter: 23% correct, patients: 23%; both $\chi^2 < 1.52$, both $p > .31$).

Table 1: *Sociodemographic, clinical and personality characteristics and experimental stimuli ratings of participants in the losartan versus placebo group (M, SD, and t-test/ X²-test p-scores).*

	Losartan (N=21)		*Placebo (N=19)*		
	M	*SD*	*M*	*SD*	*p*
Sociodemographic Data					
Gender female	33%		52%		.34
First language English	71%		83%		.46
Age	26.1	5.3	25.0	4.3	.49
Verbal intelligence (NART)	114.5	6.7	110.4	9.9	.16
Years of education	17.0	2.3	16.6	4.5	.68
Clinical and Personality Measures					
Trait Anxiety (STAIT)	35.4	8.3	36.8	7.5	.57
Anxiety Sensitivity (ASI)	16.3	10.7	13.9	10.3	.49
behavioural Activation (BAS)	24.0	4.9	25.1	5.1	.49
Behavioral Inhibition (BIS)	16.5	3.0	14.8	4.0	.14
Beck Depression Inventory (BDI)	3.6	5.7	5.2	6.5	.40
Neuroticism (EPQ)	7.0	5.2	7.6	6.2	.73
Attentional Control (ACS)					
Total	58.1	7.5	56.6	9.7	.57
Focusing	25.8	5.2	25.2	4.5	.70
Shifting	32.4	4.1	31.4	6.0	.56

Note: NART = National Adult Reading Test; STAIT = State-Trait Anxiety Inventory; ASI = Anxiety Sensitivity Index; BAS = Behavioral Activation Scale; BIS = Behavioral Inhibition Scale; BDI = Beck Depression Inventory; EPQ = Eysenck Personality Inventory; ACS = Attentional Control Scale.

Table 2 *Heart rate, blood pressure and visual analogue scale ratings in the two groups before drug intake and at drug peak-level.*

	Baseline				Drug Peak			
	Losartan		Placebo		Losartan		Placebo	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Physiological Measures								
Heart rate	75	12	74	10	67	8	67	8
Systolic blood pressure	124	15	124	15	118	15	119	14
Diastolic blood pressure	72	9	74	10	70	9	74	11
Visual Analogue Ratings								
Anxious	8	8	12	12	4	4	8	11
Tearful	2	3	5	8	2	2	4	6
Hopeless	4	9	6	11	3	5	4	8
Sad	4	6	7	9	4	7	4	5
Depressed	2	3	5	8	2	4	5	7
Sleepy	19	15	19	17	20	20	21	18
Nauseous	2	3	6	12	2	4	5	8
Dizzy	5	11	6	6	6	12	6	11
Heart racing	2	3	9	12	3	3	5	6
Alert	46	31	50	27	44	33	45	30

Fear Conditioning Task

Shock probability ratings

The mean shock probability ratings of the participants are summarized in Table 3. Tests in this section used mean subjective probability per condition as the dependent variable. Tests were performed in a per-cue manner by incorporating group, phase and visit.

Stable Cues

Testing for an effect of phase, visit and group there was no main effect or interaction in the harmful cue. In the safe cue condition, both, main effect of group, $F(1, 37) = 7.40, p = 0.010)$, and visit, $F(2, 74) = 5.89, p = 0.004$, were significant. Post-hoc tests found a significantly higher mean probability across groups in visit 1, ($\bar{x} = 0.32, \sigma = 0.14$) than in visit 2 ($\bar{x} = 0.24, \sigma = 0.14$),

$t(39) = 3.044, p = 0.004$, and in visit 3 ($\bar{x} = 0.30, \sigma = 0.15$) than in visit 2, $t(39) = -2.670, p = 0.0110$. A post-hoc analysis of the significant main effect of group found a significantly higher probability in the placebo group ($\bar{x} = 0.35, \sigma = 0.11$) than in the losartan group ($\bar{x} = 0.26, \sigma = 0.09$), $t(38) = -2.72, p = 0.010$.

Table 3 *Means and standard deviations for shock probability ratings by visit, cue and phase.*

Phase		Acquisition				Extinction			
Group		Losartan (n=21)		Placebo (n=19)		Losartan (n=21)		Placebo (n=19)	
Harmful Cue		*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
	Visit 1	0.78	0.16	0.70	0.13	0.80	0.16	0.73	0.16
	Visit 2	0.79	0.18	0.77	0.18	0.79	0.19	0.77	0.16
	Visit 3	0.73	0.16	0.73	0.18	0.74	0.16	0.74	0.15
Safe Cue									
	Visit 1	0.24	0.13	0.39	0.14	0.24	0.18	0.40	0.16
	Visit 2	0.21	0.12	0.27	0.16	0.21	0.13	0.27	0.16
	Visit 3	0.29	0.13	0.35	0.19	0.25	0.11	0.33	0.19
Reversal Cue									
	Visit 1	0.64	0.15	0.68	0.16	0.53	0.15	0.54	0.19
	Visit 2	0.65	0.10	0.70	0.13	0.55	0.13	0.51	0.16
	Visit 3	0.62	0.19	0.66	0.13	0.51	0.17	0.51	0.19

Table 4 *Mixed ANOVA of average shock probability ratings for group, phase and visit, separate for each cue.*

Stable-high Cue	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.023	1	0.023	0.277	0.602
Error	3.126	37	0.084		
Visit	0.094	2	0.047	1.444	0.243
Group * Visit	0.069	2	0.034	1.054	0.354
Error	2.407	74	0.033		
Phase	0.007	1	0.007	1.970	0.169

Group * Phase	0.001	1	0.001	0.215	0.646
Error	0.126	37	0.003		
Visit * Phase	0.003	2	0.002	0.630	0.535
Group * Visit * Phase	0.001	2	0.000	0.187	0.830
Error	0.181	74	0.002		
Stable-low Cue	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.467	1	0.467	6.531	**0.010**
Error	2.646	37	0.072		
Visit	0.252	2	0.126	4.508	**0.004**
Group * Visit	0.106	2	0.053	1.900	0.157
Error	2.072	74	0.028		
Phase	0.003	1	0.003	0.572	0.454
Group * Phase	0.001	1	0.001	0.243	0.625
Error	0.178	37	0.005		
Visit * Phase	0.013	2	0.007	2.484	0.090
Group * Visit * Phase	0.002	2	0.001	0.314	0.731
Error	0.198	74	0.003		
Reversal Cue	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.019	1	0.019	0.278	0.601
Error	2.559	37	0.069		
Visit	0.035	2	0.018	0.808	0.450
Group * Visit	0.002	2	0.001	0.054	0.947
Error	1.606	74	0.022		
Phase	1.098	1	1.098	45.519	**0.000**
Group * Phase	0.038	1	0.038	1.556	0.220
Error	0.892	37	0.024		
Visit * Phase	0.005	2	0.002	0.443	0.644
Group * Visit * Phase	0.008	2	0.004	0.738	0.481
Error	0.391	74	0.005		

Reversal Cue

In the reversal cue there was a significant main effect of phase, $F(1, 37) = 45.519, p = 0.000$. Reported probabilities in acquisition ($\bar{x} = 0.66, \sigma = 0.11$) were significantly higher than in extinction ($\bar{x} = 0.52, \sigma = 0.13$). Adding anxiety as an additional between subject factor did not identify any significant relationship between mean probability, group, phase and anxiety.

Effects of Starting Probability

To control for the potentially confounding effect of the shock probability at the beginning of each visit, I specified a model including the starting probability (SP) of the current and preceding visits as well as phase and group for each visit. Only the reversal cue was included in this analysis. This resulted in a total of six models:

i. Probability on Visit 1 ~ Starting Probability on Visit 1
ii. Probability on Visit 2 ~ Starting Probability on Visit 1
iii. Probability on Visit 2 ~ Starting Probability on Visit 2
iv. Probability on Visit 3 ~ Starting Probability on Visit 1
v. Probability on Visit 3 ~ Starting Probability on Visit 2
vi. Probability on Visit 3 ~ Starting Probability on Visit 3

Below, I report any significant main effect of starting probability or interaction of SP with group.

Table 5 Probability on visit 1 by group, phase and SPv1

	SS	*d.f.*	*MS*	*F*	*p*
(Intercept)	26.896	1	26.896	690.287	0.000
Group	0.000	1	0.000	0.005	0.946
SPv1	0.144	1	0.144	3.704	0.062
Group:SPv1	0.020	1	0.020	0.513	0.478
Error	1.364	35	0.039		
(Intercept):Phase	0.326	1	0.326	29.983	**0.000**
Group:Phase	0.003	1	0.003	0.291	0.593
SPv1	0.003	1	0.003	0.240	0.627
Group:SPv1:Phase	0.001	1	0.001	0.051	0.823
Error(Phase)	0.380	35	0.011		

ANOVA with testing the effect of SPv1 on visit 1 found no main effect of interaction of SPv1, see Table 5.

Table 6 *Probability on visit 2 by group, phase and SPv1*

	SS	*d.f.*	*MS*	*F*	*p*
(Intercept)	28.048	1	28.048	1102.976	0.000
Group	0.001	1	0.001	0.024	0.878
SPv1	0.051	1	0.051	1.995	0.166
Group:SPv1	0.021	1	0.021	0.831	0.368
Error	0.915	36	0.025		
(Intercept):Phase	0.389	1	0.389	44.102	**0.000**
Group:Phase	0.047	1	0.047	5.305	**0.027**
SPv1:Phase	0.012	1	0.012	1.412	0.243
Group:SPv1:Phase	0.005	1	0.005	0.610	0.440
Error(Phase)	0.318	36	0.009		

ANOVA testing the effect of SPv1 on visit 2 found no main effect of interaction of SPv1, see Table 6.

Table 7 *Probability on visit 2 by group, phase and SPv2*

	SS	*d.f.*	*MS*	*F*	*p*
(Intercept)	22.733	1	22.733	1195.350	0.000
Group	0.042	1	0.042	2.213	0.146
SPv2	0.244	1	0.244	12.823	**0.001**
Group:SPv2	0.082	1	0.082	4.318	**0.045**
Error	0.685	36	0.019		
(Intercept):Phase	0.363	1	0.363	40.517	**0.000**
Group:Phase	0.015	1	0.015	1.639	0.209
SPv2:Phase	0.006	1	0.006	0.723	0.401
Group:SPv2:Phase	0.005	1	0.005	0.563	0.458
Error(Phase)	0.323	36	0.009		

ANOVA testing the effect of SPv2 on visit 2 found a main effect of SPv2 and an interaction between group and SPv2 (see Table X). The SPv2 main effect was driven by significantly higher probability in the placebo group ($\bar{x} = 0.70 \sigma = 0.13$) than in the losartan group ($\bar{x} = 0.58 \sigma = 0.09$). The interaction was driven by a significantly higher mean probability in Placebo-Sp75 ($\bar{x} = 0.77 \sigma = 0.12$) than in the Placebo-Sp25 group ($\bar{x} = 0.56 \sigma = 0.10$), $t(17) = -4.07, p = 0.001$. Other post-hoc tests found no difference.

Table 8: *Probability on visit 3 by group, phase and SPv1*

	SS	*d.f.*	*MS*	*F*	*p*
(Intercept)	24.938	1	24.938	544.326	0.000
Group	0.009	1	0.009	0.187	0.668
SPv1	0.000	1	0.000	0.001	0.978
Group:SPv1	0.005	1	0.005	0.110	0.742
Error	1.649	36	0.046		
(Intercept):Phase	0.295	1	0.295	21.199	**0.000**
Group:Phase	0.018	1	0.018	1.289	0.264
SPv1:Phase	0.064	1	0.064	4.595	**0.039**
Group:SPv1:Phase	0.003	1	0.003	0.238	0.628
Error(Phase)	0.501	36	0.014		

ANOVA testing the effect of SPv1 on visit 3 found a significant interaction between phase and SPv1 (see Table 8). A set of post-hoc tests found that this interaction was driven by the s.p.25 condition in which the mean probability in acquisition ($\bar{x} = 0.66 \sigma = 0.15$) was significantly higher than in extinction ($\bar{x} = 0.49 \sigma = 0.14$), $t(22) = 5.22, p = 0.000$.

Table 9: *Probability on visit 3 by group, phase and SPv2*

	SS	*d.f.*	*MS*	*F*	*p*
(Intercept)	20.404	1	20.404	512.592	0.000
Group	0.059	1	0.059	1.483	0.231
SPv2	0.154	1	0.154	3.878	0.057
Group:SPv2	0.087	1	0.087	2.180	0.148
Error	1.433	36	0.040		
(Intercept):Phase	0.213	1	0.213	14.162	**0.001**
Group:Phase	0.000	1	0.000	0.011	0.918
SPv2:Phase	0.001	1	0.001	0.094	0.761
Group:SPv2:Phase	0.028	1	0.028	1.835	0.184
Error(Phase)	0.541	36	0.015		

Table 10: *Probability on visit 3 by group, phase and SPv3*

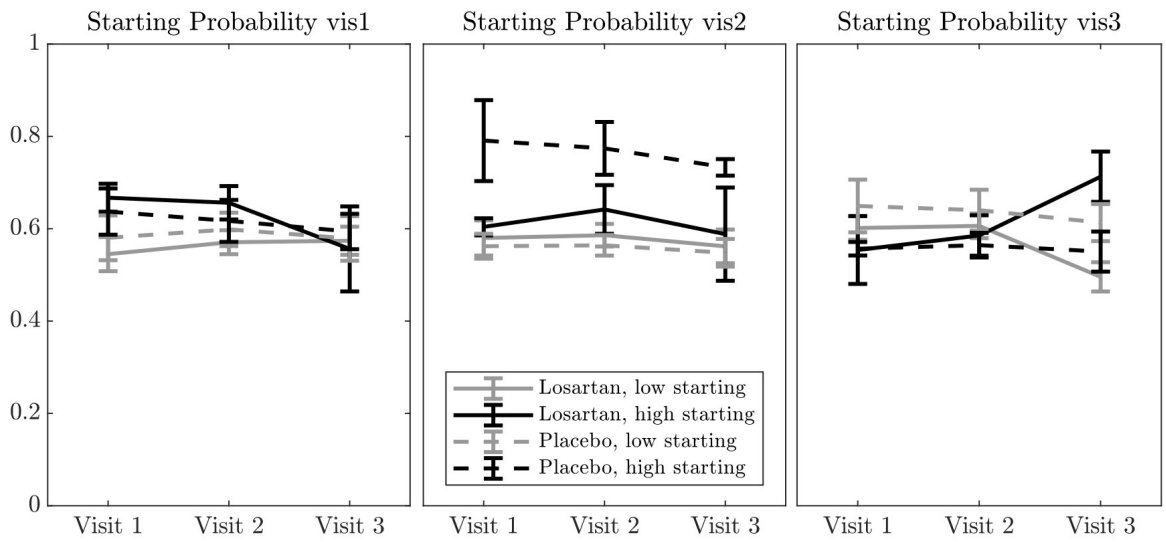
	SS	*d.f.*	*MS*	*F*	*p*
(Intercept)	26.162	1	26.162	799.008	0.000

Group	0.009	1	0.009	0.284	0.598
SPv3	0.110	1	0.110	3.375	0.074
Group:SPv3	0.364	1	0.364	11.106	**0.002**
Error	1.179	36	0.033		
(Intercept):Phase	0.399	1	0.399	32.534	**0.000**
Group:Phase	0.010	1	0.010	0.823	0.370
SPv3:Phase	0.106	1	0.106	8.649	**0.006**
Group:SPv3:Phase	0.022	1	0.022	1.780	0.191
Error(Phase)	0.441	36	0.012		

ANOVA testing the effect of SPv3 on visit 3 found a significant interaction between phase and SPv3 (see Table 10). Post-hoc test found significantly higher probability in acquisition ($\bar{x} = 0.74 \sigma = 0.12$) than in extinction ($\bar{x} = 0.51 \sigma = 0.21$) in sp75, $t(14) = 5.52, p=0.000$; a significantly higher probability in sp75 ($\bar{x} = 0.74 \sigma = 0.12$) than in sp25 ($\bar{x} = 0.58 \sigma = 0.16$) in acquisition, $t(38) = -3.24, p = 0.002$; and finally significantly higher probability in sp25 acquisition ($\bar{x} = 0.58 \sigma = 0.16$) than in sp25 extinction ($\bar{x} = 0.51 \sigma = 0.15$), $t(24) = 2.32, p = 0.029$.

Due to these significant findings, starting probability was included as a follow up in all subsequent analysed.

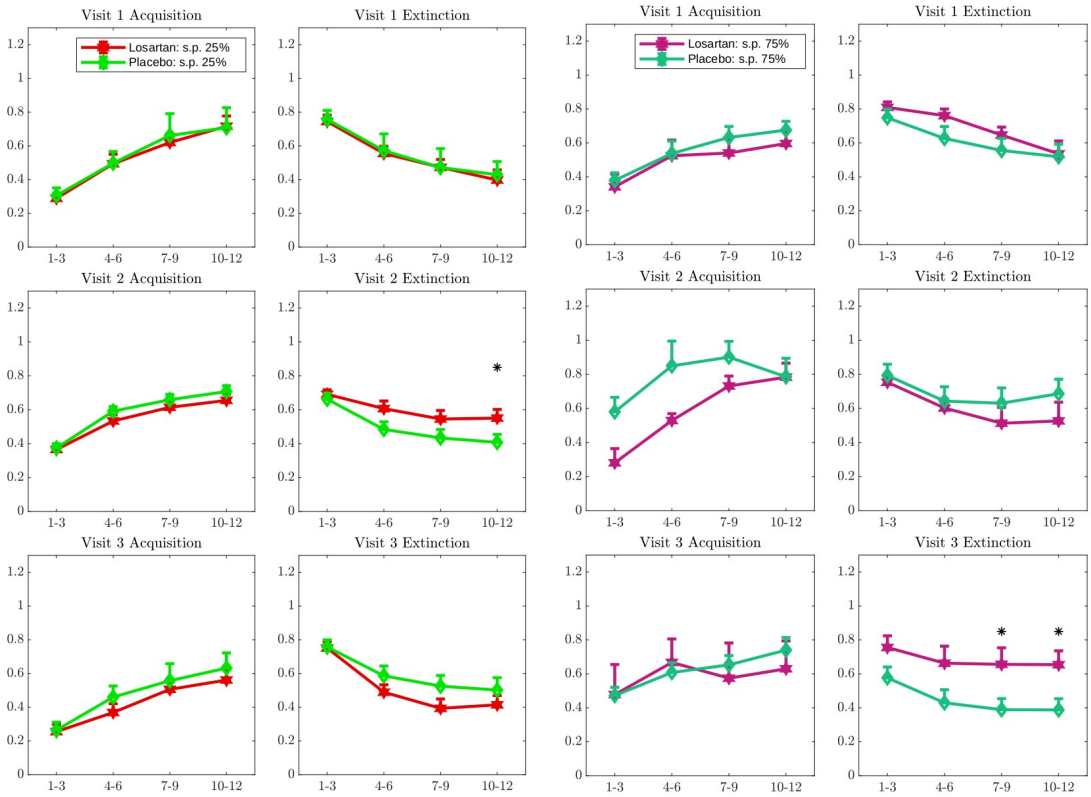
Figure 2 *Each plot represents shock probability ratings split by starting probability on a given visit (SPv1, SPv2, SPv3). While the significant interaction effect between SPv2 and group on visit 2 is likely due to pre-existing individual differences, the interaction between SPv3 and group on visit 3 can be an effect of the drug administration.*



Post-reversal learning

Next, I visualised how mean probabilities change post reversal. I split the data by low and high starting probability to clearly show its effect. The tests in this section were done by doing a t-test of each time bin between losartan and placebo. Where asterix is plotted a corrected p-value of $p<0.05$ was obtained.

