INFORMATION LEARNING BIAS IN DEPRESSION

Wanjun Lin

St. Cross College

Department of Psychiatry

University of Oxford

This dissertation is submitted for the degree of Doctor of Philosophy

Trinity Term 2021

“In times of change, learners inherit the earth, while the learned find themselves beautifully equipped to deal with a world that no longer exists.”

— Eric Hoffer

Acknowledgements

At last, I'm here overviewing my DPhil years. It had not been an easy experience for me, but it still a wonderful one. I have learnt a lot, seen a lot, grown a lot. This thesis would not be possible with the help of many people.

First and foremost, I would like to thank my primary supervisor – Dr. Michael Browning. He has been an outstanding mentor. His continuous and invaluable guidance and feedback throughout the whole process of my DPhil studies has laid the foundation of the thesis. He also encouraged me to learn more and think deeper about psychiatric disorders. Discussions with him are always really inspiring. He has opened my view and deepened my understanding of computational psychiatry. He pointed out and tolerates my perfectionism which I hadn't realized at all before.

I would also like to thank my secondary supervisor, Dr. Laurence Hunt. He is usually the tie-breaker when we disagree. His expertise, patience, and wisdom have guided me throughout my DPhil. The retreats and journal clubs with him and others have taught me a great deal about how good science should be like.

I am also very grateful to my third supervisor, Dr. Erdem Pulcu, who has been really helpful since the first day we met. He has put in a substantial amount of work on the studies included in this thesis. He has taught me how to do time management when things pile up, which I'm still not very good at but I'll keep practising his advice.

Finally, I thank my family, especially my husband, Yunzhe. He has been the beacon of my life. He helps me find directions when I'm lost, encourages me to move forward when I'm afraid. I won't even start this DPhil without him. I also should thank my parents, my younger sister and brother, who have been very supportive throughout my entire life.

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

ABSTRACT

Major depressive disorder is a mood and emotional disorder characterized by persistent low mood and lack of interest. Both reward and punishment learning deficits have been associated with depression. In this thesis, we aim to examine whether biases in reward and punishment learning adaptations to the volatility level of the environments could also be associated with depression disorders. We developed a task in which we independently manipulated the volatility of win and loss outcomes and systematically examined the underlying mechanisms of adaptive learning in depression using computational modelling, fMRI, pharmacological manipulation and pupillometry recordings. We first showed that currently depressed patients had lower loss learning rate volatility adaptations, estimated using a volatility adaptation model, compared to remitted depressed patients and healthy controls. This might be associated with reduced norepinephrine responsibility to loss outcomes, evidenced by reduced loss volatility modulation of pupil responses to loss outcomes in the currently depressed group compared to the other two groups. Using a similar volatility adaptation model, we then showed that loss learning rate adaptation estimates negatively correlated with individual differences in overall negativity levels (depression and anxiety), with might be associated with decreased segregation of win and loss volatility signals in a distributed brain network of outcome sensitive regions, especially in the ACC. Finally, we showed that manipulating norepinephrine level using an antidepressant, reboxetine, a norepinephrine reuptake inhibitor (NRIs), could increase loss learning rate adaptations to volatility. This was also associated with a reduced volatility effect on the pupil responses to loss outcomes, particularly when wins were volatile. Together, the results have provided some consistent evidence that link depression with inadequate learning adaptation to the volatility of loss outcomes which might be caused by maladaptive norepinephrine responsibility to loss outcomes.

1 INTRODUCTION

1.1 Depressive disorders

1.1.1 Definitions and diagnostic criteria

Clinical Depression, or major depressive disorder, is a common mental disorder, characterised by depressed mood and loss of interest or pleasure. People with clinical depression might also experience symptoms including (1) changes in weight or appetite, (2) psychomotor agitation or retardation, (3) persistent fatigue or loss of energy, (4) persistent feelings of worthlessness or excessive guilt, (5) having difficulties in concentrating, thinking or making decisions, (6) recurrent thoughts of death or suicidal ideation or suicidal attempts. According to the DSM-5, an individual can be diagnosed with major depressive disorder if he/she has been experienced five or more of the symptoms listed (one of them has to be depressed mood or loss interest/pleasure) consistently during a 2-week period, and the symptoms have caused significant distress or impairment in social, occupational, or other important areas of functioning (Association 2013).

1.1.2 Prevalence and socio-economic impact

Estimates of the lifetime prevalence of depressive disorder range from 2 to 21%, with the highest rates in some of the European countries, including France, Netherlands, and Belgium (Gutiérrez-Rojas et al. 2020).

Depressive disorders impose a substantial economic burden on society. And yet the impact of depressive disorder might still be increasing. For example, in the United States, the total economic costs of major depressive disorder (MDD) were estimated to be 210.5 billion US dollars in 2010, an increase of 21.5% from 2005. About 50% of the estimated economic costs were attributed to the workplace, including absenteeism (not showing up to work) and presenteeism (showing up to work but not fully functioning, due to an illness or medical condition) caused by functional impairment (Greenberg, Fournier, et al. 2015). In fact, depressive disorder has been the third leading cause of all-age years lived with disability (YLDs), determined by the number of years disabled weighted by level of disability, globally since 2007. From 2007 to 2017, the global number of all-age YLDs attributed to depressive

disorder further increased by 14.3% (Collaborators 2018). Treatment costs also constitute a large proportion of the total cost. For example, in the UK, the direct treatment costs for adult depression in 2000 was estimated to be 370 million pounds (Thomas and Morris 2003). Moreover, depression is one of the most potent risk factors for suicide (Bachmann 2018).

In summary, depression is one of the most significant public health concerns affecting not only individual lives but society as a whole. Therefore, the WHO considers depression a global crisis and calls for action including research to increase understanding of the illness and to develop more effective treatments (WFMH 2012).

1.1.3 Treatment efficacy

There are many existing effective treatments for major depressive disorder, including antidepressant medications, psychological and social interventions such as cognitive behavioural therapy (CBT), and brain stimulation therapies such as transcranial direct-current stimulation (tDCS). However, although the treatments for depression are generally statistically more effective than placebo (Khan et al. 2012; Moffa et al. 2020; Cipriani et al. 2018), their overall efficacy is moderate. A meta-analysis showed the average response rate, defined as 50% reduction of depressive symptoms, for treatment with antidepressants range from 42% to 53%, which represents a modest improvement over the response to placebo which is 35% (Cipriani et al. 2018). When adopting a more stringent, but clinically more meaningful, measure of treatment outcome, remission rate (defined as a full symptomatic recovery from depression), the efficacy is lower still. For example, in a large-scale pharmacological study, the remission rate following treatment with citalopram was 33%, based on a commonly used psychometric evaluation called the Quick Inventory of Depressive Symptomatology–Self-report (QIDS-SR) (Trivedi et al. 2006).

The low efficacy of antidepressants might be related to heterogeneity of both the antidepressants and the patients. Antidepressants vary in their mechanism of action. For example, antidepressants can be further classified into different categories based on their target neurotransmitters including selective serotonin reuptake inhibitors (SSRIs) such as citalopram, serotonin and norepinephrine reuptake inhibitors (SNRIs) such as venlafaxine, norepinephrine reuptake inhibitors (NRIs) such as reboxetine, norepinephrine and dopamine reuptake inhibitors (NDRIs) such as bupropion. It is believed that a patient might respond to one treatment but not to others. One commonly adopted strategy for this issue in the clinic is to use sequential trials of different treatments. However, in the substantial STAR*D trial (Rush et al. 2006) which examined this approach, although switching treatments to different types of antidepressants and cognitive psychotherapy was associated with some improvement in

response rate, there were still about 30% of the participants who did not remit after all the trials of treatments. From a practical perspective, each treatment must be given for 4-6 weeks before it is deemed ineffective (NICE 2009) making this sequential trial-and-error strategy time-consuming, and the inevitable delay in initiating effect therapy is likely to be very frustrating.

The heterogeneity within people who shared the diagnosis of clinical depression complicates the story even more. Diagnostic criteria for depressive disorders incorporate various symptoms from subjective feelings to physical phenomena. A study that used latent factor analysis based on various symptoms of depressive patients identified many data-driven subtypes (Van Loo et al. 2012), suggesting significant heterogeneity within depressive disorders. This heterogeneity could suggest distinct underlying mechanisms, that might also be an underlying factor of the heterogeneity in response to antidepressants.

In summary, although there are many effective existing treatments for depression, their overall efficacy is moderate. Resolution of the treatment efficacy problem might be achieved by developing novel interventions with greater efficacy, or promoting treatment to patient mapping, that is, selecting the correct treatment for the correct patient. Both of these approaches are likely to be facilitated by a better understanding of the mechanistic processes that lead to depression. In this thesis I will take a computational approach to investigate one potential mechanistic process relevant to depression—learning about affectively valanced information under uncertainty.

In the next sections I will first review the literature which has used traditional behavioural measures to study how people with depression learn from reward and punishment information and how people with depression adapt their learning in volatile environments. Then I will briefly introduce the key computational techniques that have been adopted in more recent studies of this subject and critically appraise the recent efforts to characterize depression in computational terms. Following that, I will summarize the neurobiological processes that support these learning and adaptation processes and the evidence linking abnormalities in these learning systems to depression. Finally, I will review cognitive and pharmacological interventions that have been suggested to alter the learning process. While the focus of the thesis is on preclinical, mechanistic questions, the ultimate goal is to improve our understanding of what can be a devastating illness, and thus to facilitate the development of sorely needed novel treatment options.

1.2 Reinforcement learning in depression

1.2.1 Reward and punishment processing in depression

Many emotional and cognitive symptoms of depression can be intuitively linked to aspects of reinforcement learning and decision-making. For example, anhedonia, the inability to feel pleasure, or apathy, diminished motivation for physical, cognitive or emotional activity, can be linked to deficits in reward processing (Husain and Roiser 2018). Learned helplessness, which is a common animal model of depression (Vollmayr and Henn 2001), a state exhibited by a subject after enduring repeated aversive stimuli beyond their control, can be associated with problems in learning punishment contingencies (Seligman 1972). Indecisiveness, the inability to make decisions, could be related to value learning, credit assignment or problems in using values to guide behaviour (van Randenborgh, de Jong - Meyer, and Hüffmeier 2010). Social withdrawal can be associated with impairments in processing of social reward and social pain (Porcelli et al. 2019). However, many of these intuitions bear further scientific examinations. Here I will review the literature which has mainly focused on reward and punishment learning.

1.2.1.1 Association of depression with abnormal reward processing

Several lines of research have suggested aberrant reward rather than punishment learning in depression. People with depression seem to have difficulty in learning reward contingencies. In a study adopting a probabilistic selection task (Admon et al. 2017), participants first passively learned many stimuli with varying reward and punishment contingencies and then chose between the stimuli in pairs. The results showed that depressed patients had lower accuracy for reward learning compared to non-depressed control participants, while no group difference was found for punishment learning. depressed participants showed impaired reward learning (lower accuracy in choosing the rewarding option), but not penalty learning, relative to control participants.

Depressed people also seem to make fewer behavioural adjustments in response to rewarding incentives, with studies suggesting a valence specificity of this effect: that the effects of reward incentives are diminished in depression, while punishing incentives remain effective in cognitive tasks such as signal detection (Henriques, Glowacki, and Davidson 1994) and memory (Henriques and Davidson 2000). Recent studies also support these findings. People with higher depressive symptoms (Pizzagalli, Jahn, and O'Shea 2005) and remitted major depression (Whitton et al. 2016) failed to show response bias in a signal detection task with a differential reward. However, the punishment effect was not studied in these two studies.

Moreover, people with depression showed less emotional responses to rewarding events. Comparing participants' emotional ratings before and after reward, omission of reward, punishment and successful avoidance of punishment feedback, a study showed that currently depressed participants had fewer emotion changes particularly for reward feedback compared to never depressed and previously depressed participants (McFarland and Klein 2009).

1.2.1.2 Association of depression with abnormal punishment processing

However, there is also evidence suggesting punishment learning deficits are associated with depression. Early studies of depression showed a detrimental effect of negative feedback on subsequent performance for depressed patients as well as remitted depressed patients, i.e., depressive patients were more likely to make more mistakes after a failure (Elliott et al. 1996; Elliott et al. 1997; Steffens et al. 2001). Holmes et al. (Holmes and Pizzagalli 2007) gave participants pseudo feedback about whether their performance was above average (positive) or below average (negative) at the beginning of each block of various cognitive tasks and investigated whether such feedback changed how participants adjust their performance after error and correct trials. They found that people with high depressive symptoms showed impaired post-error performance adjustment following negative feedback, but not positive feedback, and that the level of performance adjustment following negative feedback inversely correlated with depression scores, measured by Beck Depression Inventory-II, meaning that people who report higher depressive symptoms were worse at adjusting their behaviour.

While there is still dispute about whether depression is more of a reward learning deficit or punishment learning deficit, an increasing number of studies have shown co-occurring reward and punishment learning deficits in depressed patients and people with high risk for developing depression (Pizzagalli et al. 2009; Gotlib et al. 2010; McCabe et al. 2012; Ubl et al. 2015; Mukherjee et al. 2020).

However, most of these studies compared rewards and punishments indirectly, for example, in different tasks or different conditions. Whereas in real life, an event is rarely purely rewarding or punishing. Making decisions in such an environment requires people to weigh the rewards of particular actions or states against the punishments associated with the same states/actions. Therefore, it is crucial to study how people with depression combine co-occurring rewards and punishments in their decision-making process. In the following text, I will review some studies which required participants to consider both reward and punishment in their choices.

1.2.1.3 Weighing rewards against punishments

Several lines of research have investigated how people with depression might combine reward and punishment information to guide their behaviour.

One line of research studied effort-reward trade-offs in depression. A task paradigm frequently used for this line of research is the Effort-Expenditure for Rewards Task. In this task, participants chose between doing "hard-work" and "easy-work" in each trial to be eligible to obtain different magnitudes of monetary reward with some probability (Treadway et al. 2009). The "hard-work" required higher amounts of speeded manual button presses, but led to a higher reward magnitude. The "easy-work" required fewer button presses but led to lower reward magnitude. The reward probabilities were provided trial by trial, so learning of contingency was not required. They found that patients with major depression were less likely to choose the "hard-work" (high magnitude) regardless of the probabilities of obtaining the reward compared to healthy controls (Treadway et al. 2012). Interestingly, they also found that the willingness to spend effort for reward correlated with levels of anhedonia, measured by the Snaith-Hamilton Pleasure Scale (SHAPS) (Snaith et al. 1995), in a subclinical cohort (Treadway et al. 2009). People with higher anhedonia were less likely to choose the "hard-work", particularly when the probability of getting the reward was moderate (50%) or low (12%) (Treadway et al. 2009). Another study (Hershenberg et al. 2016) came to the same conclusion by adopting a different paradigm called the progressive ratio task (PRT), that examined the maximum effort a participant was willing to exert for a particular reward. The maximum effort that depressed patients were willing to make for a given reward was lower than healthy controls.

However, the exact cognitive deficits underlying the changes in effort-reward trade-offs in depression remain unanswered. Depressed patents might be less sensitive to reward so less willing to spend effort. Alternatively, they might perceive effort as more punishing. Nevertheless, it is probably worth noting that effort is a specific kind of punishment and that it is often a prerequisite for reward but does not directly affect the expected value of the outcomes. Other forms of punishment such as monetary loss has also frequently been used to study punishment–reward trade-off, for example in the Iowa gambling task (IGT). The IGT, designed to simulate learning and decision-making in real-life situations, requires participants to combine reward and punishment information (Bechara et al. 1994). In this task, participants choose from four decks of cards which vary in their amount of monetary reward and punishment. Two of the four decks (decks A and B) are disadvantageous. Because, although they generate a bigger reward at first, an even bigger punishment follows, resulting in a net loss. The other two decks (decks C and D) start with a smaller reward followed by even smaller penalty resulting in a net gain.

Similar to the results using the Effort-Expenditure for Rewards Task, people with depression also performed worse in the IGT task than healthy controls, i.e., they failed to switch to the C/D decks (better decks), when A/B decks (disadvantageous decks) became punishing (Must et al. 2006; Moniz et al. 2016; Hegedűs et al. 2018). Hegedűs et al. also found that people with depression showed inferior performance in a different version of the IGT, in which the decks

started with penalty and were followed by reward (Bechara, Tranel, and Damasio 2000), although the results were not statistically significant (Hegedűs et al. 2018). Together, this evidence suggests that depressed patients are less optimal when making behavioural adjustments after previously rewarding options become more punishing or previously punishing options become more rewarding. This was further supported by a study by Cella and colleagues (Cella, Dymond, and Cooper 2010). Using a contingency-shifting variant of the IGT, in which the reward and punishment contingencies of different decks of cards were systematically altered, they showed that the ability to adapt to the changing contingencies was impaired in people with depression. MDD patients chose more from previously good now bad decks and less from previously bad and now good decks compared to healthy controls (Cella, Dymond, and Cooper 2010).

In summary, people with depression differ from non-depressed control participants in the way they combine reward and punishment information for decisions. However, it is not clear why or how these differences arise. It could be hyposensitivity to reward, hypersensitivity to punishment or both. Alternatively, it could also result from biases in using reward and punishment information for decision making. In addition, the findings with different versions of IGT collectively suggest that depressed patients might also adapt differently when reward and punishment contingencies change. In the paradigm used in this thesis, I examine learning biases in using reward and punishment information during a learning and decision-making task in the context of changing contingencies.

In the following section, we will examine how depressed people learn in changing (i.e., volatile) environments in more detail.

1.2.2 Altered learning under uncertainty in depression

Studies have shown depressed patients have elevated level of intolerance of uncertainty, measured by the intolerance of uncertainty scale (IUS) (McEvoy and Mahoney 2012; Yook et al. 2010; Dugas, Schwartz, and Francis 2004), i.e., people with depression are more likely to report that an uncertain situation is distressing. Uncertainty is also important during reinforcement learning, where it has been classified into expected uncertainty and unexpected uncertainty (Yu and Dayan 2005). Expected uncertainty arises from the known unreliability of the predictive relationship between actions/stimuli and outcomes. Unexpected uncertainty comes from gross changes of the predictive relationships between actions/stimuli and outcomes change over time, i.e., a changing or volatile environment. While expected uncertainty makes learning more challenging it leaves the relationships of action/stimuli-outcome associations unchanged and thus learners should not adjust their estimates of such relationships in response

to it. In contrast unexpected uncertainty signals a change in these relationships and should boost learning so that the new contingency can be estimated. Learning to differentiate expected and unexpected uncertainty and adjust learning accordingly is crucial for efficient learning in many real world situations (Pulcu and Browning 2019). As shown in the studies with the contingency-shifting variants of the IGT introduces in the previous section, people with depression seem to have an impaired ability to adapt to changing environments (Cella, Dymond, and Cooper 2010), suggesting a difficulty in adapting appropriately to the volatility of associations. However, the IGT mixed the reward and punishment contingency changes with changes in reward and punishment magnitudes, making it less clear to answer how people with depression adapt their learning when facing the changes. Next, we focus on reinforcement learning tasks with changing contingencies, such as reversal learning tasks.

In reversal learning tasks, participants are first asked to learn to discriminate the values of two items, one more rewarding/less punishing than the other. After this initial phase of learning, the contingencies are reversed so that the original less rewarding/more punishing option become the more optimal one to choose. Participants need to adapt to this change and switch their choice behaviour. Studies adopting this paradigm have shown that people with depression make more errors after reversal compared to healthy controls (Dickstein et al. 2010; Hall, Milne, and Macqueen 2014). However, the studies did not further clarify the exact nature of the errors that were made.

To further examine what king of contingency changes contribute to the errors made in depression, in follow up study (Levy-Gigi and Kéri 2015), the researchers adopted a similar reversal paradigm, but instead of choosing between two options they asked participant to respond to each cue by either choosing it (if they thought it was rewarding) or skipping it (if they thought it was punishing). They found people with high depressive symptoms showed impaired learning specifically when the cue changed from negative to positive, not from positive to negative, and the level of impairment correlated with the severity of the depressive symptoms. This result suggested that depressed patients increased their estimate of reward value more slowly than healthy controls. Studies with a different version of the reversal task supported this finding (Robinson et al. 2012; Timmer et al. 2017). In this version of the reversal task, participants were asked to predict explicitly whether they thought a cue was associated with reward or punishment for each trial and then observed the outcome (i.e., this task investigated Pavlovian rather than instrumental learning). Using this task, Robinson et al. found people with MDD made more incorrect predictions when a previous punishing cue changed to rewarding, but not for the other way around (Robinson et al. 2012). Timmer et al. showed the same results in Parkinson's disease patients with symptoms of depression, suggesting it might be transdiagnostic (Timmer et al. 2017). Together, these studies seemed to suggest that people with depression were less sensitive to unexpected rewards and made more errors when reward

contingencies increased. However, in these reversal learning tasks, reward and punishment contingencies always changed at the same time making it difficult disentangle the effect of reward and punishment contingencies changes on choice accuracy.

In summary, impaired reward and punishment processing and difficulties adapting to changing environments have been described in people with depression. However, the specific cognitive processes altered in depression that produce these effects have not been clearly delineated. A deeper investigation into the cognitive and neural mechanisms that give rise to these behavioural differences is essential for improving understanding of the illness, and may go some way to explaining the inconsistences in the literature to date. To achieve this, we need to move beyond simple summary statistics of choice accuracy and reaction times, and adopt new approaches. One such approach, a branch of computational psychiatry, uses computational modelling of behaviour to identify latent cognitive processes. Recent studies adopting a computational approach have begun to help further our understanding of depressive disorder. In the next section, I will briefly review some of the computational approaches used to study learning, and then critically appraise the studies that uses computational models to investigate reward and punishment reinforcement learning in depression.

1.3 Computational modelling of depression

1.3.1 The computational approach

Theory-driven computational psychiatry, uses generative models to explain and quantify the computational mechanisms of mental illness. This approach arose from computational neuroscience. The core concept of computational neuroscience is to consider the brain as a computation unit. Therefore, to understand the brain and its function, we need to know what and how it computes. Alterations in the normative computation might lead to malfunction. Theory-driven computational psychiatry aims to identify these abnormalities and the underlying mechanisms using computational modelling (Browning et al. 2020; Huys, Maia, and Frank 2016).

The computational modelling approach has a number of potential advantages for psychiatry research. First, it could enable researchers to further investigate underlying neural mechanisms. For example, model-based fMRI analysis facilitated our understanding of the neural representation of value and prediction errors (O'Doherty et al. 2003; O'Doherty et al. 2001). Second, it has the potential to develop mechanism-based measurements (Browning et al. 2020). Theoretically, model-derived parameters may have higher validity and reliability compared to traditional behaviour measurements such as accuracy and reaction times, because models

reduce the dimensionality of the data by 'projecting' it to a few key relevant spaces (Huys, Maia, and Frank 2016). Thirdly, it may improve treatment by providing better measurements or develop new interventions based on the mechanisms identified.

1.3.2 Computational models of associative learning

While many computational models have been developed to understand human behaviours in various reinforcement learning tasks in recent years (Rescorla 1972; Sutton and Barto 1990; Daw, Niv, and Dayan 2005), I will focus on the models of associative learning of relevance to this thesis. In the following text, I will briefly introduce some key computational terms for associative learning.

In associative learning, an agent acquires knowledge about environmental contingencies between stimuli/actions and outcomes. In associative learning computational models, exemplified by the Rescorla-Wagner (RW) model (Rescorla 1972), when an agent receives new information, it updates its belief about the stimuli/actions-outcome association in proportion to its prediction error, δ, which is the difference between the expected outcome and the actual outcome received. The degree to which the new information is incorporated in the new belief is determined by learning rate, α. For example, in the RW model, the expected value (V) of a stimulus for next time point (t+1) is calculated by adding the product of learning rate(α) and prediction error (δ) to the old belief, see **Equation** 1. The prediction error (δ) is the discrepancy between the actual outcome (R) and the expected value previously assumed, see Equation 2.

$$\boldsymbol{V_{t+1} = V_t + \alpha * \delta_t} \qquad \textbf{Equation 1}$$

$$\boldsymbol{\delta_t = R_t - V_t} \qquad \textbf{Equation 2}$$

Actions/stimuli are then selected based on the expected values. One commonly used action selection function is the softmax decision rule as shown in **Equation 3**. Here, $P_{i(t)}$ the probability of choosing action i on trial t is determined by the relative value of action i over other actions (n potential actions in total) and the 'inverse temperature' parameter β, the extent to which values are used to generate actions. When β is higher, the effect that value estimates of the actions would have on choice probabilities is larger, whereas with a lower β, choices become noisier.

$$\boldsymbol{P_{i(t)} = \frac{exp\,(\beta * V_{i(t)})}{\sum_{j=1}^{n} exp\,(\beta * V_{j(t)})}} \qquad \textbf{Equation 3}$$

Despite its simplicity, the RW model is able to account for a wide range of phenomena observed for both reward and punishment learning (Siegel and Allan 1996). Many studies have tried to characterise depression using these formal computational terms in recent years. In the following text, I will review some key findings in this line of work, focusing on reward and punishment learning rates in particular.

1.3.3 Reinforcement learning models of depression

Adopting the RW model in a simple probabilistic selection task, Chase et al. (Chase et al. 2010) showed that MDD patients had lower learning rates to both positive and negative feedback than healthy controls. They also found a negative correlation between the learning rates and anhedonia, measured by the Snaith–Hamilton Pleasure Scale (SHAPS). A recent study also found a negative correlation of reward learning rates and depression scores in a more general population (Safra, Chevallier, and Palminteri 2019). This is consistent with findings reviewed in section 1.2.1.1 that suggested that people with depression are less sensitive to reward feedback and have difficulty in learning reward contingencies. This is also supported by an analysis that used a variant version of the RW model, in which reward sensitivity was estimated (Huys et al. 2013). In this meta-analysis, that included 6 datasets from a probabilistic reward task with an asymmetric reinforcement schedule to assess the effect of rewarding incentives on perceptual choice in depression, Huys et al. added a reward sensitivity term (ρ) for prediction errors (see **Equation 4**) in the RW model. The modelling results showed that people with depression had lower reward sensitivity, i.e., lower ρ. Moreover, they found that reward sensitivity ρ, rather than learning rates, negatively correlated with anhedonia (Huys et al. 2013). Note that the reward sensitivity term in this study also contributed to learning similarly to the way learning rates did, but the reward sensitivity term only further discounted how the reward for the current trial contributes to the learning not reward history. The results suggested that the reduced learning rate shown in depression might be caused by the reduced reward sensitivity. However, this bears further examination. Nevertheless, collectively, these studies suggested that depression is associated with reduced reward learning.

$$\boldsymbol{\delta_t = \rho * R_t - V_t} \qquad \textbf{Equation 4}$$

However, in contrast, higher learning rates in depression compared to healthy controls were shown in tasks where reward or punishment contingencies change. In a recent study, Aylward et al. (Aylward et al. 2019) adopted the RW model in a task that required participants to choose between four competing slot machines with fluctuating reward and punishment outcomes. They found that patients with mood disorders (including depression and anxiety) were quicker to

update their behaviour in response to negative outcomes, i.e., increased punishment learning rates, which correlated with anxiety (The State-Trait Anxiety Inventory (STAI)) and depression (Beck Depression Inventory (BDI)) symptoms. A similar effect was shown in a study using a change point detection task, that high anxious individuals had higher loss-switch rates, which corresponded to higher learning rates (Huang, Thompson, and Paulus 2017). High learning rates in depression could result from poor memory of previous values. Dombrovski et al. examined this hypothesis by introducing a memory term (M) in the RW model of a probabilistic reversal learning task. As shown in **Equation 5**, the memory parameter (M) discounts previous expected values when updating the value beliefs. Model-based analyses revealed that depressed patients with previous suicide attempts had lower memory estimates (Dombrovski et al. 2010). Note that if we rearranged the original RW model, as shown in **Equation 6**, the M in the memory model functions very similarly to 1- α. So, this result is consistent with higher learning rates shown in depression. In a more recent study, Rupprechter et al. adopted similar memory models in a probabilistic pavlovian reward learning task which required participants to learn the reward probabilities quickly (using only 4 observations). They found that patients with MDD had lower memory estimates compared to healthy volunteers, suggesting people with depression discounted previous values to a higher degree and learned more from most recent outcomes (Rupprechter et al. 2018), indicating a higher learning rate. Together, these findings suggest that people with depression have higher learning rates than healthy controls particularly when the task requires fast learning, such as in volatile environments.

$$\boldsymbol{V_{t+1} = M * V_t + R_t} \qquad \textbf{Equation 5}$$

$$\boldsymbol{V_{t+1} = (1 - \alpha) * V_t + \alpha * R_t} \qquad \textbf{Equation 6}$$

In summary, a number of studies have described aberrant learning rates in depressed patients compared to healthy controls. However, the literature is inconsistent with evidence of both increased and decreased learning rates. As touched on above, one factor that should influence learning rates is the volatility of an association, with higher learning rates being used for volatile relative to stable associations. This raises the possibility that the variable learning rates reported for depressed patients may arise due to difficulties in estimating the volatility of an association. Assessing this possibility requires a task in which the volatility of associations is manipulated. We provide an overview of tasks that have taken this approach in the next section.

1.3.4 Adaptive learning under uncertainty

As described above, learning should proceed more rapidly when associations are changeable (volatile). Within a RL framework this translates to using a higher LR in volatile than stable environments. This form of learning rate adjustment has been demonstrated in non-clinical human participants in reward (Behrens et al. 2007) and punishment learning (Browning et al. 2015), as well as social learning (Behrens et al. 2008). Building on this work, in a recent study, Pulcu and Browning developed a novel paradigm, in which stimuli were associated with both rewarding and punishing outcomes (winning and losing money). The volatility of the associations between the stimuli and the rewarding and punishing outcomes was independently manipulated across task blocks (see Figure 1.1). This work indicated that non-clinical human participants could adjust their reward and punishment learning rates flexibly to accommodate the volatility levels of each type of outcome (Pulcu and Browning 2017). This work examines how people use both reward and punishment information in their learning in environments with different volatility, which fits well with our studies goals. This task will form the basis of much of the work I will present in this thesis.

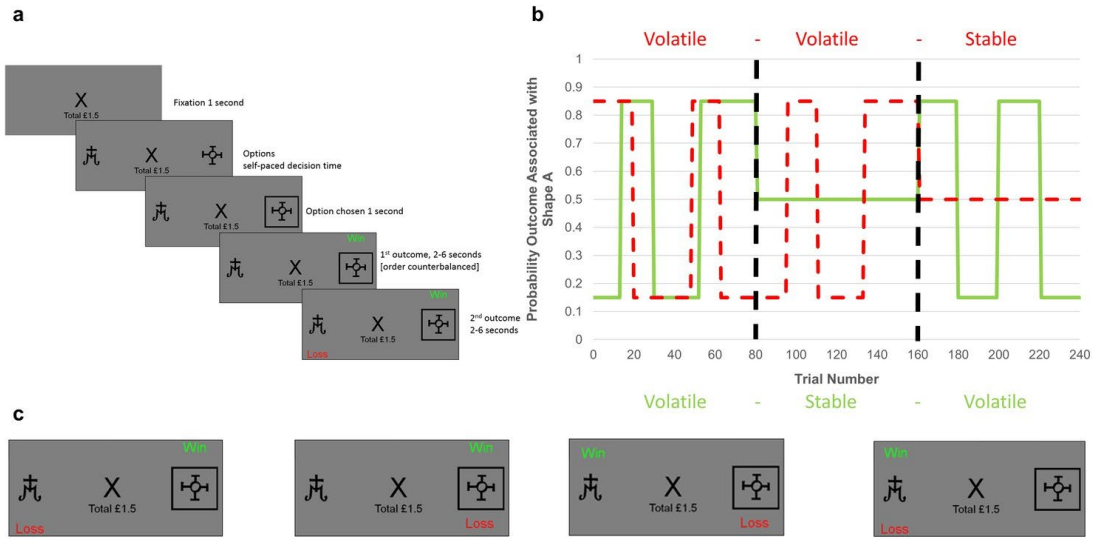

Figure 1.1 Task structure

A figure from (Pulcu and Browning 2017). (A) Timeline of one trial from the learning task used in Pulcu's study (Pulcu and Browning 2017). Participants were presented with two shapes (referred to as shape 'A' and 'B') and had to choose one. On each trial, a 'win' outcome (resulting in a win of 15 p) would be revealed on the top of one of the shapes, and a 'loss' outcome (resulting in a loss of 15 p) would be revealed on the bottom of one of the shapes. The two outcomes were independent, i.e., knowledge of the location of the win provided no information about the location of the loss. Participants had to learn where the wins and losses

were more likely to be found through trials and errors and use the information to guide their choice in order to maximise their monetary earnings. (B) Overall task structure. The blocks differed in how volatile (changeable) the outcome probabilities were. As illustrated here, in the first block both win and loss outcomes were volatile, in the second wins were stable and losses were volatile, in the third block wins were volatile and losses were stable. (C) The four potential outcomes from a trial.

However, to the best of our knowledge, no studies have been conducted to investigate how people with depression adjust learning to accommodate environmental volatility levels. Nevertheless, there have been some studies examining learning adjustments to environmental volatility in anxiety, which would have important implications for depression, given the high comorbidity of anxiety and depression disorders. One such study (Browning et al. 2015) recruited participants with a range of anxiety levels, based on the Spielberger State-Trait Anxiety Inventory (STAI) trait subscale, to perform a two-armed bandit learning task in which they had to learn the punishment contingency in order to escape electrical shocks. RW Learning rates were estimated separately for the stable block, in which one of the shapes had 75% of shock delivery and the other one 25%, and the volatile block, in which the probability of getting shocked with a given shape switched between 80% and 20% every 20 trials. People with higher anxiety levels adjusted their learning rates less between stable and volatile blocks. This is consistent with a recent study that showed people with high social anxiety, measured by the Social Interaction Anxiety Scale (SIAS), failed to lower learning rates when punishing contingencies stayed stable over a long period compared to low social anxiety participants (Beltzer et al. 2019). A possible cause of the difficulty in adjusting learning rates particularly during punishment learning could be that stress induced by aversive contexts might render individuals less sensitive to changes in the environment. Results consistent with this were found using a Bayesian model, which showed that a group exposed to acute stress, induced by performing a public speaking task, underestimated environment uncertainty and was less precise in estimating the environment volatility (Hein et al. 2019). Taken together, these studies suggested a deficit in estimating and adapting learning to different environmental volatility levels in trait or state anxiety.

In the previous section, we discussed evidence that people with depression used different learning rates than healthy controls in reward and punishment learning particularly when reward or punishment contingencies changed. In this section, we reviewed recent evidence suggesting that impaired learning rates adaptation to the volatility of the environment is associated with anxiety and stress.

However, how people with depression adapt to different volatility levels of reward and punishment contingencies has not been studied. In this thesis, we will investigate how people

with depression learn from rewarding and punishing information, how they adjust to the volatility levels of rewards and punishments, and the underlying neural mechanisms supporting this process using fMRI. In the following section, we will review model-based neuroimaging work that has sought to identify changes in functional brain activity associated with depression with a specific focus on the neural systems implicated in reinforcement learning.

1.4 The neuroanatomy and neurochemistry of disrupted reinforcement learning and volatility adaptation in depression

A distributed network of brain regions has been related to reinforcement learning. Key regions include striatum, anterior insula, OFC and dorsal ACC. In this section, we will briefly review the cognitive function of these brain regions and how they have been linked to the pathophysiology of depression.

1.4.1 Striatum

Both the ventral and dorsal striatum are essential for reward learning (Cox and Witten 2019; Robinson et al. 2012; Taswell et al. 2018). Model-based fMRI results support the striatum's role in coding reward prediction errors (RPEs) in humans (Pessiglione et al. 2006; Garrison, Erdeniz, and Done 2013). The RPE signals in the ventral striatum have been found to be reduced in patients with MDD compared to healthy controls (Rupprechter, Stankevicius, et al. 2020; Rothkirch et al. 2017). More specifically, the RPE signals in the ventral striatum were negatively correlated with the severity of the depression symptoms, such as anhedonia (measured by the Snaith–Hamilton Pleasure Scale) (Rothkirch et al. 2017; Greenberg, Chase, et al. 2015), impersonal orientation (the extent to which a person believes that attaining desired outcomes is beyond his or her control and that achievement is largely a matter of luck or fate, measured by the General Causality Orientations Scale)(Romaniuk et al. 2019), and overall depression severity (measured by the Quick Inventory of Depressive Symptomatology) (Rupprechter, Romaniuk, et al. 2020). RPEs in the dorsal striatum, including caudate and putamen, were also found to be negatively associated with QIDS scores (Rupprechter, Romaniuk, et al. 2020). A similar correlation was found for putamen activity during reward anticipation (Romaniuk et al. 2019). Moreover, the reward related activity in the putamen increased more in depressed patients who had better symptoms improvement over time (Verfaillie et al. 2016). Together these studies suggest that reward related striatal activity is reduced in depression, a finding that is theoretically similar to the reduced reward learning rates in depression reviewed in the previous section 1.3.3. However, such differences between

depressed and non-depressed participants in RPE signals in ventral striatum was not shown in a study with a lottery task (Rutledge et al. 2017), in which the probabilities of the outcomes were given to participants every trial so learning was not required in this task. The results suggested that the abnormal RPE signals in striatum in depression might be specific to reward learning. However, it is not clear whether the abnormal RPE signals in striatum might be associated with adaptive learning difficulties in depression in volatile environments.

1.4.2 Anterior Insula

The anterior insula has been suggested to be a critical brain region in processing aversive information (Simmons et al. 2004; Palminteri et al. 2012). Compared to healthy controls, the neural activity of the anterior insula to punishing events was increased in MDD patients (Remijnse et al. 2009; Johnston et al. 2015), as well as people with a family history of depression (McCabe et al. 2012). Activation in the anterior insula in response to loss events positively correlated with depression severity (Engelmann, Berns, and Dunlop 2017; Johnston et al. 2015) and anxiety level (Johnston et al. 2015) in patients with MDD.

Model-based fMRI analysis revealed the anterior insula's role in encoding punishment prediction errors (PPE) (Pessiglione et al. 2006; Fazeli and Büchel 2018). However, although studies have shown PPE signals coded in the anterior insula in both MDD and healthy controls, they failed to find significant differences between the two groups using simple probabilistic learning tasks (Rothkirch et al. 2017; Kumar et al. 2018). Insula activity may also be modulated by uncertainty. The anterior insula response to aversive pictures were larger after an uncertain cue than a certain cue (Sarinopoulos et al. 2010). This suggests that, examining the PPE signals in insula during more volatile environments may be particularly informative in terms of disrupted cognition in depression. However, to the best of our knowledge, no such studies have been completed.

