

# **Linking reward-learning and affect in health and depression**

Don Chamith Halahakoon

Department of Psychiatry

University of Oxford

Oxford, UK

St Peters College

Thesis submitted for the degree of DPhil Psychiatry

Supervisors:

Prof. Michael Browning

Prof. Catherine Harmer

Prof. John Geddes

**Declaration of Authorship:** I, Don Chamith Halahakoon, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, this has been indicated in the thesis.

**Statement of Contribution:** Chapters 2 and 3 present data from one study. I collected data for this study (with Dr Alexander Kaltenboeck, who set up the study) and solely analysed all data presented in these two chapters (under the supervision of Professor Browning). Chapters 4 and 5 present data from a separate study. I set up this study, collected data (with Ms Ryan Yan) and solely analysed all data presented in these two chapters (under the supervision of Professor Browning).

**Acknowledgements:** I would like to thank my supervisors - Professor Michael Browning, Professor Catherine Harmer and Professor John Geddes. I would especially like to acknowledge the enormous contribution that my primary supervisor, Michael Browning, has made to my development. I could always count on him to get back to me on anything, at any time of day, on any day of the week and no matter how annoying the question.

I would like to thank Alex and Ryan for their contributions to this thesis, which are outlined in the statement of contributions, above.

I would like to thank my long-suffering parents, Wilfred and Chrishanthi, for their endless support with this thesis and with everything.

I would like to thank my wife, Alice. For 3.5yrs Alice has consistently sacrificed her evenings, weekends and sleep to allow me the space to complete my work. I dedicate this thesis to her.

Finally I would like to complain about my children; Beatrix (4yrs) and Rupert (1yr). If either of you read this in the future: you made everything much, much harder.

**Changes of plan:** The thesis as originally planned was to centre solely on the D2-like receptor agonist, Pramipexole: first investigating its effect on reward learning, then (in a second study) disentangling its effect on working memory, reinforcement learning and their interaction (Collins and Frank, 2018). This second study was halted at the point of recruitment due to ongoing pandemic related restrictions. I have instead conducted the online study, investigating affective instability and its relationship to reinforcement learning, (which I report on in chapters 4 and 5).

**Funding acknowledgement:** My doctoral studies were funded by the Wellcome trust and Oxford University department of psychiatry. I thank them for their support.

## TABLE OF CONTENTS

|                                                                                                              |            |
|--------------------------------------------------------------------------------------------------------------|------------|
| <b>CHAPTER 1: INTRODUCTION .....</b>                                                                         | <b>6</b>   |
| THEORETICAL BACKGROUND: REWARD PROCESSING IN DEPRESSION .....                                                | 7          |
| THEORETICAL BACKGROUND: THE NEURAL BASIS OF REWARD PROCESSING .....                                          | 11         |
| THEORETICAL BACKGROUND: COMPUTATIONAL MODELLING IN DEPRESSION .....                                          | 14         |
| INTRODUCTION TO CHAPTERS 2 AND 3: THE EFFECT OF PRAMIPEXOLE ON BEHAVIOURAL AND NEURAL REWARD PROCESSING..... | 19         |
| INTRODUCTION TO CHAPTER 4: LINKING AFFECT INSTABILITY IN A REWARD LEARNING TASK AND IN EVERYDAY LIFE.....    | 21         |
| INTRODUCTION TO CHAPTER 5: DOES REINFORCEMENT LEARNING UNDERLIE AFFECT INSTABILITY.....                      | 27         |
| SUMMARY OF CHAPTER 1 .....                                                                                   | 29         |
| <b>CHAPTER 2: THE EFFECT OF PRAMIPEXOLE ON BEHAVIOURAL REWARD PROCESSING.....</b>                            | <b>31</b>  |
| INTRODUCTION .....                                                                                           | 31         |
| METHODS/RESULTS .....                                                                                        | 32         |
| DISCUSSION/CONCLUSION .....                                                                                  | 48         |
| <b>CHAPTER 3: THE EFFECT OF PRAMIPEXOLE ON NEURAL REWARD PROCESSING .....</b>                                | <b>51</b>  |
| INTRODUCTION .....                                                                                           | 51         |
| METHODS .....                                                                                                | 53         |
| RESULTS .....                                                                                                | 56         |
| DISCUSSION/CONCLUSION .....                                                                                  | 62         |
| <b>CHAPTER 4: LINKING MOOD VARIABILITY AT DIFFERENT TIMESCALES.....</b>                                      | <b>65</b>  |
| INTRODUCTION .....                                                                                           | 65         |
| METHODS .....                                                                                                | 69         |
| RESULTS .....                                                                                                | 76         |
| DISCUSSION/CONCLUSION .....                                                                                  | 83         |
| <b>CHAPTER 5: DOES REINFORCEMENT LEARNING UNDERLIE AFFECT INSTABILITY.....</b>                               | <b>87</b>  |
| INTRODUCTION .....                                                                                           | 87         |
| METHODS/RESULTS .....                                                                                        | 89         |
| DISCUSSION/CONCLUSION .....                                                                                  | 100        |
| <b>CHAPTER 6: GENERAL DISCUSSION .....</b>                                                                   | <b>103</b> |
| ASSESSMENT OF RESULTS AND GENERAL CONCLUSIONS FROM THE THESIS.....                                           | 103        |
| ARE COMPUTATIONAL MODELS USEFUL IN UNDERSTANDING DEPRESSION AND ITS TREATMENT..                              | 105        |
| OUTSTANDING QUESTIONS, LIMITATIONS AND CONCLUSION .....                                                      | 105        |