Figure 3 *Post-reversal subjective probabilities presented separately for starting probability of 25% (two left columns) and 75% (two right columns). Highlighted comparisons (*) were statistically significant after correction.*



Model-free learning rates

All Cues

The following section analyzes model-free learning rates as the dependent variable. First, all three cues will be analysed together and a separate analysis of the main cue of interest (reversal) will follow. Effect of starting probability is only considered in the per-visit analysis as otherwise the ANOVA model fails to estimate the parameters (this would include 4 between subject factors on 40 data points).

Means and standard deviations for this section are presented in Table 11, results of a mixed ANOVA with the factors group, cue and visit are presented in Table 9. There was a significant main effect of cue, $F(2, 76) = 25.927, p = 0.000$, on model-free learning rates. Learning was fastest for the harmful cue ($\bar{x} = 0.353, \sigma = .017$), followed by the safe cue ($\bar{x} = 0.318, \sigma = 0.019$) and the reversal cue ($\bar{x} = 0.265, \sigma = 0.017$). Bonferroni-corrected post-hoc tests found all

three differences, harmful vs safe $t(39) = 2.650, p = 0.015$, safe vs reversal $t(39) = 6.444, p = 0.000$, reversal vs harmful $t(39) = 5.657, p = 0.000$, to be significant. Further, there was a significant interaction between visit and group $F(2, 76) = 6.514, p < 0.002$.

Table 11 Means and standard deviations for learning rates by visit and cue.

Group			Losartan (n=21)		Placebo (n=19)	
	Stable-high Cue		M	SD	M	SD
		Visit 1	0.39	0.11	0.34	0.15
		Visit 2	0.35	0.10	0.36	0.15
		Visit 3	0.31	0.10	0.36	0.16
	Stable-low Cue					
		Visit 1	0.33	0.12	0.31	0.13
		Visit 2	0.31	0.11	0.35	0.15
		Visit 3	0.29	0.11	0.31	0.16
	Reversal Cue					
		Visit 1	0.26	0.11	0.28	0.11
		Visit 2	0.22	0.10	0.30	0.13
		Visit 3	0.24	0.09	0.29	0.15

Table 12 Model-free learning rates by cue, group and visit

Between-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.038	1	0.038	0.397	0.532
Within-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Cue	0.460	2	0.230	25.927	**0.000**
Cue * Group	0.030	2	0.015	1.702	0.189
Error (Cue)	0.675	76	0.009		
Visit	0.023	2	0.012	2.076	0.132

Visit * Group	0.074	2	0.037	6.514	**0.002**
Error (Visit)	0.429	76	0.006		
Cue * Visit	0.014	4	0.003	1.068	0.375
Cue * Visit * Group	0.020	4	0.005	1.516	0.200
Error (Cue * Visit)	0.489	152	0.003		

Reversal Cue

Given the focus on learning, I next tested whether there is an effect of group in the reversal cue. In included starting probability in to the analyses (I ran one test for each starting probability).

Model with SPv1 found a significant group by visit interaction, $F(2, 72) = 4.45, p = 0.015$. Model with SPv2 found a significant main effects of SPv2, $F(1, 36) = 6.02, p = 0.019$, learning rates in sp75 ($\bar{x} = 0.34 \sigma = 0.13$) were significantly higher than in sp25 ($\bar{x} = 0.24 \sigma = 0.09$). Model with Furthermore, there was a significant interaction between group and visit, $F(2, 72) = 3.70, p = 0.029$. And model with SPv3 also found a significant group by visit interaction. To investigate the group by visit interaction that was significant in all three tests I conducted a post hoc test which found a significantly higher learning rates on visit 2 in the placebo ($\bar{x} = 0.30 \sigma = 0.13$) compared to the losartan group ($\bar{x} = 0.22 \sigma = 0.09$), $t(38) = -2.18, p = 0.035$, which did not survive the correction. Data for the reversal cue are summarized in Figure 4. Figure 4, lower left, presents mean learning rates by group across the three visits, marginalized over starting probability, Figure 4, lower right, presents the learning rates by starting probability marginalized over groups.

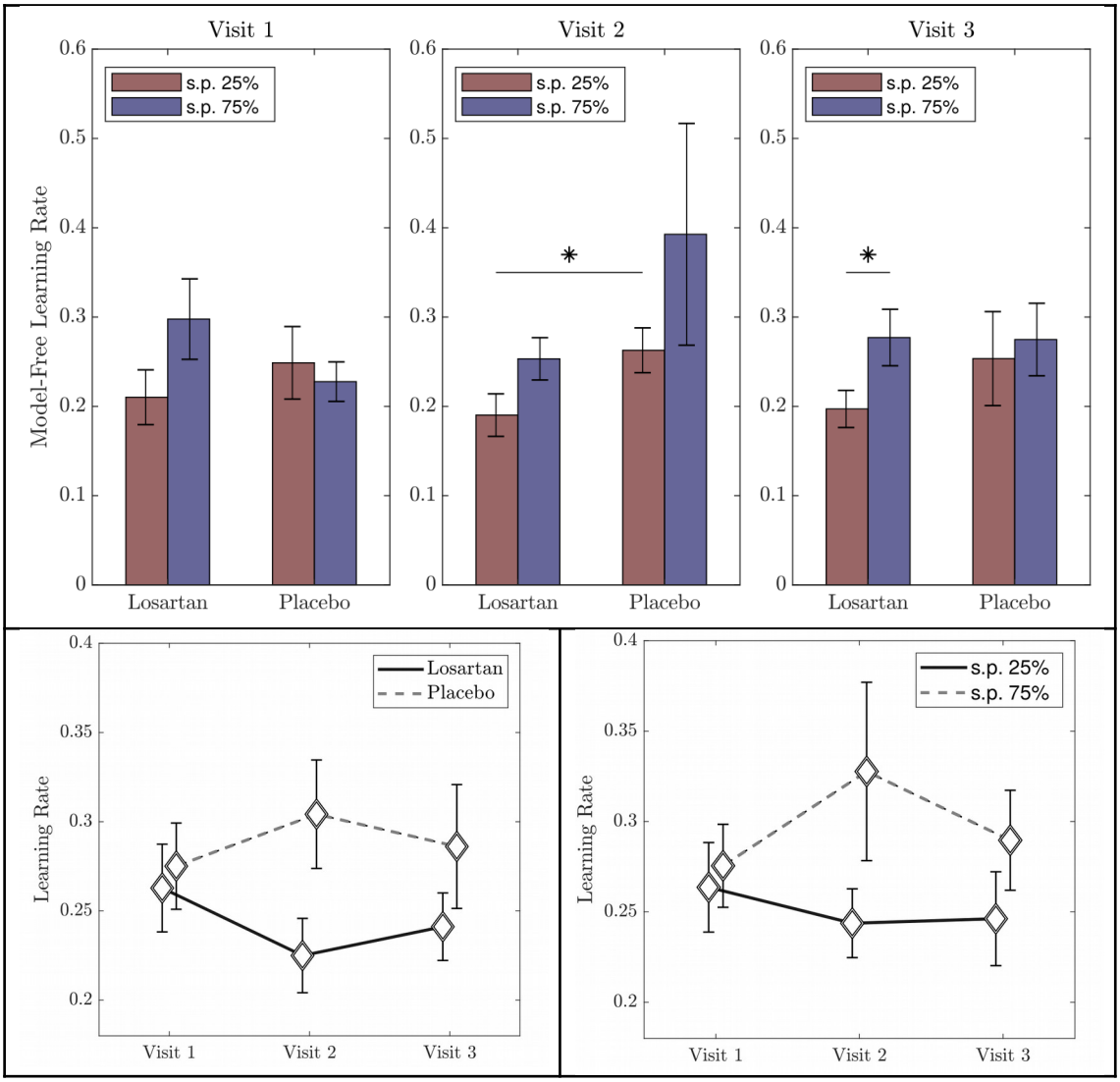


Figure 4 *Upper panel shows mean MFLRs by group, SP and visit. In the bottom panel, left plot shows MFLRs by group and visit, depicting a significantly higher learning rate on visit 2 in the placebo group; right plot shows the same data split by starting probability marginalized over groups, starting probability of 75% on visit 2 results in significantly higher learning rates than starting probability of 25%.*

Bayesian Models

To better understand how the reported ratings were generated via combination of learning, updating and perception, I next estimated the two Bayesian models

described in methods. The HBM considers the effect of group, while the HBM_sp additionally considered which of the sub-processed were impacted by starting probability. The models were subsequently compared to test whether the starting probability played a significant role on each visit.

The DIC score on visits 1 and 3 was lower for the HBM model and on visit 2 for the HBM_sp, suggesting that starting probability only mattered on visit 2 (see Table 13).

Table 13 Model Comparison results (DIC)

	HBM	HBM_sp
Visit 1	**-4626.397**	-4626.005
Visit 2	-11713.81	**-11716.23**
Visit 3	**-5277.117**	-5273.922

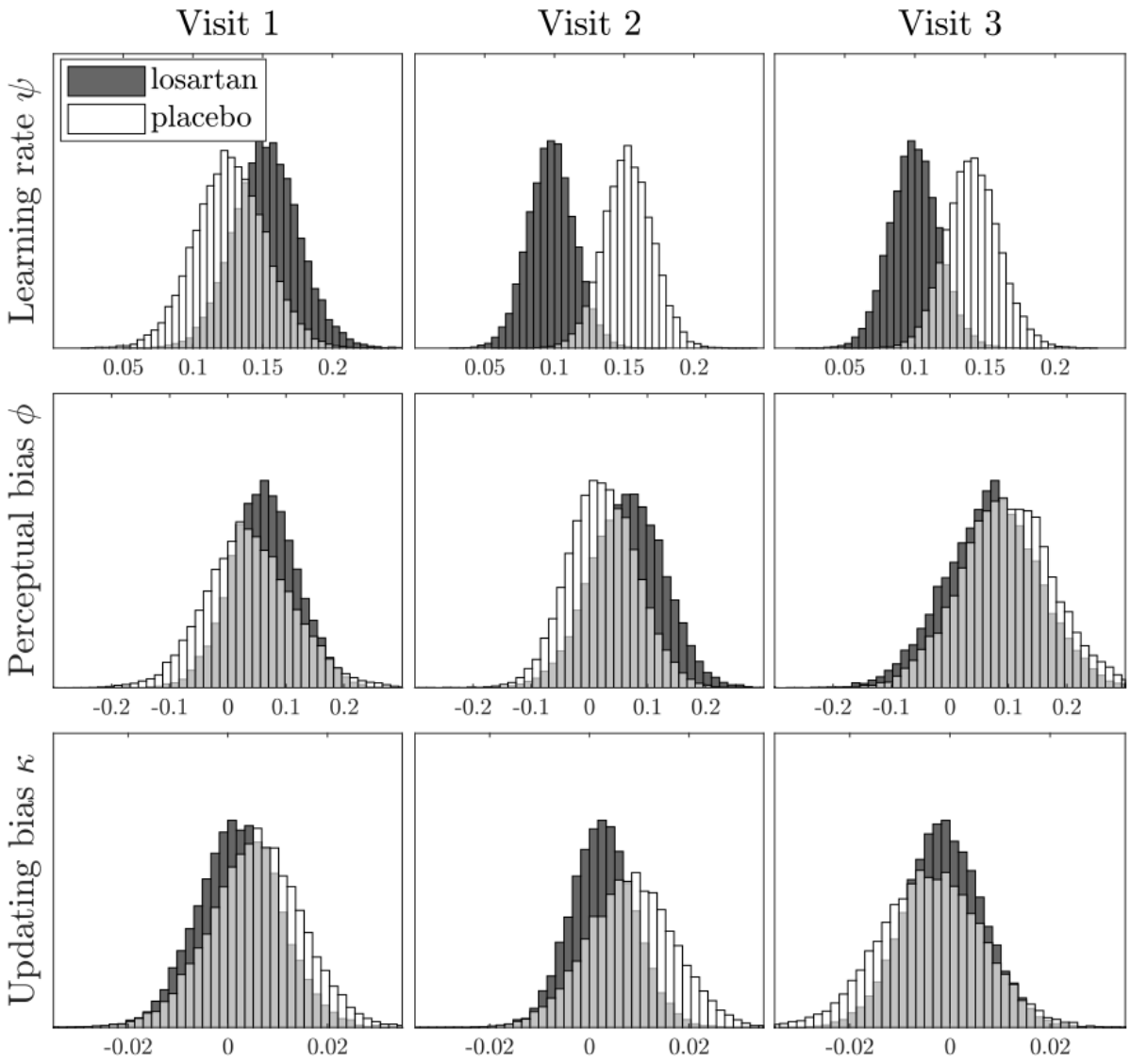


Figure 5 *Posterior distributions of hyperparameters ψ (learning rate), ϕ (perceptual bias) and κ (updating bias) for losartan and placebo groups.*

The posterior distributions of HBM hyperparameters are presented in Figure 5. The two tailed t-tests were all corrected for multiple comparison using Bonferroni correction resulting in $p_{crit} = 0.0017$ (0.05 / 15 comparison within each parameter x 2 tails of the test).

Learning rates

The losartan group had significantly higher learning rate on visit 1 than on visit

2, $t(40) = 8.25, p < .001$, and than on visit 3, $t(40) = 6.89, p < .001$. There was no significant difference between losartan and placebo on visit 1. The difference between groups was significant on both visit 2, $t(38) = -9.89, p < .001$ and visit 3, $t(38) = 4.91, p < .001$. Additionally, losartan group on visit 2 had slower learning rate than placebo on visit 3, $t(38) = -5.85, p < .001$, and placebo on visit 3 had significantly higher learning rate than losartan group on visit 3, $t(38) = 7.72, p < .001$.

To account for the effect of starting probability, I extended the above Bayesian model by including starting probability as a factor in the hyperparameter space. The results are presented in Figures 6 and 7. Figure 6 shows posterior distributions of the four hyperparameters by learning sub-component (updating, learning, perception) and visit.

Figure 6: *Estimated posterior distributions of the hyperparameters of the HBM_sp model. Each row represents a modelled candidate psychological subprocess: learning rate (ψ), perceptual bias (ϕ) and updating bias (κ). Data are presented separately for each combination of group and starting probability factors; losartan−sp25 (red), losartan−sp75 (purple), placebo-sp25 (light green), placebo-sp75 (dark green).*

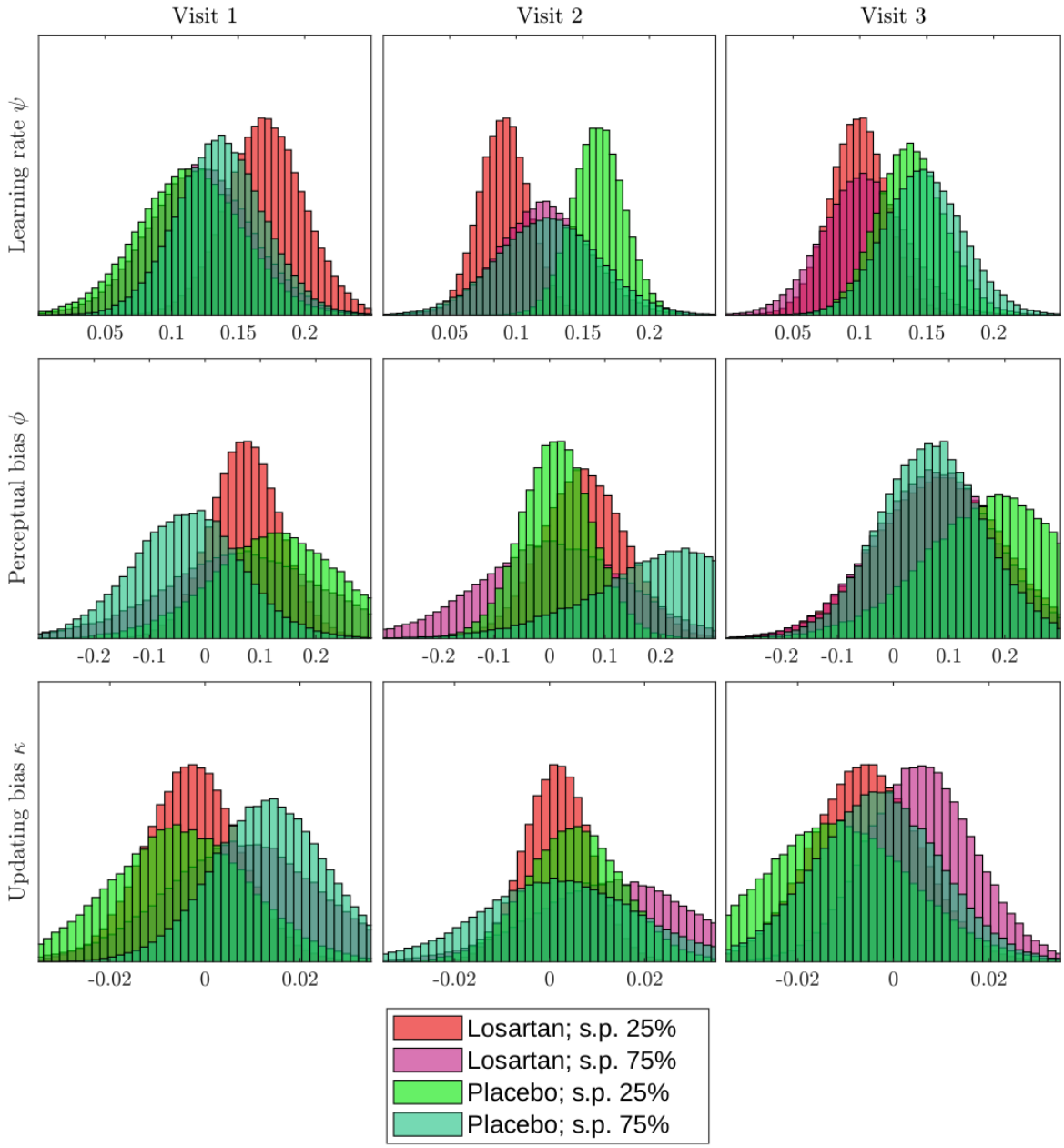
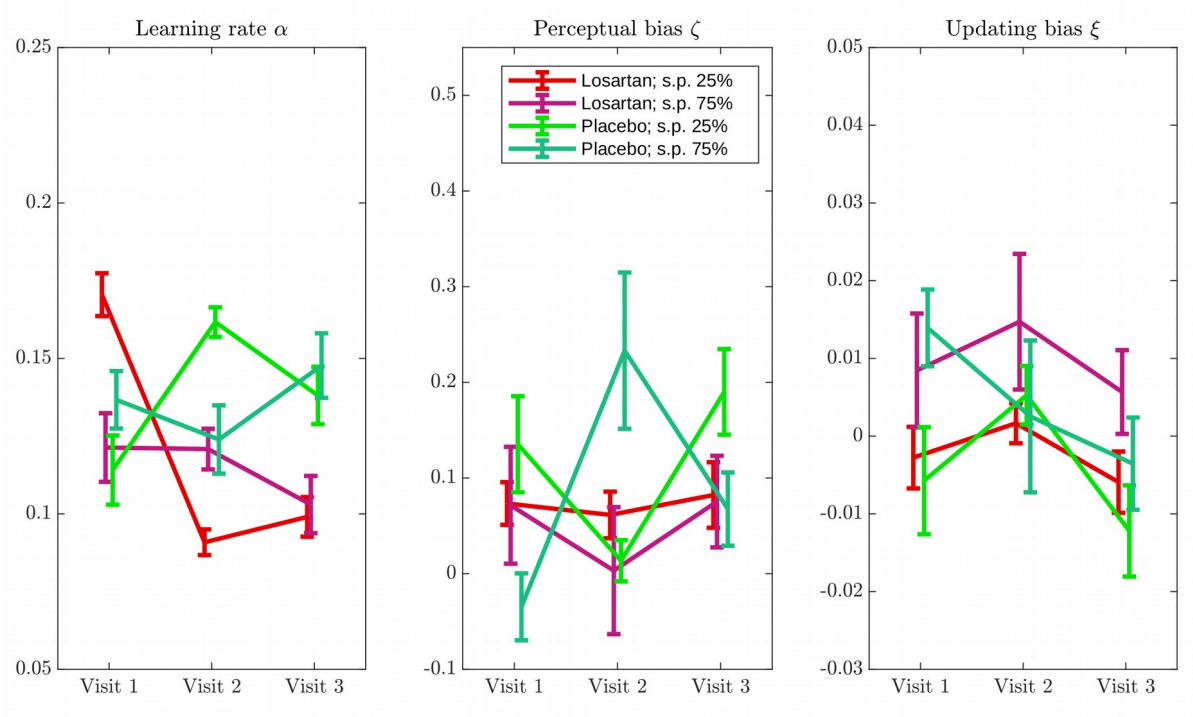


Figure 7 *HGM_sp parameter estimates with incorporated parameter uncertainty estimates (below). The error bars represent standard error of the mean that incorporates uncertainty in single parameter estimates but does not reflect correction for multiple comparisons.*



As in the previous analysis, each parameter was analyzed separately. The critical p-value for the following section was determined by using the Bonferroni correction $p_{corrected} = 0.000378$ (0.05 / number of comparisons / tails of each test = 0.05 / 66 / 2). The p-value in the following section reflects this correction.