1.4.3 Orbitofrontal cortex (OFC)

The OFC encodes stimuli-outcome associations, specifically moment-to-moment expected values of a stimuli based on current internal states (Rudebeck and Murray 2014). Specific functions have been ascribed to subregions of the OFC. A meta-analysis of fMRI data has suggested a medial-lateral valence gradient in OFC, with the more medial portion responding most often to rewards, and the more lateral portion responding most often to punishers (Kringelbach and Rolls 2004). The punishers that OFC responded to include punishment/loss as well as the omission of reward, which suggests that the OFC is a key region for reversal

learning (Tait and Brown 2007). However, others argued that the dichotomy between the medial and lateral portion of OFC was not well supported (Rushworth et al. 2011).

Various lines of evidence have suggested dysfunction of the OFC in depression. MDD patients had reduced OFC volume compared to healthy controls (Bremner et al. 2002). Altered functional connectivity of the OFC was also been described in depression. Specifically, functional connectivity between medial OFC with memory systems such as medial temporal lobe was found to be reduced while functional connectivity between lateral OFC with praecuneus, the angular gyrus, and the temporal visual cortex increased in depression (Cheng et al. 2016). In a recent study, Xie et al. showed that greater sensitivity to non-reward (No-Win) in lateral OFC and decreased sensitivity to the reward magnitudes in medial OFC were associated with higher depressive symptoms (Xie et al. 2020). Reduced activity in the lateral OFC was shown in MDD patients during reversal learning (Verfaillie et al. 2016). Fewer studies have investigated the function of the OFC using computational modelling in depression. One such study found that reward PE-related responses in medial OFC was decreased in the MDD compared to the healthy control group and the RPE signals in medial OFC correlatedly with anhedonia (Rothkirch et al. 2017). However, in this study, expected values and prediction errors were highly correlated making a simple interpretation of these findings challenging.

1.4.4 Dorsal anterior cingulate cortex (dACC)

An important function of the dACC is error monitoring (Holroyd and Coles 2002). In RL, dACC has been shown to encode reward prediction errors (Kennerley, Behrens, and Wallis 2011) or surprise (unsigned reward prediction errors) (Hayden et al. 2011). However, the dACC also tracks reward prediction errors over longer time scales compared to the ventral striatum, which encodes instantaneous prediction errors (Wittmann et al. 2016). Consistent with this, it has been shown that dACC activity increases as a function of environmental volatility which may then facilitate adaptive learning in changing environments (Behrens et al. 2007; Meder et al. 2017).

Separate lines of work have described reduced volume (van Tol et al. 2010), altered functional connectivity (Davey et al. 2012) and low glutamate concentrations (Auer et al. 2000) in the dACC in depression. MDD patients failed to increase dACC activity in response to negative feedback in a gambling task, compared to healthy controls (Steele, Kumar, and Ebmeier 2007). Model-based fMRI analysis also revealed reduced reward prediction error signals in the dACC in patients with MDD compared to healthy controls (Kumar et al. 2008). However, no previous study has examined how the dACC response to volatility varies in depressed participants.

1.4.5 Norepinephrine and depression

Norepinephrine has been theorized to signal unexpected uncertainty induced by gross changes in the environment and to boost adaptive learning (Dayan and Jyu 2003; Yu and Dayan 2005). This has been supported by studies measuring pupil dilation with electrophysiology, a commonly used measure of locus coeruleus norepinephrine (LC–NE) activity (Aston-Jones and Cohen 2005), during learning tasks. Such studies show that pupil dilations were modulated by surprise (Preuschoff, 't Hart, and Einhauser 2011), uncertainty (Lavin, San Martín, and Rosales Jubal 2014), learning noise (Findling et al. 2019) and environmental volatility (Vincent et al. 2019).

Pharmacological studies using methylphenidate, a norepinephrine and dopamine reuptake inhibitor, also partially support the role of norepinephrine in adaptive learning. These studies found that a single dose of methylphenidate increased learning rates in humans (Howlett et al. 2017), particularly when the environment was volatile (Cook et al. 2019). However, these studies did not distinguish the roles of the two catecholamines, noradrenaline and dopamine, in adaptive learning. A more recent study showed in monkeys that noradrenaline neurons rather than dopaminergic neurons were sensitive to the task state change (Jahn et al. 2020). Together, these studies provided evidence that norepinephrine modulates adaptive learning.

Norepinephrine also plays an important role in the pathophysiology and treatment of depressive disorders (Moret and Briley 2011). As listed in section 0, many effective antidepressants involve norepinephrine reuptake inhibition, such as venlafaxine, bupropion and reboxetine. However, it is not clear how the dysfunction of norepinephrine system affects reward and punishment learning in depression and how antidepressants targeting norepinephrine system, such as reboxetine, modulate this process in depression. In this thesis, I examine how an antidepressant targeting the norepinephrine system (reboxetine) affects adaptive learning.

1.5 Summary and Research Questions

In summary, behavioural and computational evidence suggests that people with depression seem to learn less from reward information and more from loss. Consistent with this, model-based fMRI results have found blunted RPE signals and heightened PPE signals in a distributed network (including striatum, anterior insula, OFC and dACC) in depressed patients. Behavioural results, such as shown in studies with the contingency-shifting variants of the IGT, suggest that the learning dysfunction in depression might be more prominent in volatile environments. However, less is known about the computational and neural mechanism of adaptive learning in depression and specifically how patients adapt to changes in volatility.

In this thesis, we will systematically examine the underlying mechanisms of adaptive learning in depression with pupillometry recordings, fMRI scanning, and pharmacological intervention.

The work is based on a task in which the volatility of rewarding and punishing outcomes is independently manipulated. More specifically, in Chapter 2, we compare the differences in volatility adaptation in volatile environments in currently depressed patients, remitted depressed patients and healthy controls using computational modelling and pupillometry responses. In Chapter 3, we assess the underlying neural mechanisms of volatility adaptation and its association with symptoms of depression using functional MRI. Finally, in Chapter 0, we investigate the role that norepinephrine plays during volatility adaptation using a pharmacological intervention (reboxetine, an antidepressant targeting the norepinephrine system).

2 CHARACTERISING INFORMATION LEARNING BIAS IN DEPRESSION

2.1 Introduction

Major depressive disorder is a mood and emotional disorder characterised by persistent low mood and lack of interest. In the framework of reinforcement learning, depressive disorder is hypothesised to be associated with deficits in value learning and decision making. Studying reinforcement learning in depression using computational models may help further understand the complex mechanism of depression.

Reward and punishment learning are key processes in reinforcement learning. Previous studies have suggested that people with depression show dysfunction in processing both reward (Henriques, Glowacki, and Davidson 1994; Whitton et al. 2016; Admon and Pizzagalli 2015) and punishment information (Holmes and Pizzagalli 2007), particularly when reward or/and punishment contingencies change (Dickstein et al. 2010; Levy-Gigi and Kéri 2015; Robinson et al. 2012; Cella, Dymond, and Cooper 2010). However, it is still not clear how people with depression learn and use reward and punishment information to make decisions, and more importantly, how they adapt learning in changing environments.

The ability to flexibly adjust learning rates in a way that reflects the volatility of the environment is crucial when learning about associations. In a stable environment, where action-outcome associations have been learned, it is less important to keep track of recent feedback as the associations are unlikely to change. In contrast, in a volatile environment, where the action-outcome associations change rapidly over time, one should pay more attention to, and be more influenced by, events that do not match current expectation. Using computational modelling, Behrens et al. have shown that healthy participants adjust their learning rates exactly as described above, with higher learning rates in volatile than stable blocks of a reward learning task (Behrens et al. 2007). Failing to adapt learning in different environments might affect well-being. A study showed poor estimates of the environment volatility levels might be associated with subjective feeling of uncertainty and stress (De Berker et al. 2016). Browning et al. found that individuals with high trait anxiety adjusted their learning rate less between the volatile and stable conditions in an aversive learning task than participants with low anxiety (Browning et

al. 2015). However, it has not been tested whether the learning adjustment to volatility level for punishment information is also impaired in depression and whether there is a similar impairment for reward learning.

Norepinephrine has been suggested as a key neuromodulator to facilitate adaptive learning in volatile environment (Yu and Dayan 2005; Dayan and Jyu 2003). Fundamental changes in the environment, i.e. volatility, induce unexpected uncertainty, which was hypothesized to boost learning through modulating phasic activity of the central norepinephrinic system (Dayan and Jyu 2003; Yu and Dayan 2005). Pharmacological manipulation of the norepinephrine system (using atomoxetine) has been shown to affect learning rates following unanticipated task changes by modulating the effect of prediction errors on catecholamines related neural activities (Jepma et al. 2016). Disturbances of norepinephrine have been suggested to play a key role in depression (Moret and Briley 2011). Many clinically proved effective antidepressants modulate the central norepinephrinic system, such as venlafaxine, and bupropion (see section 0 for more details). We hypothesized that people with depression might have difficulty in adapting learning to accommodate the volatility level of the environment. Browning et al. showed the impaired learning adaptation to volatility in individuals with high trait anxiety was associated with a diminished modulation of volatility on pupil dilation (Browning et al. 2015), which is commonly used as a measure of the activity in locus coeruleus norepinephrine (LC–NE) system (Aston-Jones and Cohen 2005). However, it has not been tested whether dysfunction of the norepinephrine system in depression is associated with aberrant learning rate adjustments when learning punishment or reward in changing environments.

In this study, we adopted a reinforcement learning paradigm in which reward and punishment volatility were manipulated. Using this paradigm, Pulcu et al. showed that healthy participants were able to independently adapt their reward and punishment learning rates in response to the volatility of reward and punishment associations (Pulcu and Browning 2017). Using computational modelling, we examine the learning adaptation to volatility both with reward and punishment learning among clinically depressed patients, remitted patients and healthy controls.

We hypothesized that people with current depression symptoms might show reduced learning rate adjustments between the volatile and stable environment for both reward and punishment learning compared to the healthy control group. We hypothesized this to be a state dependent dysfunction rather than a trait like characteristic, so expected both the remitted depressed and control groups to show greater learning rate adaptation to volatility than the currently depressed group. We also recorded pupil dilation while participants performed the task, to investigate norepinephrine activity level, and whether it was associated with learning rate adjustments. As for the learning rate adaptation, we expected a reduced pupil dilation changes in response to

environment volatility in the currently depressed group, compared to the healthy control and remitted depressed groups.

2.2 Methods

2.2.1 Experiment task and participants

2.2.1.1 Participants

2.2.1.1.1 Recruitment and screening

Participants attended a screening session during which the Structured Clinical Interview for DSM‐5 (SCID‐5) sections for mood disorders was completed. Based on this screening session participants were categorized into one of 3 groups: (1) currently depressed group, people who currently meet the DSM-V criteria for depression, (2) people who had at least two previous episodes of depression, but were not currently depressed, that is meet criteria for previous but not current depression, and (3) people who have never been clinically depressed. Participants were excluded if: (1) they were receiving any treatment for depression (including medication or talking treatments), or other psychoactive treatment, (2) they met criteria for bipolar disorder, (3) they had used any street drugs within the last 3 months, (4) they had a medical condition which affected pupil dilation, (5) they had participated in any studies using the same behavioural task as in the current study. Participants were matched for gender, age and years of education (see Table 3). The study was approved by the University of Oxford Central Research Ethics Committee. Written informed consent was obtained from all participants on the screening visit, in accordance with the Declaration of Helsinki.

2.2.1.1.2 Test session

Participants completed the testing session a maximum of four weeks after screening (if this limit was breached, participants were rescreened). During this session, participants were asked to complete a number of questionnaires, including (1) the Quick Inventory of Depressive Symptoms, 16 items self-report version (QIDS)，which measures symptoms of depression (Rush et al. 2003), (2) the Spielberger State-Trait Anxiety Inventory (STAI), which includes separate scales measuring state and trait anxiety (Spielberger 1983), (3) the Snaith–Hamilton Pleasure Scale (SHAPS), measuring the ability to experience pleasure (Nakonezny et al. 2010), (4) the Ruminative Response Scale, which measures rumination in response to depressive mood (RRS) (Nolen-Hoeksema and Morrow 1991), and the Depressive Attributions Questionnaire (DAQ), which assess the extent to which an individual had a pessimistic or depressogenic

attributional style (Kleim, Gonzalo, and Ehlers 2011). They were instructed about the information bias leaning task (see below, 2.2.1.2) and completed a brief introduction version (of 16 trials) to familiarize them with the logic of the task. Following this, participants completed the 3 blocks of the task self-paced, with rest sessions between each block. The task was presented on a VGA monitor connected to a laptop computer running Presentation software version 18.3 (Neurobehavioural Systems, Berkeley, CA).

2.2.1.2 The information bias learning task

The task used in this study was previously used in a study of non-clinical participants (Pulcu and Browning 2017). In this task, participants were instructed to learn the association between two abstract shapes and both win and loss outcomes in order to collect as much money as possible. At the beginning of each trial, participants were presented with two different abstract shapes on each side of the screen and they could press the left or right button on a mouse to choose the corresponding shape (see Figure 2.1). The choice phase was self-paced. Following this choice, a black frame would appear around the shape that was chosen. 1 second after the choice was made, on each trial, a win would appear above one of the shapes and a loss would show up below one of the shapes. Participants would receive a win if they had chosen the shape associated with the win and would receive a loss if they had chosen the shape associated with the loss. The results were revealed in randomised order for a jittered interval of 2–6 (mean 4) seconds. In half of the trials, the win outcome was presented first, with the loss outcome presented first in the other half. A 'chi-ching' sound was played if participants received a win and an error buzz if they received a loss. Auditory stimuli lasted 0.7 seconds with the same onsets as the visual feedback. The intertrial interval was fixed at 1 second.

The total money collected was displayed under the fixation cross and was updated on every trial. Participants were instructed that the order of the outcomes (i.e., win or loss first) or the side of the shapes presented were randomised from trial to trial and were not related to the outcomes. Participants were also instructed that the win association and loss association with the shapes could change over the course of a block, and that the win and loss associations were independent and could change independently. Unknown to participants, the schedule of the task differed between the testing blocks such that the volatility of the two outcomes varied by block (see Figure 2.1): In the first "both volatile" block, both win and loss outcomes were volatile, with associations between the outcomes and the stimuli switching frequently between 15 and 85%. In the "win volatile" block, the association with win outcomes was "volatile", whereas the association with loss outcomes was "stable" (i.e., fixed at 50%). Finally, in the "loss volatile" block the loss outcomes were volatile and win outcomes stable. The first block was always

"both volatile" in order to familiarise participants with the task. The win and loss volatile blocks followed this with their order counterbalanced across participants.

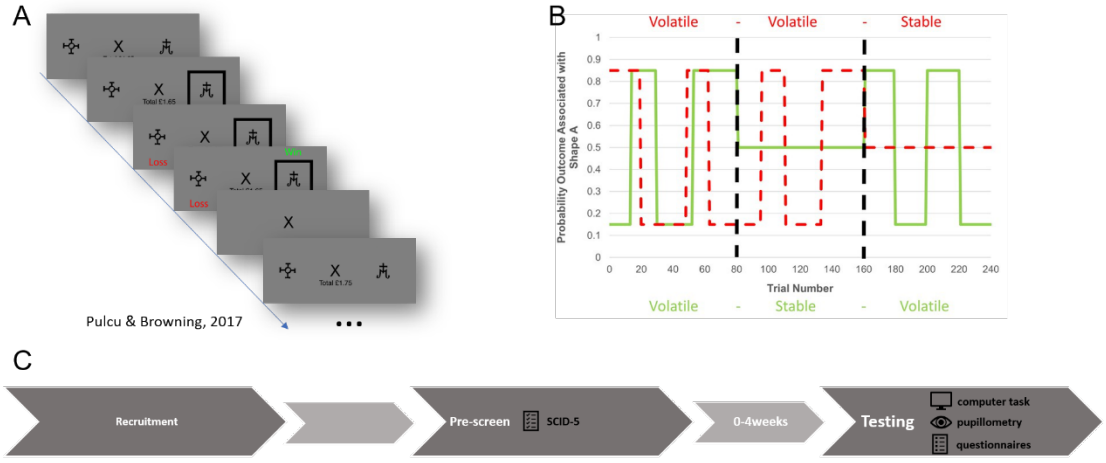

Figure 2.1 General procedures and Task structure.

(A) Timeline of one trial from the information bias learning task. Participants were presented with two abstract shapes (referred to as shape 'A' and 'B') and were asked to choose one. For each trial, a win would appear on the top of one of the shapes, choosing that shape results in a gain of 15p. Independent of the win outcome, a loss would show up on the bottom of one of the shapes for each trial with one of the shapes resulting in a loss of 15p if the shape was chosen. Participants had to learn which shape the win and loss were likely to be associated with and use this information to guide their choice in order to maximise their monetary earnings. (B) Overall task structure. The task consisted of 3 blocks of 80 trials each (i.e., vertical, dashed, dark lines separate the blocks). The y-axis represents the probability, p, that an outcome (win in solid green or loss in dashed red) would show up with shape 'A' (the probability that it was associated with shape 'B' would be 1-p). The blocks differed in how volatile (changeable) the outcome probabilities were. Participants always started with the 'both volatile' block, in which both win and loss outcomes were volatile and then half of them first did the 'win volatile' block, in which only wins were volatile while losses fixed at 50% probability, and then the 'loss volatile' block, in which wins had a fixed 50% probability and losses were volatile. (C) General procedures. The participants were first invited for the pre-screened visit, during which the Structured Clinical Interview for DSM-5 (SCID-5) sections for mood disorders was completed. Within 4 weeks of the pre-screen visit, the eligible participants were invited back for the testing session, during which they performed the information bias learning task and completed the questionnaires.

2.2.1.3 Pupillometry recordings

The participants' pupil responses when they performed the information bias learning task were recorded using an eye-tracking device. Participants' heads were stabilised using a head-and-

chin rest placed 70 cm from the screen on which an eye-tracking system was mounted (Eyelink 1000 Plus; SR Research, Ottawa, Canada). The eye-tracking device was configured to record the coordinates of both of the eyes and pupil area at a rate of 500 Hz. The abstract shapes of the learning task were drawn on either side of a fixation cross which marked the middle of the screen and was offset by around 7˚ visual angle.

2.2.2 Model-free behavioural data analysis

2.2.2.1 Task performance

The model-free behavioural data were analysed using MATLAB (version R2018b; MathWorks, MA, USA) and SPSS (version 26.0; IBM, NY, USA). The percentages of wins chosen and losses not-chosen were calculated for each block. As proportion data, such as percentages, violate the assumptions of equality of variance required by most parametric tests, the percentage data were arcsine square-root transformed (Ahrens, Cox, and Budhwar 1990) before being submitted to statistical tests.

We first examined whether the performance for each block was significantly above chance level (i.e., 50%). We then ran a 2 by 2 by 3 by 2 mixed ANOVA, with valence (win/loss) and block ("both volatile"/ "semi-volatile", i.e., win and loss performances from the both-volatile block vs. win performances from the win-volatile block and loss performances from the loss-volatile block) as repeated measures, while group (current depression group/remitted depression group/ control group) and block order ("win volatile" block first or "loss volatile" block first) as between-subjects factors. Note that in the semi-volatile blocks, only the accuracy for the volatile outcome was entered into analyses as the stable outcome was equally associated with both shapes and thus there was no "correct" answer. Block order was added to control for the block order effect.

2.2.2.2 Positive-stay-negative-switch

The switch probabilities after receiving a loss (negative outcomes for loss, $P_{switch\text{-}loss}$), not receiving a win (negative outcomes for win, $P_{switch\text{-}nowin}$), not receiving a loss (positive outcomes for loss, $P_{switch\text{-}noloss}$), and receiving a win (positive outcome for win, $P_{switch\text{-}win}$) for each block were calculated. We then calculated the differences between probabilities of switching after receiving a positive outcome and switch probabilities after receiving a negative outcome for win and loss respectively. The probabilities were first arcsine square root transformed to correct for the violation of equality of variance assumptions caused by percentage data (Ahrens, Cox,

and Budhwar 1990). We would refer the differences as the positive-stay-negative-switch (P_{psns}) indexes in the following text, see Equation 7.

$$\boldsymbol{P_{psns-win} = \sin^{-1}\sqrt{P_{switch-nowin}} - \sin^{-1}\sqrt{P_{switch-win}}}$$
$$\boldsymbol{P_{psns-loss} = \sin^{-1}\sqrt{P_{switch-loss}} - \sin^{-1}\sqrt{P_{switch-noloss}}} \qquad \textbf{Equation 7}$$

The positive-stay-negative-switch indices effectively combine the two actions prompted by an outcome (i.e., switching choice after a negative outcome and staying after a positive outcome) and thus provide a general measure of a participant's sensitivity to wins and losses. To illustrate, if a participant was particularly influenced by win outcomes, they would be more likely to switch following trials in which they do not receive a win and more likely to stay on trials in which they do receive a win and would thus have a high $P_{psns-win}$.

The positive-stay-negative-switch indexes were then submitted to a 2 by 2 by 3 by 2 mixed ANOVA, with valence (win/loss) and volatility level ("volatile"/"stable", i.e., $P_{psns-win}$ from the win-volatile block and $P_{psns-loss}$ from the loss-volatile block, vs $P_{psns-win}$ from the loss-volatile block and $P_{psns-loss}$ from the win-volatile block; note that here we only included the Ppsns from the block 2&3 (i.e., the "win volatile" block and the "loss volatile" block), as the two blocks were matched in task difficulty and order balanced) as repeated measures, and group (current depression group/remitted depression group/ control group) and block order ("win volatile" block first or "loss volatile" block first) as between-subjects factors..

2.2.3 Computational modelling of behavioural data

Volatility adaptation model:

In this analysis, we adapted a RW model to reflect the task design. More specifically, we estimated baseline learning rates for win and loss outcomes for each participant, which were used when that outcome was stable ($\propto win_{stable}, \propto loss_{stable}$). In addition, we estimated the degree to which the participant adapted their learning rates between stable and volatile blocks as separate parameters (one for win and one for loss LR, $\propto win_{adapt}, \propto loss_{adapt}$), with the learning rate used in volatile blocks being the sum of the stable and adapt LRs (see Table 1). This parameterisation provides a direct estimation of the baseline learning rate used by participants and the degree to which they adjust their learning rate in response to increased volatility.

Table 1 Equations for the volatility adaptation model for the clinical study

block	valence	Description	equation
Both Volatile	win	Win volatile learning rate	$\alpha\text{win}(1)=\alpha\text{win}_{\text{stable}}+\alpha\text{win}_{\text{adpt}}$

Win Volatile	loss	Loss volatile learning rate	$\alpha\text{loss}(1)=\alpha\text{loss}_{\text{stable}}+\alpha\text{loss}_{\text{adpt}}$
	win	Win volatile learning rate	$\alpha\text{win}(2)=\alpha\text{win}_{\text{stable}}+\alpha\text{win}_{\text{adpt}}$
Loss Volatile	loss	Loss stable learning rate	$\alpha\text{loss}(2)=\alpha\text{loss}_{\text{stable}}$
	win	Win stable learning rate	$\alpha\text{win}(3)=\alpha\text{win}_{\text{stable}}$
	loss	Loss volatile learning rate	$\alpha\text{loss}(3)=\alpha\text{loss}_{\text{stable}}+\alpha\text{loss}_{\text{adpt}}$

$$\boldsymbol{Vwin_{t+1}} = \boldsymbol{Vwin_t} + \boldsymbol{\alpha win(block)} \cdot \boldsymbol{PEwin(t)}$$

$$\boldsymbol{Vloss_{t+1}} = \boldsymbol{Vloss_t} + \boldsymbol{\alpha loss(block)} \cdot \boldsymbol{PEloss(t)}$$

Equation 8

$$\boldsymbol{PEwin(t)} = \boldsymbol{Rwin_t} - \boldsymbol{Vwin_t}$$

$$\boldsymbol{PE}loss\boldsymbol{(t)} = \boldsymbol{Rloss_t} - \boldsymbol{Vloss_t}$$

Equation 9

$$\boldsymbol{Pchoice_t} = \frac{\mathbf{1}}{\mathbf{1}+exp^{-(\beta\cdot(Vwin_t - Vloss_t))}}$$

Equation 10

The expected values for win and loss for shape A (Vwin values of shape B would be 1-$\text{Vwin}_{(\text{shape A})}$ and Vloss value would be 1-$\text{Vloss}_{(\text{shape A})}$) were updated trial by trial with the prediction errors, which is the discrepancy between the outcome received and the expected value for a trial (see Equation 9), for each trial and block-wise win and loss learning rates (see Table 1) respectively using a Rescorla-Wagner learning rule (Rescorla 1972)(**Equation 8**). The data from all 3 blocks (excluding the first 5 trials of each block) were used in the model fitting. The expected values for win and loss were reset to 0.5 at the beginning of each block. These expected values were then transformed into a single choice probability using a softmax function (Equation 10). The inverse temperature was shared across blocks for each participant, as the inverse temperate estimates did not show significant differences across blocks in the original paper (Pulcu and Browning 2017).

Models were fitted using Stan (RStan v 2.21.1, R v 3.6.0), a software package that uses the Markov chain Monte Carlo method to sample from posterior distributions of parameters (Carpenter et al. 2017). 4 parallel Monte Carlo chains were run each with 10000 iterations (5000 warm-ups), using adpt_delta of 0.8 and maximum treedepth of 15. The stable learning rate parameters and the learning rate adaptation terms were sampled from a zero-mean normal distribution with a standard deviation of 3. The stable learning rates were then transformed to the range 0-1 using logistic approximation (Bowling et al. 2009). The volatile learning rates were calculated by combining the stable learning rate estimates and volatility adaptation terms

(both zero-mean normal distributed), and then logistic transformed to the range 0-1. Therefore, theoretically, the volatile learning rates can be lower than the stable learning rates. Beta parameters were sampled from a zero-mean normal distribution with standard deviation 2 and then transformed using the exp function to a positive value. Convergence was inspected using shinyStan (v 2.5.0). No convergence problems were detected. We adopted a stringent cut-off of R-hat< 1.1 as an indicator for sufficient convergence (Brooks and Gelman 1998), and the minimum number of effective samples >50. Parameters were estimated separately for each participant with no hierarchical terms.

The statistical tests for the learning rate parameters were performed in logit space and beta parameters performed in log space using SPSS (version 26.0; IBM, NY, USA). A 2 by 3 by 2 mixed ANOVA, with valence (win/loss) as the repeated measure, and group (current depression group/remitted depression group/ control group) and block order ("win volatile" block first or "loss volatile" block first) as between-subjects factors were performed for the baseline learning rates and the volatility learning rate adaptations respectively. For any significant group interaction effect, we performed separate group one-way analysis of variance (ANOVA) and followed with post hoc Fisher Least Significant Difference tests. The beta estimates were submitted to a group one-way ANOVA.

All correlation analysis in this section were performed using Spearman correlation.

Other models:

Two other models (the betas volatility adaptation model and the fixed learning model) were also estimated using RStan with the same procedures.

The fixed learning model used only one learning rate and one beta for all the blocks. The model assumes that neither learning rates nor betas were affected by volatility manipulations.

The betas volatility adaptation model assigned different betas for each block for wins and losses respectively but with the same learning rate for wins and losses for each block (see Table 3). The model assumes that volatility only affects how participants use win and loss information to make decisions but not learning rate per se.

Table 2 Equations for the beta volatility adaptation model for the clinical study

block	valence	Description	equation
Both Volatile	win	Win volatile beta	$\boldsymbol{\beta}\text{win}(1)= \boldsymbol{\beta}\text{win}_{\text{stable}}+\boldsymbol{\beta}\text{win}_{\text{adpt}}$
	loss	Loss volatile beta	$\boldsymbol{\beta}\text{loss}(1)= \boldsymbol{\beta}\text{loss}_{\text{stable}}+\boldsymbol{\beta}\text{loss}_{\text{adpt}}$
Win Volatile	win	Win volatile beta	$\boldsymbol{\beta}\text{win}(2)= \boldsymbol{\beta}\text{win}_{\text{stable}}+\boldsymbol{\beta}\text{win}_{\text{adpt}}$
	loss	Loss stable beta	$\boldsymbol{\beta}\text{loss}(2)= \boldsymbol{\beta}\text{loss}_{\text{stable}}$
Loss Volatile	win	Win stable beta	$\boldsymbol{\beta}\text{win}(3)= \boldsymbol{\beta}\text{win}_{\text{stable}}$
	loss	Loss volatile beta	$\boldsymbol{\beta}\text{loss}(3)= \boldsymbol{\beta}\text{loss}_{\text{stable}}+\boldsymbol{\beta}\text{loss}_{\text{adpt}}$

$$Vwin_{t+1} = Vwin_t + \alpha(block) \cdot PEwin(t)$$

$$Vloss_{t+1} = Vloss_t + \alpha(block) \cdot PEloss(t)$$

Equation 11

$$PEwin(t) = Rwin_t - Vwin_t$$

$$PEloss(t) = Rloss_t - Vloss_t$$

Equation 12

$$Pchoice_t = \frac{1}{1+exp^{-(\beta win \cdot Vwin_t - \beta loss * Vloss_t)}}$$

Equation 13

We then performed model comparisons using a leave-one-out cross-validation procedure using the LOO package for R (Vehtari, Gelman, and Gabry 2017). An expected log pointwise predictive density (ELPD) and a standard error for each ELPD estimate were generated for each model as a measure of predictive accuracy.

$$\widehat{\text{elpd}}_{\text{loo},i} = \log p(y_i | y_{-i}).$$

$$\text{se}(\widehat{\text{elpd}}_{\text{loo}}) = \sqrt{n \, V_{i=1}^{n} \widehat{\text{elpd}}_{\text{loo},i}}.$$

Equation 15

2.2.4 Pupillometry data analysis

2.2.4.1 Pre-processing

Blinks were identified automatedly by the Eyelink system's built-in filter and were then removed from the data. Missing data points (including blinks) were linearly interpolated. The resulting trace was subjected to a low pass Butterworth filter with a cut-off of 3.75 Hz and then z transformed across the session (Browning et al. 2015; Pulcu and Browning 2017). The pupil response to the win and the loss outcomes were extracted separately for each trial with the epoch from 1 s prior to the presentation of the outcome to 6 s after the outcome presentation. The pupil response was baseline corrected for each outcome presentation by subtracting the mean pupil size of a one second window preceding feedback. Individual trials were excluded from the pupillometry analysis if more than 50% of the data in the epoch had been interpolated

(Browning et al. 2015; Pulcu and Browning 2017). 4 participants were excluded from the pupillometry analysis as more than 60% of their trials were excluded on this basis. Another 3 participants were excluded due to recording failure. The same number of trials (n=5) as in the model fitting (see the previous section 2.2.3) from each block were removed from the analysis to ensure consistency between the analyses and also to avoid any luminance changes that may be caused by starting a new block of the task.

2.2.4.2 Regression analysis of pupil data

Trial by trial regression analysis was performed on the pre-processed pupillometry data combining all three blocks using the ordinary least square (OLS) method for data points during -1s to +6s relative to outcome delivery for win and loss outcomes respectively. Explanatory variables (EVs) of interest included (1) self-prediction errors (self-PE): prediction errors for the current event for the current trial (i.e., win prediction errors for win outcome window, and loss prediction error for the loss outcome window. The prediction errors were calculated using the best fitting learning rates derived from the computational model (see section 2.2.3 for more details)), (2) self-volatility level: the volatility level for the current block and the current outcome event (e.g., for the win outcome delivery: 1 for blocks when win outcomes were volatile, -1 for the block when win outcomes were stable), (3) self- volatility level* self-PE.
We also added several control variables, including (4) other-prediction errors (other-PE) prior than the current event in the same trial (i.e., loss prediction errors for win outcome window, and win prediction error for the loss outcome window. For trials, when the other outcome was presented later than the current event, the value was set to 0 in the demeaned regressor for this trial) to control for the remaining effect of the other outcome on the current event; (5) self-volatility level *other-PE to control for the remaining effect of the interactions of other outcome and volatility level on the current event; (6) order: whether the current outcome event was presented first in the current trials to control for the effect of the order of the outcome presentation within the current trial; (7) trial number: the current trial number in this block, to control for drifting in pupillometry data within the block, (8) a constant term. EVs (other than the constant term) were normalised for each participant before entering into the model. The two interaction regressors, i.e., EV(3) and EV(5) were calculated using normalised relevant EVs. Then the interaction regressors were also normalised before entering into the regression model. To investigate the modulation of EVs of interests on pupil dilations, the cluster-based nonparametric tests (one sample against zero) to control for multiple comparisons were performed for each group on the parameters estimates (Maris and Oostenveld 2007). For EVs of interests showing significant modulation effects, we then follow up with one-way group

ANOVAs on the means of the beta estimates 0-6s since outcome delivery, followed with post hoc Fisher Least Significant Difference tests.

2.3 Results

2.3.1 Demographic details of the participants

In total 123 participants were recruited to this study, 41 for each group. 3 participants in total were excluded (2 from the currently depressed group, 1 from the remitted depression group) because they chose the same shape for all the trials in one of the 3 blocks. Below is a summary of the demographic details of the participants included in the analysis (Table 3).

Table 3 Demographic details of the participants

variables	**Currently Depression (n=39)**	**Remitted depression (n=40)**	**Healthy control (n=41)**	**p-value for group differenc e**
Demographic information:				
Female, n(%)	31(79.49%)	34(85.00%)	28(68.29%)	0.185[1]
Age(yrs), mean(SD)	32.28(11.38)	30.63(11.45)	31.95(11.40)	0.791[2]
YoE, mean(SD)	16.26(3.02)	17.56(2.82)	17.36(3.50)	0.151[2]
Questionnaires:				
QIDS, mean(SD)	12.92(4.52)	3.98(2.39)	2.54(2.30)	<0.001[2]
State-STAI, mean(SD)	54.51(9.56)	33.48(7.52)	28.34(7.47)	<0.001[2]
Trait-STAI, mean(SD)	62.08(9.81)	39.33(9.52)	29.02(6.64)	<0.001[2]
SHAPS, mean(SD)	38.49(5.04)	48.93(4.60)	48.88(6.99)	<0.001[2]
DAQ, mean(SD)	37.77(12.17)	15.93(9.46)	6.88(5.67)	<0.001[2]
RRS, mean(SD)	62.21(11.46)	45.70(13.80)	31.61(9.00)	<0.001[2]

Note: [1]p-values were estimated for group differences using a chi-square test. [2]p-values were estimated for group differences using one-way ANOVA. n: number of participants included in the analysis in this chapter. YoE: Years of Education. QIDS: the Quick Inventory of Depressive Symptoms. State-STAI: state anxiety from the Spielberger State-Trait Anxiety Inventory. Trait-STAI: trait anxiety from the Spielberger State-Trait Anxiety Inventory. SHAPS: the Snaith–Hamilton Pleasure Scale. RRS: the Ruminative Response Scale. SD: standard deviation.

2.3.2 Group differences in task performance

2.3.2.1 Maladaptive learning under uncertainty in depression

We first examined task performance in terms of the percentage of trials in which participants received the win and loss outcomes. As expected, across all three groups, participants' accuracies during volatile schedules were significantly above chance level (50%) both for wins (both volatile block: $t(120)=13.03$, $p<.001$, win volatile block: $t(120)=14.44$, $p<.001$) and losses (both volatile block: $t(120)=10.46.03$, $p<.001$, loss volatile block: $t(120)=14.05$, $p<.001$). While accuracies for the stable schedules, which were not learnable by design, were slightly below chance level both for wins ($t(120)=-8.93$, $p<.001$) and losses ($t(120)=-6.05$, $p<.001$) (See Figure 2.2).

We then ran a mixed ANOVA to compare participants' performance between the two volatile blocks (i.e., percentages of win chosen from the both-volatile block and the win-volatile block, and percentage of loss not chosen from the both-volatile block and the win-volatile block). The analysis revealed a main effect of valence ($F(1,114)=10.833$, $p=.001$, $\eta_p^2=.087$), indicating that the participants overall made more correct choices for wins than losses. The results also showed a marginally significant main effect of group ($F(2,114)=3.08$, $p=.050$, $\eta_p^2=.051$) and a significant interaction of block and group ($F(2,114)=4.78$, $p=.012$, $\eta_p^2=.073$). To further investigate this, we performed one-way ANOVAs for group effects on the participants' performance for each block and for wins and losses respectively. The groups did not significantly differ in the win or loss accuracies for the both-volatile blocks (both $p>.53$). The analysis revealed significant group effect for loss accuracy for the loss-volatile block ($F(2,117)=3.51$, $p=.033$, $\eta^2=.057$) and a marginally significant group effect for win accuracy for the win-volatile block ($F(2,117)=2.87$, $p=.061$, $\eta^2=.047$). Post-hoc analyses revealed that the currently depressed group had lower performances compared to the healthy control group for both losses ($p=.013$) and wins ($p=.018$), and marginally lower performances than the remitted depressed group for the loss outcomes ($p=.051$).

The mixed ANOVAs also revealed a main effect of block ($F(1,114)=27.79$, $p<.001$, $\eta_p^2=.196$), indicating lower performances for the first block (the both-volatile block) for both wins and losses (See Figure 2.2). This lower performance might be because the participants were new to the task or this block was relatively more difficult with both wins and losses being volatile. In the following section, we examine the differential performance between groups in the semi-volatile blocks in terms of participants' tendency to switch choices.

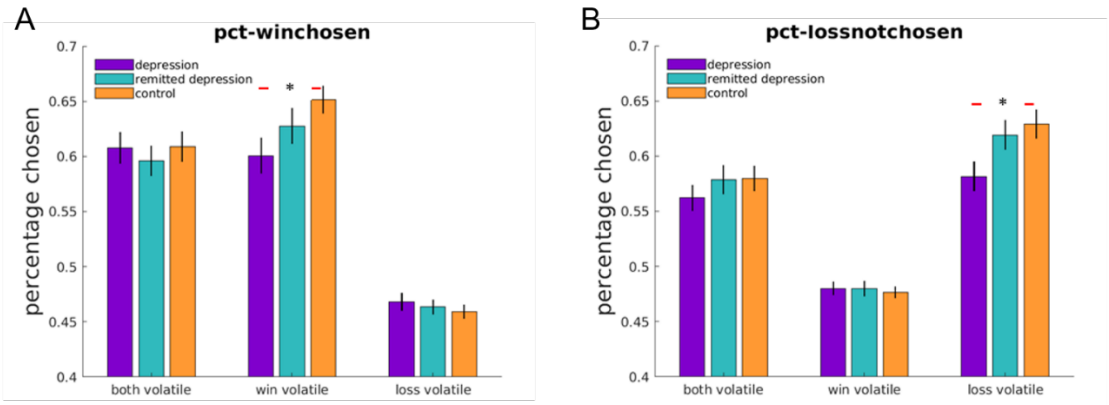

Figure 2.2 Summary of task performance for the depression, remitted depression and healthy control group.