|                               |            |
|-------------------------------|------------|
| <b>APPENDIX .....</b>         | <b>108</b> |
| APPENDICES TO CHAPTER 2 ..... | 108        |
| APPENDICES TO CHAPTER 3 ..... | 109        |
| APPENDICES TO CHAPTER 4 ..... | 118        |
| APPENDICES TO CHAPTER 5 ..... | 125        |
| <br>                          |            |
| <b>REFERENCES .....</b>       | <b>131</b> |

## Abstract

Relatively little is known about the mechanism underlying major depressive disorder (MDD), necessitating the exploration of novel investigative frameworks. In recent years reward processing has emerged as a promising theoretical framework for investigating depressive symptoms. In parallel, computational modelling has emerged as a promising analysis framework for leveraging the richness of (all manner of) psychiatric data. In this thesis I will use a reward processing framework and computational modelling to investigate a number of areas that are relevant to MDD.

My findings are from 2 studies. In the first study, I apply reinforcement learning models to behavioural and neuroimaging (functional magnetic resonance imaging; fMRI) data to investigate the effect of Pramipexole, a promising antidepressant, on behavioural and neural reward learning. The results of this study are reported in chapters 2 and 3. In chapter 2, I report on the behavioural findings from this study: Pramipexole specifically increases choice accuracy in the reward condition of a probabilistic instrumental learning task, with no effect in the loss condition. Behavioural modelling (alone) does not clearly arbitrate between potential underlying mechanisms. In Chapter 3, I report on the neuroimaging findings from this study: Pramipexole decreases the BOLD response to reward prediction errors in the ventromedial prefrontal cortex. Combined with the behavioural modelling, this finding indicates that Pramipexole enhances choice accuracy by reducing the decay of estimated values during reward learning.

In the second study, I investigate the mechanisms underlying affective instability, an emerging area of interest in depression research. I record participants' affective reports during an online reward learning task, and 'in real life' using experience sampling method (ESM). The results of this study are reported in chapters 4 and 5. In chapter 4 I separately characterise participants' task based and real-world affective profiles by applying a Bayesian filter to each dataset to calculate the parameter values that underlie participants' mean affect, the extent to which their affect fluctuates around this mean and the extent to which this mean changes over time. I then compare parameters from the two datasets and find that participants' affective profiles 'within-task' reflect their affective profiles 'in real life'. In Chapter 5, I explore a link between the (broad) reinforcement learning approach used in chapters 2/3 and the topic of chapter 4: affective instability. Specifically, I test a previously proposed model that links reinforcement learning and affective instability. I find that the model is able to replicate participant reported (within-task) affect without fitting to participant affect.

In sum, this thesis reports on a number of analyses that utilize computational modelling and a reward processing framework to link different types of data. If applied to clinical datasets, this approach may help to unpick the mechanisms underlying depression and its treatment.