In the perceptual and updating bias' there was no significant difference in any of the 66 comparisons. In learning rate, there was a significant decrease in learning in the losartan group between visit 1 ($\bar{x} = 0.17, \sigma = 0.024$) and both visit 2 ($\bar{x} = 0.09, \sigma = 0.017$), $t(26) = 10.45, p < .001$, and visit 3 ($\bar{x} = 0.10, \sigma = 0.024$), $t(24) = 7.61, p < .001$ in the $\alpha_{losartan,25}$ condition. In the placebo group, there was a significant learning rate increase in $\alpha_{placebo,25}$ from visit 1 ($\bar{x} = 0.12, \sigma = 0.029$) to visit 2 ($\bar{x} = 0.16, \sigma = 0.018$), $t(20) = -4.66, p < .05$. Furthermore, there were

several between group differences on individual visits. While there was no group difference in either starting probability, on visit 2 there was a significant difference between losartan ($\bar{x} = 0.09, \sigma = 0.017$) and placebo ($\bar{x} = 0.16, \sigma = 0.018$) in the 25% starting probability, $t(29) = -11.29, p < .001$. Full set of significant results is presented in Table 14.

Table 14 Significant differences in binary comparisons of the learning rate across all conditions

Comparison	*t*	*df*	$p_{corrected}$
v1 los25 > v2 los25	10.448	26	p<.001
v1 los25 > v3 los25	7.619	24	p<.001
v1 los25 > v3 los75	5.902	17	p<.01
v1 plac25 > v2 plac25	-4.664	20	p<.05
v1 plac75 > v2 los25	5.132	24	p<.01
v2 los25 > v2 plac25	-11.291	29	p<.001
v2 los25 > v3 plac25	-5.193	25	p<.01
v2 los25 > v3 plac75	-6.116	22	p<.001
v2 plac25 > v3 los25	7.958	27	p<.001
v2 plac25 > v3 los75	6.292	20	p<.001

Pupillometry

GLM results

Pupil data were collected only on visit 2. Raw grand averages are presented in Appendix II. The regression analysis investigated pupil encoding of expectation, outcome and unsigned prediction error. Data were binned (0.5 seconds per bin) and analyzed using a mixed ANOVA with the factors of time (binned into 500 ms post outcome), outcome and group. There was no significant effect of expectancy on pupil dilation at any time point during the trial. As expected, there was a significant main effect of outcome type (shock / shock omission), $F(3, 114) = 26.917, p = 0.000$, at the time of outcome delivery. Furthermore, we found a main effect of group on

prediction error processing, $F(3, 114) = 11.084, p = 0.000$. Prediction error encoding in time bins following outcome delivery significantly interacted with group, $F(3, 114) = 5.0385, p = .003$. Post-hoc tests found the interaction to be driven by a significantly higher prediction error beta in the losartan group 500 to 1000 ms post outcome, $t(38) = 3.159, p = 0.0032$ and significantly higher prediction error beta in the placebo group 1500 to 2000 ms post outcome, $t(38) = -2.658, p = 0.0115$.

Table 14. Mixed ANOVA results for the encoding of expectancy, outcome and prediction error by the pupil

Expectancy encoding at the time of cue onset					
Between-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.210	1	0.210	1.644	0.208
Error (Group)	4.845	38	0.127		
Within-Subject Effects	SS	d.f.	MS	F	p
Time	0.085	3	0.028	0.170	0.916
Time * Group	1.138	3	0.379	2.274	0.084
Error (Time)	19.014	114	0.167		
Outcome encoding at the time of outcome					
Between-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.108	1	0.108	4.243	**0.046**
Error (Group)	0.969	38	0.026		
Within-Subject Effects	SS	d.f.	MS	F	p
Time	5.189	3	1.730	26.917	**0.000**
Time * Group	0.325	3	0.108	1.688	0.174
Error (Time)	7.325	114	0.064		
Prediction error encoding at the time of outcome					
Between-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.094	1	0.094	2.698	0.109
Error (Group)	1.327	38	0.035		
Within-Subject Effects	SS	d.f.	MS	F	p
Time	2.180	3	0.727	11.084	**0.000**
Time * Group	0.991	3	0.330	5.039	**0.003**
Error (Time)	7.474	114	0.066		

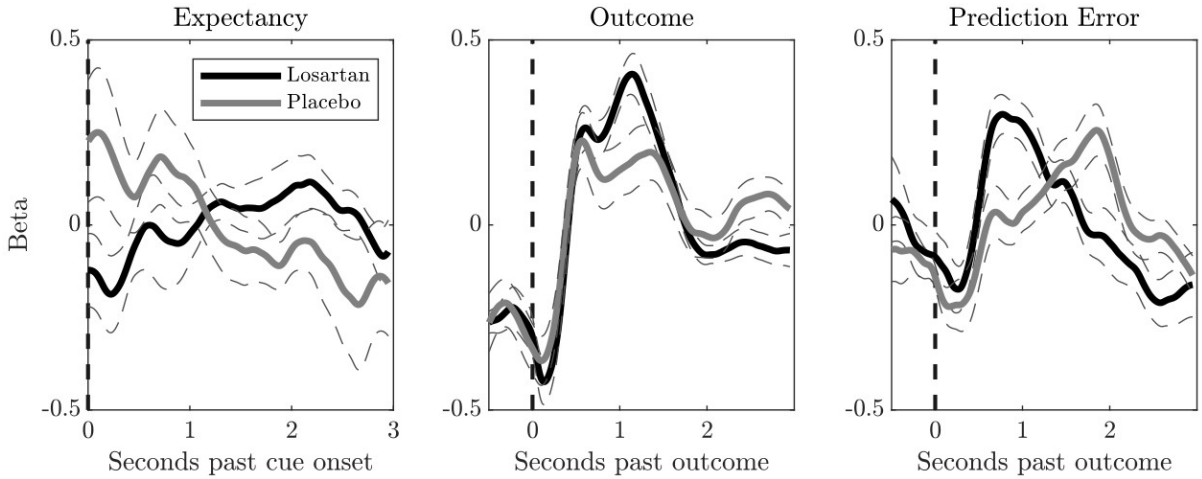


Figure 8 *Pupil encoding of shock expectancy at the time of cue onset (left), outcome at the time of outcome (middle) and prediction error at the time of outcome (right) separately for placebo (grey) and losartan (black) groups.*

To check for the effect of staring probability on pupil dilation I repeated the previous analysis with SPv2 as a between-subject factor. Full statistical results are presented in Table 15.

Table 15 Mixed ANOVA results for the encoding of expectancy, outcome and prediction error by the pupil by group and SPv2

Expectancy encoding at the time of cue onset					
Between-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.144	1	0.144	1.07	0.31
SPv2	0.000	1	0.001	0.00	0.97
Group * SPv2	0.000	1	0.000	0.000	0.99
Error (Group)	4.844	36	0.134		
Within-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Time	0.069	3	0.023	0.148	0.702
Time * Group	1.428	3	0.476	3.057	**0.031**
Time * SPv2	0.512	3	0.171	1.100	0.305
Time * Group * SPv2	1.576	3	0.525	3.373	**0.021**

Error (Time)	16.820	108	0.155		
Outcome encoding at the time of outcome					
Between-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.003	1	0.003	0.117	0.730
SPv2	0.000	1	0.000	0.037	0.842
Group * SPv2	0.161	1	0.161	7.198	**0.010**
Error (Group)	1.677	36	0.046		
Within-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Time	5.377	3	0.192	30.299	**0.000**
Time * Group	0.033	3	0.011	0.185	0.901
Time * SPv2	0.597	3	0.199	3.366	0.352
Time * Group * SPv2	0.388	3	0.129	2.191	0.091
Error (Time)	6.389	108	0.059		
Prediction error encoding at the time of outcome					
Between-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Group	0.032	1	0.032	0.910	0.353
SPv2	0.029	1	0.029	0.818	0.365
Group * SPv2	0.021	1	0.021	0.608	0.442
Error (Group)	1.271	36	0.035		
Within-Subject Effects	*SS*	*d.f.*	*MS*	*F*	*p*
Time	2.165	3	0.722	10.944	**0.000**
Time * Group	0.809	3	0.269	4.091	**0.008**
Time * SPv2	0.304	3	0.101	1.536	0.212
Time * Group * SPv2	0.054	3	0.019	0.274	0.843
Error (Time)	7.125	108	0.066		

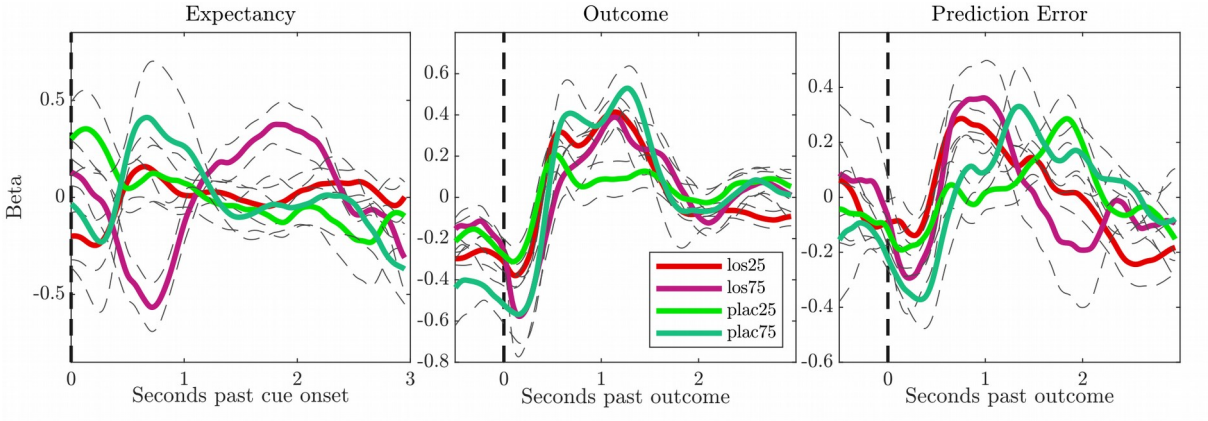


Figure 9 *Pupil encoding of shock expectancy at the time of cue onset (left), outcome at the time of outcome (middle) and prediction error at the time of outcome (right) separately for placebo-sp25 (green), placebo-sp75 (dark green), losartan-sp25 (red) and losartan-sp75 (purple) groups.*

Expectancy Encoding at the Time of Cue Onset

After accounting for the effect of starting probability there was a significant interaction between time and group, $F(3, 108) = 3.057, p = 0.031$. There was higher expectancy encoding in the 0.5 - 1s in the placebo ($\bar{x} = 0.16, \sigma = 0.55$), compared to the losartan group ($\bar{x} = -0.13, \sigma = 0.36$), $t(38) = -2.02, p = 0.025$ but this tests did not survive Bonferroni correction. Additionally, the mixed ANOVA revealed an interaction of group, SPv2 and time, $F(3, 108) = 3.373, p = 0.021$. Post-hoc tests found this effect to be driven by the difference in expectancy encoding between losartan-sp25 ($\bar{x} = -0.16, \sigma = 0.39$) and losartan-sp75 ($\bar{x} = 0.23, \sigma = 0.55$), t(29) = -2.26, p=0.031 but again this did not survive the correction.

Outcome Encoding at the Time of Outcome

At the time of outcome, there was a significant main effect of time on outcome encoding, $F(3.108) = 30.299, p = 0.000$ which was characterised by significant

differences between all time bins except between time bins 2 (0.5 - 1 s) and 3 (1 - 1.5). The tests are summarized in Table 16.

Table 16: Significant differences in binary comparisons between all time bins, outcome encoding at the time of outcome

Comparison	*t*	*df*	$p_{corrected}$
Bin1 > Bin2	-8.25	39	**p<.01****
Bin1 > Bin3	-6.39	39	**p<.01****
Bin1 > Bin4	-3.99	39	**p<.01****
Bin2 > Bin3	-0.85	39	n.s.
Bin2 > Bin4	3.11	39	**p<.05***
Bin3 > Bin4	5.20	39	**p<.01****

Furthermore, there was a significant interaction between group and and SPv2, $F(1, 36) = 7.19, p = 0.010$. Post hoc tests found a significantly higher outcome encoding in the losartan-sp25 ($\bar{x} = 0.11, \sigma = 0.05$) than in losartan-sp75 ($\bar{x} = 0.03, \sigma = 0.09$), $t(29) = 3.14, p = 0.004$, and significantly higher outcome encoding in the losartan-sp25 ($\bar{x} = 0.11, \sigma = 0.05$) than placebo-sp25 ($\bar{x} = 0.04, \sigma = 0.04$) groups, $t(19) = 2.67, p = 0.015$ which however did not survive the multiple comparison correction.

Lastly, there was a significant interaction between SPv2 and time, $F(3, 108) = 3.366, p = 0.031$ driven by significantly higher initial (0 to 0.5s post outcome) outcome encoding in the group with starting probability of 25% ($\bar{x} = -0.15, \sigma = 0.21$) compared to 75% ($\bar{x} = -0.39, \sigma = 0.27$), $t(38) = 2.85, p = 0.007$ which survived correction.

Prediction Error Encoding at the Time of Outcome

There was a significant main effect of time, $F(3, 108) = 10.94, p = 0.000$ driven by significantly lower signal in the initial 0.5s in comparison to the remaining time bins. Pairwise comparisons are presented in Table 17.

Additionally, there was a significant interaction between group and time, $F(3, 108) = 4.09, p = 0.009$ driven by a significantly higher prediction error encoding 0.5 to 1s after outcome in the losartan group ($\bar{x} = 0.25, \sigma = 0.24$) compared to the placebo group ($\bar{x} = 0.00, \sigma = 0.24$), $t(39) = 3.25, p = 0.002$ which survived correction.

Table 17: Significant differences in binary comparisons between all time bins, prediction error encoding at the time of outcome

Comparison	*t*	*df*	$p_{corrected}$
Bin1 > Bin2	-5.61	39	**p<.01****
Bin1 > Bin3	-4.53	39	**p<.01****
Bin1 > Bin4	-3.42	39	**p<.05***
Bin2 > Bin3	0.06	39	n.s.
Bin2 > Bin4	0.48	39	n.s.
Bin3 > Bin4	0.67	39	n.s.

Discussion

In this chapter I investigated the role of the Angiotensin II inhibitor losartan on human aversive learning. While the raw behavioural data don't show any strong effect of losartan on learning in contrast to placebo, and they additionally reveal a confound of initial reinforcement rate, computational modelling and pupillometry suggest an important role of losartan in aversive learning. While accounting for the effects of perceptual and updating distortions, which themselves were not modulated by either losartan or starting probability (SP), the Bayesian models comparison revealed that accounting for SP is only meaningful on the second visit. While on visits 1 and 3 there was more model evidence in favour of the HBM model, that only includes the effect of group, on visit 2 the DIC score was lower for the extended HBM_sp model, suggesting

a significant role of the SP factor on this visit. Consequently, within-visit comparisons are inferred from the winning model on a given visit, between-visits results are discussed in the context of both models.

The HBM model results show that losartan decreases overall learning rate from visit 1 to visit 2 and that this slower remains one day after on visit 3. A change in learning is not present in the placebo group where the difference between visit 1 and other visits is not significant. There is also no group difference on visit 1 while it exists on visits 2 and 3, losartan group having significantly lower learning rates than placebo. Incorporating SP into the analysis, the decrease in learning rates between visits 1 and 2 appears to be specific to low (25%) SP group. Visit 3, in contrast, doesn't have significantly lower learning rate in low SP, however, the model comparison has shown that SP doesn't play a role in the third visit. Our findings seem to agree with the result of (Pulcu et al. 2019), who also found losartan to decrease aversive learning rates.

The analysis of information encoding in the pupil dilation data revealed a significant between-group difference in prediction error processing that was not modulated by starting probability. This agrees with the modelling results, which also suggested no effect of losartan on perception and attributed the entire effect to learning. The pupil response in the losartan group seems to peak earlier compared to placebo, but it also returns to baseline more promptly. The effect of losartan on prediction error might reflect the interaction of Angiotensin II, reduced by losartan, and Norepinephrine in the locus coeruleus. (Elam et al. 1984) demonstrated how increase in Angiotensin II leads to decrease in norepinephrinergic cell firing in the LC. An alternative hypothesis would be the NE-AG interaction in the central nucleus of the amygdala which receives projections from LC and is abundant in Angiotensin II receptors. It is beyond the scope of this paper to answer these questions, however, the effect of losartan on pupil dilation suggests an interaction between Angiotensin and Norepinephrine.

Our analysis of the learning sub-processes considered perceptual and response distortions as important components of learning. The inferred effect of losartan was found to be specific to learning rates and there was no evidence of its effect on stimulus perception or expectation updating. This is conjointly supported by both Bayesian models and the pupil dilation results. The HBM found a difference between the groups' learning rates on all three visits. On visit 1, however, the effect occurred in the opposite direction than on visits 2 and 3. The to-be losartan group was learning slightly faster than the future placebo group. This was reversed by the administration of losartan on visit 2 and it remained the same on visit 3, suggesting a strong effect of losartan on learning rate reduction at least in the low SP condition.

One of the main goals of the study was to test the differential effect of losartan on acquisition and extinction. The results show an effect of losartan on learning at peak and one day after. Analysis of the behavioural data locked to the time of contingency reversal found differences between the group in late extinction. During trials 10-12 losartan participants with low SP (25%) predicted significantly higher chance of shock in second visit extinction. On the third visit, and only in high SP (75%), the losartan group showed higher anticipation of shock in trials 7 - 12. These results show that losartan appears to specifically slow down extinction, this effect, however, seems to be modulated by the initial context of the session. The analysis of model-free learning rates also revealed modulatory effects of SP. In low SP on visit 2, placebo group had significantly higher learning rate than the losartan group, however, there was no significant change in learning from visit 1 in either group. On visit 3, losartan, but not placebo, learning was again modulated by starting probability. The behavioural data suggest decreasing effect of losartan on learning, particularly in extinction. These results suggest that administration of losartan makes it hard to overcome fear anticipation likely due to decreased learning rates. A recent work in the theory of extinction, however, pointed out that gradual, rather than rapid, extinction leads to a

more permanent extinction effect and decreases the chance of relapse (Gershman et al 2013; Craske 2014). In my data set, this would mean slower rate of acquisition. While this is not observed, there is a marginal non-significant effect in the direction predicted by this hypothesis. To test explicitly this effect a follow-up study should be designed keeping in mind better balance of starting probability conditions.