(A) percentage of trials in which the participants' choices ended up with a win for each of the three task blocks. (B) percentage of trials in which the participant's choices did not end up with a loss for each of the three task blocks.

Error bars represent ±1 standard error of the mean. * $p<0.05$ for the post-hoc t-tests.

2.3.2.2 Positive-stay and negative-switch behaviours

As described in Chapter 1, agents' behaviour should be more influenced by outcomes from volatile than stable association, i.e., when associations are volatile, behaviour that had led to positive outcomes should be more likely to be repeated whereas negative outcomes should be more likely to lead to switches in behaviour. As described in the Methods section we constructed the positive-stay, negative-switch (Ppsns) index to measure this behaviour. Analysis of the Ppsns index revealed a significant main effect of volatility ($F(1,114)=88.90$, $p<.001$, $\eta_p^2=.438$), indicating that, as predicted, participants were more influenced by outcomes of volatile than stable associations.

This analysis also revealed a significant main effect of valence ($F(1,114)=22.90$, $p<.001$, $\eta_p^2=.167$) and a marginally significant valence by group interaction ($F(2,114)=3.06$, $p=.050$, $\eta_p^2=.051$). Post-hoc analysis revealed that while overall participants were more sensitive to win than loss outcomes, the currently depressed group were even less sensitive to the loss outcomes compared to the healthy control group ($p=.038$) and the remitted depressed group ($p=.024$) (see Figure 2.3).

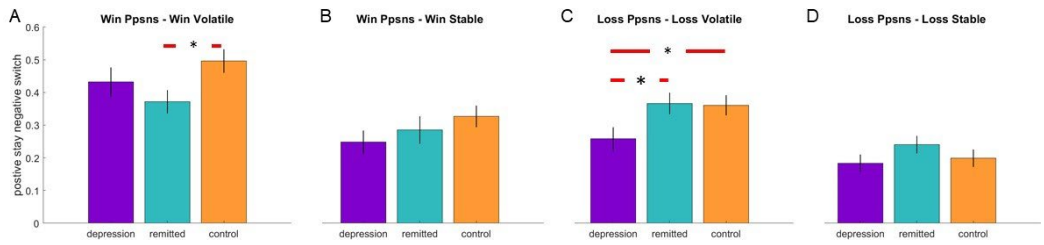

Figure 2.3 positive stay negative switch behaviour between volatile and stable environments.

(A) win Ppsns for the win-volatile block (in which win was volatile) for each group. (B) win Ppsns for the loss-volatile block (in which win was stable) for each group. (C) loss Ppsns for the loss-volatile block (in which loss was volatile) for each group. (D) loss Ppsns for the win-volatile block (in which loss was stable) for each group.

Error bars represent ±1 standard error of the mean. * $p<0.05$ for the post-hoc t-tests

The results also showed a marginally significant valence * volatility * group interaction ($F(2,114)=2.90$, $p=0.059$, $\eta_p^2=.048$), suggesting that the three groups showed different patterns of volatility behaviour adaptation to win and loss. To further investigate the volatility effect, we performed one-way ANOVAs on the Ppsns index differences between the volatile environment and stable environment for win and loss respectively. However, neither the one-way ANOVA results for win ($p=.149$) or loss ($p=.136$) reached statistical significance (see Figure 2.4). Post-hoc analysis showed a trend effect with the currently depressed group showing a greater difference in outcome sensitivity in response to volatility than the remitted depression group for wins ($p=.077$), and the currently depressed group showing reduced difference in outcome sensitivity than the healthy control group for losses ($p=.054$). Together, the results suggest that the currently depressed group might show differential behavioural adjustments to volatility for both win and loss outcomes.

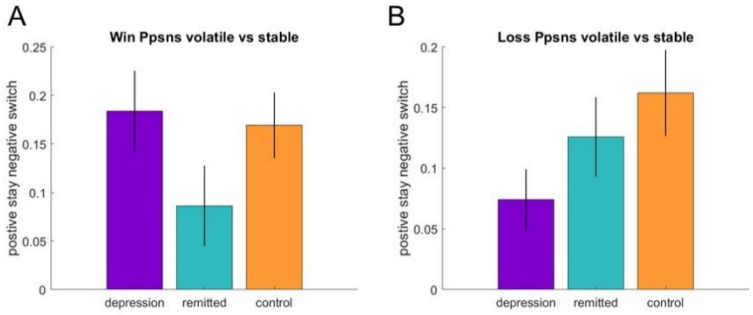

Figure 2.4 Differences of positive stay negative switch behaviour between volatile and stable environments.

(A) Difference of win Ppsns between the win-volatile block (win volatile) and the loss volatile (win stable) block. (B) Difference of loss Ppsns between the loss-volatile block (loss volatile) and the win volatile (loss stable) block.
Error bars represent ±1 standard error of the mean.

2.3.2.3 Summary

We showed that the currently depressed group performed worse than the other two groups, particularly when one information was relatively more volatile (informative) than the other. This was accompanied by reduced sensitivity to loss outcomes, indicated by Ppsns-loss, and reduced behaviour adjustment to loss volatility. However, the group effects identified are clearly not robust making it difficult to draw firm conclusions on the nature of this effect. In the following section (i.e., section 2.3.3), we examine this effect using a formal computational model.

2.3.3 Volatility adaptation model results

2.3.3.1 Stable learning rates

The mixed ANOVA for the stable (baseline) learning rates showed only a marginally significant main effect of valence ($F(1,114)=2.90$, $p=.091$, $\eta_p^2=.025$). Neither the main effect of group nor the valence * group interaction was statistically significant (both $p>.20$) (see Figure 2.5).

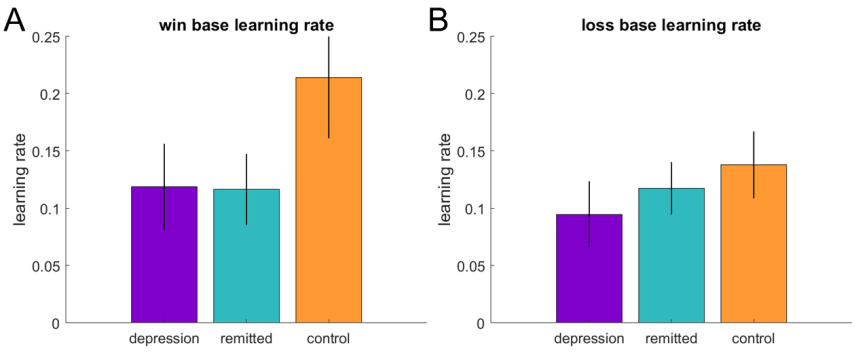

Figure 2.5 Baseline Learning rates estimates.
(A) win learning rates for the loss-volatile block (in which win was stable). (B) loss learning rates for the win-volatile block (in which loss was stable).
Error bars represent ±1 standard error of the mean.

2.3.3.2 Learning rate adaptation to volatility

The mixed ANOVA for the volatility learning rate adaptation estimates showed a significant main effect of valence ($F(1,114)=4.92$, $p=.029$, $\eta_p^2=.041$), showing in general a higher volatility learning rate adaptation for win than for loss (See Figure 2.6).

The analysis also revealed a significant group * valence interaction ($F(2,114)=3.71$, $p=.049$, $\eta_p^2=.052$). Follow up one way ANOVA for win and loss learning rate adaptation estimates respectively showed a significant group effect for loss ($F(2,117)=3.91$, $p=.023$, $\eta^2=.063$), but not for win ($p=.524$). Post-hoc analysis revealed that the current depression group had lower loss adaptation compared to the healthy control group ($p=0.011$) and the remitted depressed group ($p=0.027$) (See Figure 2.6).

We also examined whether our model estimates reflected the Ppsns volatility adaptations described in the earlier section. Correlation analysis confirmed that the win and loss learning rate adaptation estimates positively correlated with Ppsns-win (Pearson $r(120)=.60$, $p<.001$) and Ppsns-loss (Pearson $r(120)=.47$, $p<.001$) volatile vs. stable differences respectively.

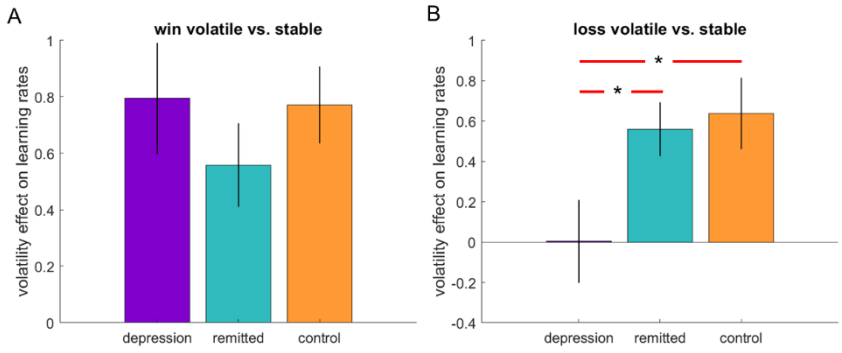

Figure 2.6 learning rate adjustments to the volatile blocks relative to the stable blocks for win and loss respectively.

(A) win learning rate adjustments to volatility (showing in logit space). (B) loss learning rate adjustments to volatility (showing in logit space).

Error bars represent ±1 standard error of the mean. * $p<0.05$ for the post-hoc t-tests.

2.3.3.3 Betas

The one-way ANOVA for beta estimates showed no significant group differences ($p=.543$). (See Figure 2.7).

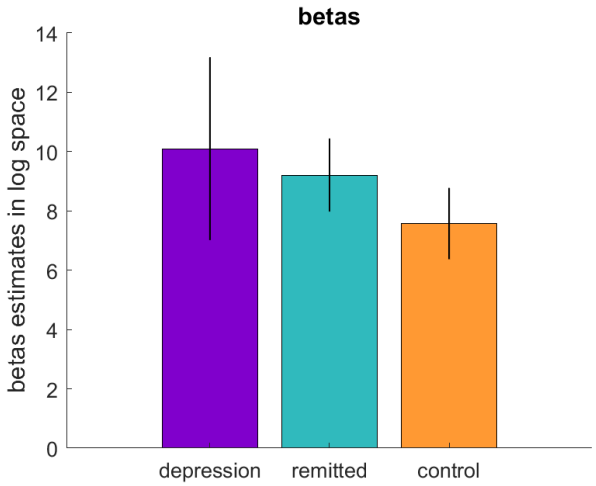

Figure 2.7 beta estimates for each group showing in log space.

Error bars represent ±1 standard error of the mean.

2.3.3.4 Correlations with symptom scores

We then checked whether the model estimates correlated with individual differences across the three groups. Consistent with the group effect we showed in the previous section, the Spearman correlation analysis confirmed significant negative correlations between loss learning rate volatility adaptation and depression/ anxiety symptoms (see Figure 2.8): with QIDs ($r(120)=-.196$, $p=.032$), state anxiety ($r(120)=-.202$, $p=0.027$), trait anxiety ($r(120)=-.179$, $p=.051$), hedonia/SHAPS ($r(120)=.193$, $p=.035$), DAQ ($r(120)=-.234$, $p=.010$), RRS ($r(120)=-.179$, $p=.051$). But the correlation between betas/base learning rates and depression/anxiety symptoms were not statistically significant (all $p>.09$). Neither did we find any statistically significant correlation between Ppsns indexes and depression/anxiety symptoms (all $p>.07$).

However, as the study was designed to find group differences, the correlation effects we showed here could be dominated by the group differences. Therefore, we examined the relationships between the loss learning rate volatility adaptation term and the questionnaires using a univariate ANCOVA model with the loss learning rate adaptation estimates as the dependent variable, group and order as fixed factors and all the questionnaire as covariates. None of the correlations remained significant ($p>.075$).

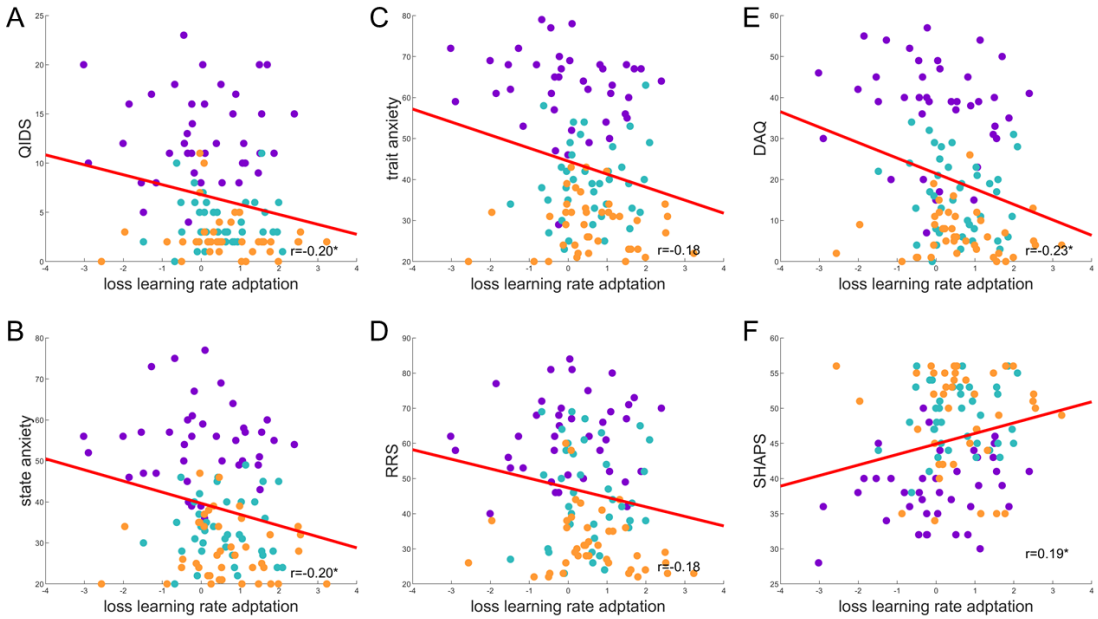

Figure 2.8 Loss learning rate adaptation correlated with depression and anxiety scores.

Loss learning rate adaptations correlated with (A) QIDs, the Quick Inventory of Depressive Symptoms (B) state anxiety from the Spielberger State-Trait Anxiety Inventory (C) trait anxiety from the Spielberger State-Trait Anxiety Inventory (D) RRS, the Ruminative Response Scale (E) DAQ, the Depressive Attributions Questionnaire (F) SHAPS, the Snaith–Hamilton Pleasure Scale.

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

2.3.3.5 Model comparisons

The learning rate volatility adaptation model had highest predictive accuracy (ELPD=, 10368.7, SE=51.8) compared to the beta volatility adaptation model (ELPD=-10852.6, SE=68.3), and the fixed learning model (ELPD=-11730.4, SE=53.8), suggesting that the hypothesis that using different learning rates across blocks captured participants behaviors better than other parameters.

2.3.3.6 Summary

The model results confirmed that the currently depressed group failed to adjust their learning rates for volatile environments, especially for losses. In the next section, we examine whether this is mirrored by a change in the degree to which volatility influences pupil responses to losses in this group.

2.3.4 Group differences in pupil responses during task

Next, we examined how pupil dilation responded to prediction errors (PE) derived from the model and whether volatility and group affected the pupil dilation effect during the task. As expected, the OLS analysis revealed a strong modulation of prediction errors on pupil dilations for both win and loss outcomes for all three groups (see Figure 2.9).

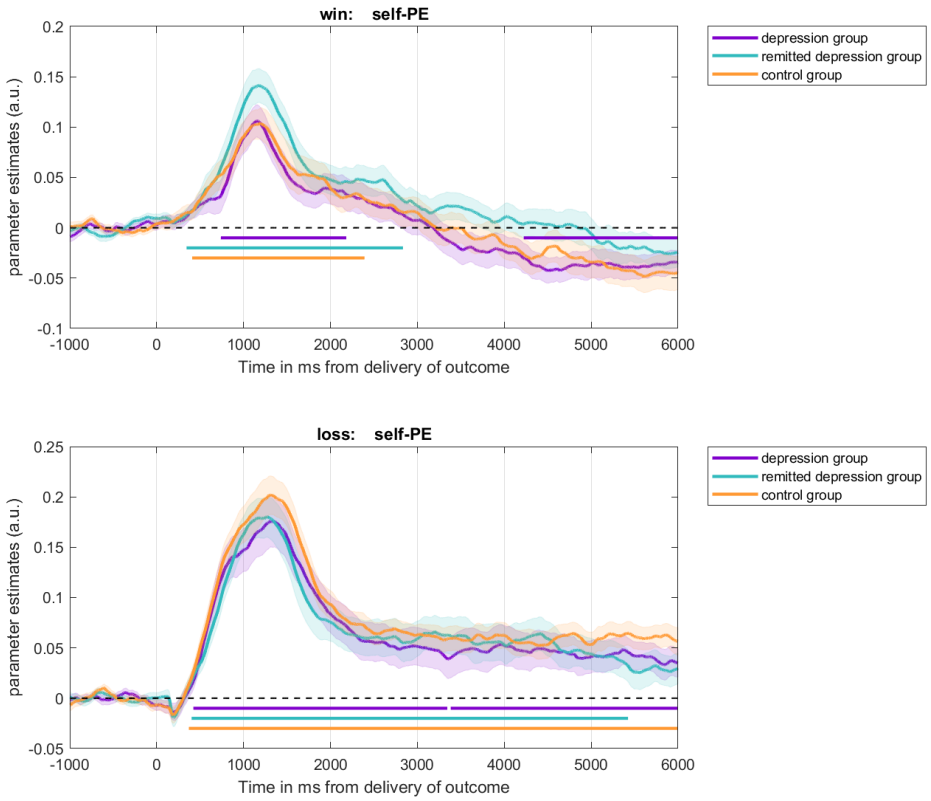

Figure 2.9 Regression analysis showing the temporal dynamics of prediction error prediction errors on pupil dilations.

The coloured solid line and shadow indicate the means ± s.e.m. calculated for each group. The coloured thick lines paralleling the x-axis indicate time windows in which coefficients were significantly different from zero after cluster-based multiple comparison correction (cluster definition threshold P = 0.05) for each group respectively. Time zero corresponds to the onset of outcome delivery.

Next, we examined whether volatility modulated the prediction error effects on pupil dilations. We found a positive volatility * PE interaction effect on pupil dilation for loss outcomes, i.e., volatility increased the modulation of PE on pupil dilations. However, only the healthy control group and the remitted depression group passed cluster-based multiple comparison correction, but not the currently depressed group (see Figure 2.10 B). To examine the group effect, we ran

one-way ANOVAs for the average coefficients (mean across 0-6s after outcome presentation) for the volatility *PE interaction. The results confirmed a significant group effect for the coefficients for loss PE and volatility interactions ($F(2,110)=3.34$, $p=.039$, $\eta^2=.057$), with lower coefficients for the currently depressed group than the remitted depressed group ($p=.015$) and the healthy control group ($p=.084$) (see Figure 2.10 D), but not for wins ($p=.223$). We also found a marginally significant positive (Pearson) correlation between the beta estimates of the volatility*PE interaction effect and loss learning rate adaption estimates ($r(113)=.183$, $p=.052$) (See Figure 2.10 F), but again not for the wins ($r(113)=-.086$) (See Figure 2.10 E).

To examine the possibility that the pupillometry finding was driven by the model estimates (i.e., the PEs used in the pupillometry analyses were calculated using the computational models fitted to participant choices), we also ran the same analysis with objective outcome results instead of model-derived win/loss prediction errors (i.e., whether receiving a win/a loss or not for each trial replaced the PE term). The analysis confirmed a strong modulation of outcomes on pupil dilations for both win and loss outcomes for all three groups and showed a qualitatively similar trend main effect of group for volatility and outcomes interaction ($p=.097$), with the currently depressed group showing lower modulation of volatility * outcomes compared to the remitted depressed group ($p=.069$) and the healthy control group ($p=.088$).

We also run the same trial by trial regression analysis for outcomes received instead of the PEs. The results were similar but slightly less significant.

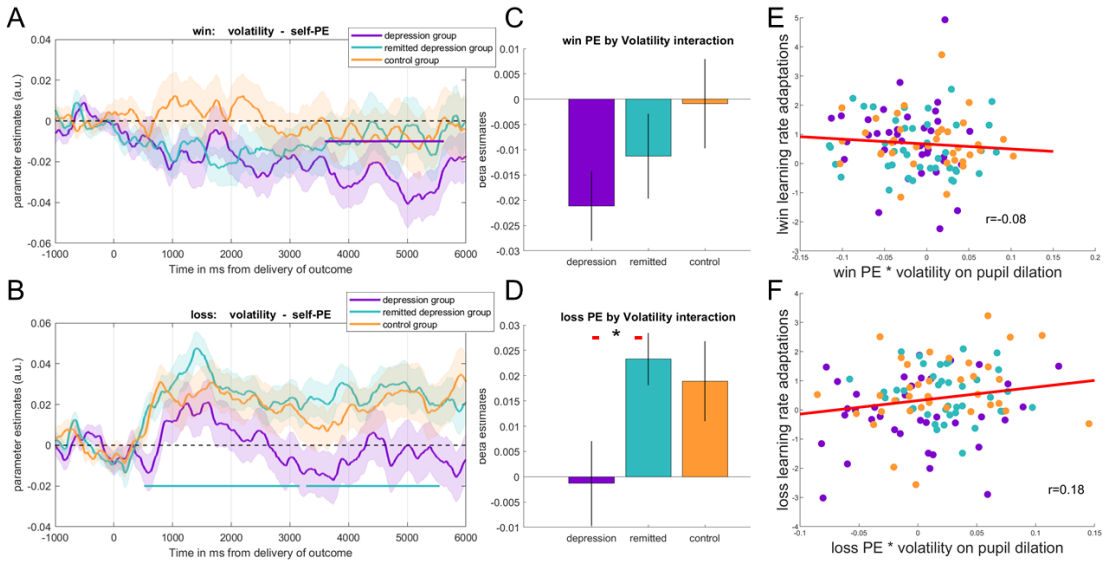

Figure 2.10 Regression analysis showing volatility by prediction error on pupil dilations. (A) The coefficients estimated for win prediction error and win volatility interaction on pupil dilations. (B) The coefficients estimated for loss prediction error and loss volatility interaction on pupil dilations. The coloured solid line and shadow indicate the means ± s.e.m. calculated for each group. The coloured thick lines paralleling the x-axis indicate time windows in which

coefficients were significantly different from zero after cluster-based multiple comparison correction (cluster definition threshold $P = 0.05$) for each group respectively. Time zero corresponds to the onset of outcome delivery.

(C) Bars represent average coefficients for win prediction errors and win volatility interaction (mean across 0-6s after the onset of outcome delivery) for each group. (D) Bars represent average coefficients for loss prediction errors and loss volatility interaction (mean across 0-6s after the onset of outcome delivery) for each group. Error bars represent ±1 standard error of the mean. * $p<0.05$ for post-hoc t-tests.

(E) Scatter plots for correlations between the win learning rate adaptation estimates from the computational model and the average coefficients for win PE and win volatility interaction. (F) Scatter plots for correlations between the loss learning rate adaptation estimates from the computational model and the average coefficients for loss PE and loss volatility interaction.

2.4 Discussion

In this chapter, we investigated the cognitive characteristics of information processing in currently depressed patients compared to remitted depressed patients and healthy controls using a learning and decision-making paradigm in which participants had to track the associations between stimuli and both reward and punishment outcomes in a changing environment. We first showed that, across all participants, task performance and choice patterns were strongly modulated by volatility. However, depressed patients had worse overall performance than the other two groups. We then examined how participants' choices were influenced by the most recent outcomes, i.e., the positive-stay and negative-switch behaviour indexes (Ppsns). Overall volatility increased the Ppsns for both win and loss across all participants. There was a marginally significant group difference for the volatility adaptation of the loss Ppsns but not for the win Ppsns. More specifically, the volatility adaptation of Ppsns was marginally less for the currently depressed group compared to the other two groups specifically to loss outcomes. We then fitted a volatility adaptation model to formalize the learning rate adjustments to the volatility. The model results confirmed that the currently depressed group had a lower adjustment in learning rates to the volatility of the environment, particularly for losses. Finally, we also showed the same effect with the pupil data, which revealed a group difference in how loss volatility modulated pupil responses to loss prediction error. Together, the results suggest that the currently depressed group had more difficulty in adjusting their learning to accommodate the volatility level of the environment for punishment but not reward.

The currently depressed group showed lower loss accuracies in the loss-volatile block compared to the other two groups. The computational model suggested a mechanism for this performance deficit-- that the depressed group had difficulty in adjusting learning rates to the

volatility of associations, particularly loss associations. This finding is consistent with a previous study, in which Browning et al. showed reduced learning rate adaptation in anxious participants during a punishment learning task (Browning et al. 2015). Recent studies have observed similar effects, with reduced loss learning rate adjustments in participants with high social anxiety (Beltzer et al. 2019) or when under acute stress (Hein et al. 2019). In a previous study with the exact same task design, Pulcu et. al (Pulcu and Browning 2017) also found that in healthy cohorts reduced learning rate adaptation to losses rather to wins was associated with trait anxiety. Our study extends this work to a clinically depressed population. Moreover, we also examined the effect between the currently depressed and the remitted depressed group. the results suggested that the reduced ability to adjust learning to loss volatility might be a state dependent effect of depression rather than a trait effect which persists between depressive episodes. Interestingly, in a recent paper, Gagne et al. showed decreased learning rate adaptation to both reward and punishment were associated with high scores on general factors across many depression and anxiety measures, but not to the depression-specific or the anxiety-specific factors (Gagne et al. 2020). In our study, we found similar results but only for the punishment learning, not the reward learning. An important difference between the two studies is that in the Gagne study they showed the correlations with the general factor, which was derived using a factor analysis on various questionnaires, while we focused on the group differences amongst clinically diagnosed groups. Another key difference between the two studies is that, in the Gagne study, reward learning and punishment learning was assessed in separate tasks, while in our study participants were required to learn about the two associations simultaneously and figured out by themselves which information was more informative to learn within a block. Therefore, it could be that when presented with both sources of information, the punishment learning problem in depressed participants becomes more apparent. In other words, there might be some interaction between valence and different sources of volatility, e.g., adjusting from win stable to win volatile might be reduced when loss is volatile. However, in the current study task we did not include a block in which both wins and losses were stable, and thus we could not examine the valence*volatility interaction effect and directly test this possibility.

In addition to the group effect on behaviour, we found that a decreased effect of volatility on pupil responses to prediction error during loss outcome delivery was associated with a reduced learning rate adaption to volatility, which is consistent with previous studies (Browning et al. 2015; Pulcu and Browning 2017). Importantly, we also found a diminished volatility effect on pupil responses to loss prediction error in the currently depressed group compared to the remitted depression and healthy control groups. This suggests that a disturbance of the norepinephrine system, or perhaps the upstream mechanism that estimates volatility before broadcasting it via the NE system, might be a candidate underlying mechanism for

impoverished learning rate adaption in currently depressed patients. However, pupil dilation might also be modulated by other factors than norepinephrine activity. Pharmacological studies, which manipulate central norepinephrine activity would be particularly useful to examine the causal effect of norepinephrine on learning adaptation to environmental volatility. We examine this further in Chapter 0.

There are a number of limitations to the current study. First, in the current design, the stable schedule was set at a fixed probability of 50%. Although the probability didn't change over a block, it is also unlearnable. Therefore, low learning rates could be because the schedule is stable and recent events are not informative, but it could also be because the outcomes were random and therefore not informative (i.e., expected uncertainty was high when unexpected uncertainty was low). This issue can be addressed by using stable periods in which the probability is not set at 50% (i.e., so unexpected uncertainty remains low but expected uncertainty is not raised). We describe a modification of the bias learning task that uses such a schedule, and that includes a both stable block in Chapters 3. Second, although we found some evidence that the loss learning rate adaption estimates from the computational model were associated with some symptoms of depression, the current study was designed to look at group difference and is not well suited to answer questions of symptom correlation. To address the above-mentioned limitations, we conducted a second study which we will discuss in detailed in the next Chapter.

In conclusion, we found evidence that currently depressed patients, relative to both previously and never depressed groups, show reduced adjustment of loss learning rates during a mixed reward/loss learning paradigm. This suggests that the difficulty in adjusting learning rates to the volatility of negative outcomes is a state rather than trait dependent facet of depression.

3 NEURAL MECHANISMS UNDERLYING INFORMATION LEARNING BIAS IN DEPRESSION

3.1 Introduction

The ability to flexibly adjust learning according to the volatility of the environment is important if one hopes to learn efficiently. Studies have shown that healthy human participants can adjust their learning rates optimally depending on the level of volatility of the environment (Behrens et al. 2007; Behrens et al. 2008; Pulcu and Browning 2017). On the other hand, failing to adapt learning to volatile environments might lead to emotional distress (De Berker et al. 2016). A recent study showed that momentary happiness was modulated by prediction errors, and crucially, people with higher depressive symptoms reported feeling less happy during a volatile environment but not a stable environment (Blain and Rutledge 2020). Together, these studies suggest that impaired learning under volatile environments might be associated with depression. Reduced punishment/loss learning rate adaptation between stable and volatile environments has been associated with high trait or social anxiety (Browning et al. 2015; Beltzer et al. 2019; Pulcu and Browning 2017). In Chapter 2, we examined this hypothesis in a clinically depressed population, and found that currently depressed patients, relative to both previously and never depressed groups, showed impaired ability to adjust learning rates to the volatility of negative outcomes. However, the neural processes underlying this effect are not clear, nor is it clear whether difficulty in adapting learning to outcome volatility is associated with individual differences in depression symptoms in a subclinical cohort (as opposed to clinical status).

Previous human and animal studies have highlighted the role of the medial prefrontal cortex (mPFC), including the anterior cingulate cortex (ACC), the orbital frontal cortex (OFC) and the ventral medial prefrontal cortex (vmPFC), in flexible learning and decision making to rewards (Behrens et al. 2007; Behrens et al. 2008; Massi, Donahue, and Lee 2018; Findling et al. 2019). More specifically, elevated activation of the ACC when learning in volatile environments has been reported both in human and primate studies (Meder et al. 2017; Behrens et al. 2008; Massi,

Donahue, and Lee 2018). Critically, in a recent study, Massi et. al, recorded single-neuron activity from the PFC in primates while the animals performed a probabilistic reversal learning task under volatile and stable environments. They showed that volatility enhanced the neural representations of recent outcomes in the PFC, including OFC and ACC (Massi, Donahue, and Lee 2018). These results suggested that volatility modulated the outcome signals represented in the brain, particularly in the mPFC.

Importantly, a rich body of research suggests that the structural and functional abnormalities in the mPFC play a crucial role in the etiology of depression (Keedwell et al. 2005; Biselli et al. 2021). A recent study showed that value signals in the vmPFC were lower in depressive patients with suicidal ideation or attempts compared to healthy controls in a volatile reward reinforcement learning task (Brown et al. 2020). However, it is not clear whether the mPFC is involved in learning adaptations to loss volatility, and if so, whether this effect is associated with depression.

To examine whether learning rate adaptation to changes in volatility is associated with individual differences in depression and study the underlying neuronal mechanism, we recruited a group of participants with a spectrum of depression symptoms. Participants performed an adapted version of the information bias learning task used in the Chapter 2. At the same time, blood oxygenation level dependent (BOLD) signal was acquired using functional magnetic resonance imaging (fMRI) while participants performed the task.

We hypothesized that participants with higher depression scores would have more difficulty in adapting to loss volatility, and that this might be accompanied by aberrant modulation of volatility on outcome signals represented in the brain, particularly for the mPFC.

3.2 Methods

3.2.1 Experiment task and participants

3.2.1.1 Participants

We recruited 30 healthy participants (14 females) with a range of depression symptoms to this study. One participant was excluded from all analyses because he chose the same shape for all the trials in one of the blocks.

Participants were asked to complete a version of the 16 items self-report version of the Quick Inventory of Depressive Symptoms (QIDS) online, which measures symptoms of depression (Rush et al. 2003). We then invited participants with a range of QIDS scores to participate in the fMRI session (i.e., by selecting a roughly equal number of participants with scores 1-5, 6-10, 11-15). Participants were also selected to ensure that age and gender did not confound the

QIDS score (i.e., selection ensured lack of significant relationship between QIDS scores and either age or gender).

During the fMRI session, participants were asked to complete the QIDS again as well as three other questionnaires; (1) the Ruminative Response Scale, which measures rumination in response to depressive mood (RRS) (Nolen-Hoeksema and Morrow 1991), (2) the Spielberger State-Trait Anxiety Inventory (STAI), which measures state and trait anxiety (Spielberger 1983), (3) The Temporal Experience of Pleasure Scale (TEPS), which measures individual trait dispositions in anticipatory and consummatory experiences of pleasure (Gard et al. 2006). These questionnaires measure different aspects of depressive experience but are generally highly correlated with each other and with the QIDS. To account for this, we additionally ran a factor analysis to derive an overall quantification of negative symptoms, using the four total questionnaire scores taken on the scan day.

3.2.1.2 The information bias learning task

The learning task used in this study was based on that described in Chapter 2, with a few important changes made to allow us to investigate how volatility adaptation differs as a function of the other-volatility levels as well as allowing us to differentiate between expected and unexpected uncertainty (i.e. the effect of win volatility on loss learning and vice versa, for more details see the discussion section (2.4) in the previous chapter).

The first change made was that the "stable" schedule of the task was not fixed at 50% anymore, but had a probability change somewhere in the middle of the block either from 15% to 85% or from 80% to 20%. This change was made to reduce the difference in expected uncertainty between the volatile and stable schedules. Second, we included a both stable block (i.e. in which both win and loss outcomes were relatively stable). This resulted in 4 different block types in this version of task (see Figure 3.1A): (1) a "both volatile" block, in which both win and loss outcomes were volatile, with associations between the outcomes and the stimuli switching frequently between 15 and 85%, (2) a "win volatile" block, in which the association with win outcomes was "volatile", whereas the association with loss outcomes had a single probability change between 80% and 20%. (3) a "loss volatile" block, in which the loss outcomes were volatile and win outcomes stable, (4) a "both stable" block, in which both win and loss outcomes were stable. Linked to this change, the order of the blocks was randomised across participants (i.e., rather than always starting with the both-volatile block). As can be seen, this allows us to test the valence*self-volatility*other-volatility effect in our analysis.

We made some further changes to the task so that it was suitable for fMRI data analysis. As before, the choice phase was self-paced. Following this choice, a black frame would appear around the shape that was chosen. The choice phase was limited to between 1s and 5s. If a

participant responded before 1 second, the black frame appeared 1 second after the onset of the choice phase. If a participant did not respond within 5 seconds, a random choice was made for the participant for this trial. Participants were informed that the money won or lost from the random choice would be included in the final payment. The chosen option was marked with a black frame for a jittered interval of 1-5 (mean 3) seconds. The win and loss outcomes were shown at the same time for each trial for a fixed duration of 3 seconds. Finally, the trials were separated with a jittered interval of 1–6 (mean 3.5) seconds (see Figure 3.1B). There were no auditory cues played for this version of the task.

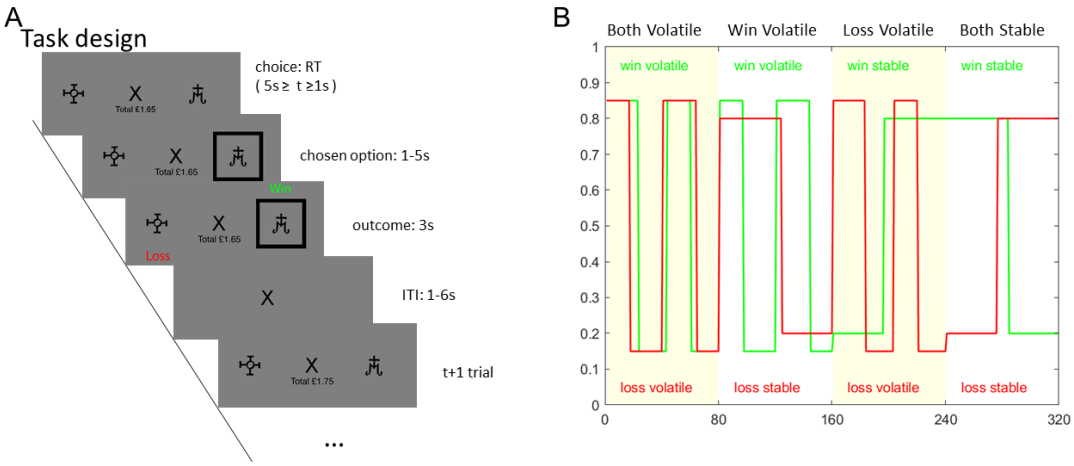

Figure 3.1 Task design

(A) Timeline of one trial from the information bias learning task. (B) Overall task structure. The task consisted of 4 blocks of 80 trials. The y-axis represents the probability, p, that an outcome (win in green or loss in red) would be associated with shape 'A' (1-p for shape 'B'). The blocks differed in how volatile (changeable) the outcome probabilities were. Block order was counterbalanced across participants.

All participants first completed a training version of the both-volatile block (120 trials) with the same timing design as the fMRI task to families them with the task before entering the scanner. They then performed the 4 different blocks (80 trials each) of the task in the MRI scanner with rest-sessions in between. The block order was counterbalanced across participants.

3.2.2 Behavioral data analysis and computational modelling

3.2.2.1.1 Task performance

The model-free behavioural data were analysed using MATLAB (version R2018b; MathWorks, MA, USA) and SPSS (version 26.0; IBM, NY, USA). The percentages of wins chosen and losses not-chosen were calculated for each block. As proportion data, such as percentages, violate the assumptions of equality of variance required by most parametric tests, the percentage data were arcsine square-root transformed before being submitted to statistical tests.