## **Chapter 1: Thesis introduction**

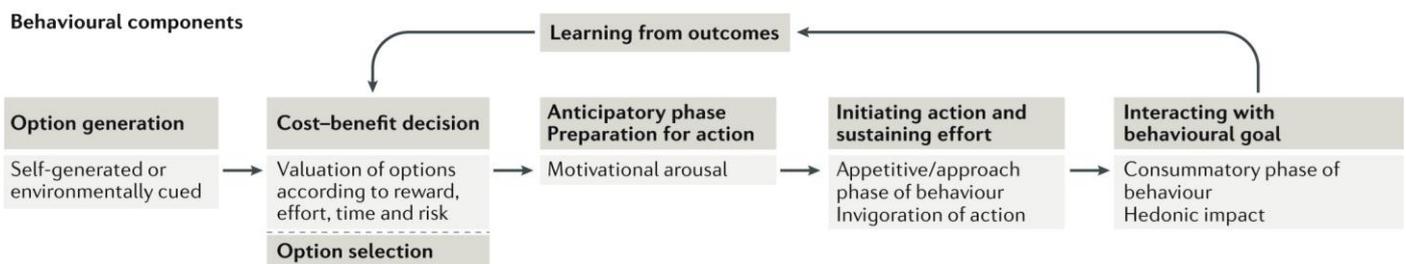
Summary of thesis introduction: In this thesis I will use a reward processing framework and computational modelling to investigate a number of areas that are relevant to major depressive disorder (MDD). The practical utility of the thesis rests on there actually being reward processing differences between MDD participants and HC (healthy controls). Therefore, the main topic addressed in the thesis introduction is whether there is compelling evidence for behavioural reward processing differences between MDD participants vs HC. This section (1.3) is based on a systematic review and meta-analysis that I published, with colleagues (Halakoon et al., 2020). Following this, I will briefly outline other topics that frame the rest of the thesis, namely, the neural basis of reward processing, the use of computational modelling in psychiatry and the computational literature on reward processing in depression. I will then outline the background and motivation for each of the four data chapters that follow.

### **1.1 Why study reward processing in depression?**

Depression is the leading cause of morbidity and mortality worldwide (Ferrari et al., 2013). Current therapeutic agents for depression are useful but limited (Casacalenda et al., 2002). 10-15% of depressed individuals do not respond, and 50-70% only partially respond to multiple pharmacological and psychological interventions (O'Reardon and Amsterdam, 1998). The imprecision of mental health nosology likely hinders the advancement of mechanistic research (and therefore the development of new treatments). The national institute of mental health is moving from studying categorical to dimensional constructs; the latter are symptom spectra with associated brain dysfunction (Cuthbert, 2014). 'Reward processing', which is relatively well characterised (McClure et al., 2004), has garnered increasing interest as a means of linking the two in depression (Treadway and Zald, 2011). A reward processing framework is especially useful for understanding symptoms related to motivation, such as reduced interest and activity (Husain and Roiser, 2018). These symptoms warrant better understanding as they are associated with poor psychosocial functioning (Vinckier et al., 2017), poorer prognosis at one year (Spijker et al., 2001), higher risk of suicide (Hall et al., 1999) and poor response to medication (Uher et al., 2012).

## 1.2 What is Reward Processing?

Reward processing describes how an organism uses reinforcement-related perceptions to guide goal-directed behaviours (Smith and Delgado, 2015) and can be divided into a number of subcomponents. According to one conceptualisation (Husain and Roiser, 2018), reward processing proceeds according to the following sequence of cognitive operations. **1) Option generation:** The generation of potentially rewarding behavioural options; **2) Decision making:** Options are subjected to a cost-benefit evaluation, which balances the utility of potential rewards against associated costs, resulting in the selection of one of the options; **3) Anticipation:** An anticipatory or preparatory phase associated with physiological arousal before the reward is obtained. **4) Action and effort:** Engagement in action in order to obtain the reward goal. **5) Consummation:** The hedonic impact arising from interacting with the reward goal (or alternatively, the frustration of an omitted outcome); **6) Reward Learning:** Learning how to modify behaviour in future interactions with similar stimuli, using an update signal.



**Figure 1. Mechanisms underlying reward-based decision making.** *Reproduced from Husain and Roiser (Husain and Roiser, 2018).* This flow diagram outlines the stages of reward processing described above.