All through the data set a strong effect of starting probability was observed. For example, the main result of learning rate decrease in the losartan group was only present in the low SP group. In mean probabilities the effect of starting probability directly interacted with the drug on visit 3: participants that received losartan the day before and started with a block with high SP reported on average significantly higher probabilities than those that started with low SP. The effect was not observed in the placebo group. The effect of SP was also observed in the pupil dilation analysis. In particular, on visit 2 outcome encoding was significantly different in the losartan group between low and high SP. This, in addition to decreased learning rates, might suggest that losartan makes people less flexible and more prone to initial anchoring effect due to differential encoding of outcome identity depending on the starting probability. In this study, the effect of SP was an undesired confound, however, the observed effect is worth subsequent investigation. Particularly, answering the question whether losartan enhances dependence on contextual inference would bear importance for clinical research as increased context-dependence has been associated with increased rates of relapse.

A question might arise why should one trust more a model estimate in contrast to ratings directly reported by the participants. It must be considered that ratings submitted by participants are a combination of numerous perceptual, cognitive and motoric processes collapsed into a single number. In fact, most studies in decision-making rely on less (i.e. binary choices). The purpose of the model is to estimate the most likely combination of these factors given the observed data. In other words, from the ratings we cannot directly tell to which degree was the model-free learning rate

influenced by reporting noise (i.e. the imprecision in slider placement) or how strong was the subjective experience of shock to the participant but the Bayesian model can provide us with the most likely scenario given the data. It is in this way that hidden cognitive variables can be studied.

While I report a specific role of losartan in learning, others have suggested that the effect of losartan is more general and due to its anti-hypertensive properties (Amenta et al., 2002; Gard 2004; Kułakowska et al. 1996). I observed no change in any of the recorded physiological and psychological measures, so my data don't support the latter view. It can however not be excluded that the role of losartan in cognition is not specific to learning. Additional experiments should be conducted to investigate its effect on other aspects of cognition such as attention or memory.

Furthermore, the employed model suffers from a number of drawbacks. Firstly, it only partially mitigates the issue with the zeta parameter estimability. In the MLE version of the same model the parameter failed to be estimated, this was one of the main reasons to employ the Bayesian framework. The Bayesian approach quantifies the uncertainty and propagates it along the hierarchical structure. While this helps to estimate the zeta, the values are still associated with relatively low certainty. Secondly, the hierarchical structure of the model makes the comparison of single subject estimates problematic. Since the hyperparameters are estimated per-group, subjects in the same group will be more similar to each other partially due to the top-down influence of common hyperdistribution. This issue can be circumvented by comparing the estimated hyperparameters directly. Thirdly, the updating parameter represents a constant bias in updates. This formulation is the most basic form of updating distortion which likely doesn't correspond to an underlying process. There are several alternative ways to model updating bias. One is to consider expected and unexpected uncertainty separately. Under this model, sudden rare events assigned to deviations from contingency (expected uncertainty) should not lead updating while rare events understood by the agent as environmental changes should lead to learning. Another

approach would be to model the bias in updating as a distortion that follows a particular distribution, here, individuals with high variability estimate of the distortion tend to suffer from higher updating bias.

Conclusions

In conclusion, I found losartan at peak level and one day after to significantly decrease aversive learning rate which resulted in decreased extinction learning. The effect of losartan on learning was supported by pupil dilation data which revealed altered encoding in prediction error processing at peak level. The peak level results were specific to context in which the session started with a block low in reinforcement rate but the effect of losartan on learning days one day after was independent of initial context.

Overall Discussion

In this doctoral thesis I set to investigate the computational and neural mechanisms of aversive learning focusing on the difference between acquisition and extinction of aversive associations. I present three studies investigating different aspects of aversive learning. In the first study, I used functional MRI, GSR and behavioural modelling to study the difference between acquisition and extinction, time-dependent effects and tested the hypothesis that high trait anxiety is associated with state learning. In the second study, I followed up on this idea investigating how environmental statistics influences the lack of extinction and whether less noisy environments encourage state switching as opposed to gradual learning. In the last study, I then explored whether a pharmacological intervention can influence aversive learning by improving extinction. The conclusions for each study are presented in their respective chapters. Here, I will revisit the seven goals set out in the introduction, state how I have addressed the key questions and discuss my answers each point. I will finish with overall conclusions, future directions and limitations of this work.

1) How do acquisition and extinction in probabilistic aversive learning differ with respect to behaviour and peripheral-physiological parameters?

The fundamental goal of this thesis was to improve our understanding of how the lack of extinction arises as a consequence of a Pavlovian learning process. Throughout all three studies, acquisition learning reached and settled close to the true reinforcement rate while extinction showed a gradual decrease and stops at about ~50% despite true probabilities being considerably lower. This is a replication of the

partial reinforcement extinction effect (PREE). It has been suggested that such lack of extinction is a result of reduced learning from shock outcome. I found that shock learning was indeed faster than no-shock learning, but this effect was modulated by the current phase: shock learning was faster in acquisition, where shock is more frequent and perhaps informative, whereas and no-shock learning was faster during extinction. In general the more confirmatory outcome was always associated with higher learning rates. This was particularly apparent in the stable cues where the harmful cue led to faster learning from shocks and the safe cue led to faster learning from no-shocks. This phenomenon has recently been described as the self-reinforcing expectancy effect (Jepma et al. 2018). It however does not explain why extinction learning is incomplete and settles at ~50%. The modulation of learning rates by the given phase demonstrates participants' knowledge of of some underlying context which suggests that learning is more than just differential learning from shock and noshock. To explore this aspect in more detail, I investigated how extinction learning changes over time. As participants experienced more switches in contingency, the lack of extinction became more pronounced. This is an important result because it provides insight into how aversive overprediction might *develop* in real-life scenarios. Although the lack of extinction has often been described, it is rarely discussed how the biased belief is formed in the first place. My data suggests that the bias is already established early on but becomes more pronounced over time. The increasing tendency to overpredict highlights the need for early intervention to prevent the development of robust overprediction belief.

2) Which role does trait anxiety play in probabilistic aversive learning?

Following up on question 1, Studies 1 and 2 explored the role of trait anxiety in extinction. Contrary to my hypothesis, high trait anxiety was associated with lower shock expectancy during extinction in study 1, indicating that anxiety *improves* extinction. In study 2 the effect was not significant across the three sessions, but high

trait anxiety was associated with a stronger difference in shock expectancy ratings between acquisition and extinction compared to low anxious group. Two key points were derived from further analyses: First, the anxiety-related difference only became apparent in the second half suggesting that the group difference is driven by a shift in strategy over time. Trait anxiety negatively correlated with mean shock probability ratings during extinction in the second half, specifically. Second, in study 2 the effect of anxiety on the difference between acquisition and extinction became more significant as the difference between the two contingency levels increased (i.e. largest effect in the 90/10 condition). The increased dissociability between acquisition and extinction was not only present in the behavioural but also in the GSR data. While a similar effect has been reported before (Indovina et al. 2011; Andreatta and Pauli 2017), other studies found a lack of conditioned response inhibition in high trait anxiety (Kindt and Soeter 2014; Haddad et al. 2012). This inconsistency might in part be due to differences in data interpretation. While some argue that the increased difference between CS+ and CS- is caused by a hypersensitivity to fear, it might also be argued that this is due an improved distinction between different internal contexts. In my analyses, I did not compare CS+ with CS- but rather acquisition and extinction. Here, an increased difference implies awareness of changes in underlying contingencies. The low anxious group seems to respond equally to acquisition and extinction which can either mean a lacking awareness of differences or a deliberate energy-saving strategy (settling on overprediction) . This raises the question whether our participants' behaviour was driven by survival or predictive optimality. The low anxious group undergoes extinction initially (at least partially) but gradually settles on consistent over-prediction. This is consistent with survival optimality: The individual samples the environment and creates a heuristic behavioural rule that minimised the amount of energy spent while maximizing the amount of predicted threat. This theory is partially supported by the observed association between the difference in positive and negative surprise and anxiety, where a portion of the low anxious group experienced decrease

in negative surprise in the second half. The computational model designed to assess whether this was a deliberate learning goal identified four participants that were driven by negative surprise minimisation *or* positive surprise maximisation.

A very useful follow-up on this finding would be to perform a long task, as I did here, but at some point change the contingencies associated with the reversal cue to stable low probability and either introduce a new reversal cue or change the safe cue into a volatile one. If the low anxious group disengaged from the task over time, they should take longer to discover this change than if the overprediction was a deliberate strategy-monitoring mechanism.

The observed improved extinction in high anxious participants constitutes a rather novel piece of evidence to which I offer two possible explanations. High trait anxious individuals might have an increased tendency to represent the environment as multiple states/contexts or they are hypervigilant and faithfully track the reported contingencies. In the next section I offer some evidence for the earlier but the latter hypothesis would have to rejected by a separate experiment.

3) Is there an association between trait anxiety and an increased tendency for state learning rather than gradual learning?

To follow up on my finding that high trait anxiety is associated with 'better' extinction (i.e., lower shock expectancy during late extinction and increased dissociability between acquisition and extinction) I tested whether this is due to an increased tendency to create internal contexts rather than gradual learning. The idea of structure learning (i.e. state learning) has been proposed previously (Redish, Jensen, et al. 2007; Gershman et al. 2013). To assess whether state learning occurs in the sample at all, I developed two computational models that have the ability to capture state switching, a Rescorla-Wagner based learner, and a beta state switcher. Fitting the models to the data, I next tested whether they are specifically associated with

anxiety. In study 1, the RW-based model provided the best fit for participants high in trait anxiety, providing strong support for the notion that state switching predominantly occurs in high trait anxiety. In study 2, the beta state switcher did not fit high trait anxious individuals better than low anxious but this was in part because it was designed to capture both gradual as well as state learning. A key parameter of the state switcher estimated a threshold necessary to create a new state. This threshold was significantly negatively correlated with trait anxiety in the extreme 90/10 condition showing that high TA participants were more likely to create new internal states. Furthermore, switch steepness in study 2 was estimated as a signature of state switching. The data show that the measure was positively associated with trait anxiety in the 90/10 condition and with the number of states internally created by the beta state switcher model. This provides evidence for an increased tendency to jump more abruptly between between different levels after reversal. While this measure on its own cannot explicitly distinguish between state switching and for example Pearce-Hall learning, it constitutes a signature suggesting more abrupt jumps in high trait anxious participants.

The behavioural and modelling data suggest that there is in fact a link between anxiety and increased tendency to represent the environment as multiple states. As discussed in Chapters 3 and 4 this mechanism is likely to be responsible for relapse phenomena seen in clinical populations (Craske et al. 2014b) and experimental studies (Lissek et al. 2005) associated with high anxiety. Ironically, better extinction might therefore mean a higher risk of relapse.

An alternative hypothesis for this phenomenon is offered by the vigilance-avoidance hypothesis (Karin Mogg, Mathews, and Weinman 1987) which argues that seemingly improved extinction in Generalised Anxiety (GAD) patients is caused by a lack of cognitive processing and so called over-inhibition. This mechanism was experimental tested by (Weinberg and Hajcak 2011) showing less elaborative processing in GAD in an ERP study. While my paradigm was not designed to explicitly

test for the difference between state learning and vigilance-avoidance (VA) hypothesis, one prediction that the VA makes is that there should be difference between anxiety groups in acquisition. However, in my data the difference lies solely in extinction. An interesting follow-up would be to design a paradigm that would attempt to directly test the difference between the two hypotheses, for example, using state decoding in the EEG or fMRI.

4) Which algorithms best explain the experimental learning data?

Throughout the three studies I employed a number of computational models to dissociate between different learning strategies or investigate whether a particular cognitive or pharmacological element (losartan, surprise) plays a role in learning. In general, most models were based on reinforcement learning which has proven to be a useful framework for modelling aversive learning. I have shown how it can be used to dissociate between models that make correlated predictions, yet are quite different their theoretical basis (e.g. state switching vs. Pearce-Hall). In this thesis, the only model that is not based on reinforcement learning is the beta state switcher which is currently being extended into a Bayesian framework.

In the current dataset the Pearce-Hall and the two state switching models seemed to fit the experimental data best. This is likely because they reflect two different hypothesis of adaptive learning. A few important points related to model definition and model comparison methods need to be discussed here. Throughout this thesis I only included models that performed well in parameter recovery test. however, in the losartan study (Chapter 5), it became apparent that the perceptual bias which was included in the prediction error term is particularly hard to estimate due to the fact that it always scales by the learning rate which is another free parameter. To ensure minimal loss of information and improve parameter estimability I employed the Bayesian formulation of the same model with hierarchical structure and population hyperpriors. When comparing models I employed the standard information criteria such

as AIC or BIC. However, in study I also used a form of cross-validation that assesses model quality based purely on the predictions it makes. I believe that the latter is a method superior to standard model comparison as it rules out the chance of overfitting. However, the method of cross-validation I used still requires further thought. An improvement could be made by fitting the data to a random selection of points across the learning and use the model to predict the gaps. While I considered this methodology, the issue is tricky when it comes to learning. In learning, unlike in other scenarios, the temporal dependence of data is crucial. A leave-one-out cross-validation might be biased towards models that provide a normative account rather than descriptive models that capture a particular cognitive phenomenon. While this is only a theoretical concern, it was outside of the scope of this doctoral work to systematically test this which is why I settled for the somewhat imperfect cross-validation method that preserves the temporal order of events.

One interesting aspect was the comparison between classical AIC model comparison and cross-validation. While the results mostly concur, there were some differences worth pointing out. In the AIC comparison, the Rescorla-Wagner model fitted best in 5, 0 and 1 subjects (respectively per environment) while when model fit was assessed by its prediction, the numbers for RW were 6, 3 and 6. The RW 'stole' some participants mostly from the PH model but not from the state-switching model. This reveals that a) the AIC, despite being penalized for the number of parameters, tends to value overfitting models, and b) that it is the PH model that tends to overfit, not the state switcher. This further supports the use of cross-validation in model comparison.

5) Which brain regions are involved in aversive acquisition and extinction learning? And where in the brain does anxiety influence the process?

The analysis of fMRI data revealed a number of regions which showed an increase in their differential response to acquisition and extinction over time.

Particularly the left hippocampus, right amygdala and the bilateral dorsal ACC were sensitive to this contrast. This provides further support that the amygdala might not be the sole critical structure driving aversive acquisition and extinction. A recent meta-analysis provided support for a key role of the dorsal ACC (Miquel A. Fullana et al. 2018) and the connectivity between dACC and the amygdala (Sevenster, Visser, and D'Hooge 2018). A novel insight into the distinction between the role of the amygdala and dACC came from my Bayesian Model Selection analysis which linked the dACC (and vmPFC) to structure learning. While this method has its limitations (discussed below), it suggests that the prefrontal regions are predominantly involved in context signalling while the amygdala might be involved in signalling of fear as reported by numerous other studies (Phelps et al, 2004). Several studies have highlighted the role of the hippocampus in context signalling to the ACC and the correlation between hippocampal and vmPFC activity. While the hippocampus was a candidate for the processing of contextual information I did not find state prediction error signalling there. One explanation might be that the hippocampus is involved in state signalling and the updating occurs via an indirect feedback mechanism. As in several previous studies, the dACC was also associated with trait anxiety. The signal in the region scaled with the difference between acquisition and extinction and trait anxiety, as well as with an increase in acquisition-extinction difference over time. This corroborates the previous finding that the dACC is involved in processing, or at least awareness of, state-related information. An interesting extension of the analysis I performed would be to investigate the connectivity between dACC, the hippocampus, amygdala, anterior insula and the vmPFC early and late in the task. This analysis could provide more information about the region(s) that first discover the state and those which subsequently hold this information. One hypothesis would be that the anterior insula is sensitive to initial state learning (as it showed increased dissociation between CS+ and CS- in the first half) while other regions such as the hippocampus, dACC and vmPFC take over once the state is discovered.

The data provides no definite answer to the questions whether the vmPFC is a region that inhibits the amygdala during extinction or whether it signals current state. While I did find SPE signalling in the vmPFC, this does not provide sufficient support for either of the hypotheses. A very interesting follow up study would be to tailor the experimental design to answering precisely this question.

Using SPE as a univariate parametric modulator I found the anterior insula to be the only region to reflect SPE processing. This analysis highlights a discrepancy between traditional methods and the novel BMC that require further investigation. The BMC method is only presented with two models and is forced to decide which model is more likely to be processed in a given voxel. This could be extended by adding a GLM without a parametric modulator that would be assigned the highest probability if neither SPE nor PHPE explained the data well.

Altogether, the neural data highlight the role of the dlPFC in guiding strategic learning based on survival optimality as well as the role of dACC and vmPFC in Pavlovian state error processing. Furthermore, dACC seems to be the key structure driving the influence of anxiety, which would support its role in state learning.

6) What is the influence of environmental statistics on learning? Do more extreme differences in contingency improve extinction?

In study 2 I investigated how the difference between the two contingency levels in acquisition and extinction influences learning. The design used three cues as in study 1, except that this time the task was split into three sessions in which probabilities between acquisition and extinction trials differed to a varying degree. The contingencies in the three environments were 60/40%, 75/25% and 90/10%, constituting a noisy, medium and clear environment. One of the goals of the second study was to investigate whether the noise introduced by the contingency levels encourages different learning strategies. In theory, the 60/40 environment presents no

clear enough states and so gradual learning should dominate whereas in the 90/10 condition, the two levels should be very clear and so state switching should be the preferred strategy. In the 75/25 condition participants were expected to be most conflicted about the optimal strategy, and so those with a tendency to represent the environment as states should refrain to that strategy while subjects with a tendency for gradual learning should follow that. My behavioural results revealed that there was no difference in switch steepness between the 60/40 and the 75/25 condition suggesting that participants used the same gradual learning throughout. Switch steepness was significantly higher in the 90/10 condition compared to both other conditions supporting the idea that state switching is promoted by the 90/10 condition. Interestingly this effect was driven entirely by the high anxious group.

Next, I used computational models to see which environments favours which strategy. The state switching model fitted best in the 60/40 and h 90/10 condition while the Pearce-Hall fitted best in the 75/25 condition. It is important to keep in mind here that the state switcher has a mechanism to learn gradually in a manner similar to RW and so it likely fitted best the subjects following RW-like learning. Apart from that, the results suggest that in the 90/10 condition state switching is the dominant strategy, the error for the states witcher is substantially lower, while in the 75/25 condition there is more competition between the state switcher and the PH model, the difference in mean errors bein marginal and amount of best fitted participants being 16 and 14.

I next looked at how different contingency levels influenced mean probability reported by subjects in acquisition and extinction. While mean probability in acquisition increased with increasing reinforcement rate, mean probability in extinction remained the same (roughly ~50%) in all three environments. This is quite a paradoxical result given that the difference in reinforcement rate in extinction between the two most extreme environments is 30% which translates to roughly 10 non-reinforced trials more in the 90/10 condition than in the 60/40 condition. The results seem to suggest that in the group as a whole the contingency difference between acquisition and extinction

doesn't seem to play a role in improving extinction. As reported previously, the only group where the 90/10 environment seemed to have played difference was the high anxious group which showed significantly association with phase dissociability.

All together, the data support the hypothesis that high trait anxiety and more extreme changes in contingencies encourage state switching.

7) How does the angiotensin-II antagonist losartan influence aversive probabilistic learning?

In the third study I found that the Angiotensin-II reuptake inhibitor losartan has a decelerating effect on learning. Furthermore, losartan seems to specifically affect learning rather than perception or updating. This was conjointly supported by the hierarchical Bayesian model and pupil data. Unfortunately, the data were strongly influenced by the starting probability of the reversal cue on phase 1 on each visit. Consequently, this factor had to be included in all tests which diluted the statistical power of many analyses (the power calculations were made for ANOVAs with one between-subject variable). Despite this complication, best effort was made to understand the effect of losartan. The hierarchical Bayesian model revealed that losartan decreased the learning rates in the group starting with 25% probability. Furthermore, on a follow-up visit this effect was still present but this time for the entire sample. Using the Bayesian model I was able to assess whether the effect of losartan was indeed driven by an influence on learning or whether it was due to a change in perception or updating. The finding that losartan influences learning and not perception was further supported by my pupillometry analysis which showed no effect of outcome encoding while there was a significant effect of losartan on prediction error processing. As discussed in the relevant chapter the decrease in learning rates resulted in less complete extinction on visit 2 in the group starting on low probability and on visit 3 in the group starting on high probability.