We first examined whether the performance for each block was significantly above chance level (i.e., 50%). We then ran a 2*2*2 Repeated measures ANOVA, with valence (win/loss), self-volatility ("self-volatile"/"self-stable", e.g. for win, the both volatile and the win-volatile block were self(win)-volatile, whereas the loss-volatile block and the both-stable block were self(win)-stable.), and other-volatility ("other volatile"/"other-stable", e.g. for win, the both volatile and the loss-volatile block were other(loss)-volatile, whereas the win-volatile block and the both-stable block were other(loss)-stable.) as repeated measures.

3.2.2.1.2 Positive-stay-negative-switch

As described in the previous chapter, switch probabilities after receiving a loss (negative outcomes for loss, $P_{switch\text{-}loss}$), not receiving a win (negative outcomes for win, $P_{switch\text{-}nowin}$), not receiving a loss (positive outcomes for loss, $P_{switch\text{-}noloss}$), and receiving a win (positive outcome for win, $P_{switch\text{-}win}$) for each block were calculated. The positive-stay-negative-switch (P_{psns}) indexes were calculated by subtracting probabilities of switching after receiving a positive outcome from switch probabilities after receiving a negative outcome (acrsin square root transformed) for win and loss respectively, see 2.2.2.2 for more details.

The positive-stay-negative-switch indexes were then submitted to a 2*2*2 Repeated measures ANOVA, with valence (win/loss), self-volatility ("self-volatile"/"self-stable", i.e. the win volatility for win outcomes and loss volatility for loss outcomes), and other-volatility ("other volatile"/"other-stable", i.e the win volatility for losses and vice versa) as repeated measures.

3.2.2.2 Computational models

Volatility adaptation model:

We developed the volatility adaptation model described in Chapter 2 (see 2.2.3 in page 27) slightly to better capture the new block design of the fMRI study. Specifically, in order to capture the effect of the other volatility on learning rates, we fitted separate stable and volatile adaptation learning rates for trials in which the other outcome was volatile vs. stable (see Table 4).

Table 4 Equations for the volatility adaptation model for the fMRI study

block	valence	Description	equation
Both Volatile	win	Win volatile learning rate	$\alpha\text{win}(1)=\alpha\text{win}_{\text{stable2}}+\alpha\text{win}_{\text{adpt2}}$
	loss	Loss volatile learning rate	$\alpha\text{loss}(1)=\alpha\text{loss}_{\text{stable2}}+\alpha\text{loss}_{\text{adpt2}}$
Win Volatile	win	Win volatile learning rate	$\alpha\text{win}(2)=\alpha\text{win}_{\text{stable1}}+\alpha\text{win}_{\text{adpt1}}$
	loss	Loss stable learning rate	$\alpha\text{loss}(2)=\alpha\text{loss}_{\text{stable2}}$
Loss Volatile	win	Win stable learning rate	$\alpha\text{win}(3)=\alpha\text{win}_{\text{stable2}}$
	loss	Loss volatile learning rate	$\alpha\text{loss}(3)=\alpha\text{loss}_{\text{stable1}}+\alpha\text{loss}_{\text{adpt1}}$
Both Stable	win	Win stable learning rate	$\alpha\text{win}(4)=\alpha\text{win}_{\text{stable1}}$
	loss	Loss stable learning rate	$\alpha\text{loss}(4)=\alpha\text{loss}_{\text{stable1}}$

Note: Stable 1 refers to stable parameters when other-volatility was stable; Stable 2 refers to stable parameters when other-volatility was volatile; adpt1 refer to adaptation terms when other-volatility was stable; adpt2 refer to adaptation terms when other-volatility was volatile;

Equation 16 :

$$\boldsymbol{Qwin_{t+1}} = \boldsymbol{Qwin_t} + \boldsymbol{\alpha win(block)} \cdot (\boldsymbol{Rwin_t} - \boldsymbol{Qwin_t})$$

Equation 17 :

$$\boldsymbol{Qloss_{t+1}} = \boldsymbol{Qloss_t} + \boldsymbol{\alpha loss(block)} \cdot (\boldsymbol{Rloss_t} - \boldsymbol{Qloss_t})$$

Equation 18 :

$$Pchoice_t = \frac{1}{1+exp^{-(\beta(block)\cdot(Qwin_t - Qloss_t)}}$$

The Expected values for win and loss were updated trial by trial with block-wise win and loss learning rate respectively using a Rescorla-Wagner learning rule(Rescorla 1972)(Equation 16, Equation 17). The Q value was set to 0.5 at the beginning of each block. These Expected values were then transformed into a single choice probability using a soft max function (Equation 18). Because the task difficulties for each block were very different in this version of the task, particularly between the "both volatile" block and the "both stable" block, we would expect the inverse temperatures might reflect the differences. Therefore, in this analysis, we fitted different inverse temperatures were fitted for each block. The first 5 trails of each block were not used in model fitting.

Model were fitted using Stan (RStan v 2.21.1, R v 3.6.0), a software package that uses Markov chain Monte Carlo method to sample from posterior distributions of parameters(Carpenter et

al. 2017). 4 parallel Monte Carlo chains were run each with 5000 iterations (5000 warm-ups), using adapt_delta of 0.99 and maximum treedepth of 10. Learning rate parameters were sampled from a zero mean normal distribution with standard deviation of 2, and then transformed using logistic approximation (Bowling et al. 2009). Beta parameters were sampled from a zero mean normal distribution with standard deviation 1, and then transformed using the exp function. Convergence was inspected using shinyStan (v 2.5.0). No convergence problems were detected. We adopted a stringent cut-off of R-hat< 1.1 as an indicator for sufficient convergence (Brooks and Gelman 1998) , and the minimum number of effective samples >50. The same procedures were also used to estimate parameters for the beta adaptation model and the fixed learning model (described in detail in 2.2.3).

We then performed model comparisons using a leave-one-out cross-validation procedure using the LOO package for R (Vehtari, Gelman, and Gabry 2017). An expected log pointwise predictive density (ELPD) and a standard error for each ELPD were estimated for each model.

To test the effect of self-volatility on learning rate adaptation (i.e. learning rates should be higher when an outcome is volatile), we ran one-sample t-tests to check whether the learning rate adaptation terms were significantly higher than zero.

We then ran a 2*2 repeated measures ANOVAs, with valence and other-volatility as repeated measures, for the learning rate adaptations and the base learning rates respectively. Note that, the learning rate adaptation terms are the degree to which learning rates change from stable to volatile blocks and thus, the analysis of these terms is equivalent to testing the valence * self-volatility * other-volatility effects of the choice data.

We then ran a 2*2 rm ANOVA for the inverse temperatures estimates, with win-volatility (win-volatile vs. win-stable) and loss volatility (loss-volatile vs. loss-stable) as repeated measures, to see if there was any modulation effect of volatility on participants' choice consistency.

To examine how the overall negativity scores (i.e., overall depression and anxiety levels, which were derived using factor analysis to various depression and anxiety questionnaires, see 3.3.1 for more details) might modulate these terms, we ran Pearson correlation analyses between parameter estimates and the overall negativity scores. To further characterise the behaviour of participants we assessed the degree of loss learning rate adaptation on participants with high vs. low negativity (calculated using a median split the participants into the high (n=15) and the low (n=14) negativity groups).

3.2.3 fMRI data acquisition and analysis

3.2.3.1 fMRI data acquisition

All image data was acquired using a Siemens Prisma 3 T scanner with a 32-channel head coil. The task-based fMRI images were acquired using a multiband-echo planar imaging (mb-EPI) sequence, with a multiband factor of 6, repetition time (TR) = 800 ms, echo time (TE) = 30 ms, 60 slices (no gap), a slice angle of 30°, interleaved acquisition, voxel size = 2.4 × 2.4 × 2.4 mm, flip angle = 52°, and field of view (FOV) = 216 mm, with local z-shimming). Field maps were obtained using a dual echo 2D gradient-echo sequence with TR = 590 ms, TE = 7.38 ms, and 4.92ms, and flip angle = 46°. The high-resolution T1-weighted anatomical MRI images were acquired using a 3D MPRAGE sequence with TR = 1900 ms, TE = 3.97 ms, 192 transversal slices, voxel size = 1 ×1 × 1 mm, and FOV = 192 mm.

3.2.3.2 fMRI data pre-processing

fMRI data were pre-processed using FEAT (FMRI Expert Analysis Tool) part of FSL (FMRIB Software Library V6 (Analysis Group, FMRIB, Oxford, UK; http://www.fmrib.ox.ac.uk/fsl)(Jenkinson et al. 2012). The procedures included (1) motion correction to a high-resolution functional image acquired at the beginning of each block using MCFLIRT (Jenkinson et al. 2002), (2) spatial correction for magnetic field inhomogeneities estimated by the B0 fieldmap (Jenkinson 2004), (3) brain extraction using BET (Smith 2002), (4) spatial smoothing using a Gaussian kernel (FWHM 5 mm), (5) high-pass temporal filtered (90 s cut-off). Then functional images were registered to both the high resolution T1w structural scan and standard space using FLIRT.

3.2.3.1 First-level and group-level analysis

The first-level and group-level GLM whole-brain and ROI-based analysis were conducted using FEAT (note, functionally defined ROIs were used, and are described below). We constructed the event-based general linear model (GLM) with the following parametric regressors at choice and outcome onset for each block. At choice onset: (1) the response (coded as 1 and -1, for the right or left response, respectively), (2) choice reaction time, and (3) the trial-by-trial decision (coded as 1 if subject switched away from the previous choice or -1 if she stayed with the same option), (4) choice mean: value set at 1. At outcome onset: (5) win outcome (coded as 1 and -1, for the win received or win not-received, respectively.), (6) loss outcome (coded as 1 and -1, for the loss received or loss not-received, respectively.), (4) outcome mean: value set at 1. Note that the trials to which participants failed to respond in time were coded using the mean of each regressor. (Note that here we used actual outcome results

but not PEs to avoid the confound of model estimates. However, we have also run a similar analysis with PEs instead of outcomes. The results were similar.) The regressors were then demeaned before entering into Feat. Standard motion parameters estimated using MCFLIRT motion correction and physiological noise regressors estimated using Physiological Noise Modelling (PNM) tool (Brooks et al. 2008) were also added in the GLM to regress out the effects of head motion and physiological noise respectively.

We first identified regions that might represent outcome signals in the brain. We estimated the mean of the parameter estimates for the contrast for positive outcomes (i.e. receiving a win and not receiving a loss) vs. negative outcomes (i.e. not receiving a win and receiving a loss) across all the 4 blocks at the individual level and then brought them to the whole-brain group-level analysis (one sample t-test). The images were cluster-thresholded at $z > 4.6$, and clusters shown survived a statistical test for extent ($p < 0.05$, fully corrected for multiple comparisons). We also did one sample t-tests for mean parameter estimates of win outcome (i.e., win vs. no-win) and loss outcome (i.e., no-loss vs. loss) across all the 4 blocks respectively, to examine the neural responses to win and loss outcome. For this follow up analyses, the images were cluster-thresholded at $z > 3.1$ and cluster-wise $p < 0.05$.

To examine whether these outcome signals in the brain were modulated by valence, self-volatility, or other volatility and how overall negativity modulated the effects, we extracted the parameter estimates for the clusters identified by the positive vs. negative outcomes contrast described above, for each block, for win and loss outcomes and for each participant respectively. An ANCOVA, with valence, self-volatility and other-volatility as repeated measures and the overall negativity scores as a covariate, were performed for each region. We follow up the significant effects with paired t-tests. For any significant effect of the negativity scores, we perform Pearson correlations to further examine the covariations.

A mediation analysis using the SOBEL test with permutations (n=5000) was performed in the SPSS (Preacher and Hayes 2004) to further examine the relationships amongst the neural signals in the dACC, learning rate estimates and the negativity traits. In the analysis the win vs. loss learning rate adaptation differences was set as the mediation factor, and the relative adaption to win vs. loss volatility signals in the dACC and the overall negativity scores as independent and dependent variables respectively.

3.3 Results

3.3.1 Questionnaires

We first assessed the relationship between participant demographics and questionnaire score (see Table 5 for a summary of the demographic details). No significant correlations between age and any of the questionnaires (all $p>0.29$) or relationship between gender and the questionnaire scores (all $p>0.57$) was found.

Table 5 Demographic details of the participants

variables	Mean	SD
Demographic information:		
Number of participants	29(14 females)	
Age(yrs)	26.34	7.09
Questionnaires:		
QIDS	8.59	5.23
STAI	86.00	20.35
TEPS	77.34	11.83
RRS	47.00	13.61

Note: QIDS: the Quick Inventory of Depressive Symptoms. STAI: Spielberger State-Trait Anxiety Inventory. TEPS: the Temporal Experience of Pleasure Scale. RRS: the Ruminative Response Scale. SD: standard deviation.

We then examined the correlations amongst the questionnaires taken on the scan day. As expected, the analyses revealed significant correlations across the questionnaires (see Figure 3.2), suggesting the presence of a common factor influencing scores.

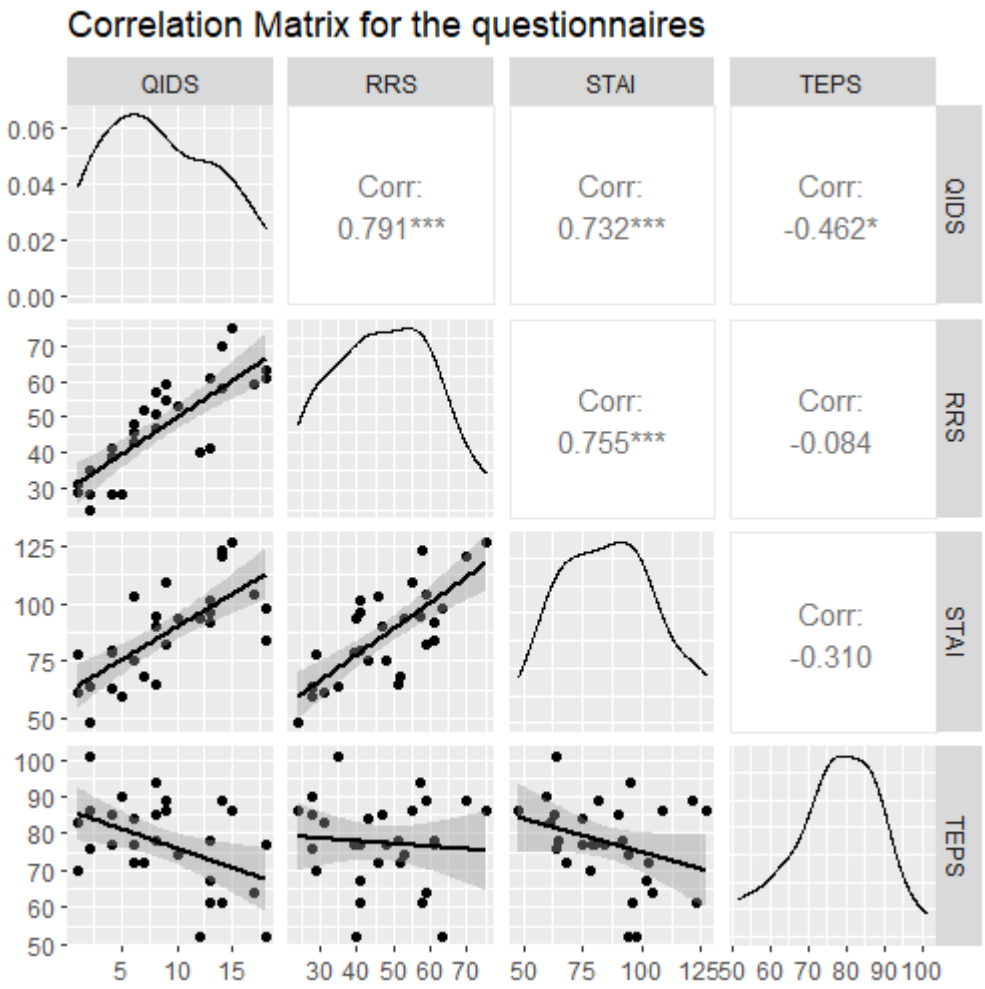

Figure 3.2 Correlation Matrix for the questionnaires**.**

To generate an estimate of this common factor we used a factor analysis. The Scree plot (see Figure 3.3) from the factor analysis showed that only one factor had an eigenvalue higher than 1, suggesting that the first factor alone explained most of the variance (Hayton, Allen, and Scarpello 2004). Therefore, we took the first factor scores as the summary measures of the participants' overall negativity level.

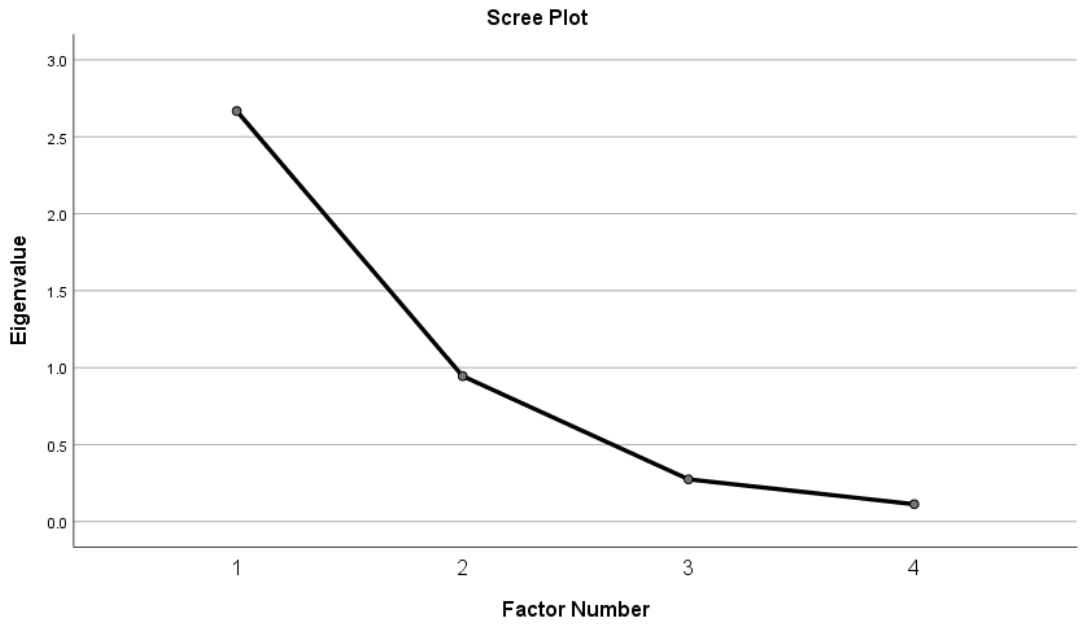

Figure 3.3 the Scree Plot of the factor analysis.

3.3.2 Behavior results

3.3.2.1 Task Performance

We next examined task performance in terms of the percentage of trials in which participants received the win and loss outcomes. One-sample t-tests against the chance level (50%) showed that all scores for both wins and losses were significant above chance level (all $t(28)>4$, $p<.001$), except for the performance for losses in the win-volatile block ($t(28)=-1.258$, $p=.219$) and the performance for the wins in the loss-volatile block ($t(28)=-2.023$, $p=.053$) (see Figure 3.4).

The valence*self-volatility*other-volatility ANOVA analysis revealed a significant main effects of self-volatility ($F(1,28)=32.52$, $p<.001$, $\eta_p^2=.537$) and an effect of other-volatility ($F(1,28)=105.55$, $p<.001$, $\eta_p^2=.790$), and a significant interaction effect of self-volatility * other volatility ($F(1,28)=102.10$, $p<.001$, $\eta_p^2=.785$) (see Figure 3.4). The results suggested that in general participants selected the best options both when outcomes were volatile and when they were stable, but in semi-volatile blocks (i.e. when one outcome was volatile and the other stable), participants relied mostly on the volatile outcomes rather than the stable one to make choices.

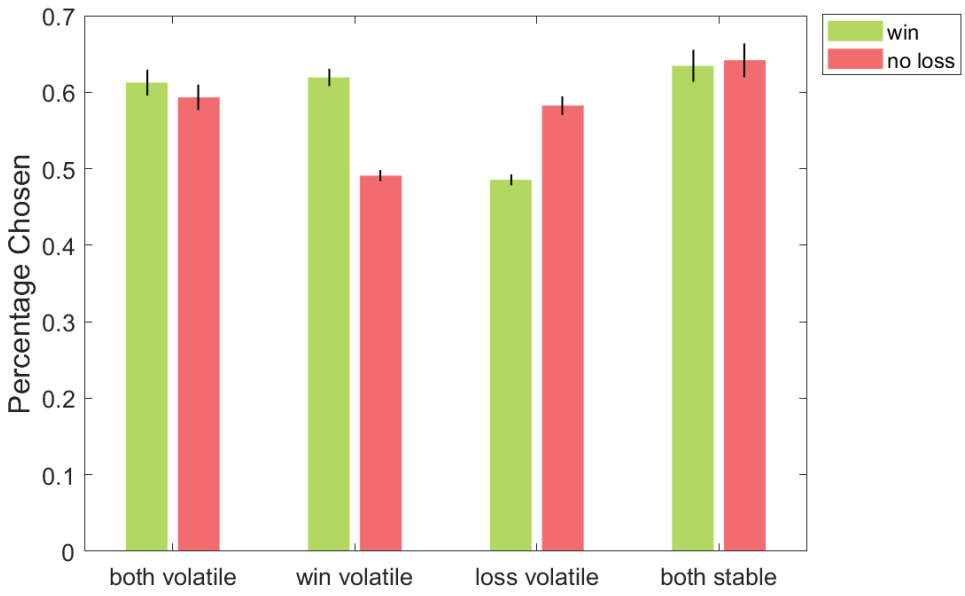

Figure 3.4 Summary of task performance for each block for all the participants.
Error bars represent ±1 standard error of the mean. * $p<0.05$

3.3.2.2 Positive-stay and negative-switch behaviours

We ran a valence*self-volatility*other-volatility ANOVA on the positive-stay, negative-switch (Ppsns) index, which measures how a participant's choices were influenced by the most recent feedback. This analysis revealed a significant main effect of valence ($F(1,28)=5.93$, $p=.022$, $\eta_p2=.175$), indicating that participants were generally more sensitive to win than to loss outcome (see Figure 3.5). The results also showed a significant main effect of self-volatility ($F(1,28)=47.39$, $p<.001$, $\eta_p2=.629$), indicating that, as predicted, participants were more influenced by outcomes of volatile than stable associations, and a significant main effect of other-volatility ($F(1,28)=5.76$, $p=.023$, $\eta_p2=.170$), showing that overall the other outcome being volatile reduced the influence of the current outcome.

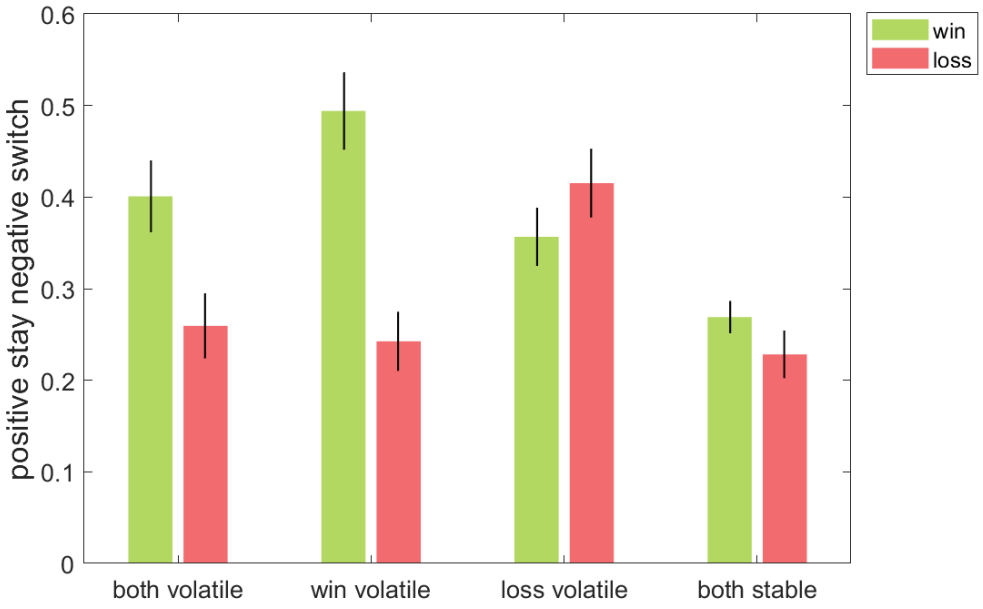

Figure 3.5 positive stay negative switch behaviour summaries for each block across all participants.

Error bars represent ±1 standard error of the mean.

Interestingly, the analysis also revealed a significant self-volatility by other-volatility interaction ($F(1,28)=25.24$, $p<.001$, $\eta_p2=.474$). To further examine this effect, we ran separate valence*self-volatility ANOVAs for when the other-volatility was volatile and when the other-volatility was stable. These analyses revealed a robust effect of self-volatility when the other-volatility was stable ($F(1,28)=68.17$, $p<.001$, $\eta_p2=.709$), but not when the other-volatility was volatile ($p=.214$), suggesting that the overall environment being more volatile hindered participant's volatility adaptations for both wins and losses (see Figure 3.6, in which we rearranged Figure 3.5 to show the effect of other-volatility more clearly).

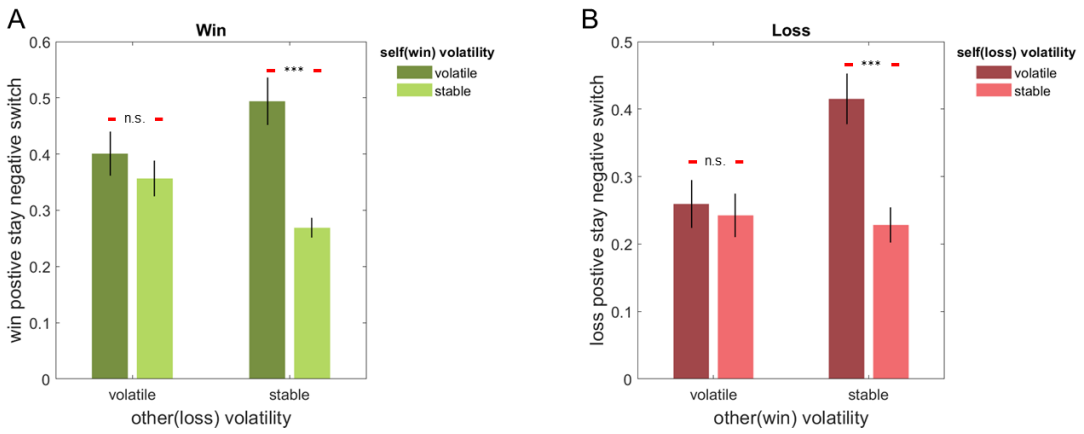

Figure 3.6 positive stay negative switch behaviour summaries showing the self-volatility and other-volatility interaction.

(A)Win Ppsns and (B)loss Ppsns for each of the four task blocks.

Error bars represent ±1 standard error of the mean. *** p<.001 n.s.: not significant

3.3.2.3 Modelling results

We fitted a volatility adaptation model with independent win and loss volatility adaptation terms, which captured the self-volatility adaptations (i.e., self-volatile vs. self-stable), for when the other-volatility was volatile and stable respectively (see Figure 3.7 for the win and loss learning rates estimates from this model for each task block).

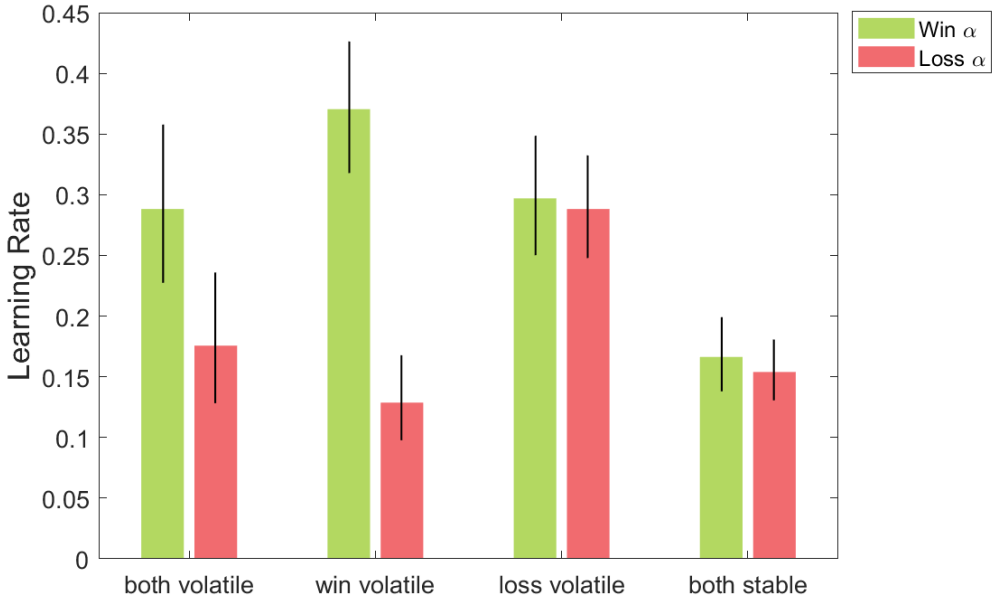

Figure 3.7 win and loss learning rates estimates for the 4 blocks from the volatility adaptation model.

Error bars represent ±1 standard error of the mean.

3.3.2.3.1 Learning rate adaptation to volatility

We first examined whether participants overall increased their learning rates under volatile environments. One sample t-tests on the volatility learning rate adaptation terms showed the same result pattern as the Ppsns indexes shown earlier. More specifically, learning rate adaptations were significant above zero when the other-volatility was stable for both wins ($t(28)=3.92$, $p=.001$) and losses ($t(28)=3.76$, $p=.001$), but not when other-volatility was volatile (both $p>.23$) (see Figure 3.8). The valence*other-volatility ANOVA on the learning rate volatility adaptation terms confirmed a main effect of other-volatility ($F(1,28)=8.91$, $p=.006$, $\eta_p2=.241$), and a marginally significant valence *other-volatility interaction ($F(1,28)=3.53$, $p=.071$, $\eta_p2=.112$). Post-hoc analysis revealed that other-volatility significantly reduced self-volatility adaptation for win ($t(28)=3.32$, $p=.002$) but not for loss ($p=.220$). The results suggest that the other-volatility diminished the self-volatility adaptation for both win and loss, however, loss was slightly less subject to the effect than win.

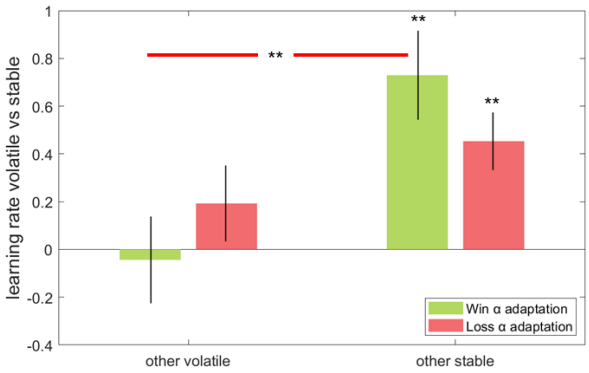

Figure 3.8 win and loss learning rate volatility adaptation estimates under different environments.

Error bars represent ±1 standard error of the mean. ** p<.01

We then examined whether there were any correlations between the learning rate adaptation estimates and participants' overall negativity scores. From the results of the Chapter 2, we

hypothesised that the overall negativity scores might be negatively correlated with loss learning rate adaptation estimates. Indeed, we found a significant negative correlation (r(29)=-0.37, p=.045) between the overall negativity scores and the mean loss learning rate adaptations (see Figure 3.9 A). This appeared to be driven by the correlation between the negativity scores and the loss learning rate adaption when win was volatile (r(29)=-0.39, p=.037, see Figure 3.9 B left) but not when win was stable (r(29)=-.081, p=.677, see Figure 3.9 C left). These results suggested that participants with higher depression and anxiety had more difficulty adapting to the loss volatility level particularly when win was volatile. To further illustrate this effect, we performed a median split on participants (into high negativity and low negativity groups) and then tested whether they showed significant loss volatility adaptation when win was volatile and stable respectively. The one sample t-tests revealed that while the low negativity group showed significant loss learning rate adaptation under both environments (both p<.048), the high negativity group only showed significant loss adaptation when win was stable (p=.023), but not when win was volatile (p=.577) (see Figure 3.9 B&C right). These results suggest that people with high negativity scores have more difficulty adapting their learning when there was more than one source of volatility in the environment.

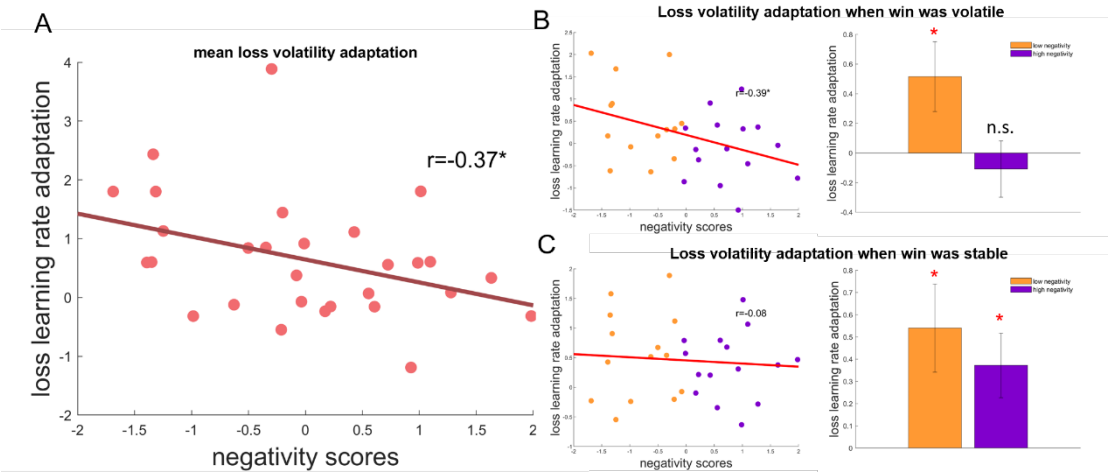

Figure 3.9 loss learning rate adaptations negatively correlated with overall negativity scores.

(A) a scatter plot showing the negative correlation between the overall loss learning rate volatility adaptation (mean across when win volatile and stable conditions) and the overall negativity scores.

(B) Left: a scatter plot showing the negative correlation between the loss learning rate volatility adaptation when win was volatile and the overall negativity scores. Right: a bar graph showing loss volatility adaptation when win was volatile for the high (purple) and the low (orange) negativity groups respectively.

(C) Left: a scatter plot showing that there was no significant correlation between the loss learning rate volatility adaptation when win was stable and the overall negativity scores. Right: a bar graph showing loss volatility adaptation when win was stable for the high (purple) and the low (orange) negativity groups respectively.
* p<0.05, n.s. not significant.

3.3.2.3.2 Stable learning rates

Next, we examined the effect of other-volatility on the base learning rates (i.e. when the "self-volatility" was low) for win and loss. The valence*other-volatility ANOVA revealed a significant main effect of valence ($F(1,28)=4.62$, $p=.040$, $\eta_p2=.142$), and a marginally significant valence * other-volatility interaction ($F(1,28)=4.13$, $p=.052$, $\eta_p2=.128$). Post-hoc analysis revealed that the win base learning rates were higher when loss was volatile than loss stable ($t(28)=2.31$, $p=.028$), while the loss base learning rates were less subject to the influence of win volatility ($p=.555$) (see Figure 3.10). There was no significant correlation between the overall depression scores and any of the base learning rates (all $p>.656$).

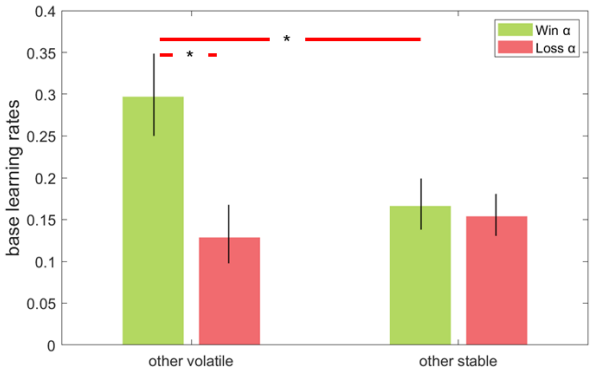

Figure 3.10 The effect of the volatility of the other outcomes on win and loss base learning rates (i.e., estimated learning rate when an outcome is stable).
Error bars represent ±1 standard error of the mean. * p<.05

3.3.2.3.3 The inverse temperature

The win-volatility*loss-volatility ANOVA analysis revealed a significant main effect of loss-volatility ($F(1,28)=4.41$, $p=.045$, $\eta_p2=.136$) and a significant win-volatility and loss-volatility interaction($F(1,28)=35.41$, $p<.001$, $\eta_p2=.558$). Post-hoc paired t-tests showed reduced inverse

temperature in the semi-volatile blocks compared to the both-volatile block and the both-stable block (see Figure 3.11), suggesting the choice consistency was reduced in the semi-volatile blocks. However, there was no significant correlation between the overall depression scores and any of the inverse temperatures (all p>.099).

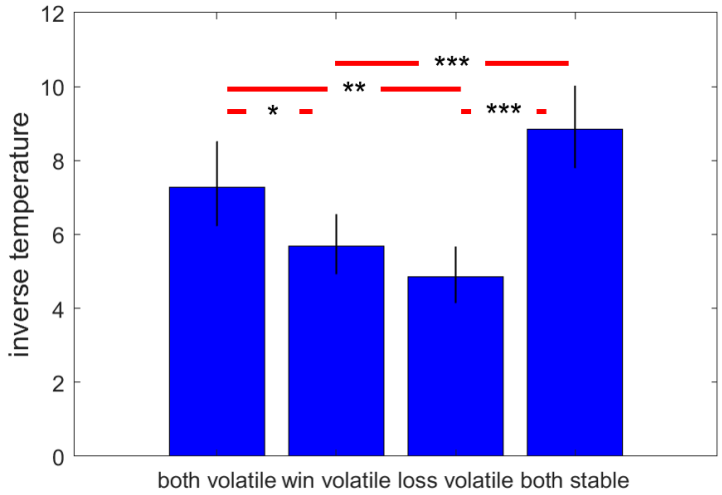

Figure 3.11 the inverse temperature estimates for each block.