The key finding that losartan influences learning is encouraging but further

studies are needed to explore its potential in more detail.

Overall Conclusions

In my doctoral thesis I investigated the behavioural, computational and neural differences between aversive acquisition and extinction. Across studies, extinction was found to be incomplete compared to acquisition, a finding that was not modulated by the level of reinforcement. The lack of extinction increased over time, suggesting a possible mechanism for the development of overprediction bias. The right dlPFC was associated with this effect. Furthermore, trait anxiety played a significant role in extinction, high trait anxious individuals consistently reported lower probability. I proposed and tested that this was due to increased tendency of the high anxiety group to organise learning environment into distinct states, and switch between them, finding a support for the hypothesis in behavioural, GSR and neural data. I demonstrate that activity in the dorsal anterior cingulate scales with anxiety and that the same region together with the vmPFC and IPL process information about current state. This finding constitutes a major outcome of the thesis, and proposes a novel computational mechanism behind the increased tendency to relapse in highly anxious people. Lastly, I tested whether aversive learning can be affected by pharmacological intervention for which I found a moderately strong support.

References

Acerbi, Luigi, and Wei Ji Ma. 2017. “Practical Bayesian Optimization for Model Fitting with Bayesian Adaptive Direct Search.” In *Advances in Neural Information Processing Systems 30*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, 1836–46. Curran Associates, Inc.

Allen, A. M., J. Zhuo, and F. A. Mendelsohn. 1999. “Localization of Angiotensin AT1 and AT2 Receptors.” *Journal of the American Society of Nephrology: JASN* 10 Suppl 11 (January): S23–29.

Alvarez, R. P., L. Johnson, and C. Grillon. 2007. “Contextual-Specificity of Short-Delay Extinction in Humans: Renewal of Fear-Potentiated Startle in a Virtual Environment.” *Learning & Memory*. https://doi.org/10.1101/lm.493707.

Andersson, Erik, Erik Hedman, Jesper Enander, Diana Radu Djurfeldt, Brjánn Ljótsson, Simon Cervenka, Josef Isung, et al. 2015. “D-Cycloserine vs Placebo as Adjunct to Cognitive Behavioral Therapy for Obsessive-Compulsive Disorder and Interaction With Antidepressants: A Randomized Clinical Trial.” *JAMA Psychiatry* 72 (7): 659–67.

Andreatta, Marta, and Paul Pauli. 2017. “Learning Mechanisms Underlying Threat Absence and Threat Relief: Influences of Trait Anxiety.” *Neurobiology of Learning and Memory* 145 (November): 105–13.

Ansari, Tahereh L., Nazanin Derakshan, and Anne Richards. 2008. “Effects of Anxiety on Task Switching: Evidence from the Mixed Antisaccade Task.” *Cognitive, Affective & Behavioral Neuroscience* 8 (3): 229–38.

Arch, Joanna J., and Michelle G. Craske. 2009. “First-Line Treatment: A Critical Appraisal of Cognitive Behavioral Therapy Developments and Alternatives.” *The Psychiatric Clinics of North America* 32 (3): 525–47.

Armando, Ines, Simona Volpi, Greti Aguilera, and Juan M. Saavedra. 2007. “Angiotensin II AT1 Receptor Blockade Prevents the Hypothalamic Corticotropin-Releasing Factor Response to Isolation Stress.” *Brain Research* 1142 (April): 92–99.

Aston-Jones, Gary, and Jonathan D. Cohen. 2005. “An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance.” *Annual Review of Neuroscience* 28: 403–50.

Baas, Johanna M. P., and Ivo Heitland. 2015. “The Impact of Cue Learning, Trait Anxiety and Genetic Variation in the Serotonin 1A Receptor on Contextual Fear.” *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology* 98 (3 Pt 2): 506–14.

Bach, Dominik R. 2014. “A Head-to-Head Comparison of SCRalyze and Ledalab, Two Model-Based Methods for Skin Conductance Analysis.” *Biological Psychology* 103 (December): 63–68.

Bach, Dominik R., Jean Daunizeau, Karl J. Friston, and Raymond J. Dolan. 2010. “Dynamic Causal Modelling of Anticipatory Skin Conductance Responses.” *Biological Psychology*. https://doi.org/10.1016/j.biopsycho.2010.06.007.

Bach, Dominik R., Guillaume Flandin, Karl J. Friston, and Raymond J. Dolan. 2009. “Time-Series Analysis for Rapid Event-Related Skin Conductance Responses.” *Journal of Neuroscience Methods* 184 (2): 224–34.

Bach, Dominik R., and Karl J. Friston. 2013. “Model-Based Analysis of Skin Conductance Responses: Towards Causal Models in Psychophysiology.” *Psychophysiology* 50 (1): 15–22.

Bach, Dominik R., Christoph W. Korn, Johanna Vunder, and Antonia Bantel. 2018. “Effect of Valproate and Pregabalin on Human Anxiety-like Behaviour in a Randomised Controlled Trial.” *Translational Psychiatry*. https://doi.org/10.1038/s41398-018-0206-7.

Bach, D. R., A. Tzovara, and J. Vunder. 2018. “Blocking Human Fear Memory with the

Matrix Metalloproteinase Inhibitor Doxycycline." *Molecular Psychiatry* 23 (7): 1584–89.
Barrett, J., and J. L. Armony. 2009. "Influence of Trait Anxiety on Brain Activity during the Acquisition and Extinction of Aversive Conditioning." *Psychological Medicine* 39 (2): 255–65.
Baum, Morrie. 1972. "Repeated Acquisition and Extinction of Avoidance in Rats Using Flooding (response Prevention)." *Learning and Motivation*. https://doi.org/10.1016/0023-9690(72)90023-9.
Beck, Aaron T., R. A. Steer, and G. Brown. 1996. "Beck Depression Inventory–II." *PsycTESTS Dataset*. https://doi.org/10.1037/t00742-000.
Behrens, Timothy E. J., Mark W. Woolrich, Mark E. Walton, and Matthew F. S. Rushworth. 2007. "Learning the Value of Information in an Uncertain World." *Nature Neuroscience*. https://doi.org/10.1038/nn1954.
Belleau, Emily L., Walker S. Pedersen, Tara A. Miskovich, Fred J. Helmstetter, and Christine L. Larson. 2018. "Cortico-Limbic Connectivity Changes Following Fear Extinction and Relationships with Trait Anxiety." *Social Cognitive and Affective Neuroscience* 13 (10): 1037–46.
Berger, James O. 1985. "Statistical Decision Theory and Bayesian Analysis." *Springer Series in Statistics*. https://doi.org/10.1007/978-1-4757-4286-2.
Berker, Archy O. de, Robb B. Rutledge, Christoph Mathys, Louise Marshall, Gemma F. Cross, Raymond J. Dolan, and Sven Bestmann. 2016. "Computations of Uncertainty Mediate Acute Stress Responses in Humans." *Nature Communications* 7 (March): 10996.
Beylergil, Sinem Balta, Anne Beck, Lorenz Deserno, Robert C. Lorenz, Michael A. Rapp, Florian Schlagenhauf, Andreas Heinz, and Klaus Obermayer. 2017. "Dorsolateral Prefrontal Cortex Contributes to the Impaired Behavioral Adaptation in Alcohol Dependence." *NeuroImage. Clinical* 15 (April): 80–94.
Birrell, Jane, Kevin Meares, Andrew Wilkinson, and Mark Freeston. 2011. "Toward a Definition of Intolerance of Uncertainty: A Review of Factor Analytical Studies of the Intolerance of Uncertainty Scale." *Clinical Psychology Review*. https://doi.org/10.1016/j.cpr.2011.07.009.
Bishop, Sonia J. 2009. "Trait Anxiety and Impoverished Prefrontal Control of Attention." *Nature Neuroscience* 12 (1): 92–98.
Bonardi, Charlotte, R. C. Honey, and Geoffrey Hall. 1990. "Context Specificity of Conditioning in Flavor-Aversion Learning: Extinction and Blocking Tests." *Animal Learning & Behavior* 18 (3): 229–37.
Botvinick, Matthew M., Yael Niv, and Andrew C. Barto. 2009. "Hierarchically Organized Behavior and Its Neural Foundations: A Reinforcement Learning Perspective." *Cognition* 113 (3): 262–80.
Boughner, Robert L., and Mauricio R. Papini. 2006. "Survival of the Partial Reinforcement Extinction Effect after Contextual Shifts." *Learning and Motivation*. https://doi.org/10.1016/j.lmot.2005.09.001.
Bouton, Mark E. 1994. "Context, Ambiguity, and Classical Conditioning." *Current Directions in Psychological Science*. https://doi.org/10.1111/1467-8721.ep10769943.
Bouton. 2002. "Context, Ambiguity, and Unlearning: Sources of Relapse after Behavioral Extinction." *Biological Psychiatry* 52 (10): 976–86.
Bouton, Mark E., and Robert C. Bolles. 1979. "Contextual Control of the Extinction of Conditioned Fear." *Learning and Motivation*. https://doi.org/10.1016/0023-9690(79)90057-2.
Bouton, Mark E., and Douglas C. Brooks. 1993. "Time and Context Effects on Performance in a Pavlovian Discrimination Reversal." *Journal of Experimental Psychology: Animal Behavior Processes*. https://doi.org/10.1037//0097-7403.19.2.165.
Bouton, Mark E., and David A. King. 1983. "Contextual Control of the Extinction of

Conditioned Fear: Tests for the Associative Value of the Context." *Journal of Experimental Psychology: Animal Behavior Processes*. https://doi.org/10.1037//0097-7403.9.3.248.
Bouton, M. E. and King, P. 1986. "Effect of Context on Performance to Conditioned Stimuli with Mixed Histories of Reinforcement and Nonreinforcement." *Journal of Experimental Psychology: Animal Behavior Processes*. https://doi.org/10.1037/0097-7403.12.1.4.
Brooks, D. C., and M. E. Bouton. 1993. "A Retrieval Cue for Extinction Attenuates Spontaneous Recovery." *Journal of Experimental Psychology. Animal Behavior Processes* 19 (1): 77–89.
Brooks, Steve, Andrew Gelman, Galin Jones, and Xiao-Li Meng. 2011. *Handbook of Markov Chain Monte Carlo*. CRC Press.
Browning, Michael, Timothy E. Behrens, Gerhard Jocham, Jill X. O'Reilly, and Sonia J. Bishop. 2015. "Anxious Individuals Have Difficulty Learning the Causal Statistics of Aversive Environments." *Nature Neuroscience* 18 (4): 590–96.
Brown, Vanessa M., Kevin S. LaBar, Courtney C. Haswell, Andrea L. Gold, Mid-Atlantic MIRECC Workgroup, Gregory McCarthy, and Rajendra A. Morey. 2014. "Altered Resting-State Functional Connectivity of Basolateral and Centromedial Amygdala Complexes in Posttraumatic Stress Disorder." *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology* 39 (2): 351–59.
Burri, Andrea, and Andreas Maercker. 2014. "Differences in Prevalence Rates of PTSD in Various European Countries Explained by War Exposure, Other Trauma and Cultural Value Orientation." *BMC Research Notes* 7 (June): 407.
Carver, Charles S., and Teri L. White. 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): 319–33.
Chang, Chun-Hui, and Stephen Maren. 2011. "Medial Prefrontal Cortex Activation Facilitates Re-Extinction of Fear in Rats." *Learning & Memory* 18 (4): 221–25.
Chib, Vikram S., Antonio Rangel, Shinsuke Shimojo, and John P. O'Doherty. 2009. "Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29 (39): 12315–20.
Colizoli, Olympia, Jan Willem de Gee, Anne E. Urai, and Tobias H. Donner. 2018. "Task-Evoked Pupil Responses Reflect Internal Belief States." *Scientific Reports* 8 (1): 13702.
Corcoran, Kevin A., Timothy J. Desmond, Kirk A. Frey, and Stephen Maren. 2005. "Hippocampal Inactivation Disrupts the Acquisition and Contextual Encoding of Fear Extinction." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 25 (39): 8978–87.
Craske, Michelle G., Michael Treanor, Christopher C. Conway, Tomislav Zbozinek, and Bram Vervliet. 2014a. "Maximizing Exposure Therapy: An Inhibitory Learning Approach." *Behaviour Research and Therapy*. 58 (July): 10–23.
Culver, Najwa C., Bram Vervliet, and Michelle G. Craske. 2015. "Compound Extinction." *Clinical Psychological Science*. https://doi.org/10.1177/2167702614542103.
Daunizeau, Jean, Vincent Adam, and Lionel Rigoux. 2014. "VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data." *PLoS Computational Biology* 10 (1): e1003441.
Daw, Nathaniel D., Yael Niv, and Peter Dayan. 2005. "Uncertainty-Based Competition between Prefrontal and Dorsolateral Striatal Systems for Behavioral Control." *Nature Neuroscience* 8 (12): 1704–11.
Dawson, M. E., A. M. Schell, and H. T. Banis. 1986. "Greater Resistance to Extinction of Electrodermal Responses Conditioned to Potentially Phobic CSs: A Noncognitive Process?" *Psychophysiology* 23 (5): 552–61.

Dayan, Peter, and Kent C. Berridge. 2014. “Model-Based and Model-Free Pavlovian Reward Learning: Revaluation, Revision, and Revelation.” *Cognitive, Affective & Behavioral Neuroscience* 14 (2): 473–92.
Dayan, Peter, and Angela Jyu. 2003. “Uncertainty and Learning.” *IETE Journal of Research*. https://doi.org/10.1080/03772063.2003.11416335.
Delgado, Mauricio R., Jian Li, Daniela Schiller, and Elizabeth A. Phelps. 2008. “The Role of the Striatum in Aversive Learning and Aversive Prediction Errors.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363 (1511): 3787–3800.
Derryberry, Douglas, and Marjorie A. Reed. 2002. “Anxiety-Related Attentional Biases and Their Regulation by Attentional Control.” *Journal of Abnormal Psychology* 111 (2): 225–36.
Dickey, James M. 1971. “The Weighted Likelihood Ratio, Linear Hypotheses on Normal Location Parameters.” *The Annals of Mathematical Statistics*. https://doi.org/10.1214/aoms/1177693507.
Dunsmoor, Joseph E., Marijn C. W. Kroes, Jian Li, Nathaniel D. Daw, Helen B. Simpson, and Elizabeth A. Phelps. 2019. “Role of Human Ventromedial Prefrontal Cortex in Learning and Recall of Enhanced Extinction.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, February. https://doi.org/10.1523/JNEUROSCI.2713-18.2019.
Dunsmoor, Joseph E., Yael Niv, Nathaniel Daw, and Elizabeth A. Phelps. 2015. “Rethinking Extinction.” *Neuron* 88 (1): 47–63.
Edwards, Mark S., Elizabeth J. Edwards, and Michael Lyvers. 2017. “Cognitive Trait Anxiety, Stress and Effort Interact to Predict Inhibitory Control.” *Cognition and Emotion*. https://doi.org/10.1080/02699931.2016.1152232.
Elam, M., T. Yao, T. H. Svensson, and P. Thoren. 1984. “Regulation of Locus Coeruleus Neurons and Splanchnic, Sympathetic Nerves by Cardiovascular Afferents.” *Brain Research* 290 (2): 281–87.
Eysenck, H. J., and S. B. G. Eysenck. 1975. “Junior Eysenck Personality Questionnaire.” *PsycTESTS Dataset*. https://doi.org/10.1037/t05462-000.
Feng, Pan, Yong Zheng, and Tingyong Feng. 2016. “Resting-State Functional Connectivity between Amygdala and the Ventromedial Prefrontal Cortex Following Fear Reminder Predicts Fear Extinction.” *Social Cognitive and Affective Neuroscience* 11 (6): 991–1001.
First, Michael B., Robert L. Spitzer, Janet B. W. Williams, and Rhonda S. Karg. 2015. *Structured Clinical Interview for Dsm-5 Disorders (Scid-5-Cv): Clinician Version*.
Fitzgerald, Paul J., Jocelyn R. Seemann, and Stephen Maren. 2014. “Can Fear Extinction Be Enhanced? A Review of Pharmacological and Behavioral Findings.” *Brain Research Bulletin*. https://doi.org/10.1016/j.brainresbull.2013.12.007.
Forstmann, B. U., R. Ratcliff, and E-J Wagenmakers. 2016. “Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions.” *Annual Review of Psychology* 67: 641–66.
Fullana, M. A., B. J. Harrison, C. Soriano-Mas, B. Vervliet, N. Cardoner, A. Àvila-Parcet, and J. Radua. 2016. “Neural Signatures of Human Fear Conditioning: An Updated and Extended Meta-Analysis of fMRI Studies.” *Molecular Psychiatry* 21 (4): 500–508.
Fullana, Miquel A., Anton Albajes-Eizagirre, Carles Soriano-Mas, Bram Vervliet, Narcís Cardoner, Olívia Benet, Joaquim Radua, and Ben J. Harrison. 2018. “Fear Extinction in the Human Brain: A Meta-Analysis of fMRI Studies in Healthy Participants.” *Neuroscience & Biobehavioural Reviews*. https://doi.org/10.1016/j.neubiorev.2018.03.002.
Gard, Paul R. 2004. “Angiotensin as a Target for the Treatment of Alzheimer’s Disease, Anxiety and Depression.” *Expert Opinion on Therapeutic Targets* 8 (1): 7–14.
Gazendam, Femke J., Jan H. Kamphuis, and Merel Kindt. 2013. “Deficient Safety

Learning Characterizes High Trait Anxious Individuals." *Biological Psychology* 92 (2): 342–52.
Gershman, Samuel J., David M. Blei, and Yael Niv. 2010. "Context, Learning, and Extinction." *Psychological Review*. https://doi.org/10.1037/a0017808.
Gershman, Samuel J., Carolyn E. Jones, Kenneth A. Norman, Marie-H Monfils, and Yael Niv. 2013. "Gradual Extinction Prevents the Return of Fear: Implications for the Discovery of State." *Frontiers in Behavioral Neuroscience* 7 (November): 164.
Gerster, Samuel, Barbara Namer, Mikael Elam, and Dominik R. Bach. 2018. "Testing a Linear Time Invariant Model for Skin Conductance Responses by Intraneural Recording and Stimulation." *Psychophysiology* 55 (2). https://doi.org/10.1111/psyp.12986.
Gläscher, Jan, Nathaniel Daw, Peter Dayan, and John P. O'Doherty. 2010. "States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning." *Neuron* 66 (4): 585–95.
Gong, Wan-Kun, Jun Lü, Feng Wang, Bin Wang, Meng-Ya Wang, and Hong-Ping Huang. 2015. "Effects of Angiotensin Type 2 Receptor on Secretion of the Locus Coeruleus in Stress-Induced Hypertension Rats." *Brain Research Bulletin* 111 (February): 62–68.
Gorka, Stephanie M., Lynne Lieberman, Brady D. Nelson, Casey Sarapas, and Stewart A. Shankman. 2014. "Aversive Responding to Safety Signals in Panic Disorder: The Moderating Role of Intolerance of Uncertainty." *Journal of Anxiety Disorders* 28 (7): 731–36.
Gottlieb, Daniel A., and Robert A. Rescorla. 2010. "Within-Subject Effects of Number of Trials in Rat Conditioning Procedures." *Journal of Experimental Psychology. Animal Behavior Processes* 36 (2): 217–31.
Haaker, J., T. B. Lonsdorf, D. Schümann, M. Menz, S. Brassen, N. Bunzeck, M. Gamer, and R. Kalisch. 2015. "Deficient Inhibitory Processing in Trait Anxiety: Evidence from Context-Dependent Fear Learning, Extinction Recall and Renewal." *Biological Psychology* 111 (October): 65–72.
Haddad, Anneke D. M., David Pritchett, Shmuel Lissek, and Jennifer Y. F. Lau. 2012. "Trait Anxiety and Fear Responses to Safety Cues: Stimulus Generalization or Sensitization?" *Journal of Psychopathology and Behavioral Assessment*. https://doi.org/10.1007/s10862-012-9284-7.
Hall, Geoffrey, and R. C. Honey. 1989. "Contextual Effects in Conditioning, Latent Inhibition, and Habituation: Associative and Retrieval Functions of Contextual Cues." *Journal of Experimental Psychology: Animal Behavior Processes*. https://doi.org/10.1037/0097-7403.15.3.232.
Harlow, Harry F. 1949. "The Formation of Learning Sets." *Psychological Review* 56 (1): 51–65.
Harris, Justin A., Megan L. Jones, Glynis K. Bailey, and R. Frederick Westbrook. 2000. "Contextual Control over Conditioned Responding in an Extinction Paradigm." *Journal of Experimental Psychology: Animal Behavior Processes*. https://doi.org/10.1037/0097-7403.26.2.174.
Hartley, Catherine A., and Elizabeth A. Phelps. 2010. "Changing Fear: The Neurocircuitry of Emotion Regulation." *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology* 35 (1): 136–46.
Hermann, Andrea, Tanja Keck, and Rudolf Stark. 2014. "Dispositional Cognitive Reappraisal Modulates the Neural Correlates of Fear Acquisition and Extinction." *Neurobiology of Learning and Memory* 113 (September): 115–24.
Hermann, Andrea, Rudolf Stark, Carlo R. Blecker, Mohammed R. Milad, and Christian J. Merz. 2017. "Brain Structural Connectivity and Context-Dependent Extinction Memory." *Hippocampus* 27 (8): 883–89.
Hofmann, Stefan G. 2008. "Cognitive Processes during Fear Acquisition and Extinction in Animals and Humans: Implications for Exposure Therapy of Anxiety Disorders."

Clinical Psychology Review 28 (2): 199–210.
Hsu, Ming, Meghana Bhatt, Ralph Adolphs, Daniel Tranel, and Colin F. Camerer. 2005. “Neural Systems Responding to Degrees of Uncertainty in Human Decision-Making.” *Science* 310 (5754): 1680–83.
Indovina, Iole, Trevor W. Robbins, Anwar O. Núñez-Elizalde, Barnaby D. Dunn, and Sonia J. Bishop. 2011. “Fear-Conditioning Mechanisms Associated with Trait Vulnerability to Anxiety in Humans.” *Neuron* 69 (3): 563–71.
Ironside, Maria, Michael Browning, Tahereh L. Ansari, Christopher J. Harvey, Mama N. Sekyi-Djan, Sonia J. Bishop, Catherine J. Harmer, and Jacinta O’Shea. 2018. “Effect of Prefrontal Cortex Stimulation on Regulation of Amygdala Response to Threat in Individuals With Trait Anxiety: A Randomized Clinical Trial.” *JAMA Psychiatry* , October. https://doi.org/10.1001/jamapsychiatry.2018.2172.
Jenkinson, Mark, Christian F. Beckmann, Timothy E. J. Behrens, Mark W. Woolrich, and Stephen M. Smith. 2012. “FSL.” *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2011.09.015.
Jenkins, William O., and Julian C. Stanley. 1950. “Partial Reinforcement: A Review and Critique.” *Psychological Bulletin*. https://doi.org/10.1037/h0060772.
Jensen, Jimmy, Anthony R. McIntosh, Adrian P. Crawley, David J. Mikulis, Gary Remington, and Shitij Kapur. 2003. “Direct Activation of the Ventral Striatum in Anticipation of Aversive Stimuli.” *Neuron*. https://doi.org/10.1016/s0896-6273(03)00724-4.
Jepma, Marieke, Leonie Koban, Johnny van Doorn, Matt Jones, and Tor D. Wager. 2018. “Behavioural and Neural Evidence for Self-Reinforcing Expectancy Effects on Pain.” *Nature Human Behaviour*. https://doi.org/10.1038/s41562-018-0455-8.
Joshi, Siddhartha, Yin Li, Rishi M. Kalwani, and Joshua I. Gold. 2016. “Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex.” *Neuron* 89 (1): 221–34.
Kahneman, Daniel, and Amos Tversky. 2000. “Prospect Theory: An Analysis of Decision under Risk.” *Choices, Values, and Frames*. https://doi.org/10.1017/cbo9780511803475.003.
Kakade, Sham, and Peter Dayan. 2002. “Acquisition and Extinction in Autoshaping.” *Psychological Review* 109 (3): 533–44.
Kalisch, Raffael, Elian Korenfeld, Klaas E. Stephan, Nikolaus Weiskopf, Ben Seymour, and Raymond J. Dolan. 2006. “Context-Dependent Human Extinction Memory Is Mediated by a Ventromedial Prefrontal and Hippocampal Network.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 26 (37): 9503–11.
Khoury, Nayla M., Paul J. Marvar, Charles F. Gillespie, Aliza Wingo, Ann Schwartz, Bekh Bradley, Michael Kramer, and Kerry J. Ressler. 2012. “The Renin-Angiotensin Pathway in Posttraumatic Stress Disorder: Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers Are Associated with Fewer Traumatic Stress Symptoms.” *The Journal of Clinical Psychiatry* 73 (6): 849–55.
Kindt, Merel, and Marieke Soeter. 2014. “Fear Inhibition in High Trait Anxiety.” *PloS One* 9 (1): e86462.
Kleine, Rianne A. de, Barbara O. Rothbaum, and Agnes van Minnen. 2013. “Pharmacological Enhancement of Exposure-Based Treatment in PTSD: A Qualitative Review.” *European Journal of Psychotraumatology* 4 (October). https://doi.org/10.3402/ejpt.v4i0.21626.
Knight, David C., Christine N. Smith, Dominic T. Cheng, Elliot A. Stein, and Fred J. Helmstetter. 2004. “Amygdala and Hippocampal Activity during Acquisition and Extinction of Human Fear Conditioning.” *Cognitive, Affective & Behavioral Neuroscience* 4 (3): 317–25.
Knill, David C., and Alexandre Pouget. 2004. “The Bayesian Brain: The Role of Uncertainty in Neural Coding and Computation.” *Trends in Neurosciences* 27 (12): 712–19.

Koenig, Stephan, Metin Uengoer, and Harald Lachnit. 2018. “Pupil Dilation Indicates the Coding of Past Prediction Errors: Evidence for Attentional Learning Theory.” *Psychophysiology* 55 (4). https://doi.org/10.1111/psyp.13020.
Krause, Eric G., Annette D. de Kloet, Karen A. Scott, Jonathan N. Flak, Kenneth Jones, Michael D. Smeltzer, Yvonne M. Ulrich-Lai, et al. 2011. “Blood-Borne Angiotensin II Acts in the Brain to Influence Behavioral and Endocrine Responses to Psychogenic Stress.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31 (42): 15009–15.
Kret, Mariska E., and Elio E. Sjak-Shie. 2019. “Preprocessing Pupil Size Data: Guidelines and Code.” *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-1075-y.
Kuchibhotla, Kishore V., Tom Hindmarsh Sten, Eleni S. Papadoyannis, Sarah Elnozahy, Kelly A. Fogelson, Rupesh Kumar, Yves Boubenec, Peter C. Holland, Srdjan Ostojic, and Robert C. Froemke. 2019. “Dissociating Task Acquisition from Expression during Learning Reveals Latent Knowledge.” *Nature Communications* 10 (1): 2151.
Kułakowska, A., W. Karwowska, K. Wiśniewski, and J. J. Braszko. 1996. “Losartan Influences Behavioural Effects of Angiotensin II in Rats.” *Pharmacological Research: The Official Journal of the Italian Pharmacological Society* 34 (3-4): 109–15.
LaBar, Kevin S., J. Christopher Gatenby, John C. Gore, Joseph E. LeDoux, and Elizabeth A. Phelps. 1998. “Human Amygdala Activation during Conditioned Fear Acquisition and Extinction: A Mixed-Trial fMRI Study.” *Neuron*. https://doi.org/10.1016/s0896-6273(00)80475-4.
LaBar, Kevin S., and Elizabeth A. Phelps. 2005. “Reinstatement of Conditioned Fear in Humans Is Context Dependent and Impaired in Amnesia.” *Behavioral Neuroscience* 119 (3): 677–86.
Lavín, Claudio, René San Martín, and Eduardo Rosales Jubal. 2014. “Pupil Dilation Signals Uncertainty and Surprise in a Learning Gambling Task.” *Frontiers in Behavioral Neuroscience* 7. https://doi.org/10.3389/fnbeh.2013.00218.
Lawson, Rebecca P., Christoph Mathys, and Geraint Rees. 2017. “Adults with Autism Overestimate the Volatility of the Sensory Environment.” *Nature Neuroscience* 20 (9): 1293–99.
Lazaroni, Thiago Luiz do Nascimento, Cristiane Perácio Bastos, Márcio Flávio Dutra Moraes, Robson Souza Santos, and Grace Schenatto Pereira. 2016. “Angiotensin-(1-7)/Mas Axis Modulates Fear Memory and Extinction in Mice.” *Neurobiology of Learning and Memory* 127 (January): 27–33.
LeDoux, J. E. 2000. “Emotion Circuits in the Brain.” *Annual Review of Neuroscience* 23: 155–84.
LeDoux, Joseph. 2012. “Rethinking the Emotional Brain.” *Neuron* 73 (4): 653–76.
Lee, Michael D., and Eric-Jan Wagenmakers. 2014. *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.
Levy, Dino J., and Paul W. Glimcher. 2012. “The Root of All Value: A Neural Common Currency for Choice.” *Current Opinion in Neurobiology* 22 (6): 1027–38.
Li, Jian, Daniela Schiller, Geoffrey Schoenbaum, Elizabeth A. Phelps, and Nathaniel D. Daw. 2011. “Differential Roles of Human Striatum and Amygdala in Associative Learning.” *Nature Neuroscience* 14 (10): 1250–52.
Lissek, Shmuel, Alice S. Powers, Erin B. McClure, Elizabeth A. Phelps, Girma Woldehawariat, Christian Grillon, and Daniel S. Pine. 2005. “Classical Fear Conditioning in the Anxiety Disorders: A Meta-Analysis.” *Behaviour Research and Therapy* 43 (11): 1391–1424.
Li, Yuzhe, Ken Nakae, Shin Ishii, and Honda Naoki. 2016. “Uncertainty-Dependent Extinction of Fear Memory in an Amygdala-mPFC Neural Circuit Model.” *PLOS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1005099.
Lockwood, Patricia L., Marco K. Wittmann, Matthew A. J. Apps, Miriam C. Klein-

Flügge, Molly J. Crockett, Glyn W. Humphreys, and Matthew F. S. Rushworth. 2018. “Neural Mechanisms for Learning Self and Other Ownership.” *Nature Communications*. https://doi.org/10.1038/s41467-018-07231-9.
Lodewyckx, Tom, Woojae Kim, Michael D. Lee, Francis Tuerlinckx, Peter Kuppens, and Eric-Jan Wagenmakers. 2011. “A Tutorial on Bayes Factor Estimation with the Product Space Method.” *Journal of Mathematical Psychology*. https://doi.org/10.1016/j.jmp.2011.06.001.
Lo, M. W., M. R. Goldberg, J. B. McCrea, H. Lu, C. I. Furtek, and T. D. Bjornsson. 1995. “Pharmacokinetics of Losartan, an Angiotensin II Receptor Antagonist, and Its Active Metabolite EXP3174 in Humans.” *Clinical Pharmacology and Therapeutics* 58 (6): 641–49.
Lovibond, Peter F., and David R. Shanks. 2002. “The Role of Awareness in Pavlovian Conditioning: Empirical Evidence and Theoretical Implications.” *Journal of Experimental Psychology. Animal Behavior Processes* 28 (1): 3–26.
MacArdy, E. A., and D. C. Riccio. 1991. “Increased Generalization between Drug-Related Interoceptive Stimuli with Delayed Testing.” *Behavioral and Neural Biology* 56 (2): 213–19.
Mackintosh, N. J. 1975. “A Theory of Attention: Variations in the Associability of Stimuli with Reinforcement.” *Psychological Review*. https://doi.org/10.1037/h0076778.
MacLeod, C., and A. Mathews. 1988. “Anxiety and the Allocation of Attention to Threat.” *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology* 40 (4): 653–70.
Marek, Roger, Jingji Jin, Travis D. Goode, Thomas F. Giustino, Qian Wang, Gillian M. Acca, Roopashri Holehonnur, et al. 2018. “Author Correction: Hippocampus-Driven Feed-Forward Inhibition of the Prefrontal Cortex Mediates Relapse of Extinguished Fear.” *Nature Neuroscience* 21 (9): 1291.
Marin, Marie-France, Rachel G. Zsido, Huijin Song, Natasha B. Lasko, William D. S. Killgore, Scott L. Rauch, Naomi M. Simon, and Mohammed R. Milad. 2017. “Skin Conductance Responses and Neural Activations During Fear Conditioning and Extinction Recall Across Anxiety Disorders.” *JAMA Psychiatry* 74 (6): 622–31.
Marshall, Louise, Christoph Mathys, Diane Ruge, Archy O. de Berker, Peter Dayan, Klaas E. Stephan, and Sven Bestmann. 2016. “Pharmacological Fingerprints of Contextual Uncertainty.” *PLoS Biology* 14 (11): e1002575.
Marvar, Paul J., Jared Goodman, Sebastien Fuchs, Dennis C. Choi, Sunayana Banerjee, and Kerry J. Ressler. 2014. “Angiotensin Type 1 Receptor Inhibition Enhances the Extinction of Fear Memory.” *Biological Psychiatry* 75 (11): 864–72.
Mataix-Cols, David, Cynthia Turner, Benedetta Monzani, Kayoko Isomura, Caroline Murphy, Georgina Krebs, and Isobel Heyman. 2014. “Cognitive–behavioural Therapy with Post-Session D-Cycloserine Augmentation for Paediatric Obsessive–compulsive Disorder: Pilot Randomised Controlled Trial.” *British Journal of Psychiatry* 204 (01): 77–78.
Mathys, Christoph, Jean Daunizeau, Karl J. Friston, and Klaas E. Stephan. 2011. “A Bayesian Foundation for Individual Learning under Uncertainty.” *Frontiers in Human Neuroscience* 5 (May): 39.
Mathys, Christoph D., Ekaterina I. Lomakina, Jean Daunizeau, Sandra Iglesias, Kay H. Brodersen, Karl J. Friston, and Klaas E. Stephan. 2014a. “Uncertainty in Perception and the Hierarchical Gaussian Filter.” *Frontiers in Human Neuroscience*. (November): 825.
McNally, Gavan P., Joshua P. Johansen, and Hugh T. Blair. 2011. “Placing Prediction into the Fear Circuit.” *Trends in Neurosciences* 34 (6): 283–92.
Mechaeil, Rasha, Paul Gard, Anne Jackson, and Jennifer Rusted. 2011. “Cognitive Enhancement Following Acute Losartan in Normotensive Young Adults.” *Psychopharmacology* 217 (1): 51–60.
Michael, Tanja, Jens Blechert, Noortje Vriends, Jürgen Margraf, and Frank H. Wilhelm. 2007. “Fear Conditioning in Panic Disorder: Enhanced Resistance to Extinction.”

Journal of Abnormal Psychology 116 (3): 612–17.
Milad, Mohammed R., Gregory J. Quirk, Roger K. Pitman, Scott P. Orr, Bruce Fischl, and Scott L. Rauch. 2007. “A Role for the Human Dorsal Anterior Cingulate Cortex in Fear Expression.” *Biological Psychiatry* 62 (10): 1191–94.
Mineka, Susan, and Katherine Oehlberg. 2008. “The Relevance of Recent Developments in Classical Conditioning to Understanding the Etiology and Maintenance of Anxiety Disorders.” *Acta Psychologica* 127 (3): 567–80.
Mineka, Susan, and Richard Zinbarg. 2006. “A Contemporary Learning Theory Perspective on the Etiology of Anxiety Disorders: It’s Not What You Thought It Was.” *The American Psychologist* 61 (1): 10–26.
Mobbs, Dean, Predrag Petrovic, Jennifer L. Marchant, Demis Hassabis, Nikolaus Weiskopf, Ben Seymour, Raymond J. Dolan, and Christopher D. Frith. 2007. “When Fear Is near: Threat Imminence Elicits Prefrontal-Periaqueductal Gray Shifts in Humans.” *Science* 317 (5841): 1079–83.
Mogg, Karin, Andrew Mathews, and John Weinman. 1987. “Memory Bias in Clinical Anxiety.” *Journal of Abnormal Psychology*. https://doi.org/10.1037//0021-843x.96.2.94.
Mogg, K., A. Mathews, and J. Weinman. 1987. “Memory Bias in Clinical Anxiety.” *Journal of Abnormal Psychology* 96 (2): 94–98.
Morriss, Jayne, Shannon Hoare, and Carien M. van Reekum. 2018. “It’s Time: A Commentary on Fear Extinction in the Human Brain Using fMRI.” *Neuroscience & Biobehavioural Reviews*. https://doi.org/10.1016/j.neubiorev.2018.06.025.
Morriss, Jayne, Francesco Saldarini, and Carien M. van Reekum. 2019. “The Role of Threat Level and Intolerance of Uncertainty in Extinction.” *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, May. https://doi.org/10.1016/j.ijpsycho.2019.05.013.
Moser, Jason S., Mark W. Becker, and Tim P. Moran. 2012. “Enhanced Attentional Capture in Trait Anxiety.” *Emotion* 12 (2): 213–16.
Myers, Karyn M., Kerry J. Ressler, and Michael Davis. 2006. “Different Mechanisms of Fear Extinction Dependent on Length of Time since Fear Acquisition.” *Learning & Memory* 13 (2): 216–23.
Mystkowski, Jayson L., Michelle G. Craske, Aileen M. Echiverri, and Jennifer S. Labus. 2006. “Mental Reinstatement of Context and Return of Fear in Spider-Fearful Participants.” *Behavior Therapy* 37 (1): 49–60.
Napier, Renée M., Michaela Macrae, and E. James Kehoe. 1992. “Rapid Reacquisition in Conditioning of the Rabbit’s Nictitating Membrane Response.” *Journal of Experimental Psychology: Animal Behavior Processes*. https://doi.org/10.1037//0097-7403.18.2.182.
Nassar, Matthew R., Katherine M. Rumsey, Robert C. Wilson, Kinjan Parikh, Benjamin Heasly, and Joshua I. Gold. 2012. “Rational Regulation of Learning Dynamics by Pupil-Linked Arousal Systems.” *Nature Neuroscience* 15 (7): 1040–46.
“National Adult Reading Test.” n.d. In *SpringerReference*.
Nelson, Hazel E. 1982. *National Adult Reading Test (NART): For the Assessment of Premorbid Intelligence in Patients with Dementia*.
Nelson, James Byron. 2002. “Context Specificity of Excitation and Inhibition in Ambiguous Stimuli.” *Learning and Motivation*. https://doi.org/10.1006/lmot.2001.1112.
O’Doherty, John, Peter Dayan, Johannes Schultz, Ralf Deichmann, Karl Friston, and Raymond J. Dolan. 2004. “Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning.” *Science* 304 (5669): 452–54.
Ohtawa, M., F. Takayama, K. Saitoh, T. Yoshinaga, and M. Nakashima. 1993. “Pharmacokinetics and Biochemical Efficacy after Single and Multiple Oral Administration of Losartan, an Orally Active Nonpeptide Angiotensin II Receptor Antagonist, in Humans.” *British Journal of Clinical Pharmacology* 35 (3): 290–97.
O’Reilly, Jill X., Urs Schüffelgen, Steven F. Cuell, Timothy E. J. Behrens, Rogier B.