Error bars represent ±1 standard error of the mean. * p<.05 **p<.01 ***<.001

3.3.2.3.4 Model comparisons

The learning rate volatility adaptation model had highest predictive accuracy (ELPD=-3555.8, SE=48.7) compared to the beta volatility adaptation model (ELPD=-3720.9, SE=73.0), and the fixed learning model (ELPD=-3865.9, SE=50.0), again suggesting that the learning rate volatility model fitted better to participants' behavior than the other two models.

3.3.3 fMRI results

We first examined the positive vs. negative outcomes contrast, i.e., positive outcomes (win and no-loss) minus negative outcomes (no-win and loss). A classical network of brain regions implicated in positive emotions (Burgdorf and Panksepp 2006) showed increased BOLD responses to positive outcomes (see Figure 3.12 top) compared to negative outcomes, including the ventral striatum (VS), the ventral medial prefrontal cortex (vmPFC), the rostral anterior cingulate cortex (rACC) and the frontal polar cortex (FPC). In contrast, a network of brain regions implicated in negative emotions (Shin and Liberzon 2010) showed increased BOLD responses to negative outcomes, including the dorsal anterior cingulate cortex (dACC), the

insula, the middle frontal gyrus (MFG), and the intra-calcarine cortex. Similar brain regions were shown for the win (see Figure 3.12 middle) and loss outcomes (see Figure 3.12 bottom) respectively.

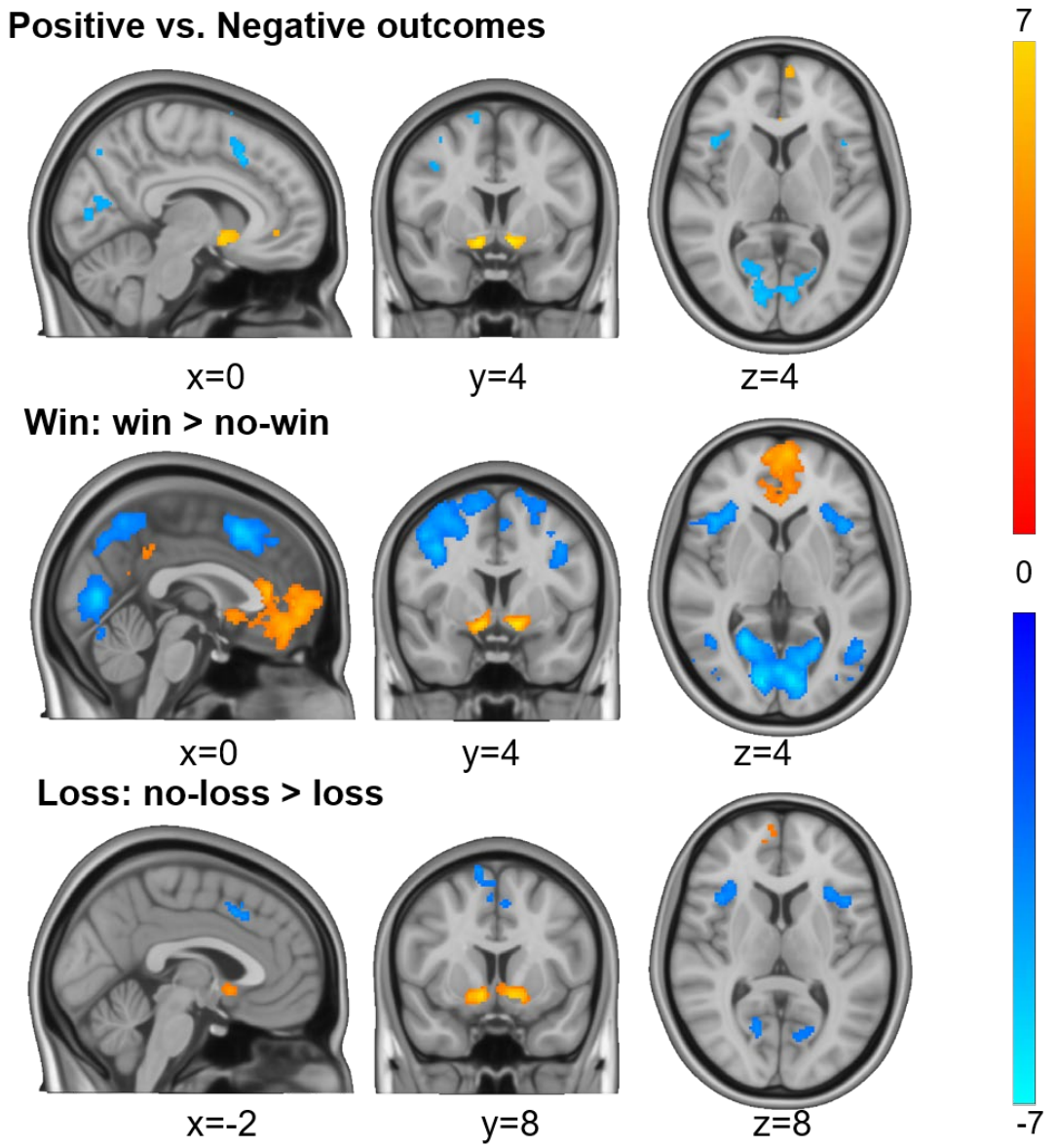

Figure 3.12 the brain regions showing changed BOLD response to outcomes

(top) brain regions obtained for the positive minus negative outcomes contrast, i.e., win and no-loss minus no-win and loss ($z>4.6$, cluster-wise $P<0.05$). (middle) brain regions obtained for the win minus no-win contrast ($z>3.1$, cluster-wise $P<0.05$). (bottom) brain regions obtained for the no-loss minus loss contrast ($z>3.1$, cluster-wise $P<0.05$). Brain coordinates are expressed in Montreal Neurological Institute (MNI) coordinate space.

To examine whether the outcome signals in these regions were modulated by valance (i.e win vs. no win relative to loss vs. no loss), self-volatility, or other-volatility, we extracted parameter

estimates for win and loss outcomes contrasts respectively from each region that showed significant response to outcomes (i.e. for clusters that showed a significant positive vs. negative outcome contrast in either direction, $z>4.6$, cluster-wise $P<0.05$). More specifically, for areas that showed higher activation to positive outcomes than negative outcomes, including the rACC, the vmPFC, the FPC, and the VS, we extracted parameter estimates from the win minus no-win contrast and the no-loss minus loss contrast respectively. For areas that showed higher activation to negative outcomes than positive outcomes, including the dACC, the MFG, the insula and the intra-calcarine cortex, we extracted parameter estimates from the no-win minus win contrast and the loss minus no-loss contrast respectively.

We then ran valence*self-volatility*other-volatility*negativity repeated measures ANCOVAs for each region respectively. The results showed significant main effects of valence in all these regions (all $p<.034$), except for the rACC ($p=.149$) and the VS ($p=.345$), indicating that in general win (relative to no win) outcome signals were stronger than loss (relative to no loss) outcome signals in these regions (see Figure 3.12). This is consistent with the finding introduced in the previous sections (see 3.3.2.2 and 3.3.2.3.2) that the participants' behaviors were generally more influenced by win outcomes than loss outcomes.

Self (but not Other) Volatility Modifies BOLD Response to Task Outcomes

Importantly, the results also revealed main effects of self-volatility in the vmPFC ($F(1,27)=6.48$, $p=.017$, $\eta_p2=.194$) and the rACC ($F(1,27)=9.97$, $p=.004$, $\eta_p2=.270$), and marginally significantly in the dACC ($F(1,27)=3.66$, $p=.066$, $\eta_p2=.119$), the VS ($F(1,27)=3.50$, $p=.072$, $\eta_p2=.115$) and the MFG ($F(1,27)=2.51$, $p=.125$, $\eta_p2=.085$). As can be seen from Figure 3.13 and Figure 3.14 activity related outcome (positively coded in Figure 3.13 and negatively coded in Figure 3.14) was higher when self-volatility was high relative to low (see the second column of each figure). However, none of these 8 ROI regions showed significant main effects of other-volatility or significant modulation of other-volatility on self-volatility (all $p>.261$). These results suggested that self-volatility modulated the outcome signals particularly in the mPFC, but that other-volatility had less influence on the outcome signals in these regions. The self-volatility * valance interactions were not significant in these regions (all $p>.217$), except for the VS ($F(1,27)=5.30$, $p=.029$, $\eta_p2=.164$), suggesting that the self-volatility effects were generally consistent across the two valences. However, the post hoc tests suggested that the self-volatility effects in the vmPFC and rACC was driven more by a volatility effect for win outcomes (rACC $p=.016$, see Figure 3.13 A2; vmPFC: $p=.064$, see Figure 3.13 B2) than loss outcomes. In contrast, the self-volatility effect in the dACC was driven more by a volatility effect for loss outcomes ($p=.030$, see Figure 3.14 A2). Together, these results suggest that regions which show increased activation to positive relative to negative outcomes, such as the rACC and the vmPFC maybe somewhat more sensitive to win volatility, whereas regions which

show increased activation to negative outcomes, such as the dACC maybe more sensitive to loss volatility.

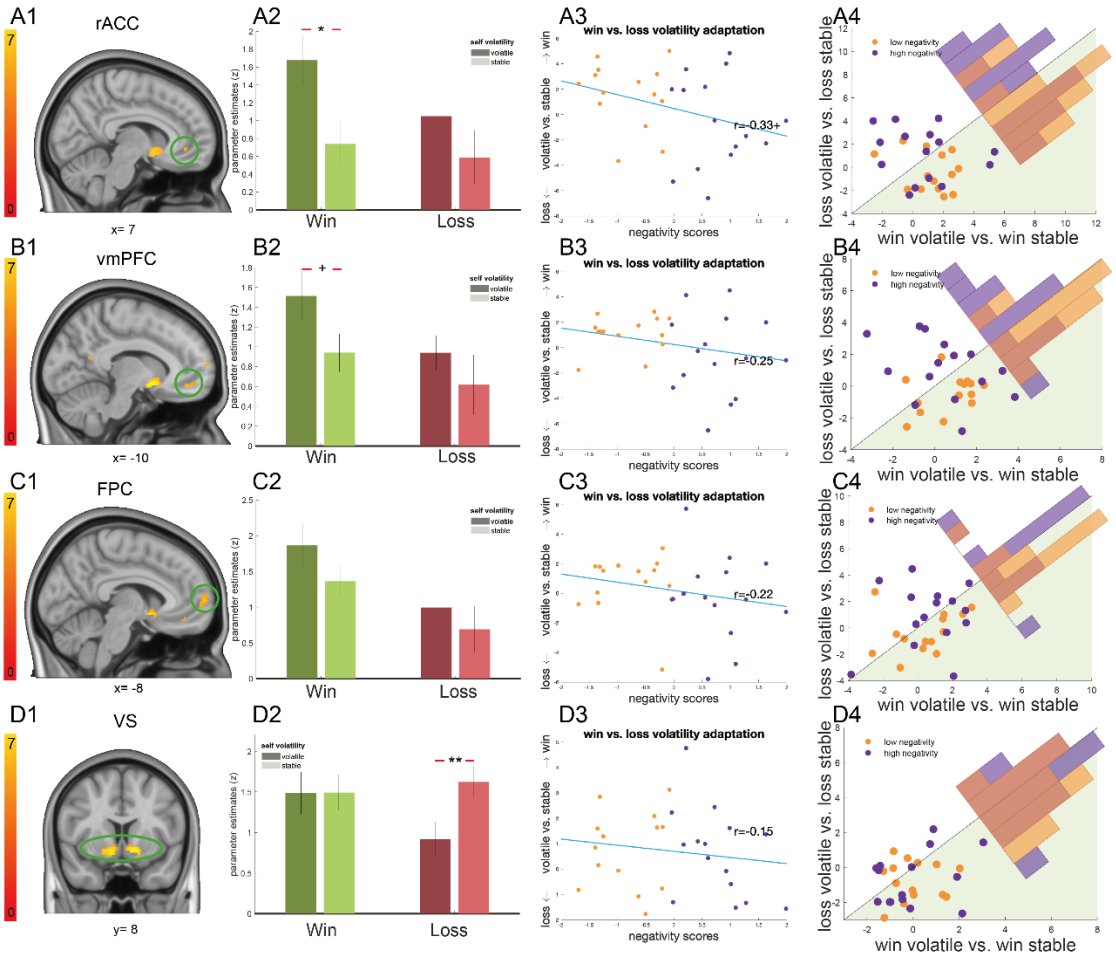

Figure 3.13 Clusters from the whole brain analysis that showed significant higher activation for positive outcomes than negative outcomes.

[A] (1) a sagittal view of the rostral anterior cingulate cortex (rACC) (circled in green). (2) BOLD signals extracted from this region for win and loss self-volatility adaptations. (3) the relative adaption to win vs. loss volatility (win volatile vs. stable – loss volatile vs. stable) in this region marginally negatively correlated with the overall negativity scores. (4) a scatter plot showing win volatility adaptation (x-axis) and loss volatility adaptation (y-axis) in this region for the high (purple) and low (orange) negativity (median split) groups respectively. The low negativity group were mostly below the identity line (areas shown in green), suggesting relatively more win volatility adaptation than loss in this group in this region. The histogram shows the distributions of the differences between win and loss self-volatility adaptations for the high and low group respectively.

[B] (1) a sagittal view of the ventral medial prefrontal cortex (vmPFC) (circled in green). (2) BOLD signals extracted from this region for win and loss self-volatility adaptations. (3) the relative adaptation to win vs. loss volatility (win volatile vs. stable – loss volatile vs. stable) in this region does not correlate with the overall negativity scores. (4) a scatter plot showing win volatility adaptation (x-axis) and loss volatility adaptation (y-axis) in this region for the high (purple) and low (orange) negativity (median split) groups respectively. The low negativity group were mostly below the identity line (areas shown in green), suggesting a relatively more win volatility adaptation than loss in this group in this region. The histogram shows the distributions of the differences between win and loss self-volatility adaptations for the high and low group respectively.

[C] (1) a sagittal view of the frontal polar cortex (FPC) (circled in green). (2) BOLD signals extracted from this region for win and loss self-volatility adaptations. (3) the relative adaptation to win vs. loss volatility (win volatile vs. stable – loss volatile vs. stable) in this region does not correlate with the overall negativity scores. (4) a scatter plot showing win volatility adaptation (x-axis) and loss volatility adaptation (y-axis) in this region for the high (purple) and low (orange) negativity (median split) groups respectively. The histogram shows the distributions of the differences between win and loss self-volatility adaptations for the high and low group respectively.

[D] (1) a coronal view of the ventral striatum (VS) (circled in green). (2) BOLD signals extracted from this region for win and loss self-volatility adaptations. (3) the relative adaptation to win vs. loss volatility (win volatile vs. stable – loss volatile vs. stable) in this region does not correlate with the overall negativity scores. (4) a scatter plot showing win volatility adaptation (x-axis) and loss volatility adaptation (y-axis) in this region for the high (purple) and low (orange) negativity (median split) groups respectively. The histogram shows the distributions of the differences between win and loss self-volatility adaptations for the high and low group respectively.

Error bars represent ±1 standard error of the mean. + p<.10; * p<.05; **p<.01; ***<.001.

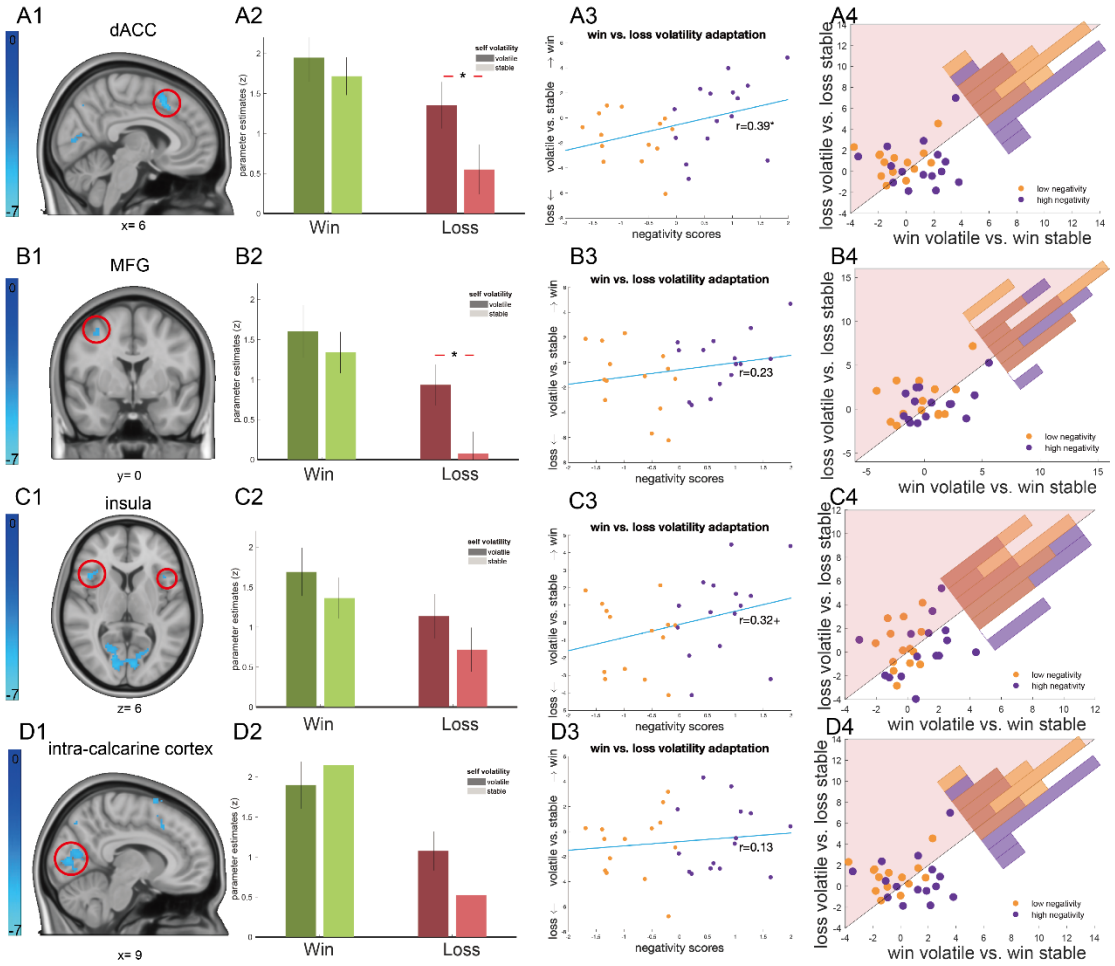

Figure 3.14 Clusters from the whole brain analysis that showed significant stronger deactivation for positive outcomes than negative outcomes.

(A) (1) a sagittal view of the dorsal anterior cingulate cortex (dACC) (circled in red). (2) BOLD signals extracted from this region for win and loss self-volatility adaptations. (3) the relative adaption to win vs. loss volatility (win volatile vs. stable – loss volatile vs. stable) in this region positively correlated with the overall negativity scores. (4) a scatter plot showing win volatility adaptation (x-axis) and loss volatility adaptation (y-axis) in this region for the high (purple) and low (orange) negativity (median split) groups respectively. The low negativity group were mostly above the identity line (areas shown in red), suggesting a relatively more loss volatility adaptation than loss in this group in this region. The histogram shows the distributions of the differences between win and loss self-volatility adaptations for the high and low group respectively.

(B) (1) a coronal view of the middle frontal gyrus (MFG) (circled in red). (2) BOLD signals extracted from this region for win and loss self-volatility adaptations. (3) the relative adaption to win vs. loss volatility (win volatile vs. stable – loss volatile vs. stable) in this region does not correlate with the overall negativity scores. (4) a scatter plot showing win volatility adaptation (x-axis) and loss volatility adaptation (y-axis) in this region for the high (purple) and low

(orange) negativity (median split) groups respectively. The histogram shows the distributions of the differences between win and loss self-volatility adaptations for the high and low group respectively.
(C) (1) an axial view of the insula (circled in red). (2) BOLD signals extracted from this region for win and loss self-volatility adaptations. (3) the relative adaption to win vs. loss volatility (win volatile vs. stable – loss volatile vs. stable) in this region marginally positively correlated with the overall negativity scores. (4) a scatter plot showing win volatility adaptation (x-axis) and loss volatility adaptation (y-axis) in this region for the high (purple) and low (orange) negativity (median split) groups respectively. The histogram shows the distributions of the differences between win and loss self-volatility adaptations for the high and low group respectively.
(D) (1) a sagittal view of the intra-calcarine cortex (circled in red). (2) BOLD signals extracted from this region for win and loss self-volatility adaptations. (3) the relative adaption to win vs. loss volatility (win volatile vs. stable – loss volatile vs. stable) in this region does not correlate with the overall negativity scores. (4) a scatter plot showing win volatility adaptation (x-axis) and loss volatility adaptation (y-axis) in this region for the high (purple) and low (orange) negativity (median split) groups respectively. The histogram shows the distributions of the differences between win and loss self-volatility adaptations for the high and low group respectively.
Error bars represent ±1 standard error of the mean. + p<.10; * p<.05; **p<.01; ***<.001.

Higher Negativity Scores are Associated with Decreased Anatomical Specificity of Volatility Signals

Although the self-volatility*valence interactions were not significant in these regions, there was a self-volatility*valence*negativity interaction in the dACC ($F(1,27)=4.76$, $p=.038$, $\eta_p2=.150$) with a marginally significant effect in the rACC ($F(1,27)=3.37$, $p=.077$, $\eta_p2=.111$). Together, these results suggested that the overall negativity scores might modulate the self-volatility effect in the ACC in a valence specific manner. To investigate this, we ran Pearson correlations between the overall negativity scores and the relative win vs. loss self-volatility adaptation (i.e. win self-volatile vs. win self-stable minus loss self-volatile vs. loss self-stable) for each region. The results showed that, in the dACC, which was sensitive to loss rather than win volatility at the group level (see Figure 3.14 A2), overall negativity scores correlated with relatively more win volatility adaptation than loss volatility adaptation ($r(29)=.39$, $p=.038$, see Figure 3.14 A3). In contrast, in the rACC, which was modulated more by win volatility a the group level (see Figure 3.13 A2), we found a marginally significant negative correlation between the overall negativity scores and relatively more loss volatility adaptation signals than win volatility adaptation ($r(29)=-.33$, $p=.077$, see Figure 3.13 A3). Together, this pattern of results raises the

possibility that people with higher depression and anxiety levels have less anatomically distinct representations of win-volatility and loss-volatility in the ACC.

To further illustrate this effect, we again split the group into high and low negativity participants based on a median split and tested the degree to which each group showed a differential response to the volatility of win and loss outcomes in the identified regions. The low negativity group (see Figure 3.14 A4&B4) had significantly higher win volatility adaptation in the vmPFC ($t(13)=3.12$, $p=.008$) and rACC ($t(13)=2.29$, $p=.039$) and significantly higher loss volatility adaption in the dACC ($t(13)=.288$, $p=.013$, see Figure 3.14 A4), while the high negativity group did not show significant differences between win and loss volatility adaptations in these regions ($p=. 687$). These results further showed that people with higher negativity scores might have less distinct win and loss volatility adaptation signals in the mPFC.

The relative win vs. loss volatility adaptation in the dACC could be associated with the overall negativity scores through the mediation of the win vs. loss learning rate adaptation differences

In the previous section, we showed that overall negativity scores correlated with relatively more win volatility adaptation signals than loss volatility adaptation in the dACC ($r(29)=.39$, $p=.038$, see Figure 3.14 A3). We then examined whether the relative adaption to win vs. loss volatility signals in the dACC could also be linked to the difference in learning rate adaptation between win and loss. The results revealed positive correlations between the win vs. loss learning rate adaptation differences and the relative win vs. loss volatility adaptation signals in the dACC ($r(29)=.41$, $p=.026$). We also found that the overall negativity scores were correlated with the win vs. loss learning rate adaptation differences ($r(29)=.48$, $p=.009$). Together, the results seem to suggest close associations amongst volatility signals in the dACC, volatility learning rate adaptation and the negativity traits.

To further examine the relationships, we ran a mediation analysis with win vs. loss learning rate adaptation differences as the mediation factor, and the relative adaption to win vs. loss volatility signals in the dACC and the overall negativity scores as independent and dependent variables respectively. The mediation analysis suggested that the association between the relative adaption to win vs. loss volatility signals in the dACC and the overall negativity scores could be partially mediated by the win vs. loss learning rate adaptation differences (bootstrap mean = 0.058, 95% CI: 0.003, 0.136). Together, these results further support the associations between characteristic in learning and the negativity traits and suggested a potential underlying neural mechanism.

3.4 Discussion

In this chapter, we investigated the association between the ability to flexibly learn and negative symptoms, as well as the underlying neural mechanism using fMRI and a learning and decision-making paradigm, in which we manipulated the volatility of win and loss outcomes independently.

We showed that participants' performance and choices following win and loss outcomes were not only modulated by win and loss volatility respectively, but also were influenced by the volatility of the other outcome. We first showed that participants' win and loss performance was strongly modulated by both self and other volatility. We then further examined how participants' choices were influenced by the most recent outcomes, i.e., the positive-stay and negative-switch behaviour indexes (Ppsns). We showed a robust increased in Ppsns to self-volatility when other-volatility was stable, but not when other information was also volatile. We then fitted a volatility adaptation model to formalize the learning rate adjustments to the volatility. The model results confirmed that volatility adaptation was impaired when the other outcome was volatile. Together, the results suggest that a high overall environmental volatility impairs participants' ability to adapt learning to the specific volatility of individual outcomes.

We then showed that the loss learning rate adaptation was again impaired in people with higher depression. More specifically, we found a negative correlation between depression and loss learning rate adaptation, particularly when wins were volatile. This is consistent with the results we showed in the Chapter 2, in which currently depressed patients had lower loss learning rate volatility adaptation compared to the remitted depressed patients and healthy controls. Our results are also consistent with previous studies which showed a negative correlation between punishment learning volatility adaptations and anxiety (Browning et al. 2015; Pulcu and Browning 2017; Beltzer et al. 2019; Gagne et al. 2020). But importantly, we further demonstrated that adapting learning to loss volatility is particularly impaired when other aspects of the environment (in this case, the win outcomes) are volatile. Together, the results of this and previous chapters provide consistent evidence that impaired ability to adapt learning to loss volatility, particularly in very volatile environments, might be crucially related to mood disorders.

For the neural data, we first identified regions that represent general positive vs. negative outcome signals in the brain and examined whether the outcome signals in these regions were also modulated by valence, self-volatility and other-volatility. The results first showed a general modulation by valence. The win outcome signals were generally stronger than loss outcome signals across these regions. This is consistent with our behavioural finding showing that participants were in general more influenced by win than loss outcomes. This is also consistent with findings that showed reward seemed to be more effective than punishment in changing

animal's behaviours (Nakatani et al. 2009; Thorndike 1932). Using a reward and punishment mixed paradigm, we showed that, when given with both information, humans' behaviour was also, in general, driven more by reward than punishment.

More importantly, we also found main effects of self-volatility on the outcome signals in the brain, especially in the mPFC. More specifically, we found that win volatility dominates in prefrontal regions that show increased activation for positive outcomes, i.e., the rACC and the vmPFC. In contrast, loss volatility dominates in prefrontal regions that show increased activity for negative outcomes, i.e., the dACC and the MFG. Together, we showed that areas in the ventral prefrontal cortex were modulated by win-volatility for win outcomes, while those in the dorsal prefrontal cortex were modulated by loss-volatility for loss outcomes. It might be worth highlighting that the ROIs were defined in terms of subjective valence, i.e., positive (getting a win or not getting a loss) vs. negative (getting a loss or not getting a win) outcomes, whereas the volatility signals were for the objective outcomes (i.e. whether the win or loss outcomes were volatile/stable, regardless of the outcome received by the participant). These factors were relatively orthogonal. The functional and anatomical specificity of the prefrontal cortex shown might suggest some level of interactions of subjective and objective valence information in the brain.

Previous studies have shown that the ACC tracks volatility using reward learning tasks (Behrens et al. 2007; Silvetti, Seurinck, and Verguts 2013; Rushworth and Behrens 2008). However, most previous studies have identified dACC rather than rACC. In this study, we found that, in a mixed reward and punishment learning paradigm, the rACC was more associated with win-volatility whereas the dACC was more related to loss-volatility. The cause of this difference is not clear, although the current study is unlike those previously reported in that both win and loss outcomes were presented (and their volatility manipulated) within a single task, whereas previous work presented one outcome at a time. We also found that vmPFC activity tracked win-volatility. A previous study has shown in primates that OFC also showed increased value signals in volatile environments (Massi, Donahue, and Lee 2018). The OFC has been shown to encode value during learning in primate studies, while vmPFC rather than OFC is associated with reward value in human studies (Wallis 2011). Therefore, we think the vmPFC could reflect volatility signals in humans as the OFC does in primates. Consistent with this, a recent fMRI study showed vmPFC activity reflects noisy computation in volatile environments (Findling et al. 2019).

In contrast to our analyses of participant choice, it seemed that outcomes signals in the brain were less affected by the volatility of the other outcome. This may be caused by limited power in this data to test the other-volatility effect, which might assert only a mild, if any, modulation on the outcome signals in the brain. Further studies might be able to examine the other-volatility

effect by increasing the statistical power, for example, by increasing the participant samples, introducing longer tasks, or heightening the volatility levels of the task.

Crucially, we found that the overall negativity was associated with relatively more win volatility adaptations than loss volatility adaptations in dACC, which was dominated by loss volatility at the group level, and relatively more loss volatility adaptations than win volatility adaptations in the rACC, which was dominated by win volatility at the group level. These results suggest that activity in these areas reflects the volatility of both wins and losses in people with higher levels of depression and anxiety, while the win-volatility and loss-volatility signals were more distinctly represented in the brain in people with low depression and anxiety. Supporting this idea, we showed that, using a median-split, the low negativity group had significantly greater win volatility adaptation than loss in the vmPFC and the rACC, and had significantly more loss volatility adaptation than win in the dACC. However, such distinction was not found in the high depression group. Potentially these results may suggest a less precise separation between different sources of volatility in participants with higher levels of anxiety and depression. This may be consistent with the modelling results that showed the loss-volatility learning rate adaptations were impaired particularly when win was volatile in people with higher negativity scores. The mechanism underlying any difficulty in information segregation is not clear, although it is perhaps consistent with the more general observation that people with higher levels of anxiety find it difficult to segregate ambiguous and threatening cues during conditioning (Lissek et al. 2014). Interestingly, further mediation analysis suggested the association between the unbalanced neural volatility adaptation in the dACC and the overall negativity scores could be partially mediated by the biased learning rate adaptations. Together, the results suggested an important role of dACC in modulating learning rate adaptation to the loss volatility, particularly when wins were also volatile. Failure to adjust learning to loss volatility might be associated with higher depression or anxiety levels.

In conclusion, we showed that general negativity in a subclinical cohort is associated with impaired ability to adjust loss learning rates in volatile environments during a mixed reward/loss learning paradigm, which is accompanied by higher co-representation of win and loss volatility signals in the brain. Consistent with this, we also showed in Chapter 2 that currently depressed patients showed diminished loss learning rate volatility adaptations compared to the remitted depressed patients. Together, these results suggest that the ability to adjust loss learning rate in volatile environments might be a state-dependent characteristic and the severity of it might be associated with severity of the overall anxiety and depression levels. A potential next question arises from this; whether antidepressants are able to modify loss learning rate adaptation. In the next Chapter, we investigate whether reboxetine, a selective norepinephrine reuptake inhibitor antidepressant, could change how people adapt their learning in changing environments.

4 PHARMACOLOGICAL MODIFICATION OF INFORMATION LEARNING

4.1 Introduction

In the previous Chapters, we showed a strong association between the loss learning adaptation in volatile environments and depression symptoms. More specifically, in Chapter 2, we found an impairment in adjusting loss learning to volatility in currently depressed patients, whereas the remitted depressed patients showed unimpaired volatility learning adaptation as in the healthy controls. Therefore, the difficulty in adapting learning to loss-volatility could be a state-dependent cognitive impairment present in people during episodes of depression. In Chapter 3, we further showed a correlation between reduced loss learning rate adaptation to volatility and individual differences in depression symptoms. If, as suggested by these results, the ability to adjust learning in volatile environments is a cognitive marker of depression, an obvious next question is whether antidepressants modify the adaptation to loss outcomes in volatile environments.

Norepinephrine has been theorized as a key neuromodulator to facilitate adaptive learning in a volatile environment (Dayan and Jyu 2003; Yu and Dayan 2005). Activation of the Locus coeruleus, the principal site for brain synthesis of norepinephrine, has been shown to accelerate learning in rodents (Glennon et al. 2019). In Chapter 2, we also showed an association between the loss learning rate volatility adaption and the modulation of loss volatility on pupil responses, which were thought to reflect the activity of the locus coeruleus norepinephrine (LC–NE) activity (Aston-Jones and Cohen 2005). This result was also consistent with other studies that showed pupil responses could be linked to rates of belief updating (Nassar et al. 2012; O'Reilly et al. 2013). Together, the results suggested that norepinephrine may play a key role in adaptive learning.

Interestingly, norepinephrine also plays an important role in the treatment of depressive disorders (Moret and Briley 2011). Antidepressants such as reboxetine, a norepinephrine reuptake inhibitor (NRI), act by binding to the norepinephrine transporter and blocking the reuptake of extracellular norepinephrine (Hajós et al. 2004). However, the cognitive mechanisms of antidepressants that target the norepinephrine system are poorly understood.

Here, we test a potential cognitive mechanism of NRI antidepressants-- that they facilitate adaptive learning under uncertainty by increasing extracellular norepinephrine.

To date, only a few studies have examined whether manipulating norepinephrine level could modulate adaptive learning in a changing environment in humans and the results have been mixed. For example, Jepma et. al, showed that increasing norepinephrine signalling, using atomoxetine, a selective norepinephrine reuptake inhibitor, influenced learning rates following unanticipated task changes in a complex manner—increasing learning rates in those participants with low pre-treatment learning rates, and decreasing it in those with high pre-treatment learning rates (Jepma et al. 2016). Second, decreasing norepinephrine transmission, using prazosin, a noradrenergic (α1-receptor) antagonist was found to increase the rate at which individuals updated their volatility estimates estimated using reaction times in a reward learning task (Marshall et al. 2016). In summary, the evidence for the effect of pharmacological manipulation of central norepinephrine activity on learning rates is mixed with no clear effect emerging. To date, the evidence has been limited to reward-based tasks rather than tasks incorporating both rewarding and punishing outcomes.

In the current study, we examined the effect of a single dose reboxetine, which is a NRI antidepressant, on volatility learning adaptation in healthy participants. We used the information bias learning task described in the previous chapter (i.e. including the both stable block), which allowed us to assess learning rate adaptation to changes in volatility for both win and loss outcomes, and collected pupillometry data during the task.

We hypothesized that administration of reboxetine would increase learning rate adjustments between the volatile and stable environment for both reward and punishment learning compared to the placebo group. Accompanied by the enhanced adaptive learning to volatility, we expected that reboxetine would also increase the modulation of volatility on pupil responses. Following on from the results of Chapter 3 (where the behavioural effect of negativity scores was predominantly found when the other volatility was high relative to low) we were specifically interested in whether the effects of reboxetine differed between trials in which other volatility was high vs. low, hypothesising that the effect might be greater when other volatility was high.

4.2 Methods

4.2.1 Experiment task and participants

4.2.1.1 Participants

4.2.1.1.1 Recruitment and screening

50 (30 female) healthy participants (mean age 26.9 years, range=19-44 years) were recruited in this study. Participants attended a screening session during which a structured interview was completed to assess if participants met the relevant psychiatric and physical conditions for the study. Participants were excluded if: (1) they had any history of psychiatric disorders (2) their BMIs at the time of the experiment were outside of range 18.5 and 30, (3) they had any severe medical condition not stabilized at the time of the experiment that, in the opinion of the study medic, would compromise the safety or conduct of the study including significant hypertension (diastolic pressure > 100mmHg) or bradycardia (pulse less than 50 bpm), (4) they had any history of seizures, glaucoma or pancreatitis, (5) they had lactose intolerance, (6) they had any current or past physical illness that has the potential to significantly affect mental functioning (e.g. epilepsy, hypothyroidism, Parkinson's disease, multiple sclerosis etc.), (7) if they were taking or had taken (last month) any medication that has a significant potential to affect mental functioning (e.g. benzodiazepines, antidepressants, neuroleptics etc.), (8) they had taken any recreational drugs in the last 3 months prior to the experiment (e.g. marijuana, ecstasy etc.), (9) they had harmful alcohol use in the last 6 months prior to the experiment, (10) they had any history of allergic reactions to relevant substances (e.g. reboxetine), (11) they had participated in a study using the same or similar tasks. Female participants were also subject to exclusion if (1) they were pregnant, or lactating, (2) they were sexually active but not using any medically accepted method of contraception. The study was approved by the University of Oxford Central Research Ethics Committee. Written informed consent was obtained from all participants on the screening visit, in accordance with the Declaration of Helsinki.

4.2.1.1.2 Test session

The participants completed a testing session (baseline testing) immediately following the screening. During this session, participants were asked to complete a number of questionnaires, including (1) the Quick Inventory of Depressive Symptoms, 16 items self-report version (QIDS)，which measures symptoms of depression (Rush et al. 2003), (2) the Spielberger State-Trait Anxiety Inventory (STAI), which includes separate scales measuring state and trait anxiety (Spielberger 1983), (3) the Befindlichkeits Scale (BFS), which assess transient

subjective affective state (Von Zerssen, Strian, and Schwarz 1974). They were instructed about the information bias leaning task and completed a brief introduction version (of 16 trials) to familiarize them with the logic of the task. Following this, participants completed the 4 blocks of the task self-paced, with rest sessions between each block. The task was presented on a VGA monitor connected to a laptop computer running PsychoPy2 software version 1.90.3 (Peirce et al. 2019).

The participants were invited back for a second testing within 4 weeks after the first testing. The participants were randomly assigned to either the reboxetine group or the placebo group. Both participants and researchers were blind to which participants were allocated to drug or the placebo condition. On arrival at the second visit, the reboxetine group was given a single dose of reboxetine (4mg) and the placebo group was given a lactose tablet. 1.5 hours after the drug administration, the participants completed the same questionnaires and tasks as in the first testing (see Figure 4.1A).