Mars, and Matthew F. S. Rushworth. 2013. “Dissociable Effects of Surprise and Model Update in Parietal and Anterior Cingulate Cortex.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (38): E3660–69.

Otto, Michael W., Teresa M. Leyro, Kelly Christian, Christen M. Deveney, Hannah Reese, Mark H. Pollack, and Scott P. Orr. 2007. “Prediction of ‘Fear’ Acquisition in Healthy Control Participants in a De Novo Fear-Conditioning Paradigm.” *Behavior Modification*. https://doi.org/10.1177/0145445506295054.

Otto, Michael W., David F. Tolin, Naomi M. Simon, Godfrey D. Pearlson, Shawnee Basden, Suzanne A. Meunier, Stefan G. Hofmann, Katherine Eisenmenger, John H. Krystal, and Mark H. Pollack. 2010. “Efficacy of D-Cycloserine for Enhancing Response to Cognitive-Behavior Therapy for Panic Disorder.” *Biological Psychiatry* 67 (4): 365–70.

Park, Soyoung Q., Thorsten Kahnt, Anne Beck, Michael X. Cohen, Raymond J. Dolan, Jana Wrase, and Andreas Heinz. 2010. “Prefrontal Cortex Fails to Learn from Reward Prediction Errors in Alcohol Dependence.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 30 (22): 7749–53.

Pavlatou, Maria G., George Mastorakos, Ioannis Lekakis, Stavros Liatis, Georgia Vamvakou, Emmanuel Zoumakis, Ioannis Papassotiriou, Andreas D. Rabavilas, Nikolaos Katsilambros, and George P. Chrousos. 2008. “Chronic Administration of an Angiotensin II Receptor Antagonist Resets the Hypothalamic-Pituitary-Adrenal (HPA) Axis and Improves the Affect of Patients with Diabetes Mellitus Type 2: Preliminary Results.” *Stress* 11 (1): 62–72.

Pavlov, P. Ivan. 2010. “Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex.” *Annals of Neurosciences* 17 (3): 136–41.

Payzan-LeNestour, Elise, and Peter Bossaerts. 2011. “Risk, Unexpected Uncertainty, and Estimation Uncertainty: Bayesian Learning in Unstable Settings.” *PLoS Computational Biology* 7 (1): e1001048.

Pearce, J. M., and M. E. Bouton. 2001. “Theories of Associative Learning in Animals.” *Annual Review of Psychology* 52: 111–39.

Pearce, John M., and Geoffrey Hall. 1980. “A Model for Pavlovian Learning: Variations in the Effectiveness of Conditioned but Not of Unconditioned Stimuli.” *Psychological Review*. https://doi.org/10.1037//0033-295x.87.6.532.

Penny, William D., Karl J. Friston, John T. Ashburner, Stefan J. Kiebel, and Thomas E. Nichols. 2011. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier.

Phelps, Elizabeth A., Mauricio R. Delgado, Katherine I. Nearing, and Joseph E. LeDoux. 2004a. “Extinction Learning in Humans.” *Neuron,* (6): 897–905.

Pineles, Suzanne L., Dawne S. Vogt, and Scott P. Orr. 2009. “Personality and Fear Responses during Conditioning: Beyond Extraversion.” *Personality and Individual Differences*. https://doi.org/10.1016/j.paid.2008.09.003.

Piray, Payam, Verena Ly, Karin Roelofs, Roshan Cools, and Ivan Toni. 2019. “Emotionally Aversive Cues Suppress Neural Systems Underlying Optimal Learning in Socially Anxious Individuals.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 39 (8): 1445–56.

Pruim, Raimon H. R., Maarten Mennes, Daan van Rooij, Alberto Llera, Jan K. Buitelaar, and Christian F. Beckmann. 2015. “ICA-AROMA: A Robust ICA-Based Strategy for Removing Motion Artifacts from fMRI Data.” *NeuroImage* 112 (May): 267–77.

“PTSD Statistics | PTSD United.” n.d. Accessed February 1, 2019. http://www.ptsdunited.org/ptsd-statistics-2/.

Pulcu, Erdem, and Michael Browning. 2017a. “Using Computational Psychiatry to Rule Out the Hidden Causes of Depression.” *JAMA Psychiatry* 74 (8): 777–78.

Pulcu, Erdem, and Michael Browning. 2017b. “Correction: Affective Bias as a Rational Response to the Statistics of Rewards and Punishments.” *eLife* 6 (October).

https://doi.org/10.7554/eLife.32902.
Pulcu, Erdem, Lorika Shkreli, Carolina Guzman Holst, Marcella Woud, Michelle Craske, Michael Browning, and Andrea Reinecke. 2018. “The Effects of the Angiotensin II Receptor Antagonist Losartan on Appetitive versus Aversive Learning.” https://doi.org/10.1101/472050.
Pulcu, Erdem, Lorika Shkreli, Carolina Guzman Holst, Marcella L. Woud, Michelle G. Craske, Michael Browning, and Andrea Reinecke. 2019. “The Effects of the Angiotensin II Receptor Antagonist Losartan on Appetitive Versus Aversive Learning: A Randomized Controlled Trial.” *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2019.04.010.
Quirk, Gregory J. 2002. “Memory for Extinction of Conditioned Fear Is Long-Lasting and Persists Following Spontaneous Recovery.” *Learning & Memory* 9 (6): 402–7.
Rachman, S., and A. Arntz. 1991. “The Overprediction and Underprediction of Pain.” *Clinical Psychology Review*. https://doi.org/10.1016/0272-7358(91)90112-8.
Raes, An K., Rudi De Raedt, Bruno Verschuere, and Jan De Houwer. 2009. “Failure to Loose Fear: The Impact of Cognitive Load and Trait Anxiety on Extinction.” *Behaviour Research and Therapy* 47 (12): 1096–1101.
Redish, A. David, A. David Redish, Steve Jensen, Adam Johnson, and Zeb Kurth-Nelson. 2007. “Reconciling Reinforcement Learning Models with Behavioral Extinction and Renewal: Implications for Addiction, Relapse, and Problem Gambling.” *Psychological Review*. https://doi.org/10.1037/0033-295x.114.3.784.
Redish, A. David, Steve Jensen, Adam Johnson, and Zeb Kurth-Nelson. 2007. “Reconciling Reinforcement Learning Models with Behavioral Extinction and Renewal: Implications for Addiction, Relapse, and Problem Gambling.” *Psychological Review* 114 (3): 784–805.
Rescorla, Robert A., and C. Donald Heth. 1975. “Reinstatement of Fear to an Extinguished Conditioned Stimulus.” *Journal of Experimental Psychology: Animal Behavior Processes*. https://doi.org/10.1037//0097-7403.1.1.88.
Ressler, Kerry J., Barbara O. Rothbaum, Libby Tannenbaum, Page Anderson, Ken Graap, Elana Zimand, Larry Hodges, and Michael Davis. 2004. “Cognitive Enhancers as Adjuncts to Psychotherapy: Use of D-Cycloserine in Phobic Individuals to Facilitate Extinction of Fear.” *Archives of General Psychiatry* 61 (11): 1136–44.
Robbins, Steven J. 1990. “Mechanisms Underlying Spontaneous Recovery in Autoshaping.” *Journal of Experimental Psychology: Animal Behavior Processes*. https://doi.org/10.1037/0097-7403.16.3.235.
Roesch, Matthew R., Guillem R. Esber, Jian Li, Nathaniel D. Daw, and Geoffrey Schoenbaum. 2012. “Surprise! Neural Correlates of Pearce-Hall and Rescorla-Wagner Coexist within the Brain.” *The European Journal of Neuroscience* 35 (7): 1190–1200.
Rosas, Juan M., Travis P. Todd, and Mark E. Bouton. 2013. “Context Change and Associative Learning.” *Wiley Interdisciplinary Reviews: Cognitive Science*. https://doi.org/10.1002/wcs.1225.
Rosen, Jeffrey B., and Melanie P. Donley. 2006. “Animal Studies of Amygdala Function in Fear and Uncertainty: Relevance to Human Research.” *Biological Psychology* 73 (1): 49–60.
Roy, Mathieu, Daphna Shohamy, Nathaniel Daw, Marieke Jepma, G. Elliott Wimmer, and Tor D. Wager. 2014. “Representation of Aversive Prediction Errors in the Human Periaqueductal Gray.” *Nature Neuroscience* 17 (11): 1607–12.
Saavedra, Juan M., Enrique Sánchez-Lemus, and Julius Benicky. 2011. “Blockade of Brain Angiotensin II AT1 Receptors Ameliorates Stress, Anxiety, Brain Inflammation and Ischemia: Therapeutic Implications.” *Psychoneuroendocrinology* 36 (1): 1–18.
Schiller, Daniela, Christopher K. Cain, Nina G. Curley, Jennifer S. Schwartz, Sarah A.

Stern, Joseph E. Ledoux, and Elizabeth A. Phelps. 2008. “Evidence for Recovery of Fear Following Immediate Extinction in Rats and Humans.” *Learning & Memory* 15 (6): 394–402.
Schiller, Daniela, Ifat Levy, Yael Niv, Joseph E. LeDoux, and Elizabeth A. Phelps. 2008. “From Fear to Safety and Back: Reversal of Fear in the Human Brain.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 28 (45): 11517–25.
Schlagenhauf, Florian, Quentin J. M. Huys, Lorenz Deserno, Michael A. Rapp, Anne Beck, Hans-Joachim Heinze, Ray Dolan, and Andreas Heinz. 2014. “Striatal Dysfunction during Reversal Learning in Unmedicated Schizophrenia Patients.” *NeuroImage* 89 (April): 171–80.
Schuck, Nicolas W., Ming Bo Cai, Robert C. Wilson, and Yael Niv. 2016. “Human Orbitofrontal Cortex Represents a Cognitive Map of State Space.” *Neuron* 91 (6): 1402–12.
Schuck, Nicolas W., Robert C. Wilson, and Yael Niv. n.d. “A State Representation for Reinforcement Learning and Decision-Making in the Orbitofrontal Cortex.” https://doi.org/10.1101/210591.
Schultz, Wolfram, Kerstin Preuschoff, Colin Camerer, Ming Hsu, Christopher D. Fiorillo, Philippe N. Tobler, and Peter Bossaerts. 2008. “Explicit Neural Signals Reflecting Reward Uncertainty.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363 (1511): 3801–11.
Sehlmeyer, C., U. Dannlowski, S. Schöning, H. Kugel, M. Pyka, B. Pfleiderer, P. Zwitserlood, et al. 2011. “Neural Correlates of Trait Anxiety in Fear Extinction.” *Psychological Medicine* 41 (4): 789–98.
Sehlmeyer, Christina, Sonja Schöning, Pienie Zwitserlood, Bettina Pfleiderer, Tilo Kircher, Volker Arolt, and Carsten Konrad. 2009. “Human Fear Conditioning and Extinction in Neuroimaging: A Systematic Review.” *PLoS ONE*. https://doi.org/10.1371/journal.pone.0005865.
Seligman, Martin E. P. 2016. “Phobias and Preparedness - Republished Article.” *Behavior Therapy* 47 (5): 577–84.
Sevenster, Dieuwke, Renée M. Visser, and Rudi D’Hooge. 2018. “A Translational Perspective on Neural Circuits of Fear Extinction: Current Promises and Challenges.” *Neurobiology of Learning and Memory* 155 (November): 113–26.
Seymour, Ben, Nathaniel Daw, Peter Dayan, Tania Singer, and Ray Dolan. 2007. “Differential Encoding of Losses and Gains in the Human Striatum.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 27 (18): 4826–31.
Seymour, Ben, John P. O’Doherty, Martin Koltzenburg, Katja Wiech, Richard Frackowiak, Karl Friston, and Raymond Dolan. 2005. “Opponent Appetitive-Aversive Neural Processes Underlie Predictive Learning of Pain Relief.” *Nature Neuroscience* 8 (9): 1234–40.
Shannon, C. E. 1948. “A Mathematical Theory of Communication.” *Bell System Technical Journal*. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.
Shekhar, Anantha. 2014. “Angiotensin Type 1 Receptor Antagonists—A Novel Approach to Augmenting Posttraumatic Stress Disorder and Phobia Therapies?” *Biological Psychiatry* 75 (11): 836–37.
Smits, Jasper A. J., David Rosenfield, Michael W. Otto, Mark B. Powers, Stefan G. Hofmann, Michael J. Telch, Mark H. Pollack, and Candyce D. Tart. 2013. “D-Cycloserine Enhancement of Fear Extinction Is Specific to Successful Exposure Sessions: Evidence from the Treatment of Height Phobia.” *Biological Psychiatry* 73 (11): 1054–58.
Soch, Joram, and Carsten Allefeld. 2018. “MACS – a New SPM Toolbox for Model Assessment, Comparison and Selection.” *Journal of Neuroscience Methods*. https://doi.org/10.1016/j.jneumeth.2018.05.017.
Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde.

2002. “Bayesian Measures of Model Complexity and Fit.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. https://doi.org/10.1111/1467-9868.00353.
Squire, Larry R. 1992. “Declarative and Nondeclarative Memory: Multiple Brain Systems Supporting Learning and Memory.” *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.1992.4.3.232.
Staib, Matthias, Giuseppe Castegnetti, and Dominik R. Bach. 2015. “Optimising a Model-Based Approach to Inferring Fear Learning from Skin Conductance Responses.” *Journal of Neuroscience Methods* 255 (November): 131–38.
Staples-Bradley, Lindsay K., Michael Treanor, and Michelle G. Craske. 2018. “Discrimination between Safe and Unsafe Stimuli Mediates the Relationship between Trait Anxiety and Return of Fear.” *Cognition and Emotion*. https://doi.org/10.1080/02699931.2016.1265485.
Starosta, Sarah, Metin Uengoer, Isabelle Bartetzko, Sara Lucke, Onur Güntürkün, and Maik C. Stüttgen. 2016. “Context Specificity of Both Acquisition and Extinction of a Pavlovian Conditioned Response.” *Learning & Memory* 23 (11): 639–43.
Sussman, Tamara J., Akos Szekely, Greg Hajcak, and Aprajita Mohanty. 2016. “It’s All in the Anticipation: How Perception of Threat Is Enhanced in Anxiety.” *Emotion*. https://doi.org/10.1037/emo0000098.
Swartzentruber, Dale, and Mark E. Bouton. 1992. “Context Sensitivity of Conditioned Suppression Following Preexposure to the Conditioned Stimulus.” *Animal Learning & Behavior*. https://doi.org/10.3758/bf03200406.
Takezawa, Kunio. 2012. “A Revision of *AIC* for Normal Error Models.” *Open Journal of Statistics*. https://doi.org/10.4236/ojs.2012.23038.
Taylor, Steven, and Brian J. Cox. 1998. “An Expanded Anxiety Sensitivity Index.” *Journal of Anxiety Disorders* 12 (5): 463–83.
Thaler, R. H., A. Tversky, D. Kahneman, and A. Schwartz. 1997. “The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test.” *The Quarterly Journal of Economics*. https://doi.org/10.1162/003355397555226.
Torrents-Rodas, David, Miquel A. Fullana, Albert Bonillo, Xavier Caseras, Oscar Andión, and Rafael Torrubia. 2013. “No Effect of Trait Anxiety on Differential Fear Conditioning or Fear Generalization.” *Biological Psychology*. https://doi.org/10.1016/j.biopsycho.2012.10.006.
Tzourio-Mazoyer, N., B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. 2002. “Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain.” *NeuroImage*. https://doi.org/10.1006/nimg.2001.0978.
Tzovara, Athina, Christoph W. Korn, and Dominik R. Bach. 2018. “Human Pavlovian Fear Conditioning Conforms to Probabilistic Learning.” *PLoS Computational Biology* 14 (8): e1006243.
Venkatraman, Vinod, and Scott A. Huettel. 2012. “Strategic Control in Decision-Making under Uncertainty.” *The European Journal of Neuroscience* 35 (7): 1075–82.
Von Neumann, John, and Oskar Morgenstern. 1947. *Theory of Games and Economic Behavior*.
Wager, Tor D., Lauren Y. Atlas, Martin A. Lindquist, Mathieu Roy, Choong-Wan Woo, and Ethan Kross. 2013. “An fMRI-Based Neurologic Signature of Physical Pain.” *The New England Journal of Medicine* 368 (15): 1388–97.
Wagner, A. R., L. S. Siegel, and G. G. Fein. 1967. “Extinction of Conditioned Fear as a Function of Percentage of Reinforcement.” *Journal of Comparative and Physiological Psychology* 63 (1): 160–64.
Watson, John B., and Rosalie Rayner. 1920. “Conditioned Emotional Reactions.” *Journal of Experimental Psychology* 3 (1): 1–14.
Weinberg, Anna, and Greg Hajcak. 2011. “Electrocortical Evidence for Vigilance-Avoidance in Generalized Anxiety Disorder.” *Psychophysiology* 48 (6): 842–51.
Zhang, Meng, Chao Ma, Yanyan Luo, Ji Li, Qingwei Li, Yijun Liu, Cody Ding, and Jiang

Qiu. 2016. “Neural Basis of Uncertain Cue Processing in Trait Anxiety.” *Scientific Reports* 6 (February): 21298.
Zhang, Suyi, Hiroaki Mano, Gowrishankar Ganesh, Trevor Robbins, and Ben Seymour. 2016. “Dissociable Learning Processes Underlie Human Pain Conditioning.” *Current Biology: CB* 26 (1): 52–58.

Appendix I

Comparison of model-based and model-free learning rates

A Rescorla-Wagner model with separate learning rate for shock and for no-shock was fitted to the data and the results were compared to model-free learning rates acquired for the same conditions.

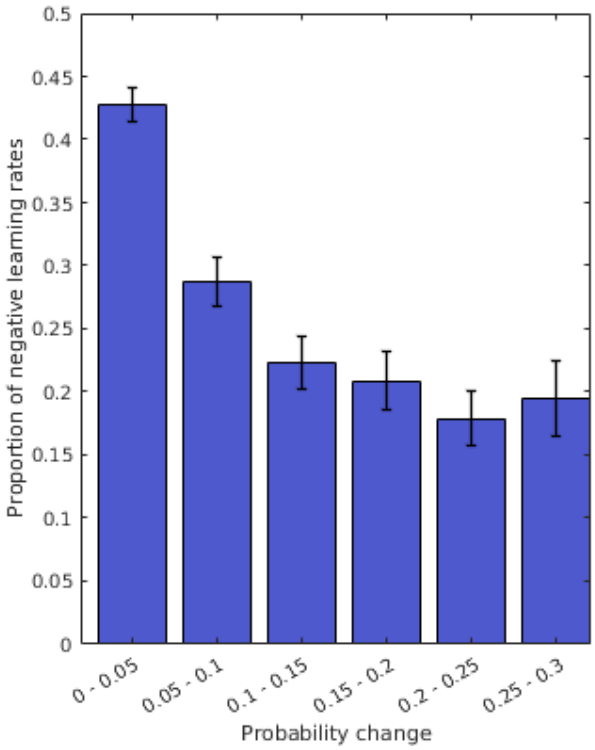
The model-free learning rates with different cutoff boundaries (0 to 1; -1 to 1; and no boundaries) were correlated with the model-based learning rates. Across all visits and cues the learning rates with cutoff boundaries between 0 and 1 correlated highest with the estimated learning rates (mean r=.504), followed by model-free learning rates cut off at –1 and 1 (mean r=0.419). The learning rates without any restriction included large negative outliers. There was not a relationship between unrestricted model-free learning rates and model-fitted learning rates (mean r=-.048). Our main analysis focuses on learning rates in the reversal cue. Here the correlation between model-based and 0 - 1 restricted model-free learning rates reached .065.