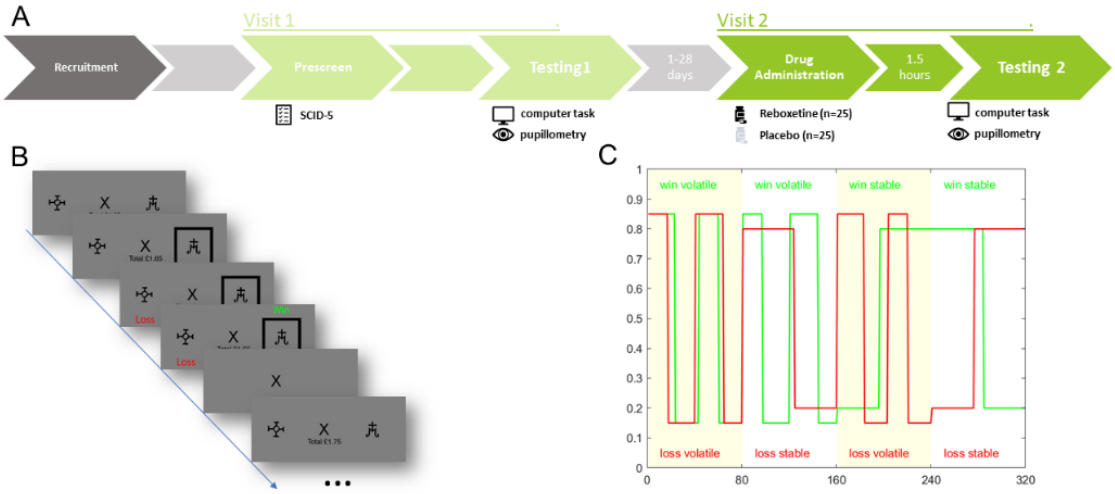

Figure 4.1 experiment task and testing procedure

(A) General procedures. The participants were first invited for the first visit. On this visit, the participants were checked if they met relevant psychiatric and physical conditions. The eligible participants then completed the task without drug administration. They were invited back for a second visit 1 to 28 days after the visit 1. At the arrival of the visit 2, the participants were given either a single dose of reboxetine or placebo. 1.5 hours after the drug administration, the participants performed the same task as in the visit 1. (B) Timeline of one trial from the information bias learning task. This was similar to the task timeline in Chapter 2. (C) Overall task structure. This was the same as in Chapter 3.

The task used in this study were the same task used in Chapter 2&3. The overall structure of the task was the same as described in section 3.2.1.2 in Chapter 3 (also see Figure 4.1). The block orders were counterbalanced across participants and groups. A participant started with

either a balanced block (the both-volatile block or the both-stable block) or a semi-volatile block (the win-volatile block or the loss-volatile block) on the first visit, then he/she would start with a semi-volatile block or a balanced block respectively on the second visit.
However, we adopted a similar timeline as in Chapter 2 for pupillometry data (see section 2.2.1.2 in Chapter 2). First, the choice phase was self-paced. The chosen option was cued with a black frame for 1 second before the outcome presentation. Secondly, the win and loss outcomes were presented one by one with randomised order across trials with a jittered interval of 2–5 (mean 3.5) seconds (Note that we reduced the jitter a little to shorten the duration of the whole task, which was 2–6 (mean 4) seconds in Chapter 2). In half of the trials, the win outcome was presented first, with the loss outcome presented first in the other half. Thirdly, auditory cues were used accompanying the visual presentations of the outcomes. More specifically, a 'chi-ching' sound was played if participants received a win and an error buzz if they received a loss. Auditory stimuli lasted 0.7 seconds with the same onsets as the visual feedback. Finally, the intertrial interval was fixed at 1 second.

4.2.1.2 Pupillometry recordings

The participants' pupil responses when they performed the information bias learning task were recorded using an eye-tracking device. Participants' heads were stabilised using a head-and-chin rest placed 70 cm from the screen on which an eye-tracking system was mounted (Eyelink 1000 Plus; SR Research, Ottawa, Canada). The eye-tracking device was configured to record the coordinates of both eyes and pupil area at a rate of 500 Hz. The abstract shapes of the learning task were drawn on either side of a fixation cross which marked the middle of the screen and was offset by around 7˚ visual angle.

4.2.2 Model-free behavioural data analysis

4.2.2.1 Task performance

The model-free behavioural data were analysed using MATLAB (version R2018b; MathWorks, MA, USA) and SPSS (version 26.0; IBM, NY, USA). The percentages of wins chosen and losses not-chosen were calculated for each block. We ran a 5-way mixed ANOVA, with valence (win/loss) and self-volatility (volatile/stable) other-volatility (volatile/stable) * visit (visit1 and visit2) as repeated measures, while treatment (reboxetine/placebo) was a between-subjects factor. We also examined whether the performances for each block were significantly above chance level (i.e., 50%) for wins and losses respectively using one-sample t-tests. Note that for

these (and all other) analyses, an effect of treatment on the dependent variables is captured by the treatment * visit interaction terms.

4.2.2.2 Positive-stay-negative-switch

The positive-stay-negative-switch (P_{psns}) indexes were calculated for win and loss respectively in the same way as in Chapter 2 & 3 (see section 2.2.2.2 for more details). The positive-stay-negative-switch indexes were then submitted to a 5-way mixed ANOVA, with valence (win/loss) and self-volatility (volatile/stable) other-volatility (volatile/stable) * visit (visit1 and visit2) as repeated measures, while treatment (reboxetine/placebo) was a between-subjects factor.

As we were specifically interested in how the drug treatment modifies self-volatility adaptation under different levels of the other-volatility (see results from Chapter 3), we also examined the self-volatility adaptation when the other-volatility was volatile or stable respectively. More specifically, we calculated self-volatility adaptation for when other-volatility was volatile or stable and for wins and losses separately, i.e., (1) win self-volatility Ppsns adaptation when loss was volatile (Ppsns-win from the both-volatile block minus Ppsns-win from the loss-volatile block), (2) win self-volatility Ppsns adaptation when loss was stable (Ppsns-win from the win-volatile block minus Ppsns-win from the both-stable block), (3) loss self-volatility Ppsns adaptation when win was volatile (Ppsns-loss from the both-volatile block minus Ppsns-loss from the win-volatile block), and (4) loss self-volatility Ppsns adaptation when win was stable(Ppsns-loss from the loss-volatile block minus Ppsns-win from the both-stable block). Then we calculated differences across the two visits for these 4 variables. These were then submitted to a 2*2 model for when other-volatility was either volatile or stable, with valence (win/loss) as a repeated measure and treatment (reboxetine/placebo) as a between-subjects factor. We followed the significant interaction effects found here with two-sample t-tests to examine the treatment effects.

4.2.3 Computational modelling of behavioural data

Volatility adaptation model:

The same model adopted in Chapter 3 was used here (see section 3.2.2.2 for more details). Models were fitted using Stan (RStan v 2.21.1, R v 3.6.0). The data from all 4 blocks (excluding the first 5 trials of each block) from both visits were used in the model fitting. The expected values for win and loss were reset to 0.5 at the beginning of each block. 4 parallel Monte Carlo chains were run each with 20000 iterations (10000 warm-ups), using adapt_delta of 0.8 and maximum treedepth of 15. Learning rate parameters were sampled from a zero-mean normal

distribution with a standard deviation of 2, and then transformed to the range 0-1 using logistic approximation (Bowling et al. 2009). Beta parameters were sampled from a zero-mean normal distribution with a standard deviation of 1 and then transformed using the exp function to a positive value. Convergence was inspected using shinyStan(v 2.5.0). No convergence problems were detected. We adopted a stringent cut-off of R-hat< 1.1 as an indicator for sufficient convergence (Brooks and Gelman 1998), and the minimum number of effective samples >40. Parameters were estimated separately for each participant with no hierarchical term. The same procedures were also used to estimate parameters for the beta adaptation model and the fixed learning model (described in detail in 2.2.3).
We then performed model comparisons using a leave-one-out cross-validation procedure using the LOO package for R (Vehtari, Gelman, and Gabry 2017). An expected log pointwise predictive density (ELPD) and a standard error for each ELPD were estimated for each model.

The statistical tests for the learning rate parameters were performed using SPSS (version 26.0; IBM, NY, USA). We first examined whether the learning rate adaptation estimates were statistically above zero using one-sample t-tests against zero for estimates from all participants for each visit respectively.
A 4-way mixed ANOVA, with valence (win/loss), other-volatility (volatile/stable) and visit (visit1 and visit2) as repeated measures, while treatment (reboxetine/placebo) as a between-subjects factor, was performed for the baseline learning rates and the volatility learning rate adaptations respectively. Note that the statistical tests for the learning rate volatility adaptation term and the baseline learning rate estimates were done in logit space. The betas were log-transformed and then submitted to a 4-way mixed ANOVA, with win-volatility (win volatile/win stable) and loss-volatility (loss volatile/loss stable) * visit (visit1 and visit2) as repeated measures, and treatment (reboxetine/placebo) as a between-subjects factor. To further examined the effects for the treatment induced changes in learning rate volatility adaptations between the two visits, we then performed valence * treatment model for when other-volatility was volatile or stable respectively, with valence (win/loss) as a repeated measure and treatment (reboxetine/placebo) as a between-subjects factor. Finally, we followed up the treatment effects with two-sample t-tests between the reboxetine and placebo group.

4.2.4 Pupillometry data analysis

4.2.4.1 Pre-processing

Blinks were identified automatedly by the Eyelink system's built-in filter and were then removed from the data. Missing data points (including blinks) were linearly interpolated. The

resulting trace was subjected to a low pass Butterworth filter with a cut-off of 3.75 Hz and then z transformed across the session (Browning et al. 2015; Pulcu and Browning 2017). The pupil response to the win and the loss outcomes were extracted separately for each trial with the epoch from 1 s prior to the presentation of the outcome to 5 s after the outcome presentation. The pupil response was baseline corrected for each outcome presentation by subtracting the mean pupil size of a one second window preceding feedback. Individual trials were excluded from the pupillometry analysis if more than 50% of the data in the epoch had been interpolated (Browning et al. 2015; Pulcu and Browning 2017). 3 participants were excluded from the pupillometry analysis as more than 60% of their trials were excluded on this basis. The same number of trials (n=5) as in the model fitting (see the previous section 0) from each block were removed from the analysis to ensure consistency between the analyses and also to avoid any luminance changes that may be caused by starting a new block of the task.

4.2.4.2 Regression analysis of pupil data

Trial by trial regression analysis was performed on the pre-processed pupillometry data combining all three blocks using the ordinary least square (OLS) method for data points during -1s to +5s relative to outcome delivery for win and loss outcomes respectively. Explanatory variables (EVs) of interest included (1) self-prediction errors (self-PE): prediction errors for the current event for the current trial (i.e., win prediction errors for win outcome window, and loss prediction error for the loss outcome window. The prediction errors were calculated using the best fitting learning rates derived from the computational model, (2) self-volatility level: the volatility level for the current block and the current outcome event (e.g., for the win outcome delivery: 1 for blocks when win outcomes were volatile, -1 for the block when win outcomes were stable), (3) other-volatility level: the other volatility level for the current block and the current outcome event (e.g., for the win outcome delivery: 1 for blocks when loss outcomes were volatile, -1 for the blocks when loss outcomes were stable), (4) self- volatility level* self-PE, (5) other- volatility level* self-PE, (6) self- volatility level * other-volatility level* self-PE.

We also added several control variables, including (7) other-prediction errors (other-PE) prior to the current event in the same trial (i.e., loss prediction errors for win outcome window, and win prediction error for the loss outcome window. For trials, when the other outcome was presented later than the current event, the value was set to 0 in the demeaned regressor for this trial.) to control for the remaining effect of the other outcome on the current event; (8) order: whether the current outcome event was presented first in the current trials to control for the effect of the order of the outcome presentation within the current trial; (9) trial number: the current trial number in this block, to control for drifting in pupillometry data within the block,

(10) a constant term. EVs (other than the constant term) were normalised for each participant before entering into the model. The three interaction regressors, i.e., EV(4), EV(5) and EV(6) were calculated using normalised relevant EVs. Then the interaction regressors were also normalised before entering into the regression model.

To investigate the modulation of EVs of interests on pupil dilations, one sample (against zero) cluster-wised multiple comparisons correction were performed for each group on the parameters estimates (Maris and Oostenveld 2007).

To further examine the effects of other volatility, we ran separate OLS analyses for when other volatility was volatile or stable respectively. More specifically, data for win or loss outcomes were split into two halves based on whether other-volatility was volatile or stable (i.e., (1) the "both volatile" block and the "win volatile" block for win outcomes, (2) the "loss volatile" block and the "both stable" block for the win outcomes, (3) the "both volatile" block and the "loss volatile" block for the loss outcomes, (4) the "win volatile" block and the "both stable" block for loss outcomes). The OLS model included the same regressors described above except for any regressors involved other-volatility (i.e., EV(3), and EV(5) and EV(6)). We then took the means of the beta estimates 0-5s since outcome delivery and calculated the differences between the 2 visits for each condition and group. Then the differences were submitted to a mixed ANOVA (valence * * other volatility * treatment). For any significant effect, we follow up with 2-sample t-tests.

4.3 Results

4.3.1 Model-free behaviour results

4.3.1.1 Task performance

We first examined task performance in terms of the percentage of trials in which participants received the win and loss outcomes. As expected, one-sample t-tests against the chance level (50%) showed that performances both for wins and losses were significantly above chance level (all $t>7$, $p<.001$), other than the performances for wins in the loss-volatile block for either of the two visits (not significant, both $p>.010$) and the performances for losses in the win-volatile block for both visits (significantly below chance level for visit1 ($t(49)=-3.86$, $p<.001$) and visit2 ($t(49)=-3.63$, $p<=.001$)). The results suggested that in general participants used both the volatile and stable information effectively in the task, but in semi-volatile blocks, participants relied mostly on the volatile outcome at the expense of the stable outcome to make choices.

The valence*visit*self-volatility*other-volatility*treatment ANOVA analysis revealed a significant main effect of valence ($F(1,48)=5.77$, $p=.020$, $\eta_p^2=.107$), indicating a general effect

of higher accuracies for wins than losses across participants and a significant main effect of visits ($F(1,48)=7.70$, $p=.008$, $\eta_p^2=.138$), indicating that in general participants' accuracies were higher in the second visit.

The 5-way mixed ANOVA also revealed significant main effects of self-volatility ($F(1,48)=85.50$, $p<.001$, $\eta_p^2=.640$) and other-volatility ($F(1,48)=385.57$, $p<.001$, $\eta_p^2=.889$), and a significant interaction effect of self-volatility * other volatility ($F(1,48)=408.89$, $p<.001$, $\eta_p^2=.895$), suggesting block specific effect for accuracies (see Figure 4.2).

The analysis also revealed a significant interaction of valence* self-volatility ($F(1,48)=7.39$, $p=.009$, $\eta_p^2=.133$), suggesting the self-volatility effect was different for wins and losses. To further examine the effects for wins and losses respectively, we then follow this up by running self-volatility*other-volatility*visit*treatment ANOVAs for wins and losses separately. The results showed a significant visit*treatment interaction for loss accuracies ($F(1,48)=4.95$, $p=.031$, $\eta_p^2=.094$) but not for the wins ($p>.84$), indicating an higher increase in accuracies for losses across the visits in the reboxetine group compared to the placebo group ($t(48)=2.23$, $p=.031$) (See Figure 4.3).

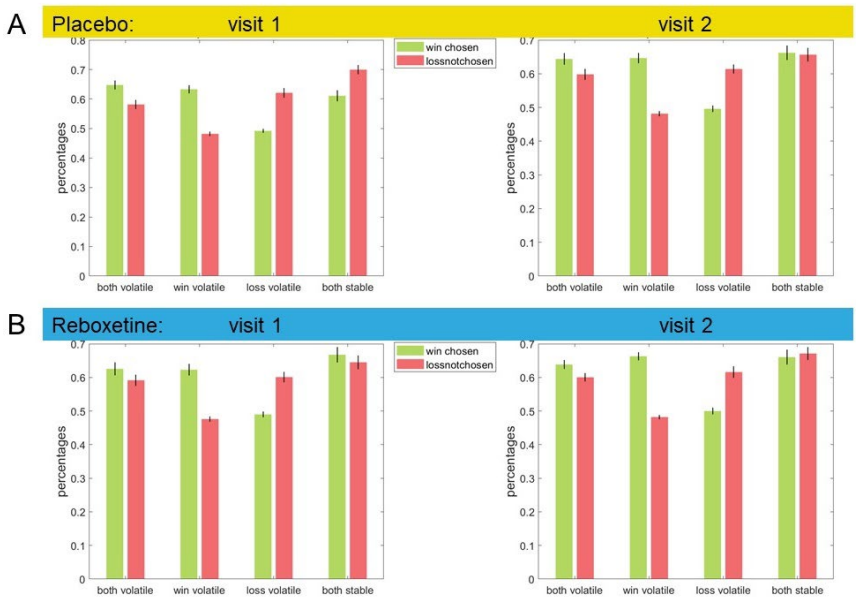

Figure 4.2 Summary of task performance for each visit for the reboxetine and placebo groups.

Percentage of trials in which the participant selected the option associated with a win (win chosen) or with no loss (loss not chosen) for each of the four task blocks for each visit for the placebo group (A) and for the reboxetine group (B).

Error bars represent ±1 standard error of the mean.

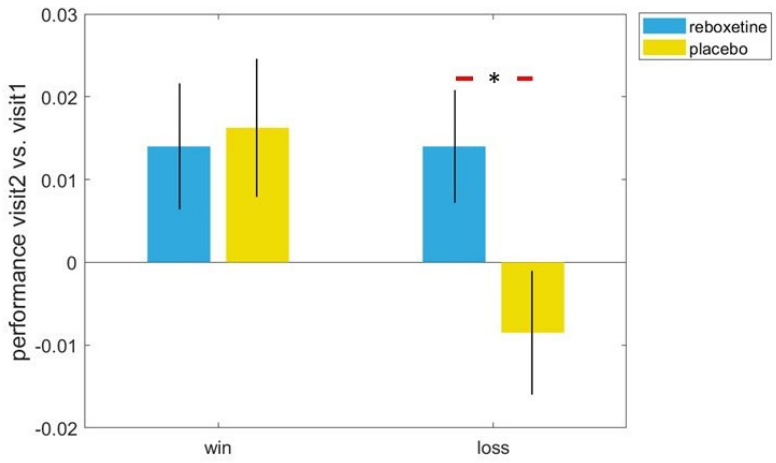

Figure 4.3 Reboxetine selectively increased loss performance.

Changes in accuracies for win and loss from visit1 to visit2 for each group.

Error bars represent ±1 standard error of the mean. * p<0.05 for post-hoc t-tests.

4.3.1.2 Positive-stay and negative-switch behaviours

We ran a valence*visit*self-volatility*other-volatility*treatment ANOVA analysis on the positive-stay, negative-switch (Ppsns) index to see whether reboxetine affected how participants weighted the most recent feedback in their decisions. The results showed a significant main effect of self-volatility ($F(1,48)=115.18$, $p<.001$, $\eta_p2=.706$), indicating that, as predicted, participants were more influenced by outcomes of volatile than stable associations, and a significant self-volatility by other-volatility interaction ($F(1,48)=97.78$, $p<.001$, $\eta_p2=.671$), suggesting higher modulation of self-volatility on Ppsns when other-volatility were stable. This analysis also revealed a significant main effect of valence ($F(1,48)=5.81$, $p=.020$, $\eta_p2=.108$), indicating that participants were generally more sensitive to win than loss outcomes, and a significant interaction of valence * self-volatility ($F(1,48)=8.04$, $p=.007$, $\eta_p2=.144$), indicating higher modulation of self-volatility on Ppsns for wins than losses.

Importantly, the results also showed a significant visit*treatment*self-volatility*other-volatility effect ($F(1,48)=4.52$, $p=.039$, $\eta_p2=.086$), suggesting the modulation of reboxetine on self-volatility behaviour adaptation differed when other-volatility was volatile relative to stable. To further characterise this effect, we ran valence* treatment ANOVAs for the Ppsns self-volatility adaptation (i.e., self-volatile – self-stable) changes between the two visits (i.e., visit2 vs visit1) for when other-volatility was volatile or stable respectively. The analysis revealed a significant treatment effect ($F(1,48)=5.02$, $p=.030$, $\eta_p2=.095$) when the other-volatility was volatile, but the treatment effect was not statistically significant when the other volatility was stable ($p= .377$), indicating that the reboxetine in general facilitated volatility adaptation when the other-volatility was also volatile. The post-hoc analysis confirmed that reboxetine increased

the volatility adaptation for loss when the wins were volatile compared to the placebo (t(48)=2.79, p=.008), but not when wins were stable (p=.661). However, the treatment effect did not reach statistical significance for wins (when losses were volatile: p=.381, when losses were stable: p=.547) (see Figure 4.5).

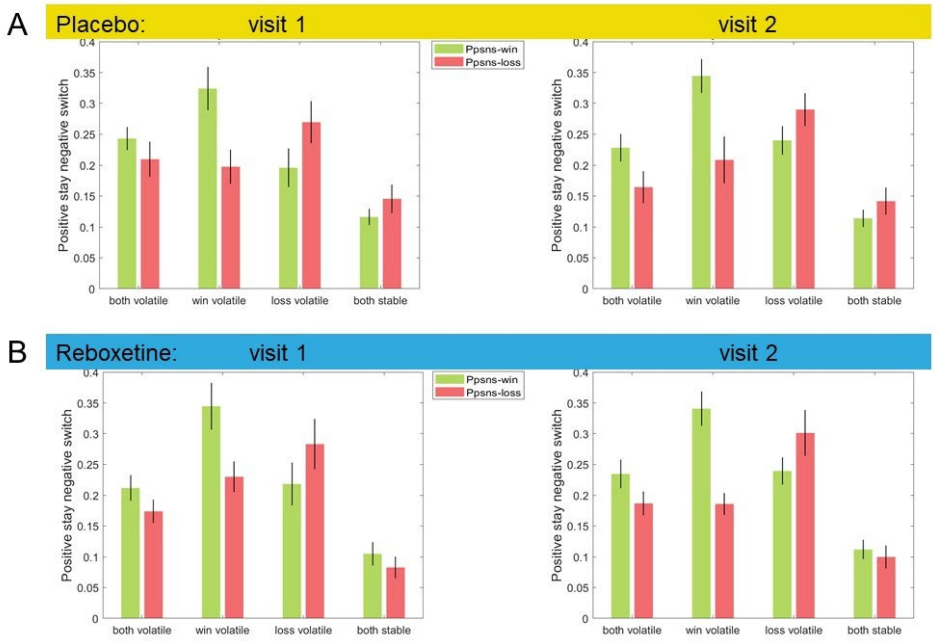

Figure 4.4 Summary of positive stay negative switch indices for each visit for reboxetine and placebo group.

Win Ppsns and loss Ppsns for each of the four task blocks for each visit for the placebo group (A) and for the reboxetine group (B).
Error bars represent ±1 standard error of the mean.

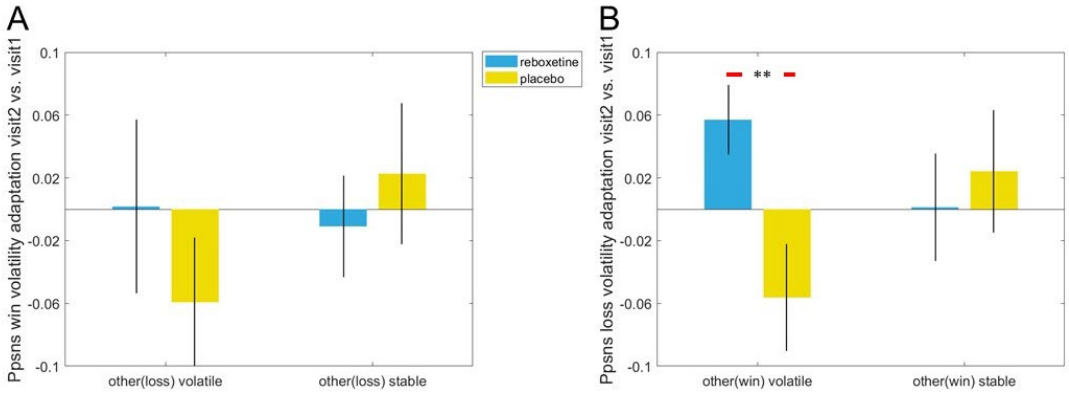

Figure 4.5 Reboxetine increased loss Ppsns volatility adaptation when win was volatile.
(A) Changes in win Ppsns volatility adaptation (win volatile vs. win stable) from visit1 to visit2 for when loss was volatile or stable respectively for each group. (A) Changes in loss Ppsns volatility adaptation (loss volatile vs. loss stable) from visit1 to visit2 for when win was volatile or stable respectively for each group.

Error bars represent ±1 standard error of the mean. ** p<0.01 for post-hoc t-tests.

4.3.1.3 Summary

We found that reboxetine seemed to increase task performance, particularly for loss, compared to placebo. This was accompanied increased sensitivity to loss outcomes, indexed by Ppsns. In the following section, we used the same computational model applied in Chapter 3, to see if reboxetine increased learning rate adaptation to volatility.

4.3.2 Modelling results

4.3.2.1 learning rates

The learning rate estimates showed the same pattern as in Chapter 3 for each group and each visit (see Figure 4.6). One sample t-tests on the volatility learning rate adaptation terms confirmed that self-volatility in general increased learning rates. More specifically, all learning rate adaption terms were significantly higher than zero for visit1(all p<.016). The visit2 showed the sample pattern (all p<.044), although marginally for the win learning rate adaption when losses were volatile(p=.090).

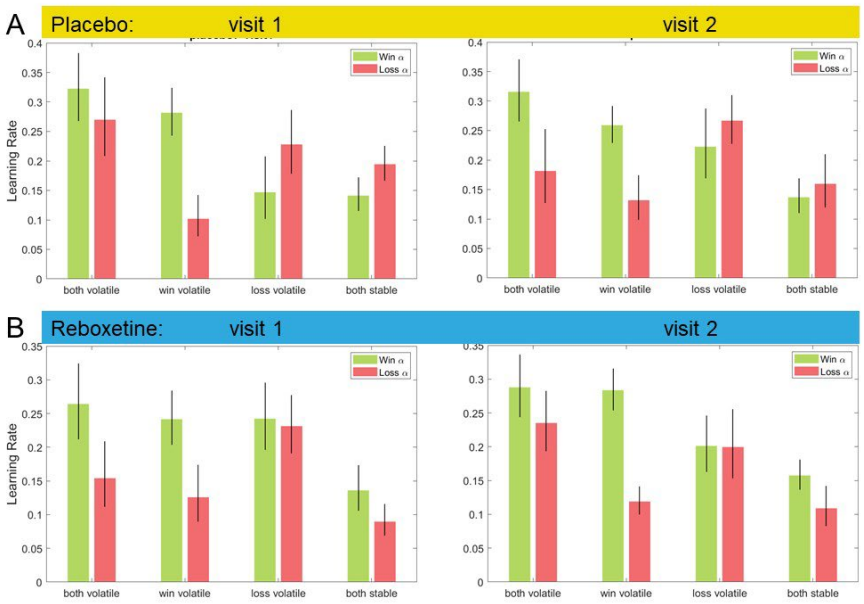

Figure 4.6 Summary of learning rate estimates for each visit for reboxetine and placebo group.

Win and loss learning rates for each of the four task blocks for each visit for the placebo group (A) and for the reboxetine group (B).
Error bars represent ±1 standard error of the mean.

We then examined the effect of reboxetine administration on the volatility learning rate adaptation estimates with a mixed ANOVA (valence*other volatility*visit*treatment). The results showed a marginally significant visit * treatment interaction ($F(1,48)=3.92$, $p=.053$, $\eta_p^2=.076$), indicating that overall the reboxetine group showed an increment in volatility adaptation overtime while the placebo group showed decreased volatility adaptation overtime. The results also revealed a significant other-volatility * visit * treatment interaction ($F(1,48)=7.07$, $p=.011$, $\eta_p^2=.128$), showing that the visit * treatment effect was more robust when the other volatility was volatile than stable. Critically, the result revealed a significant valence * other-volatility * visit * treatment interaction ($F(1,48)=4.97$, $p=.030$, $\eta_p^2=.094$), suggesting the effect was mainly driven by losses rather than wins.
We then followed this up by running valence * treatment ANOVAs for the learning rate self-volatility adaptation (i.e., self-volatile – self-stable) changes between the two visits (i.e., visit2 vs visit1) for when other-volatility was volatile or stable respectively. This post-hoc analyses revealed that the treatment effect was significant when other-volatility was high ($p=.007$), but not statistically significant when other-volatility was low ($p=.516$). Independent t-tests confirmed that reboxetine significantly facilitated the learning rate volatility adaptations for loss when win was volatile ($p=.005$), and marginally increased the win learning rate volatility adaptations when loss was volatile ($p=.081$), but not for learning rate volatility adaptations neither for win ($p=.128$) or loss ($p=.483$), when the other information was stable (see Figure 4.7).

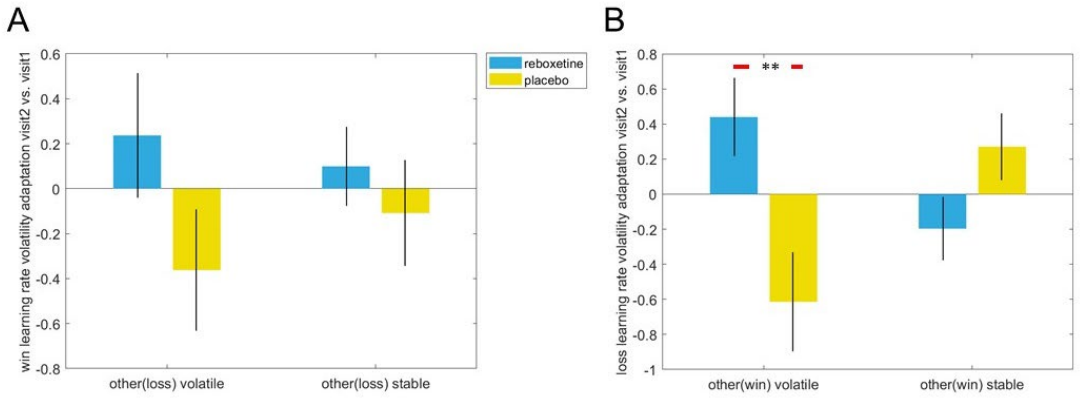

Figure 4.7 The reboxetine increased loss learning rate volatility adaptation when win was volatile.

(A) Changes in win learning rate volatility adaptation (win volatile vs. win stable) from visit1 to visit2 for when loss was volatile or stable respectively for each group. (B) Changes in loss learning rate volatility adaptation (loss volatile vs. loss stable) from visit1 to visit2 for when win was volatile or stable respectively for each group.
Error bars represent ±1 standard error of the mean. ** p<0.01 for post-hoc t-tests.

The mixed ANOVA (valence*other volatility*visit*treatment) for the baseline learning rates revealed a significant main effect of valence ($F(1,48)=4.82$, $p=.033$, $\eta^2=.091$), indicating that overall win baseline learning rates were higher than loss. The results also showed a significant valence * other volatility interaction ($F(1,48)=6.37$, $p=.015$, $\eta^2=.117$), suggesting that win baseline learning rates were more affected by other(loss) volatility than loss baseline learning rates by win volatility. However, there was no significant main effects or interaction for treatment or visit for the baseline learning rates (all $p>.103$).

4.3.2.2 Betas

The win volatility * loss volatility * visit * treatment ANOVA for beta estimates showed significant effects of win volatility ($F(1,48)=7.97$, $p=.007$, $\eta^2=.142$) and loss volatility ($F(1,48)=15.13$, $p<.001$, $\eta^2=.240$), indicating higher betas for stable schedules than volatile ones. The results also revealed a significant win and loss volatility interaction, suggesting that the effects were mostly driven by higher betas for the both-stable block compared to the other three blocks (See Figure 4.8). However, the analysis did not reveal any treatment effect or interactions (all $p>.09$), suggesting reboxetine had little effect on choice stochasticity.

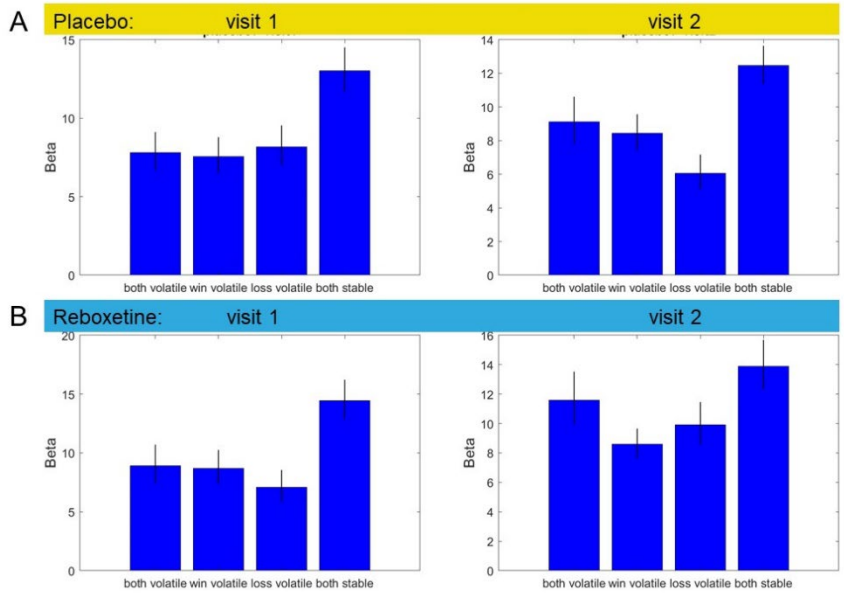

Figure 4.8 beta estimates for each visit for reboxetine and placebo group

Error bars represent ±1 standard error of the mean.

4.3.2.3 Model comparisons

Similar to the previous studies, the learning rate volatility adaptation model had highest predictive accuracy (ELPD=-5732.3, SE=47.8) compared to the beta volatility adaptation model (ELPD=-6213.7, SE=60.1), and the fixed learning model (ELPD=-6314.4, SE=55.6), suggesting that the learning rate volatility adaptation model fitted the data better than the other two models.

4.3.2.4 Summary

With the volatility adaptation model, we showed that reboxetine seemed to increase learning rate volatility adaptation, particularly to loss, compared to placebo. In the following section, we examined whether the effect was mirrored by changes in pupil responses during the task.

4.3.3 Pupil responses during task

We first checked whether reboxetine caused a general pupil dilation change on visit2 compared to visit1. We examined whether there was any visit * treatment effect for the mean pupil dilation (without any normalization) during a 1s window right before delivery of each outcome, which was used as baselines in the following analyses, for each block and for wins and losses respectively. The ANOVA results did not show any significant main effects or interactions. Nor did pair t-tests show any significant differences between the baselines. This suggested that reboxetine might not affect general pupil dilations during the study (see Figure 4.9).

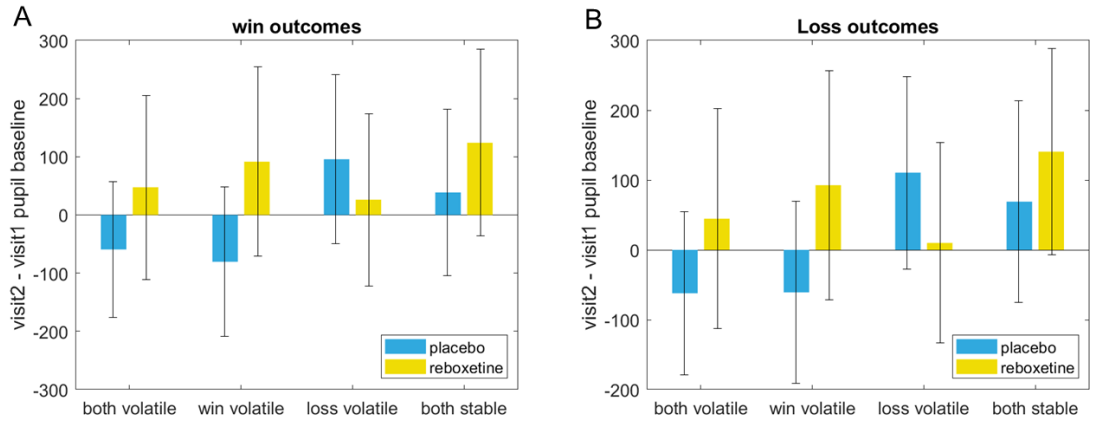

Figure 4.9 baseline pupil dilation was not affected by the treatment.

(A) Changes in mean baseline pupil dilation from visit1 to visit2 for win outcomes for each group. (B) Changes in mean baseline pupil dilation from visit1 to visit2 for loss outcomes for each group.

Next, we examined whether and how reboxetine affected the modulation of prediction errors (PE) on pupil dilation. Consistent with results from Chapter 2, the OLS analysis showed a strong modulation of prediction errors on pupil dilations for both win and loss outcomes (see Figure 4.10 A-D & Figure 2.9 in Chapter 2). The results also revealed reduced win and loss PE modulations of pupil dilations for the 2nd visit compared to the 1st visit for the reboxetine group, while such differences were not shown for the placebo group (see Figure 4.10 E&F).

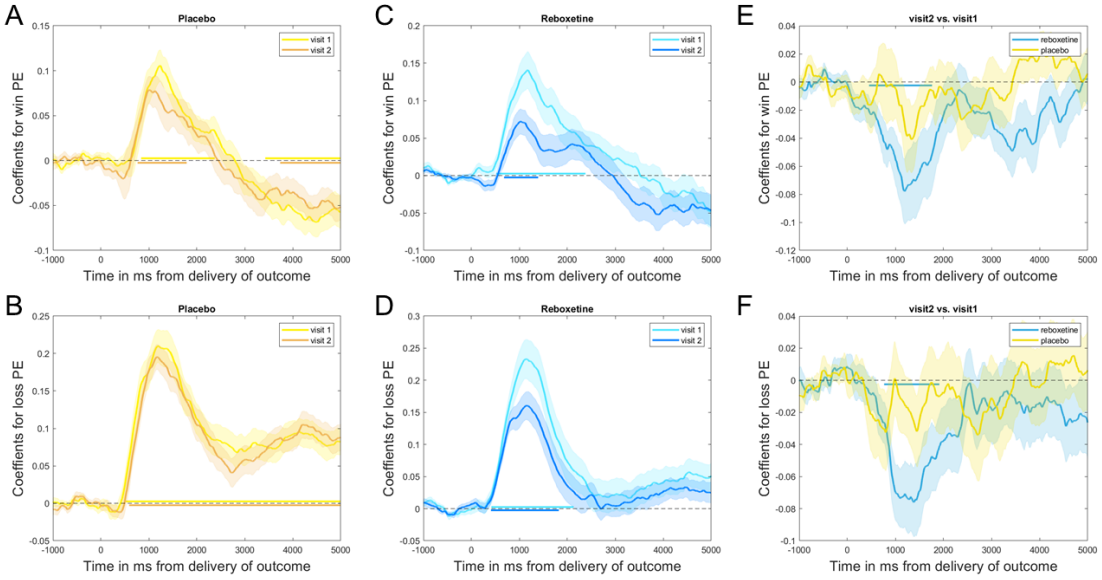

Figure 4.10 Regression analysis showing the temporal dynamics of prediction errors on pupil dilations.

(A&C) The coefficients estimated for the effect of win prediction errors on pupil dilations for the placebo (A) and the reboxetine group (C) respectively. (B&D) The coefficients estimated for the effect of loss prediction errors on pupil dilations for the placebo (B) and the reboxetine group (D) respectively. (E) Changes in the coefficients estimated from visit1 to visit2 for the effect of win prediction errors on pupil dilations for the placebo group and reboxetine group respectively. (F) Changes in the coefficients estimated from visit1 to visit2 for the effect of loss prediction errors on pupil dilations for the placebo group and reboxetine group respectively. The coloured solid line and shadow indicate the means ± s.e.m. calculated for each group. The coloured thick lines paralleling the x-axis indicate time windows in which coefficients were significantly different from zero after cluster-based multiple comparison correction (cluster definition threshold $P = 0.05$) for each visit or each group. Time zero corresponds to the onset of outcome delivery.

Interestingly, consistent with the reboxetine effect we showed with the behaviour data, we also found a significant other-volatility * self-volatility * PE * visit * treatment interaction on pupil dilations. More specifically, compared to the placebo group, reboxetine seemed to reduce the effect of other-volatility*self-volatility*PE on pupil dilations in general (see Figure 4.11). To further investigate the 3-way interaction, we ran OLS analyses separately for when other-volatility was stable and when other-volatility was volatile. The results suggested that reboxetine reduced the self-volatility effect on loss PE pupil responses particularly when wins were volatile (see Figure 4.12 B). To formally examine the interactions, we took the mean of

the coefficient estimates across 0-5s from the outcome delivery, calculated the difference in the beta estimates between the 2 visits, and then submitted them to a mixed ANOVA. The analyses confirmed a significant valence * other-volatility * treatment interaction ($F(1,48)=4.96$, $p=.031$, $\eta^2=.099$). Post-hoc 2-sample t-tests confirmed that the placebo and reboxetine group differed in the self-volatility * self-PE modulation on pupil responses particularly for loss outcomes when wins were volatile ($p=.007$) (see Figure 4.12 F).

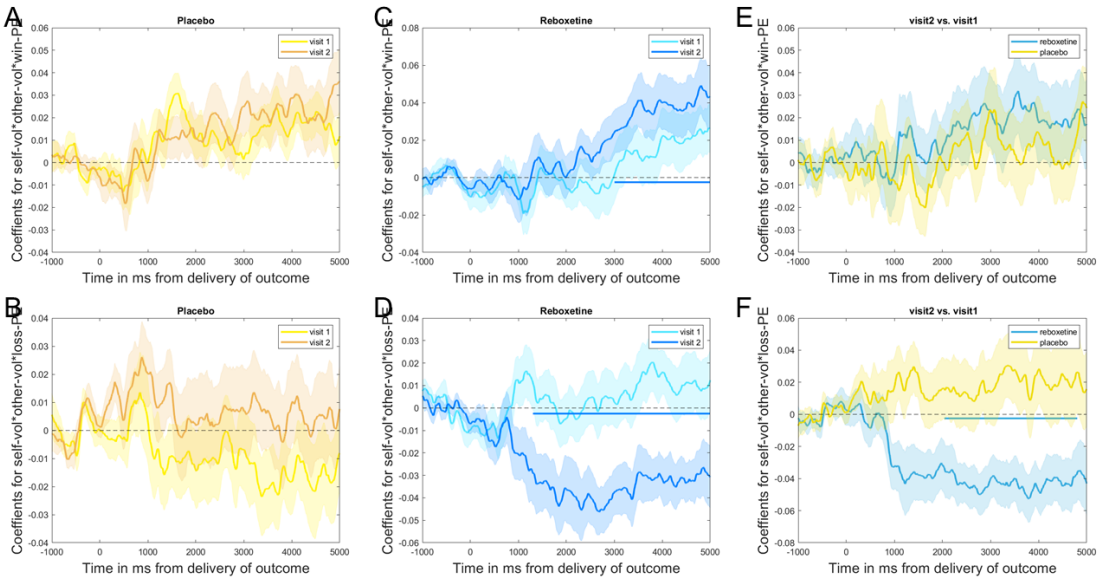

Figure 4.11 Regression analysis showing the temporal dynamics of self-volatility * other-volatility * prediction errors interaction effect on pupil dilations.

(A&C) The coefficients estimated for the 3-way interaction of win prediction errors, win volatility and loss volatility on pupil dilations for the placebo (A) and the reboxetine group (C) respectively. (B&D) The coefficients estimated for the 3-way interaction of loss prediction errors, loss volatility and win volatility on pupil dilations for the placebo (B) and the reboxetine group (D) respectively. (E) Changes in the coefficients estimated from visit1 to visit2 for the 3-way interaction of win prediction errors, win volatility and loss volatility on pupil dilations for the placebo group and reboxetine group respectively. (F) Changes in the coefficients estimated from visit1 to visit2 for the 3-way interaction of loss prediction errors, loss volatility and win volatility on pupil dilations for the placebo group and reboxetine group respectively. The coloured solid line and shadow indicate the means ± s.e.m. calculated for each group. The coloured thick lines paralleling the x-axis indicate time windows in which coefficients were significantly different from zero after cluster-based multiple comparison correction (cluster definition threshold $P = 0.05$) for each visit or each group. Time zero corresponds to the onset of outcome delivery.

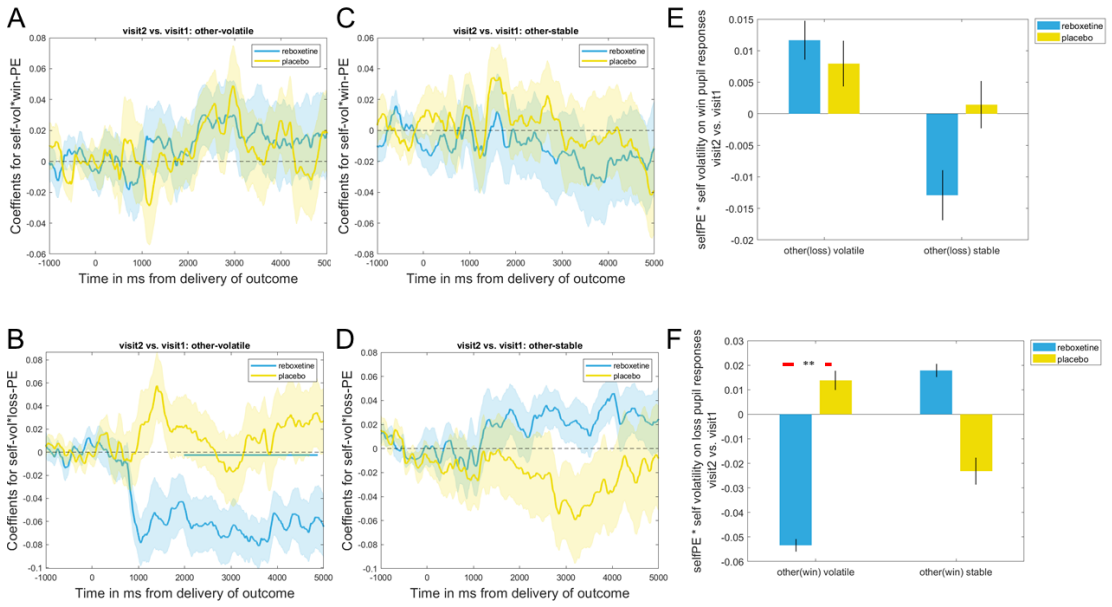

Figure 4.12 reboxetine reduced the self-volatility effect on pupil responses to loss prediction errors when win was volatile but not when win was stable.

(A&C) Changes in the coefficients estimated from visit1 to visit2 for win prediction error and win volatility interaction on pupil dilations for when loss was volatile (A) and for when loss was stable (C) respectively. (B&D) Changes in the coefficients estimated from visit1 to visit2 for loss prediction error and loss volatility interaction on pupil dilations for when win was volatile (B) and for when win was stable (D) respectively.

The coloured solid line and shadow indicate the means ± s.e.m. calculated for each group. The coloured thick lines paralleling the x-axis indicate time windows in which coefficients were significantly different from zero after cluster-based multiple comparison correction (cluster definition threshold P = 0.05) for each group respectively. Time zero corresponds to the onset of outcome delivery.

(E) Bars represent average delta coefficients between the two visits for win PE and win volatility interaction (mean across 0-6s after the onset of outcome delivery) for when loss was volatile and for when loss was stable, and for each group respectively. (F) Bars represent average delta coefficients between the two visits for loss PE and loss volatility interaction (mean across 0-6s after the onset of outcome delivery) for when win was volatile and for when win was stable, and for each group respectively.

Error bars represent ±1 standard error of the mean. * $p<0.05$ for post-hoc t-tests.

Finally, we also found evidence that the behavioural and pupillometry measures were related: the decrease in the modulation of self-volatility on loss pupil responses to self-PE on visit 2 compared to visit 1 was associated with increased loss learning rate adaption estimates on visit

2 compared to visit 1 across all participants (r(46)=-0.29, p=.049, see Figure 4.13). However, such correlation was not found within the reboxetine group (r(23)=-.042).

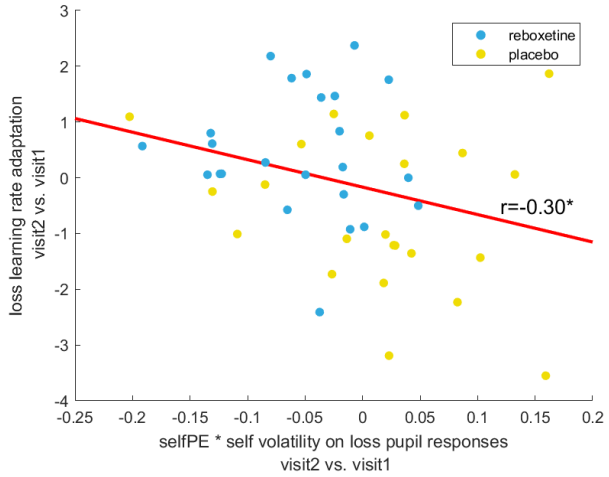

Figure 4.13 Across the two groups, decreased in the modulation of the loss volatility on loss pupil responses to self-PE from visit 1 to visit 2 was associated with increased loss learning rate adaption estimates from visit1 to visit2 when win was volatile.

The Red line indicates the best fitting line across the two groups. Blue dots: the reboxetine group, yellow dots: the placebo group. * $p<0.05$ for the Pearson Correlation.

4.4 Discussion

In this study, we adopted a randomised double-blind placebo-controlled design to investigate the effect of a single dose of reboxetine on volatility adaptation to win and loss outcomes. We first showed that loss accuracies in participants with reboxetine treatment increased relative to their baseline loss performances compared to the placebo group. We then examined the impact of reboxetine on the participants' choices as a function of the most recent outcome, i.e., the positive-stay and negative-switch behaviour indexes (Ppsns). We found that reboxetine increased the effect of loss-volatility on positive-stay negative-switch behaviour particularly when wins were volatile. We then fitted a volatility adaptation model to formalize the learning rate adjustments to volatility. The model results confirmed that the reboxetine group had an increased learning rate adjustment to changes in loss-volatility, particularly when wins were volatile. However, contrary to our expectation, we found that reboxetine reduced the effect of loss volatility on the pupil response to loss prediction errors particularly when wins were volatile. Nevertheless, the level of reduced loss-volatility effect on the pupillometry outcomes negatively correlated with the enhanced loss-volatility learning rate adjustments across all participants. Together, the results suggest that reboxetine might be able to boost learning adaptation, particularly when the general environment is volatile.

Our results showed a marginally significant effect of reboxetine on overall learning rate volatility adaptations, with a more pronounced effect on loss learning rates when wins were volatile (overall highly volatile environment). This finding has some similarities with recent studies which showed that a single dose of methylphenidate, a norepinephrine and dopamine reuptake inhibitor, increased learning rates in humans compared to the placebo group (Howlett et al. 2017), particularly when the environment was volatile (Cook et al. 2019). An obvious difference between these previous findings, and the current results, are that we found an increase in learning rate adaptation, rather than in overall learning rates deployed. Given our findings in Chapters 2 and 3, where symptoms of depression were associated with decreased learning rate adaptation to losses, rather than a generalised reduction of learning rates, these results suggest that the effects of reboxetine found here, may be more closely linked to the antidepressant effect of reboxetine. From a neurochemical perspective, methylphenidate modulates both noradrenaline and dopamine, so these previous studies could not distinguish the roles of the two catecholamines in adaptive learning. In our study, we showed that modulation of norepinephrine (using reboxetine, a selective norepinephrine reuptake inhibitors (SNRIs)) alone could influence learning adaption to volatility. Interestingly, in a more recent study, Lawson et. al. showed that administration of propranolol, a β-adrenergic receptor antagonist, had the opposite effect, i.e., slowing learning in volatile environments (Lawson et al. 2021). Propranolol, which reduces norepinephrine signalling, has been shown to exert an opposite effect on cognition and neural activity than reboxetine (De Martino, Strange, and Dolan 2008; Mizoguchi et al. 2008). Together, these findings support the role of norepinephrine in adaptive learning, which has been hypothesized as a key neuromodulator in the neural implementation of volatility (Dayan and Jyu 2003; Yu and Dayan 2005).

In this study, we also found that administration of reboxetine did not affect baseline pupil size, or dilation, but specifically reduced the modulation effect of loss volatility on pupil responses to loss prediction error, particularly when wins were volatile. We further showed that the reduced modulation of volatility on pupil responses correlated with the increased loss learning rate adaptation to volatility across the reboxetine and placebo groups. However, in Chapter 2, we showed a positive correlation between loss learning rate adaptation to volatility and volatility modulation of pupil responses to loss prediction error. One speculative possibility is that reboxetine might lead to increased tonic pupil diameter and decreased phasic pupil responses. However, we did not find treatment effect for the baseline during the 1s period prior to the outcome delivery. However, our design that there was no visual separation between the first and the second outcome delivery within a trial might hinder this test. Future studies are needed to examine how reboxetine changes tonic and phasic pupil diameter, and the changes in pupil relate to the changes in learning rate adaptation to volatility.

Interestingly, Lawson et. al. showed that propranolol, which had the opposite effect on learning rates to that we found for reboxetine, also reduced the modulation of volatility and prediction error on pupil dilations (Lawson et al. 2021). Another speculative possibility is that there is an inverted U-shaped relationship between the norepinephrine level and pupil responses to volatility, so reducing or increasing norepinephrine levels in healthy participants both lead to reduced modulation of volatility on pupil dilations. More studies are needed to further investigate this effect of manipulating norepinephrine on pupil dilations during a learning task and whether the effect is different for healthy participants and depressed patients.

We found that reboxetine reduced the pupil responses to both win and loss outcome. However, the volatility adaptation effect was more prominent for loss volatility adaptation particularly when win was volatile. In this Chapter and Chapter 2, as well as in the previous work (Pulcu and Browning 2017), loss/punishing outcome seemed to elicit stronger pupil responses than win/rewarding outcome, and the association between pupil volatility adaptation and learning rate volatility adaptation were also stronger for loss than win. Browning et. al (Browning et al. 2015) also showed an association between learning rate volatility adaptation and pupil volatility adaptation to shock (punishing stimuli). Together, the results suggest that norepinephrine system is more sensitive to punishing outcome. Therefore, the effect of reboxetine might be more prominent for punishing outcome than rewarding outcome.

Nevertheless, we showed that a single dose of reboxetine affected adaptive learning particularly for punishment and that this effect was accompanied by changes in pupil response during the task. Together with the findings from previous Chapters, our results have important implications in understanding the psychopharmacological mechanism of reboxetine on depression. In Chapter 2, we found an impairment in adjusting learning to the volatility of losses in currently depressed patients. In Chapter 3, we further showed that reduced loss learning rate adaptation to volatility was associated with depressive symptoms. In the current Chapter, we showed that reboxetine, a SNRIs antidepressant, may increase loss learning rate adaptation in volatile environments. This raises the possibility that reboxetine might help alleviate depression by improving adaptive learning of punishment information in depressed patients. Further studies are needed to better understand the effect of reboxetine on adaptive learning in clinically depressed patients. For example, it would be interesting to test whether changes in loss learning rate adaptation to volatility mediated the symptomatic effects of reboxetine in clinical treatment studies. If successful, this would suggest that these cognitive effects might usefully be used to predict patients who might benefit more from taking reboxetine.

In the next chapter we summarise the studies and findings described in this thesis and consider their broader implications.

5 CONCLUSION AND GENERAL DISCUSSION

In this thesis, we used a task in which we independently manipulated the volatility of win and loss outcomes and systematically examined the underlying mechanisms of adaptive learning in depression using computational modelling, fMRI, pharmacological manipulation and pupillometry recordings. The results have provided some consistent evidence that link depression with inadequate learning adaptation to the volatility of loss outcomes. In Chapter 2, we showed that currently depressed patients had lower loss learning rate volatility adaptations, estimated using a volatility adaptation model, compared to remitted depressed patients and healthy controls. This was accompanied by reduced loss volatility modulation of pupil responses to loss outcomes (loss prediction errors) in the currently depressed group compared to the other two groups. In Chapter 3, using a similar volatility adaptation model, we showed that loss learning rate adaptation estimates negatively correlated with individual differences in overall negativity levels. This was paralleled by decreased segregation of win and loss volatility signals in a distributed brain network of outcome sensitive regions, especially in the ACC. In Chapter 4, we showed that a single dose of a norepinephrine reuptake inhibitors (NRIs) antidepressant (reboxetine) could increase loss learning rate adaptations to volatility. This was also associated with a reduced volatility effect on the pupil responses to loss outcomes (loss prediction errors), particularly when wins were volatile.

In this final chapter, we discuss some general conclusions that can be drawn from these three studies, how they might be interpreted, some limitations of the studies, as well as how future work might help clarify and expand on the ideas.

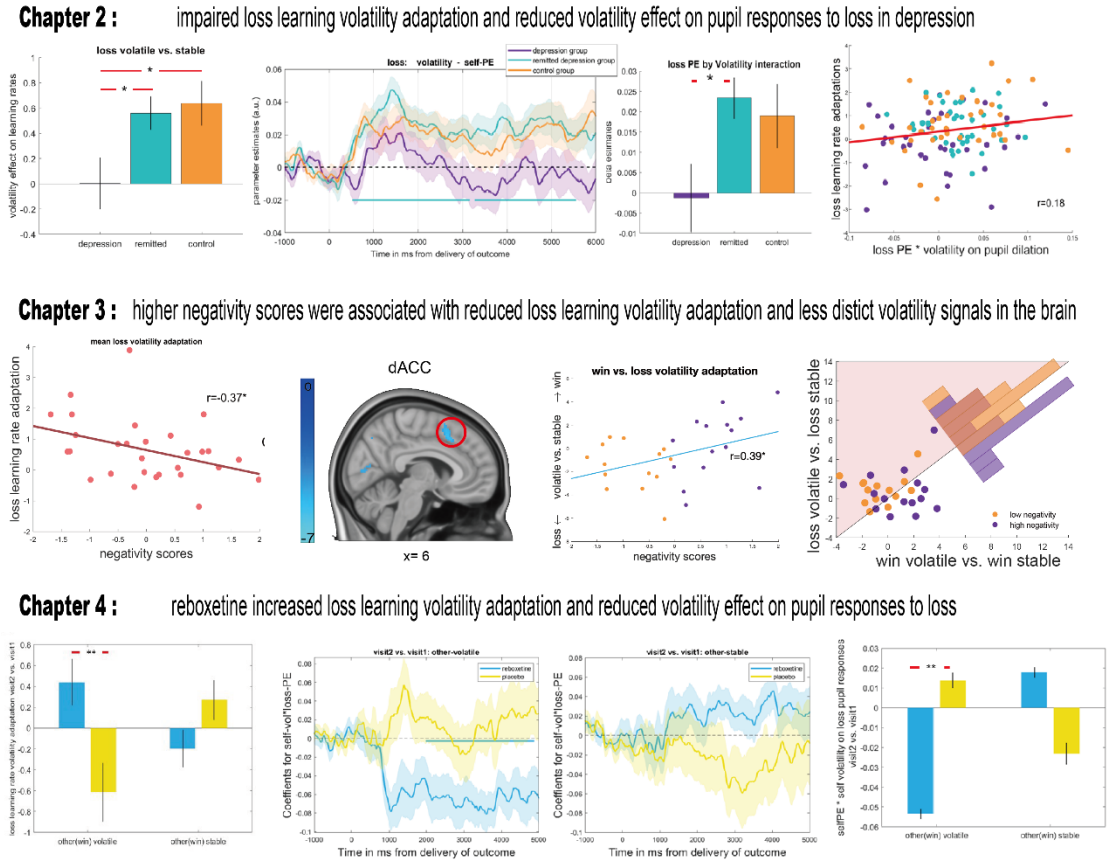

Figure 5.1 A summary of the main results for each Chapter in this thesis

The association between symptoms of depression and reduced learning rate adaptations to changes in loss volatility

First and foremost, our results suggest that the impaired ability to adapt learning to the volatility level of loss outcomes in depression might be state-dependent, and probably linked to individual differences in negativity generally, rather than to anxiety or depression specifically. In Chapter 2, we showed a significant group difference in loss learning rate adaptation between patients who were currently depressed and patients who had remitted from depression at the time of the testing. In that chapter, we also showed correlations between the loss learning rate adaptation estimates and various depression and anxiety questionnaires scores. In Chapter 3, we showed such correlation again, i.e., the loss learning rate adaptation estimates negatively correlated with participants' overall negativity levels in a subclinical cohort, recruited to have a range of depression scores. These results are consistent with previous studies showing reduced punishment learning rate volatility adaptations in people with higher anxiety and/or general negativity levels (Browning et al. 2015; Pulcu and Browning 2017; Beltzer et al. 2019; Gagne et al. 2020). Together, the results suggest that impaired loss/punishment learning adaptation to volatile environment might be a common deficit in people with mood disorders.

However, it is still not clear whether the difficulty in adjusting loss learning causally contributes to the symptoms or whether the symptoms cause the learning deficit. Our finding in Chapter 2, that showed only the currently depressed group had reduced loss learning rate adaptations but not in the remitted depressed group, suggests that this cognitive characteristic might be state-dependent. Nevertheless, studies are needed to examine whether there is a causal relationship between depression and difficulties in loss learning rate volatility adaptation. For example, future studies can adopt a longitudinal design to check whether improvement in learning adaptation precedes depression symptom alleviation. Or more directly, running experimental studies, which manipulate loss learning rate adaptation using cognitive or pharmacological intervention and see whether this improves depression. In Chapter 4, we showed that an antidepressant (reboxetine) could increase loss learning rate adaptation, suggesting that antidepressants could help alleviate depression symptoms by changing how individuals adapt to changes in loss volatility in the environment. However, we did not use clinical depression sample in this study, and did assess the changes in depression symptoms. Further studies could better examine this hypothesis by formally assessing whether changes in loss learning rate adaption mediate the impact of treatments like reboxetine on symptoms in a clinical population. For example, a randomised controlled trial (e.g., reboxetine vs. placebo) for depressed patients, where changes in symptoms and learning rate adaptation are measured over many weeks.

It is also not clear why the impaired learning volatility adaptation is more specific for loss and why the reboxetine effect is also more prominent for loss. However, the association between negative mood and loss learning rate volatility adaptation were consistent across many studies (Beltzer et al. 2019; Hein et al. 2019; Pulcu and Browning 2017; Browning et al. 2015; Gagne et al. 2020). In Chapter 2&4, we showed that loss/punishing outcome seemed to elicit stronger pupil responses than win/rewarding outcome, and the association between pupil volatility adaptation and learning rate volatility adaptation for loss/punishing outcome. This is also consistent with previous work (Pulcu and Browning 2017; Browning et al. 2015). Together, this suggests that the norepinephrine system may be more sensitive to punishing than rewarding outcomes. Therefore, the effect of reboxetine (targeting the norepinephrine system) would also be expected to be more prominent for punishing than rewarding outcomes. This raises the, admittedly speculative, possibility that people with depression might show impairment in loss adaptive learning because of their dysfunctional norepinephrine system.

However, it is not clear whether the deficit in loss learning adaptation is specific to negative moods, or whether it might also be associated with other psychiatric symptoms, such as psychosis. Studies with different patient cohorts are needed to clarify the specificity of reduced loss learning rate volatility adaptation to different types of psychiatric symptoms.

Finally, the cognitive and neural mechanisms underlying the association between negative emotions and loss learning adaptation deficits are not clear. In the following sections, we discuss some potential cognitive and neural mechanisms suggested by the studies in this thesis.

The effect of distractor volatility on learning rate adaptation

The results from the three studies suggest a potential underlying cognitive mechanism for the impaired loss adaptive learning in depression; that people with higher depression levels might be more affected by distractor volatility. In Chapter 3, we showed that higher overall negativity scores were associated with reduced loss-volatility learning rate adaptations particularly when wins were volatile, suggesting that win distractor volatility might hinder the loss-volatility learning adaptation in people with higher negativity scores. In Chapter 3, we also showed less segregated win-volatility and loss-volatility signals in the brain for people with higher negativity levels, suggesting that having to deal with two sources of volatile information might be particularly difficult for people with higher negativity scores. Consistent with this, the impaired loss learning rate adaptations in currently depressed patients we showed in Chapter 2 were also apparent when wins were volatile. In Chapter 4, we showed that an antidepressant (reboxetine) increased loss learning rate volatility adaptations in healthy participants, particularly when wins were volatile.

One interpretation of these results is that people with depression might have more difficulty in differentiating distinct sources of volatility and, as a result, find it more difficult to adapt their learning optimally to the volatility level of a specific association, when other associations are volatile. If correct, this would suggest that the most efficient route to ameliorating these cognitive deficits would be helping patients segregate different associations rather than specifically enhancing the degree to which they adapt to changes in volatility. However, the evidence for this conclusion is indirect, and more studies are needed to further examine this possibility. For example, in the current work only two levels of volatility were employed (high/low) leaving open the possibility that the findings are limited to situations in which volatility is treated in such a binary fashion, rather than as a continuous quantity. Future studies might address this possibility using a magnitude learning task and have participants learn about one option while varying the volatility of the other outcome across a broad range of levels to examine how parametric modulation of distractor volatility levels affects participants' volatility adaptation and how this might be associated with depression or anxiety.

The Role of Norepinephrine in Learning Adaptation to Loss Volatility

The results indicated that norepinephrine may play an important role in maladaptive learning adjustment to volatility, particular for punishment information, in depression. In Chapter 2, we found that the pupil responses to loss outcomes were diminished in the currently depressed

group compared to the remitted depressed group and the healthy control group, suggesting that the impaired loss learning adaptation might be caused by (or at least, related to) reduced norepinephrine (Aston-Jones and Cohen 2005; Joshi et al. 2016) or other neurotransmitters (de Gee et al. 2017) that modulate pupil size. In Chapter 4, we manipulated the norepinephrine system using a single dose of reboxetine, a norepinephrine reuptake inhibitor, and showed that the drug reduced the effect of win-volatility on how loss-volatility modulated the pupil responses to loss outcomes. More importantly, we showed that reboxetine increased loss learning adaptations in volatile environments. In other words, we showed a causal relationship between the norepinephrine system and loss learning adjustment to volatility.

However, it is not clear why in Chapter 2 we found a positive correlation between the volatility effect on pupil responses to loss PEs and loss learning rate adaptation. However, in Chapter 4, the reduced volatility effect on pupil responses to loss PEs induced by reboxetine was associated with increase in loss learning rate adaptation. One speculative possibility is that reboxetine might lead to increased tonic pupil diameter and decreased phasic pupil responses. Future studies are needed to examine how reboxetine changes tonic and phasic pupil diameter, and the changes in pupil relate to the changes in learning rate adaptation to volatility.

Nevertheless, these results raise the possibility that reboxetine may alleviate depression symptoms by helping patients to better adapt to the volatility level of punishments in their environment. By adjusting learning rates to the volatility of punishing outcomes, patients might be more efficient at tracking the environmental statistics (i.e., the associations), so that they could better avoid the negative affect of experiencing surprising punishing events (which lead to learned helplessness). In Chapter 4, we showed that loss learning rate adaptation could be modified by an NRIs antidepressant (reboxetine) in healthy participants, which suggests the ability to adapt learning to the environment might be modifiable. Clearly however, more studies are needed to examine whether reboxetine treatment over a meaningful treatment period in currently depressed patients could increase their loss learning volatility adaptation, and whether the effect of reboxetine on learning rate adaptation mediates the treatment effect, i.e., alleviation of depression symptoms. Even if the effect of reboxetine on learning rate adaptation does not mediate its clinical effect, it might be useful as a behavioural marker of the effect of reboxetine and, potentially, a predictor of treatment outcome. This could be tested, for example, in a cohort of patients treated with reboxetine, by testing whether those patients who showed early enhancement of volatility adaptations to reboxetine were more likely to respond clinically to treatment. Another question that arises is whether the observed effect is specific to reboxetine or NRIs antidepressants. Future studies could examine whether other types of antidepressants or other interventions that target other neurotransmitters that also modulate pupil dilation, such as acetylcholine (Larsen and Waters 2018), have similar effects on loss learning rates.

Our results are in general consistent with recent studies that showed increasing norepinephrine level increased learning rates (Howlett et al. 2017), particularly when the environment was volatile (Cook et al. 2019), while decreasing norepinephrine reduced learning rates in volatile environments (De Martino, Strange, and Dolan 2008; Mizoguchi et al. 2008; Lawson et al. 2021). However, it is still not clear how norepinephrine affects loss volatility adaptation particularly when presented with distractor volatility. Increasing norepinephrine level might increase loss adaptive learning by allowing better volatility information processing through increasing attention and volatility signal segregation in the brain.

The Potential Neural Mechanism of the Impaired Loss Volatility Learning Adaptation in Depression

In Chapter 3, we showed that win volatility predominantly increased win outcome signals in prefrontal regions which had increased responses to positive outcomes, such as rACC and vmPFC, while loss volatility predominantly increased loss outcome signals in prefrontal regions which had increased responses to negative outcomes, such as dACC and MFG. Moreover, we showed that the overall negativity was associated with relatively more win volatility adaptations than loss volatility adaptations in dACC (dominated by loss volatility), and relatively more loss volatility adaptations than win volatility adaptations in the rACC (dominated by win volatility). The results suggest that people with higher overall negativity scores might have less distinct win-volatility and loss-volatility in the prefrontal cortex. Interestingly, a mediation analysis suggested that the ratio of win vs. loss volatility adaptation signals in the dACC might affect the negativity scores through the mediation effect of biased loss volatility learning rate adaptation. The results suggested a role of dACC in loss volatility adaptation and that less distinct loss volatility representation in the dACC could be associated with impaired loss volatility adaptation in depression.

The results are consistent with previous work that showed the dACC tracks volatility in reinforcement learning tasks (Behrens et al. 2007; Silvetti, Seurinck, and Verguts 2013; Rushworth and Behrens 2008). Our results further suggested that a possible dysfunction in the dACC might be associated impaired loss volatility learning adaptation in depression. The dACC could help reduce distractor interference by supporting a more precise segregation of distinct associations in the environment which allow more effective suppression of conflicting information, perhaps via modulation of the NE system and pupil responses (Ebitz and Platt 2015). Therefore, less distinct loss volatility representation in people with depression might interfere with their ability to adjust learning optimally to the volatility level in the environment, potentially leading to emotional distress (Blain and Rutledge 2020).

The mechanism underlying any difficulty in segregating win and loss volatility in depression is not clear. It is perhaps one aspect of a more general effect, e.g., depressed patients are just less

good at representing things. Consistent with this a number of studies have shown that people with higher levels of anxiety find it difficult to segregate ambiguous and threatening cues during conditioning (Lissek et al. 2014). It would be interesting to test whether reduced adaptation to volatility was associated with these other potential markers of reduced cognitive segregation.

Some Pros and Cons of the New Information Bias Learning Task

We adopted a new information bias learning task in Chapter 3&4, which was developed based on the task used in Chapter 2. We made two important changes to the task to enable us to better investigate the effect of other-volatility. The first change made was that the "stable" schedule of the task was not fixed at 50% anymore, but had a probability change somewhere in the middle of the block either from 15% to 85% or from 80% to 20%, which reduced the difference in expected uncertainty between the volatile and stable schedules. Second, we included a both stable block (i.e., in which both win and loss outcomes were relatively stable), which allows us to test the valence*self-volatility*other-volatility effect in our analysis.

With the new task, we demonstrated clear main effects of other volatility and self-volatility*other-volatility interactions in behaviour data and model results in Chapter 3&4, which it was not possible to examine with the old version of the task. Importantly, in Chapter 3, we showed that the overall negativity scores were associated with loss learning rate adaptation when wins were volatile, but not when wins were stable. Similarly, we showed in Chapter 4 that a single dose of reboxetine increased loss learning adaptation when wins were volatile, but not when wins were stable. Such differences were not found in Chapter 2 because the old version of task in Chapter 2 only studied learning rate adaptations when other-volatility was volatile. Future studies are needed to examine whether the loss learning rate adaptations in clinically depressed participants were specifically impaired when wins were volatile, but not when wins were stable, with the new version of the task.

However, we did not find much evidence of the effect of other-volatility nor the interaction of self and other volatility on the neural data in Chapter 3, although these effects were shown in the behavioural results in Chapter 3. A possible reason might be that the study might be statistically underpowered to detect the other-volatility effects, which, suggested by the behavioural results, are likely to be smaller than those for the self-volatility. Nevertheless, we showed that overall negativity scores were associated with a reduced segregation of win and loss volatility signals in the brain. Future studies should increase the statistical power and further examine how other-volatility might affect the neural computations of the self-volatility. For example, increasing participant sample size might increase the statistical power in general. Further, studies could also try to increase volatility level contrasts between win and loss volatility, for example, introducing more rapid changes in the task.

A second limitation to the new task was that the stable condition had one change in association in it. In other words, it was less stable than in the original task. This aspect of the task design was included as it allowed us to include the stable/stable block (without it, there would have been no reversal of optimal choice, making it difficult to measure participant behaviour) and also prevented the stable schedules from being heavily biased to one particular shape. However, the inclusion of a change in association during the stable block necessarily reduced the volatility contrast in the task (i.e., the difference in volatility between volatile and stable periods was reduced) raising the possibility that some aspects of volatility adaptation may not have been detected by the new task.

In conclusion, using computational modelling approaches, we have shown consistent evidence in this thesis that depression is associated with impaired learning adaptation to the volatility of loss outcomes. The pupillometry and fMRI results highlighted the important role of the dACC and its interaction with the norepinephrine system in the control of learning and in relation to symptoms of depression. The pharmacological study provided important information about the causal role of norepinephrine in the adaptation to uncertainty and suggested a potential novel cognitive mechanism of NRIs antidepressants. Together, this thesis showed consistent evidence of the biased information processing in depression and suggested potential underlying cognitive and neural mechanisms.

REFERENCES

Admon, Roee, Roselinde H Kaiser, Daniel G Dillon, Miranda Beltzer, Franziska Goer, David P Olson, Gordana Vitaliano, and Diego A Pizzagalli. 2017. 'Dopaminergic enhancement of striatal response to reward in major depression', *American Journal of Psychiatry*, 174: 378-86.

Admon, Roee, and Diego A Pizzagalli. 2015. 'Dysfunctional reward processing in depression', *Current Opinion in Psychology*, 4: 114-18.

Ahrens, William H, Darrell J Cox, and Girish Budhwar. 1990. 'Use of the arcsine and square root transformations for subjectively determined percentage data', *Weed Science*, 38: 452-58.

Association, American Psychiatric. 2013. *Diagnostic and statistical manual of mental disorders (DSM-5®)* (American Psychiatric Pub).

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.

Auer, Dorothee P., Benno Pütz, Eduard Kraft, Bernadette Lipinski, Julia Schill, and Florian Holsboer. 2000. 'Reduced glutamate in the anterior cingulate cortex in depression: an in vivo proton magnetic resonance spectroscopy study', *Biological psychiatry*, 47: 305-13.

Aylward, Jessica, Vincent Valton, Woo-Young Ahn, Rebecca L. Bond, Peter Dayan, Jonathan P. Roiser, and Oliver J. Robinson. 2019. 'Altered learning under uncertainty in unmedicated mood and anxiety disorders', *Nature Human Behaviour*, 3: 1116-23.

Bachmann, S. 2018. 'Epidemiology of Suicide and the Psychiatric Perspective', *Int J Environ Res Public Health*, 15.

Bechara, Antoine, Antonio R Damasio, Hanna Damasio, and Steven W Anderson. 1994. 'Insensitivity to future consequences following damage to human prefrontal cortex', *Cognition*, 50: 1-3.

Bechara, Antoine, Daniel Tranel, and Hanna Damasio. 2000. 'Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions', *Brain*, 123: 2189-202.

Behrens, Timothy E. J., Laurence T. Hunt, Mark W. Woolrich, and Matthew F. S. Rushworth. 2008. 'Associative learning of social value', *Nature*, 456: 245-49.

Behrens, Timothy EJ, Mark W Woolrich, Mark E Walton, and Matthew FS Rushworth. 2007. 'Learning the value of information in an uncertain world', *Nature neuroscience*, 10: 1214-21.

Beltzer, Miranda L., Stephen Adams, Peter A. Beling, and Bethany A. Teachman. 2019. 'Social Anxiety and Dynamic Social Reinforcement Learning in a Volatile Environment', *Clinical Psychological Science*, 7: 1372-88.