To further compare MB and MF learning rates I used the data set from Study 3 to see if they result in the same set of statistical results. As reported in the main analysis in Chapter 5, in the model-based learning rates there was a significant interaction between visit and group. Subjecting the model-free learning rates to the same analysis, identical set of statistical results was obtained (reported in the main text). These results show that the model-free learning rates are a valid metric comparable to the more commonly used model-based learning rates.

To fully understand the difference between model-based and model-free learning rates, we examined the cases where model-free learning rates were negative. The Rescorla-Wagner model defines the range of the learning rate parameter to lie between 0 and 1. In our trial-by-trial treatment, however, we found that 31% of all learning rates were outside of that range (33% in stable-high (harmful); 32% in stable-low (safe) and 29% in the reversal cue). This explains why fitted learning rates do not perfectly explain the behaviour of participants.

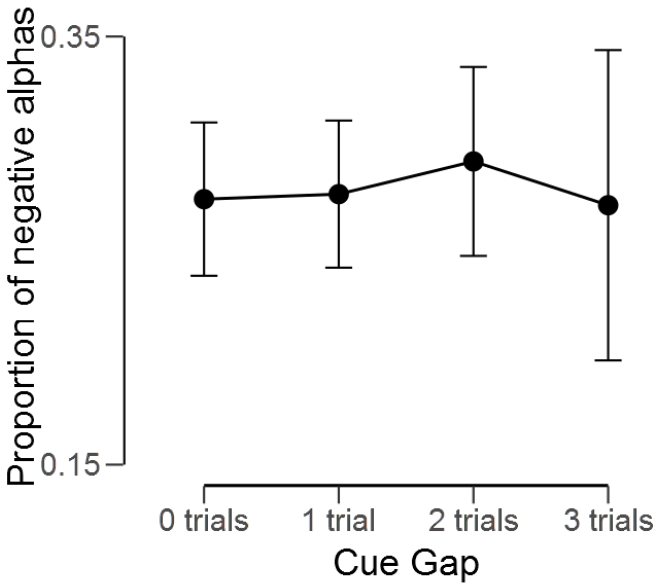
On some occasions, participants receive a shock but decrease their probability on the next trial. This leads to a paradoxical situation in which they seemingly unlearn, i.e. the learning rate is negative. This also means that one of the assumptions of the Rescorla-Wagner model is violated. Negative learning rates are likely a consequence of two mechanisms: a) reporting noise; and b) memory decay. Participants are not shown the exact probability they are submitting, so in ratings that near the scale

extrema or where participant wishes to stay at the same level there is a higher chance that they will report lower probability despite receiving a shock, simply due to imprecision of the reporting scale. To check whether this is the case we calculated the distribution of negative learning rates as a function of update magnitude $|P_{t+1} - P_t|$. The proportion of negative learning rate trials was highest in updates of magnitude between 0 and 0.05 (See Appendix I Figure 1). This indicates that at least some of the negative learning rates can be attributed to reporting noise.



Appendix I Figure 1: *Proportion of negative model-free learning rates as a function of update magnitude.*

Next, we tested whether memory also plays a role in negative learning rates. A specific cue can be presented on two subsequent trials but there may also be a gap of up to four trials, as the two other cues are presented. If memory decay causes negative learning rates, they should be observed more frequently with increasing time from when the cue was last presented. Repeated measures ANOVA found no significant effect of memory on the negative learning rates, $F(3,111)=.0142$, *n. s.*



Importantly, the model-based learning rates also should not be considered the golden standard. While their recoverability from artificial data is nearly 100% in most flavours of the RW model participants often behave in way that cannot be captured by the RW model, i.e. have negative learning rate even when reporting noise is accounted for. In these cases the α will be estimated as 0 - this was the case in several of our participants, especially for no-shock learning. One reason for such scenario is a case where the participant does not receive a shock on several subsequent trials and increases their probability of shock on the next trial reasoning that every no-shock increases the chance of shock. In fact, this was reported by several participants in post-experiment interviews. This bias can be introduced to the RW model as a sort of “hazard rate”.

The comparison of model-free and model-based learning rates found partial correspondence between the two. The comparison also identified strengths and weaknesses of the two methods. While model-based learning rates are less susceptible to noise such as imperfection in reporting, model-free learning rates provide a direct measure of participants’ trial-by-trial behaviour that contains learning itself as well as additional factors, such as attention, memory and other biases which need to be considered during the analysis.

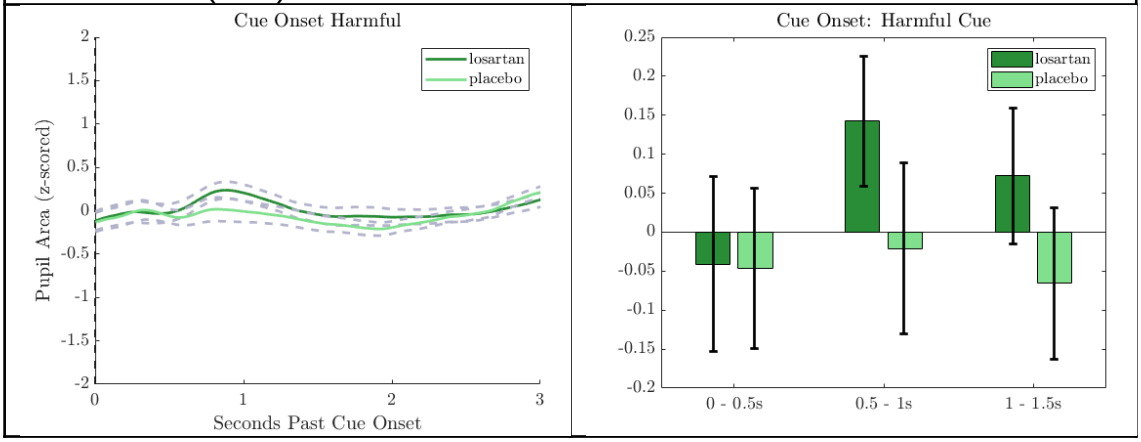
Appendix II

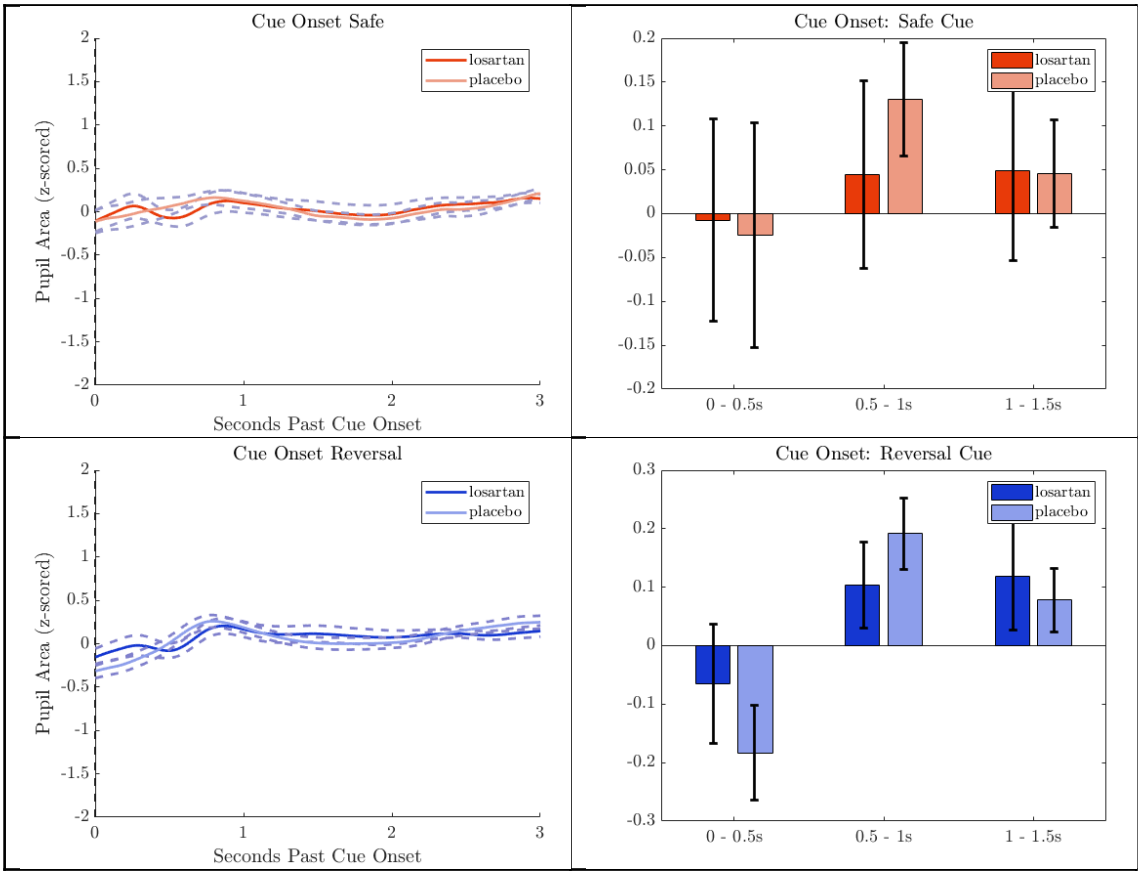
Average pupil response

Grand averages for pupil response are presented in Appendix IV. Pupil dilation following the presentation of the cue was not modulated by the drug group in any of the three cues. At the time of response (rating submission), the losartan group had a stronger initial pupil response, F(1,37) = 5.652, p<.05. This was identified by a significant interaction between group and time bin past response. Applying Bonferroni correction to a set of post-hoc tests, this interaction effect remained significant in the safe, F(1,37)=6.399, p<.0167, and harmful, F(1,37)=9.317, p<.01, cues, but not in the reversal cue, F(1,37) = 3.852, *n.s*.

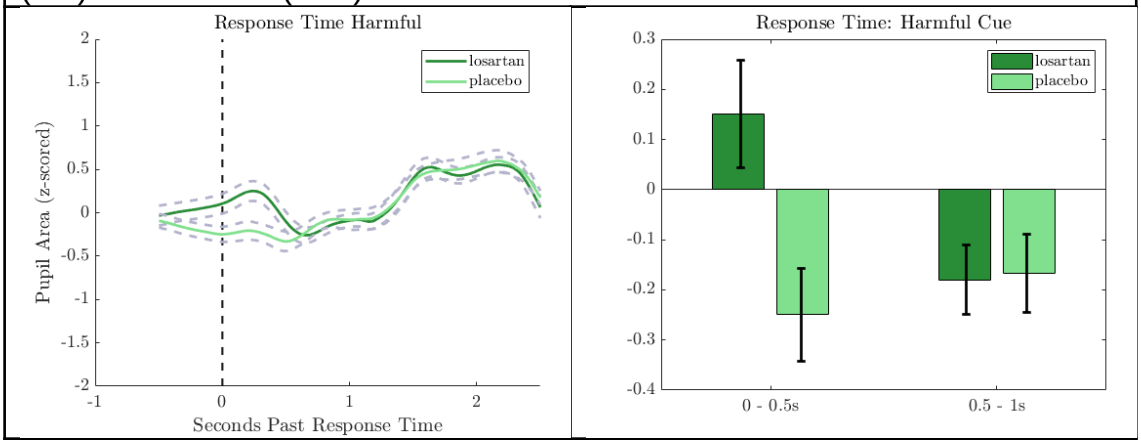
Outcome type (shock / shock omission) was added as a within-subject factor to the analysis of outcome-locked pupil response. There was no main effect or interaction of the group in any of the three cues.

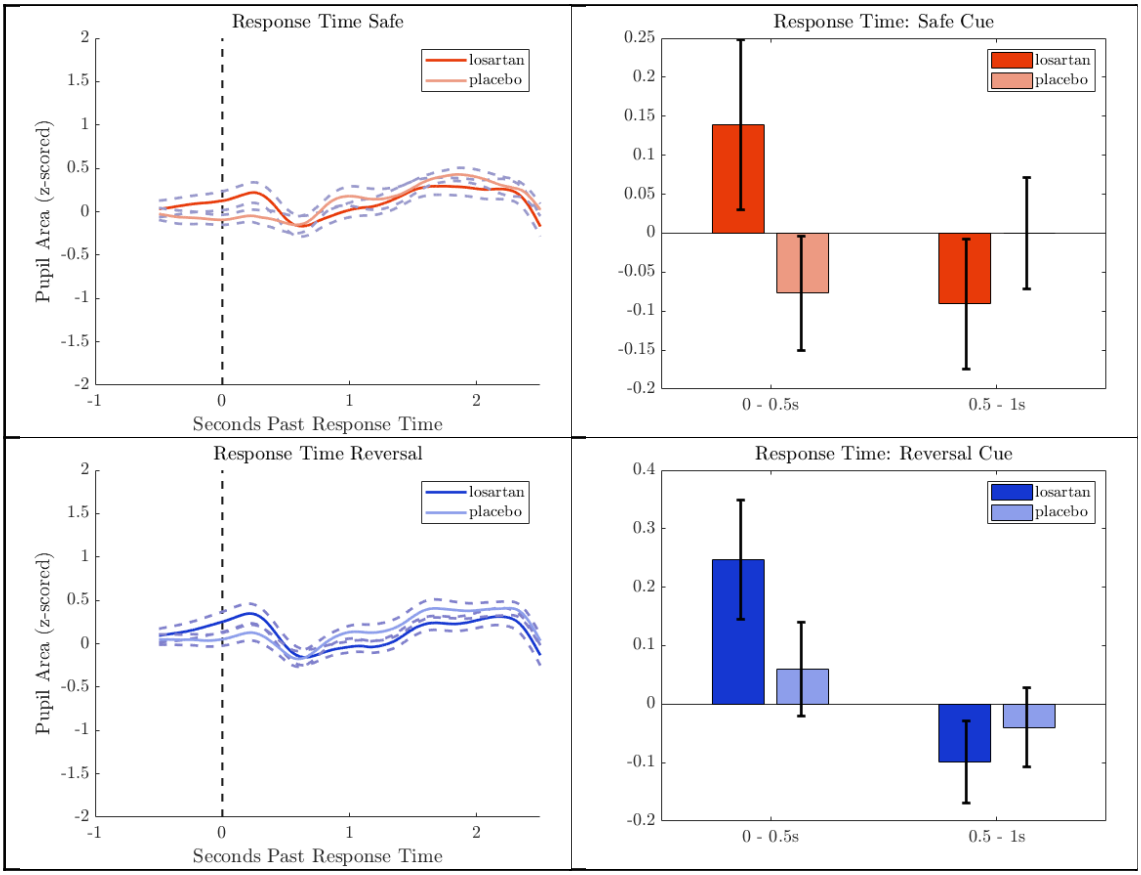
Supplementary Table 2: Pupil response to onset of harmful (green), safe (red) and reversal (blue) cues



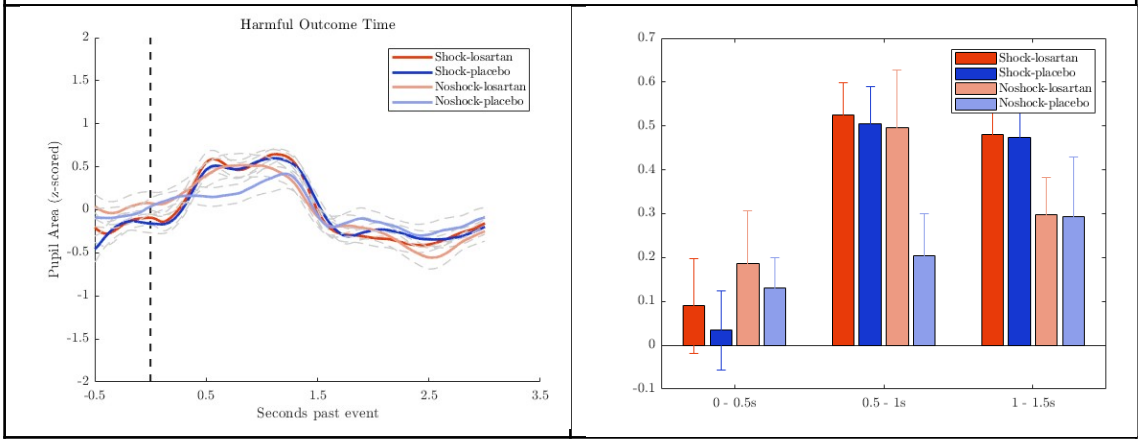


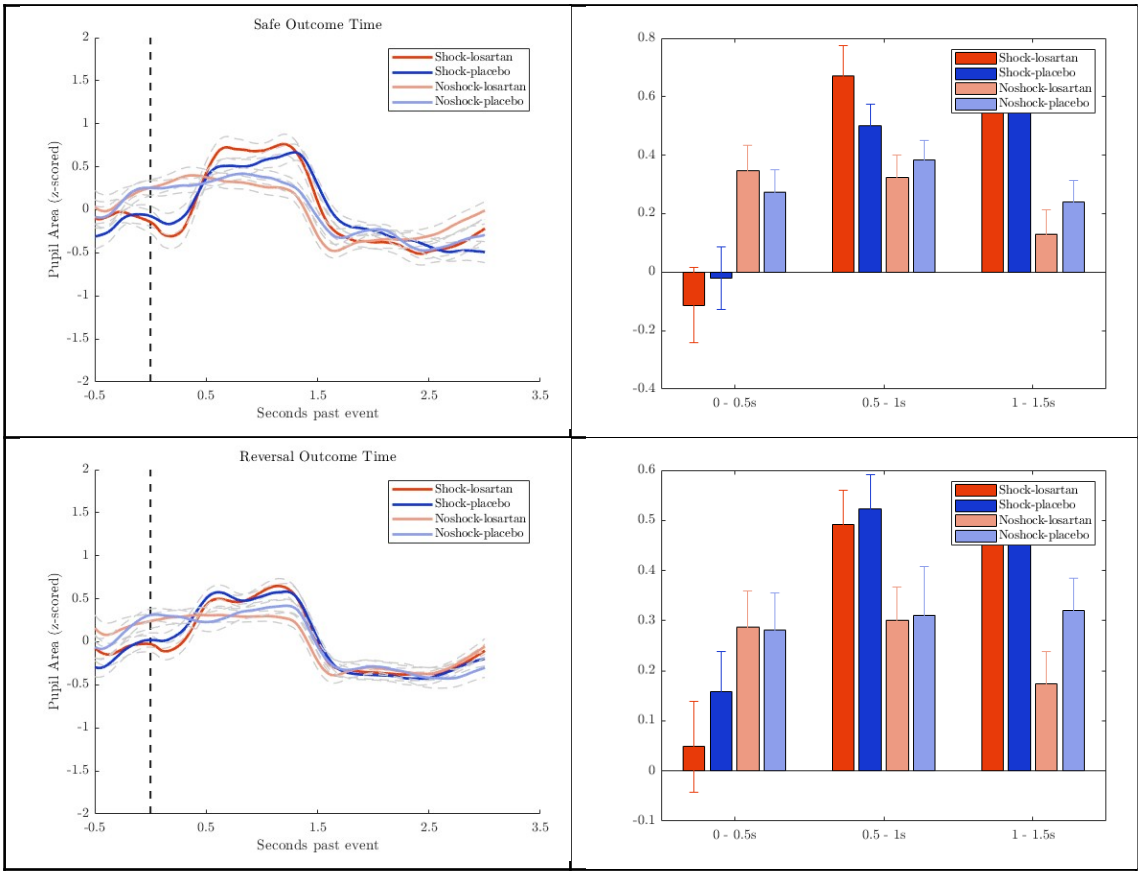
Supplementary Table 3: Pupil response to rating submission for (green), safe (red) and reversal (blue) cues





Supplementary Table 4: Pupil response to outcome




Safe Outcome Time
Shock-losartan
Shock-placebo
Noshock-losartan
Noshock-placebo
Pupil Area (z-scored)
Seconds past event
0 - 0.5s
0.5 - 1s
1 - 1.5s
Reversal Outcome Time