Biselli, T., S. S. Lange, L. Sablottny, J. Steffen, and A. Walther. 2021. 'Optogenetic and chemogenetic insights into the neurocircuitry of depression-like behaviour: A systematic review', *Eur J Neurosci*, 53: 9-38.

Blain, Bastien, and Robb B. Rutledge. 2020. 'Momentary subjective well-being depends on learning and not reward', *Elife*, 9: e57977.

Bowling, Shannon R, Mohammad T Khasawneh, Sittichai Kaewkuekool, and Byung Rae Cho. 2009. 'A logistic approximation to the cumulative normal distribution', *Journal of industrial engineering and management*, 2: 114-27.

Bremner, J. Douglas, Meena Vythilingam, Eric Vermetten, Ahsan Nazeer, Jahangir Adil, Sarfraz Khan, Lawrence H. Staib, and Dennis S. Charney. 2002. 'Reduced volume of orbitofrontal cortex in major depression', *Biological psychiatry*, 51: 273-79.

Brooks, Jonathan CW, Christian F Beckmann, Karla L Miller, Richard G Wise, Carlo A Porro, Irene Tracey, and Mark Jenkinson. 2008. 'Physiological noise modelling for spinal functional magnetic resonance imaging studies', *Neuroimage*, 39: 680-92.

Brooks, Stephen P, and Andrew Gelman. 1998. 'General methods for monitoring convergence of iterative simulations', *Journal of computational and graphical statistics*, 7: 434-55.

Brown, Vanessa M., Jonathan Wilson, Michael N. Hallquist, Katalin Szanto, and Alexandre Y. Dombrovski. 2020. 'Ventromedial prefrontal value signals and functional connectivity during decision-making in suicidal behavior and impulsivity', *Neuropsychopharmacology*, 45: 1034-41.

Browning, M., C. S. Carter, C. Chatham, H. Den Ouden, C. M. Gillan, J. T. Baker, A. M. Chekroud, R. Cools, P. Dayan, J. Gold, R. Z. Goldstein, C. A. Hartley, A. Kepecs, R. P. Lawson, J. Mourao-Miranda, M. L. Phillips, D. A. Pizzagalli, A. Powers, D. Rindskopf, J. P. Roiser, K. Schmack, D. Schiller, M. Sebold, K. E. Stephan, M. J. Frank, Q. Huys, and M. Paulus. 2020. 'Realizing the Clinical Potential of Computational Psychiatry: Report From the Banbury Center Meeting, February 2019', *Biol Psychiatry*, 88: e5-e10.

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: 590-96.

Burgdorf, Jeffrey, and Jaak Panksepp. 2006. 'The neurobiology of positive emotions', *Neuroscience & Biobehavioral Reviews*, 30: 173-87.

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. 'Stan: A probabilistic programming language', *Journal of statistical software*, 76.
Cella, Matteo, Simon Dymond, and Andy Cooper. 2010. 'Impaired flexible decision-making in major depressive disorder', *Journal of affective disorders*, 124: 207-10.
Chase, H. W., M. J. Frank, A. Michael, E. T. Bullmore, B. J. Sahakian, and T. W. Robbins. 2010. 'Approach and avoidance learning in patients with major depression and healthy controls: relation to anhedonia', *Psychol Med*, 40: 433-40.
Cheng, W., E. T. Rolls, J. Qiu, W. Liu, Y. Tang, C. C. Huang, X. Wang, J. Zhang, W. Lin, L. Zheng, J. Pu, S. J. Tsai, A. C. Yang, C. P. Lin, F. Wang, P. Xie, and J. Feng. 2016. 'Medial reward and lateral non-reward orbitofrontal cortex circuits change in opposite directions in depression', *Brain*, 139: 3296-309.
Cipriani, A., T. A. Furukawa, G. Salanti, A. Chaimani, L. Z. Atkinson, Y. Ogawa, S. Leucht, H. G. Ruhe, E. H. Turner, J. P. T. Higgins, M. Egger, N. Takeshima, Y. Hayasaka, H. Imai, K. Shinohara, A. Tajika, J. P. A. Ioannidis, and J. R. Geddes. 2018. 'Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis', *Lancet*, 391: 1357-66.
Collaborators, GBD 2017 Disease and Injury Incidence and Prevalence. 2018. 'Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017', *The Lancet*, 392: 1789-858.
Cook, Jennifer L., Jennifer C. Swart, Monja I. Froböse, Andreea O. Diaconescu, Dirk E. M. Geurts, Hanneke E. M. den Ouden, and Roshan Cools. 2019. 'Catecholaminergic modulation of meta-learning', *Elife*, 8: e51439.
Cox, Julia, and Ilana B. Witten. 2019. 'Striatal circuits for reward learning and decision-making', *Nature Reviews Neuroscience*, 20: 482-94.
Davey, C. G., B. J. Harrison, M. Yuecel, and N. B. Allen. 2012. 'Regionally specific alterations in functional connectivity of the anterior cingulate cortex in major depressive disorder', 42: 2071-81.
Daw, Nathaniel D, Yael Niv, and Peter Dayan. 2005. 'Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control', *Nature neuroscience*, 8: 1704-11.
Dayan, Peter, and Angela Jyu. 2003. 'Uncertainty and Learning', *IETE Journal of Research*, 49: 171-81.
De Berker, Archy O, 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: 10996.
de Gee, Jan Willem, Olympia Colizoli, Niels A Kloosterman, Tomas Knapen, Sander Nieuwenhuis, and Tobias H Donner. 2017. 'Dynamic modulation of decision biases by brainstem arousal systems', *Elife*, 6: e23232.
De Martino, Benedetto, Bryan A Strange, and Raymond J Dolan. 2008. 'Noradrenergic neuromodulation of human attention for emotional and neutral stimuli', *Psychopharmacology*, 197: 127-36.
Dickstein, D. P., E. C. Finger, M. A. Brotman, B. A. Rich, D. S. Pine, J. R. Blair, and E. Leibenluft. 2010. 'Impaired probabilistic reversal learning in youths with mood and anxiety disorders', *Psychol Med*, 40: 1089-100.
Dombrovski, A. Y., L. Clark, G. J. Siegle, M. A. Butters, N. Ichikawa, B. J. Sahakian, and K. Szanto. 2010. 'Reward/Punishment reversal learning in older suicide attempters', *Am J Psychiatry*, 167: 699-707.
Dugas, Michel J, Andrea Schwartz, and Kylie Francis. 2004. 'Brief report: Intolerance of uncertainty, worry, and depression', *Cognitive Therapy and Research*, 28: 835-42.
Ebitz, R. Becket, and Michael L Platt. 2015. 'Neuronal Activity in Primate Dorsal Anterior Cingulate Cortex Signals Task Conflict and Predicts Adjustments in Pupil-Linked Arousal', *Neuron*, 85: 628-40.
Elliott, R, BJ Sahakian, JJ Herrod, TW Robbins, and ES Paykel. 1997. 'Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment', *Journal of Neurology, Neurosurgery & Psychiatry*, 63: 74-82.
Elliott, Rebecca, Barbara J Sahakian, AP McKay, JJ Herrod, Trevor W Robbins, and ES Paykel. 1996. 'Neuropsychological impairments in unipolar depression: the influence of perceived failure on subsequent performance', *Psychological medicine*, 26: 975-89.
Engelmann, JB, GS Berns, and BW Dunlop. 2017. 'Hyper-responsivity to losses in the anterior insula during economic choice scales with depression severity', *Psychological medicine*, 47: 2879.
Fazeli, Sepideh, and Christian Büchel. 2018. 'Pain-Related Expectation and Prediction Error Signals in the Anterior Insula Are Not Related to Aversiveness', *The Journal of Neuroscience*, 38: 6461.
Findling, Charles, Vasilisa Skvortsova, Rémi Dromnelle, Stefano Palminteri, and Valentin Wyart. 2019. 'Computational noise in reward-guided learning drives behavioral variability in volatile environments', *Nature neuroscience*, 22: 2066-77.
Gagne, Christopher, Ondrej Zika, Peter Dayan, and Sonia J. Bishop. 2020. 'Impaired adaptation of learning to contingency volatility in internalizing psychopathology', *Elife*, 9: e61387.
Gard, David E., Marja Germans Gard, Ann M. Kring, and Oliver P. John. 2006. 'Anticipatory and consummatory components of the experience of pleasure: A scale development study', *Journal of Research in Personality*, 40: 1086-102.
Garrison, Jane, Burak Erdeniz, and John Done. 2013. 'Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies', *Neuroscience & Biobehavioral Reviews*, 37: 1297-310.
Glennon, E., I. Carcea, A. R. O. Martins, J. Multani, I. Shehu, M. A. Svirsky, and R. C. Froemke. 2019. 'Locus coeruleus activation accelerates perceptual learning', *Brain Res*, 1709: 39-49.
Gotlib, I. H., J. P. Hamilton, R. E. Cooney, M. K. Singh, M. L. Henry, and J. Joormann. 2010. 'Neural processing of reward and loss in girls at risk for major depression', *Arch Gen Psychiatry*, 67: 380-7.

Greenberg, Paul E, Andree-Anne Fournier, Tammy Sisitsky, Crystal T Pike, and Ronald C Kessler. 2015. 'The economic burden of adults with major depressive disorder in the United States (2005 and 2010)', *The Journal of clinical psychiatry*.
Greenberg, T., H. W. Chase, J. R. Almeida, R. Stiffler, C. R. Zevallos, H. A. Aslam, T. Deckersbach, S. Weyandt, C. Cooper, M. Toups, T. Carmody, B. Kurian, S. Peltier, P. Adams, M. G. McInnis, M. A. Oquendo, P. J. McGrath, M. Fava, M. Weissman, R. Parsey, M. H. Trivedi, and M. L. Phillips. 2015. 'Moderation of the Relationship Between Reward Expectancy and Prediction Error-Related Ventral Striatal Reactivity by Anhedonia in Unmedicated Major Depressive Disorder: Findings From the EMBARC Study', *Am J Psychiatry*, 172: 881-91.
Gutiérrez-Rojas, Luis, Alejandro Porras-Segovia, Henry Dunne, Nelson Andrade-González, and Jorge A. Cervilla. 2020. 'Prevalence and correlates of major depressive disorder: a systematic review', *Brazilian Journal of Psychiatry*.
Hajós, Mihály, Joseph C Fleishaker, Jacqueline K Filipiak‐Reisner, Mark T Brown, and Erik HF Wong. 2004. 'The selective norepinephrine reuptake inhibitor antidepressant reboxetine: pharmacological and clinical profile', *CNS drug reviews*, 10: 23-44.
Hall, G. B., A. M. Milne, and G. M. Macqueen. 2014. 'An fMRI study of reward circuitry in patients with minimal or extensive history of major depression', *Eur Arch Psychiatry Clin Neurosci*, 264: 187-98.
Hayden, B. Y., S. R. Heilbronner, J. M. Pearson, and M. L. Platt. 2011. 'Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior', *J Neurosci*, 31: 4178-87.
Hayton, James C, David G Allen, and Vida Scarpello. 2004. 'Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis', *Organizational research methods*, 7: 191-205.
Hegedűs, Klára M, Andrea Szkaliczki, Bernadett I Gál, Bálint Andó, Zoltán Janka, and Péter Z Álmos. 2018. 'Decision-making performance of depressed patients within 72 h following a suicide attempt', *Journal of affective disorders*, 235: 583-88.
Hein, Thomas P., Lilian A. Weber, Jan de Fockert, and Maria Herrojo Ruiz. 2019. 'State anxiety biases estimates of uncertainty during reward learning in volatile environments', *bioRxiv*: 809749.
Henriques, J. B., J. M. Glowacki, and R. J. Davidson. 1994. 'Reward fails to alter response bias in depression', *J Abnorm Psychol*, 103: 460-6.
Henriques, Jeffrey B, and Richard J Davidson. 2000. 'Decreased responsiveness to reward in depression', *Cognition & Emotion*, 14: 711-24.
Hershenberg, Rachel, Theodore D. Satterthwaite, Aylin Daldal, Natalie Katchmar, Tyler M. Moore, Joseph W. Kable, and Daniel H. Wolf. 2016. 'Diminished effort on a progressive ratio task in both unipolar and bipolar depression', *Journal of affective disorders*, 196: 97-100.
Holmes, A. J., and D. A. Pizzagalli. 2007. 'Task feedback effects on conflict monitoring and executive control: relationship to subclinical measures of depression', *Emotion*, 7: 68-76.
Holroyd, C. B., and M. G. H. Coles. 2002. 'The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity', *Psychol Rev*, 109: 679-709.
Howlett, Jonathon R., He Huang, Cédric M. Hysek, and Martin P. Paulus. 2017. 'The effect of single-dose methylphenidate on the rate of error-driven learning in healthy males: a randomized controlled trial', *Psychopharmacology*, 234: 3353-60.
Huang, H., W. Thompson, and M. P. Paulus. 2017. 'Computational Dysfunctions in Anxiety: Failure to Differentiate Signal From Noise', *Biol Psychiatry*, 82: 440-46.
Husain, Masud, and Jonathan P. Roiser. 2018. 'Neuroscience of apathy and anhedonia: a transdiagnostic approach', *Nature Reviews Neuroscience*, 19: 470-84.
Huys, Quentin J. M., Tiago V. Maia, and Michael J. Frank. 2016. 'Computational psychiatry as a bridge from neuroscience to clinical applications', *Nature neuroscience*, 19: 404-13.
Huys, Quentin JM, Diego A Pizzagalli, Ryan Bogdan, and Peter Dayan. 2013. 'Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis', *Biology of mood & anxiety disorders*, 3: 12.
Jahn, Caroline I., Chiara Varazzani, Jérôme Sallet, Mark E. Walton, and Sébastien Bouret. 2020. 'Noradrenergic But Not Dopaminergic Neurons Signal Task State Changes and Predict Reengagement After a Failure', *Cerebral Cortex*, 30: 4979-94.
Jenkinson, Mark. 2004. 'Improving the registration of B0-disorted Epi images using calculated cost function weights: we 202', *Neuroimage*, 22.
Jenkinson, Mark, Peter Bannister, Michael Brady, and Stephen Smith. 2002. 'Improved optimization for the robust and accurate linear registration and motion correction of brain images', *Neuroimage*, 17: 825-41.
Jenkinson, Mark, Christian F Beckmann, Timothy EJ Behrens, Mark W Woolrich, and Stephen M Smith. 2012. 'Fsl', *Neuroimage*, 62: 782-90.
Jepma, Marieke, Peter R Murphy, Matthew R Nassar, Mauricio Rangel-Gomez, Martijn Meeter, and Sander Nieuwenhuis. 2016. 'Catecholaminergic regulation of learning rate in a dynamic environment', *PLOS Computational Biology*, 12: e1005171.
Johnston, B. A., S. Tolomeo, V. Gradin, D. Christmas, K. Matthews, and J. D. Steele. 2015. 'Failure of hippocampal deactivation during loss events in treatment-resistant depression', *Brain*, 138: 2766-76.
Joshi, S., Y. Li, R. M. Kalwani, and J. I. Gold. 2016. 'Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex', *Neuron*, 89: 221-34.
Keedwell, Paul A., Chris Andrew, Steven C. R. Williams, Mick J. Brammer, and Mary L. Phillips. 2005. 'A Double Dissociation of Ventromedial Prefrontal Cortical Responses to Sad and Happy Stimuli in Depressed and Healthy Individuals', *Biological psychiatry*, 58: 495-503.
Kennerley, Steven W., Timothy E. J. Behrens, and Jonathan D. Wallis. 2011. 'Double dissociation of value computations in orbitofrontal and anterior cingulate neurons', *Nature neuroscience*, 14: 1581-89.

Khan, A., J. Faucett, P. Lichtenberg, I. Kirsch, and W. A. Brown. 2012. 'A systematic review of comparative efficacy of treatments and controls for depression', *PLOS ONE*, 7: e41778.
Kleim, Birgit, Désirée Gonzalo, and Anke Ehlers. 2011. 'The Depressive Attributions Questionnaire (DAQ): Development of a short self-report measure of depressogenic attributions', *Journal of Psychopathology and Behavioral Assessment*, 33: 375-85.
Kringelbach, Morten L., and Edmund T. Rolls. 2004. 'The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology', *Progress in neurobiology*, 72: 341-72.
Kumar, P., G. Waiter, T. Ahearn, M. Milders, I. Reid, and J. D. Steele. 2008. 'Abnormal temporal difference reward-learning signals in major depression', *Brain*, 131: 2084-93.
Kumar, Poornima, Franziska Goer, Laura Murray, Daniel G. Dillon, Miranda L. Beltzer, Andrew L. Cohen, Nancy H. Brooks, and Diego A. Pizzagalli. 2018. 'Impaired reward prediction error encoding and striatal-midbrain connectivity in depression', *Neuropsychopharmacology*, 43: 1581-88.
Larsen, Rylan S., and Jack Waters. 2018. 'Neuromodulatory Correlates of Pupil Dilation', *Frontiers in Neural Circuits*, 12.
Lavin, 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.
Lawson, Rebecca P., James Bisby, Camilla L. Nord, Neil Burgess, and Geraint Rees. 2021. 'The Computational, Pharmacological, and Physiological Determinants of Sensory Learning under Uncertainty', *Current Biology*, 31: 163-72.e4.
Levy-Gigi, E., and S. Kéri. 2015. 'The interactive effect of negative reversal learning and age on depression: Possible cognitive mechanisms underlying the elevated depressive symptoms in older adults', *Psychol Aging*, 30: 341-7.
Lissek, Shmuel, Antonia N. Kaczkurkin, Stephanie Rabin, Marilla Geraci, Daniel S. Pine, and Christian Grillon. 2014. 'Generalized Anxiety Disorder Is Associated With Overgeneralization of Classically Conditioned Fear', *Biological psychiatry*, 75: 909-15.
Maris, Eric, and Robert Oostenveld. 2007. 'Nonparametric statistical testing of EEG- and MEG-data', *Journal of Neuroscience Methods*, 164: 177-90.
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: e1002575.
Massi, Bart, Christopher H. Donahue, and Daeyeol Lee. 2018. 'Volatility Facilitates Value Updating in the Prefrontal Cortex', *Neuron*, 99: 598-608.e4.
McCabe, C., C. Woffindale, C. J. Harmer, and P. J. Cowen. 2012. 'Neural processing of reward and punishment in young people at increased familial risk of depression', *Biol Psychiatry*, 72: 588-94.
McEvoy, Peter M., and Alison E. J. Mahoney. 2012. 'To Be Sure, To Be Sure: Intolerance of Uncertainty Mediates Symptoms of Various Anxiety Disorders and Depression', *Behavior Therapy*, 43: 533-45.
McFarland, B. R., and D. N. Klein. 2009. 'Emotional reactivity in depression: diminished responsiveness to anticipated reward but not to anticipated punishment or to nonreward or avoidance', *Depress Anxiety*, 26: 117-22.
Meder, David, Nils Kolling, Lennart Verhagen, Marco K. Wittmann, Jacqueline Scholl, Kristoffer H. Madsen, Oliver J. Hulme, Timothy E. J. Behrens, and Matthew F. S. Rushworth. 2017. 'Simultaneous representation of a spectrum of dynamically changing value estimates during decision making', *Nature communications*, 8: 1942.
Mizoguchi, N., T. Saigusa, Y. Aono, R. Sekino, K. Takada, Y. Oi, K. Ueda, N. Koshikawa, and A. R. Cools. 2008. 'The reboxetine-induced increase of accumbal dopamine efflux is inhibited by l-propranolol: a microdialysis study with freely moving rats', *Eur J Pharmacol*, 601: 94-8.
Moffa, Adriano H., Donel Martin, Angelo Alonzo, Djamila Bennabi, Daniel M. Blumberger, Isabela M. Benseñor, Zafiris Daskalakis, Felipe Fregni, Emmanuel Haffen, Sarah H. Lisanby, Frank Padberg, Ulrich Palm, Lais B. Razza, Bernardo Sampaio-Jr, Colleen Loo, and Andre R. Brunoni. 2020. 'Efficacy and acceptability of transcranial direct current stimulation (tDCS) for major depressive disorder: An individual patient data meta-analysis', *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 99: 109836.
Moniz, Marco, Saul Jesus, Eduardo Gonçalves, Andreia Pacheco, and Joao Viseu. 2016. 'Decision-making in adult unipolar depressed patients and healthy subjects: Significant differences in net score and in non-traditional alternative measures', *Neuropsychological Trends*: 7-15.
Moret, Chantal, and Mike Briley. 2011. 'The importance of norepinephrine in depression', *Neuropsychiatric disease and treatment*, 7: 9-13.
Mukherjee, Dahlia, Sangil Lee, Rebecca Kazinka, Theodore Sattherwaite, and Joe Kable. 2020. 'The Depressed Decision-Maker: How Value Based Decision Making Differs in Major Depressive Disorder'.
Must, Anita, Zoltán Szabó, Nikoletta Bódi, Anna Szász, Zoltán Janka, and Szabolcs Kéri. 2006. 'Sensitivity to reward and punishment and the prefrontal cortex in major depression', *Journal of affective disorders*, 90: 209-15.
Nakatani, Yoshihiro, Yukihisa Matsumoto, Yasuhiro Mori, Daisuke Hirashima, Hiroshi Nishino, Kentaro Arikawa, and Makoto Mizunami. 2009. 'Why the carrot is more effective than the stick: Different dynamics of punishment memory and reward memory and its possible biological basis', *Neurobiology of Learning and Memory*, 92: 370-80.
Nakonezny, Paul A, Thomas J Carmody, David W Morris, Benji T Kurian, and Madhukar H Trivedi. 2010. 'Psychometric evaluation of the Snaith–Hamilton Pleasure Scale (SHAPS) in adult outpatients with major depressive disorder', *International clinical psychopharmacology*, 25: 328.
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: 1040-46.
NICE. 2009. 'Clinical guideline for Depression in adults: recognition and management'. https://www.nice.org.uk/guidance/cg90.

Nolen-Hoeksema, Susan, and Jannay Morrow. 1991. 'A prospective study of depression and posttraumatic stress symptoms after a natural disaster: the 1989 Loma Prieta Earthquake', *Journal of personality and social psychology*, 61: 115.
O'Doherty, J., M. L. Kringelbach, E. T. Rolls, J. Hornak, and C. Andrews. 2001. 'Abstract reward and punishment representations in the human orbitofrontal cortex', *Nature Neuroscience*, 4: 95-102.
O'Doherty, John P., Peter Dayan, Karl Friston, Hugo Critchley, and Raymond J. Dolan. 2003. 'Temporal Difference Models and Reward-Related Learning in the Human Brain', *Neuron*, 38: 329-37.
O'Reilly, Jill X, Urs Schüffelgen, Steven F Cuell, Timothy EJ Behrens, Rogier B Mars, and Matthew FS Rushworth. 2013. 'Dissociable effects of surprise and model update in parietal and anterior cingulate cortex', *Proceedings of the National Academy of Sciences*, 110: E3660-E69.
Palminteri, Stefano, Damian Justo, Céline Jauffret, Beth Pavlicek, Aurélie Dauta, Christine Delmaire, Virginie Czernecki, Carine Karachi, Laurent Capelle, Alexandra Durr, and Mathias Pessiglione. 2012. 'Critical Roles for Anterior Insula and Dorsal Striatum in Punishment-Based Avoidance Learning', *Neuron*, 76: 998-1009.
Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. 'PsychoPy2: Experiments in behavior made easy', *Behavior Research Methods*, 51: 195-203.
Pessiglione, Mathias, Ben Seymour, Guillaume Flandin, Raymond J. Dolan, and Chris D. Frith. 2006. 'Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans', *Nature*, 442: 1042-45.
Pizzagalli, D. A., A. J. Holmes, D. G. Dillon, E. L. Goetz, J. L. Birk, R. Bogdan, D. D. Dougherty, D. V. Iosifescu, S. L. Rauch, and M. Fava. 2009. 'Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder', *Am J Psychiatry*, 166: 702-10.
Pizzagalli, D. A., A. L. Jahn, and J. P. O'Shea. 2005. 'Toward an objective characterization of an anhedonic phenotype: a signal-detection approach', *Biol Psychiatry*, 57: 319-27.
Porcelli, Stefano, Nic Van Der Wee, Steven van der Werff, Moji Aghajani, Jeffrey C. Glennon, Sabrina van Heukelum, Floriana Mogavero, Antonio Lobo, Francisco Javier Olivera, Elena Lobo, Mar Posadas, Juergen Dukart, Rouba Kozak, Estibaliz Arce, Arfan Ikram, Jacob Vorstman, Amy Bilderbeck, Ilja Saris, Martien J. Kas, and Alessandro Serretti. 2019. 'Social brain, social dysfunction and social withdrawal', *Neuroscience & Biobehavioral Reviews*, 97: 10-33.
Preacher, Kristopher J, and Andrew F Hayes. 2004. 'SPSS and SAS procedures for estimating indirect effects in simple mediation models', *Behavior research methods, instruments, & computers*, 36: 717-31.
Preuschoff, Kerstin, Bernard 't Hart, and Wolfgang Einhauser. 2011. 'Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making', *Frontiers in Neuroscience*, 5.
Pulcu, Erdem, and Michael Browning. 2017. 'Affective bias as a rational response to the statistics of rewards and punishments', *Elife*, 6: e27879.
———. 2019. 'The misestimation of uncertainty in affective disorders', *Trends in Cognitive Sciences*, 23: 865-75.
Remijnse, PL, MMA Nielen, AJLM Van Balkom, G-J Hendriks, WJ Hoogendijk, HBM Uylings, and DJ Veltman. 2009. 'Differential frontal–striatal and paralimbic activity during reversal learning in major depressive disorder and obsessive–compulsive disorder', *Psychological medicine*, 39: 1503-18.
Rescorla, Robert A. 1972. 'A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement', *Current research and theory*: 64-99.
Robinson, O. J., R. Cools, C. O. Carlisi, B. J. Sahakian, and W. C. Drevets. 2012. 'Ventral striatum response during reward and punishment reversal learning in unmedicated major depressive disorder', *Am J Psychiatry*, 169: 152-9.
Romaniuk, L., A. L. Sandu, G. D. Waiter, C. J. McNeil, S. Xueyi, M. A. Harris, J. A. Macfarlane, S. M. Lawrie, I. J. Deary, A. D. Murray, M. R. Delgado, J. D. Steele, A. M. McIntosh, and H. C. Whalley. 2019. 'The Neurobiology of Personal Control During Reward Learning and Its Relationship to Mood', *Biol Psychiatry Cogn Neurosci Neuroimaging*, 4: 190-99.
Rothkirch, M., J. Tonn, S. Köhler, and P. Sterzer. 2017. 'Neural mechanisms of reinforcement learning in unmedicated patients with major depressive disorder', *Brain*, 140: 1147-57.
Rudebeck, Peter H, and Elisabeth A Murray. 2014. 'The Orbitofrontal Oracle: Cortical Mechanisms for the Prediction and Evaluation of Specific Behavioral Outcomes', *Neuron*, 84: 1143-56.
Rupprechter, S., A. Stankevicius, Q. J. M. Huys, P. Series, and J. D. Steele. 2020. 'Abnormal reward valuation and event-related connectivity in unmedicated major depressive disorder', *Psychological medicine*: 1-9.
Rupprechter, Samuel, Liana Romaniuk, Peggy Series, Yoriko Hirose, Emma Hawkins, Anca-Larisa Sandu, Gordon D. Waiter, Christopher J. McNeil, Xueyi Shen, Mathew A. Harris, Archie Campbell, David Porteous, Jennifer A. Macfarlane, Stephen M. Lawrie, Alison D. Murray, Mauricio R. Delgado, Andrew M. McIntosh, Heather C. Whalley, and J. Douglas Steele. 2020. 'Blunted medial prefrontal cortico-limbic reward-related effective connectivity and depression', *Brain*, 143: 1946-56.
Rupprechter, Samuel, Aistis Stankevicius, Quentin J. M. Huys, J. Douglas Steele, and Peggy Seriès. 2018. 'Major Depression Impairs the Use of Reward Values for Decision-Making', *Scientific Reports*, 8: 13798.
Rush, A John, Madhukar H Trivedi, Hicham M Ibrahim, Thomas J Carmody, Bruce Arnow, Daniel N Klein, John C Markowitz, Philip T Ninan, Susan Kornstein, and Rachel Manber. 2003. 'The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression', *Biological psychiatry*, 54: 573-83.
Rush, A. J., M. H. Trivedi, S. R. Wisniewski, J. W. Stewart, A. A. Nierenberg, M. E. Thase, L. Ritz, M. M. Biggs, D. Warden, J. F. Luther, K. Shores-Wilson, G. Niederehe, and M. Fava. 2006. 'Bupropion-SR, sertraline, or venlafaxine-XR after failure of SSRIs for depression', *N Engl J Med*, 354: 1231-42.

Rushworth, Matthew F. S., and Timothy E. J. Behrens. 2008. 'Choice, uncertainty and value in prefrontal and cingulate cortex', *Nature neuroscience*, 11: 389-97.
Rushworth, Matthew FS, MaryAnn P Noonan, Erie D Boorman, Mark E Walton, and Timothy E Behrens. 2011. 'Frontal cortex and reward-guided learning and decision-making', *Neuron*, 70: 1054-69.
Rutledge, Robb B, Michael Moutoussis, Peter Smittenaar, Peter Zeidman, Tanja Taylor, Louise Hrynkiewicz, Jordan Lam, Nikolina Skandali, Jenifer Z Siegel, and Olga T Ousdal. 2017. 'Association of neural and emotional impacts of reward prediction errors with major depression', *JAMA Psychiatry*, 74: 790-97.
Safra, Lou, Coralie Chevallier, and Stefano Palminteri. 2019. 'Depressive symptoms are associated with blunted reward learning in social contexts', *PLOS Computational Biology*, 15: e1007224.
Sarinopoulos, I., D. W. Grupe, K. L. Mackiewicz, J. D. Herrington, M. Lor, E. E. Steege, and J. B. Nitschke. 2010. 'Uncertainty during Anticipation Modulates Neural Responses to Aversion in Human Insula and Amygdala', *Cerebral Cortex*, 20: 929-40.
Seligman, Martin EP. 1972. 'Learned helplessness', *Annual review of medicine*, 23: 407-12.
Shin, Lisa M., and Israel Liberzon. 2010. 'The Neurocircuitry of Fear, Stress, and Anxiety Disorders', *Neuropsychopharmacology*, 35: 169-91.
Siegel, Shepard, and Lorraine G. Allan. 1996. 'The widespread influence of the Rescorla-Wagner model', *Psychonomic Bulletin & Review*, 3: 314-21.
Silvetti, Massimo, Ruth Seurinck, and Tom Verguts. 2013. 'Value and prediction error estimation account for volatility effects in ACC: A model-based fMRI study', *Cortex*, 49: 1627-35.
Simmons, Alan, Scott C. Matthews, Murray B. Stein, and Martin P. Paulus. 2004. 'Anticipation of emotionally aversive visual stimuli activates right insula', *NeuroReport*, 15.
Smith, Stephen M. 2002. 'Fast robust automated brain extraction', *Human brain mapping*, 17: 143-55.
Snaith, RP, M Hamilton, S Morley, A Humayan, D Hargreaves, and P Trigwell. 1995. 'A scale for the assessment of hedonic tone the Snaith–Hamilton Pleasure Scale', *The British Journal of Psychiatry*, 167: 99-103.
Spielberger, Charles D. 1983. 'State-trait anxiety inventory for adults'.
Steele, J. D., P. Kumar, and K. P. Ebmeier. 2007. 'Blunted response to feedback information in depressive illness', *Brain*, 130: 2367-74.
Steffens, David C, H Ryan Wagner, Robert M Levy, Kimberly A Horn, and K Ranga Rama Krishnan. 2001. 'Performance feedback deficit in geriatric depression', *Biological psychiatry*, 50: 358-63.
Sutton, Richard S, and Andrew G Barto. 1990. 'Time-derivative models of pavlovian reinforcement'.
Tait, David Scott, and Verity J. Brown. 2007. 'Difficulty Overcoming Learned Non-reward during Reversal Learning in Rats with Ibotenic Acid Lesions of Orbital Prefrontal Cortex', *Annals of the New York Academy of Sciences*, 1121: 407-20.
Taswell, Craig A., Vincent D. Costa, Elisabeth A. Murray, and Bruno B. Averbeck. 2018. 'Ventral striatum's role in learning from gains and losses', *Proceedings of the National Academy of Sciences*, 115: E12398.
Thomas, C. M., and S. Morris. 2003. 'Cost of depression among adults in England in 2000', *Br J Psychiatry*, 183: 514-9.
Thorndike, Edward L. 1932. 'Reward and punishment in animal learning', *Comparative Psychology Monographs*.
Timmer, M. H. M., G. Sescousse, M. E. van der Schaaf, R. A. J. Esselink, and R. Cools. 2017. 'Reward learning deficits in Parkinson's disease depend on depression', *Psychological medicine*, 47: 2302-11.
Treadway, Michael T., Nicholas A. Bossaller, Richard C. Shelton, and David H. Zald. 2012. 'Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia', *Journal of abnormal psychology*, 121: 553-58.
Treadway, Michael T., Joshua W. Buckholtz, Ashley N. Schwartzman, Warren E. Lambert, and David H. Zald. 2009. 'Worth the 'EEfRT'? The Effort Expenditure for Rewards Task as an Objective Measure of Motivation and Anhedonia', *PLOS ONE*, 4: e6598.
Trivedi, Madhukar H., A. John Rush, Stephen R. Wisniewski, Andrew A. Nierenberg, Diane Warden, Louise Ritz, Grayson Norquist, Robert H. Howland, Barry Lebowitz, Patrick J. McGrath, Kathy Shores-Wilson, Melanie M. Biggs, G. K. Balasubramani, and Maurizio Fava. 2006. 'Evaluation of Outcomes With Citalopram for Depression Using Measurement-Based Care in STAR*D: Implications for Clinical Practice', *American Journal of Psychiatry*, 163: 28-40.
Ubl, B., C. Kuehner, P. Kirsch, M. Ruttorf, C. Diener, and H. Flor. 2015. 'Altered neural reward and loss processing and prediction error signalling in depression', *Soc Cogn Affect Neurosci*, 10: 1102-12.
Van Loo, Hanna M, Peter De Jonge, Jan-Willem Romeijn, Ronald C Kessler, and Robert A Schoevers. 2012. 'Data-driven subtypes of major depressive disorder: a systematic review', *BMC medicine*, 10: 156.
van Randenborgh, Annette, Renate de Jong‐Meyer, and Joachim Hüffmeier. 2010. 'Decision making in depression: differences in decisional conflict between healthy and depressed individuals', *Clinical Psychology & Psychotherapy*, 17: 285-98.
van Tol, Marie-José, Nic J. A. van der Wee, Odile A. van den Heuvel, Marjan M. A. Nielen, Liliana R. Demenescu, André Aleman, Remco Renken, Mark A. van Buchem, Frans G. Zitman, and Dick J. Veltman. 2010. 'Regional Brain Volume in Depression and Anxiety Disorders', *Archives of General Psychiatry*, 67: 1002-11.
Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017. 'Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC', *Statistics and Computing*, 27: 1413-32.
Verfaillie, Sander C. J., Stella J. de Wit, Chris Vriend, Peter L. Remijnse, Dick J. Veltman, and Odile A. van den Heuvel. 2016. 'The course of the neural correlates of reversal learning in obsessive–compulsive disorder and major depression: A naturalistic follow-up fMRI study', *Journal of Obsessive-Compulsive and Related Disorders*, 9: 51-58.

Vincent, Peter, Thomas Parr, David Benrimoh, and Karl J. Friston. 2019. 'With an eye on uncertainty: Modelling pupillary responses to environmental volatility', *PLOS Computational Biology*, 15: e1007126.
Vollmayr, Barbara, and Fritz A Henn. 2001. 'Learned helplessness in the rat: improvements in validity and reliability', *Brain Research Protocols*, 8: 1-7.
Von Zerssen, D, F Strian, and D Schwarz. 1974. 'Evaluation of depressive states, especially in longitudinal studies.' in, *Psychological measurements in psychopharmacology* (Karger Publishers).
Wallis, Jonathan D. 2011. 'Cross-species studies of orbitofrontal cortex and value-based decision-making', *Nature neuroscience*, 15: 13-19.
WFMH. 2012. 'DEPRESSION: A Global Crisis'. https://www.who.int/mental_health/management/depression/wfmh_paper_depression_wmhd_2012.pdf?Ua=1.
Whitton, A. E., P. Kakani, D. Foti, A. Van't Veer, A. Haile, D. J. Crowley, and D. A. Pizzagalli. 2016. 'Blunted neural responses to reward in remitted major depression: A high-density event-related potential study', *Biol Psychiatry Cogn Neurosci Neuroimaging*, 1: 87-95.
Wittmann, Marco K., Nils Kolling, Rei Akaishi, Bolton K. H. Chau, Joshua W. Brown, Natalie Nelissen, and Matthew F. S. Rushworth. 2016. 'Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex', *Nature communications*, 7: 12327.
Xie, Chao, Tianye Jia, Edmund T. Rolls, Trevor W. Robbins, Barbara J. Sahakian, Jie Zhang, Zhaowen Liu, Wei Cheng, Qiang Luo, Chun-Yi Zac Lo, He Wang, Tobias Banaschewski, Gareth J. Barker, Arun L. W. Bokde, Christian Büchel, Erin Burke Quinlan, Sylvane Desrivières, Herta Flor, Antoine Grigis, Hugh Garavan, Penny Gowland, Andreas Heinz, Sarah Hohmann, Bernd Ittermann, Jean-Luc Martinot, Marie-Laure Paillère Martinot, Frauke Nees, Dimitri Papadopoulos Orfanos, Tomáš Paus, Luise Poustka, Juliane H. Fröhner, Michael N. Smolka, Henrik Walter, Robert Whelan, Gunter Schumann, and Jianfeng Feng. 2020. 'Reward vs Non-reward Sensitivity of the Medial vs Lateral Orbitofrontal Cortex Relates to the Severity of Depressive Symptoms', *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*.
Yook, Keunyoung, Keun-Hyang Kim, Shin Young Suh, and Kang Soo Lee. 2010. 'Intolerance of uncertainty, worry, and rumination in major depressive disorder and generalized anxiety disorder', *Journal of Anxiety Disorders*, 24: 623-28.
Yu, Angela J., and Peter Dayan. 2005. 'Uncertainty, Neuromodulation, and Attention', *Neuron*, 46: 681-92.