## 1.3 Reward processing in depression

A number of studies have examined behavioural and neural reward/punishment processing in depression. In this section I explore the question: ‘are there reward processing differences between MDD participants and HC?’ The mainstay of my literature review is presented under the heading ‘**1.4 Are there reward processing differences between depressed individuals and healthy controls?**’ Some of the constraints I have placed on the scope of the review warrant explanation: 1) I have opted to focus on behavioural reward processing as behavioural findings generally provide more definitive answers to this question than neural findings (especially neural differences found in the absence of behavioural findings). I have however reviewed studies of neural reward processing in MDD that are relevant to this thesis in **section 1.8, ‘regions of interest’**. 2) I have restricted the

review to non-social reward. I have done this because the data chapters in this thesis focus on non-social reward and though there are many overlaps between social and non-social reward-processing, there are also important differences (Ruff and Fehr, 2014). Reviewing both literatures risks not covering either in sufficient detail. 3) I have reviewed response to reward, but not also punishment. This is an important omission given the importance of punishment sensitivity to prominent cognitive models of depression (Elliott et al., 1997b). However, response to punishment is sufficiently different to response to reward, and sufficiently out-with the focus of this thesis that I felt it reasonable to omit. 4) I have (for the most part) restricted my review to studies in which cases qualify as clinical MDD, according to DSM or ICD criteria (as opposed to e.g. a ‘high depressive symptoms’ group). I could alternatively have taken a more inclusive, (even cross-diagnostic) approach and explored reward processing and its relationship to depressive symptoms in general/subclinical populations, or even across categorical diagnoses. Arguably my approach is not in keeping with the ultimate goal of dimensional research (to move away from categorical diagnoses) and risks circularity. However, I believe that restricting my focus to a population that matches our current best understanding of the disease construct is a useful and (importantly) manageable first step in what will likely be an iterative process that (hopefully) moves us from a categorical to dimensional understanding of depressive symptoms. All this said, where there is a dearth of studies involving clinical MDD (i.e. sections 1.8: ‘regions of interest’ and 1.10: ‘What does computational modelling tell us about depression?’) I have strayed outside of this boundary to include studies of e.g. low vs high depressive symptoms.

#### **1.4 Are there reward processing differences between depressed individuals and healthy controls?**

Past studies of behavioural reward processing in depression fall into 4 categories; *option valuation*, *reward bias*, *reward response vigour* and *reward learning*. I summarise my findings in each category before reporting on my overall findings regarding reward processing in depression.

**Option Valuation:** Part of subcomponent 2 from section 1.2 above (‘decision making’): *Option valuation* describes the process by which individuals evaluate reward-related options when given explicit information about possible options (e.g. reward, cost and probability). An individual’s choice is assumed to reflect the weights that they place on potential rewards and costs (costs may include a potential loss of points/money, or the effort needed to obtain the reward) (Husain and Roiser, 2018). *Option valuation* can be measured using gambling tasks, for example the Cambridge gambling task (CGT) (Clark et al., 2011). In this task, HC generally modulate their bet-sizes based on reward probabilities more than MDD participants, (Clark et al., 2011; Murphy et al., 2001) though not always (Dombrowski et al., 2012). Other gambling tasks similarly show that depressed individuals show more

risk aversion than controls (Baek et al., 2017; Charpentier et al., 2017; Chung et al., 2017). One can alternatively investigate *option valuation* by measuring the amount of effort a person *chooses to attempt* (as opposed to the effort eventually taken) in order to obtain reward. The effort expenditure for reward task (Treadway et al., 2009) allows us to measure this and generally indicates that MDD participants decide to increase effort for increased reward less often than do HC. Effect sizes range from small (Yang et al., 2014) to medium (Subramaniapillai et al., 2019; Zou et al., 2020) to large (Treadway et al., 2012), though a couple of studies show no effect (Wang et al., 2022; Yang et al., 2021). Overall, depressed individuals appear to have small, but significant differences to HC in option valuation such that they are less likely to choose to endure costs to obtain reward (Halahakoon et al., 2020).

**Reward Bias:** Also thought to reflect subcomponent 2 from section 1.2 above ('decision making'): *Reward bias* is measured while individuals make difficult decisions (most often perceptual) that are rewarded asymmetrically, distinguishing this process from *option valuation*. Information relating to potential rewards, losses and probabilities is typically not provided explicitly. The reward bias measure, derived from signal detection theory, reflects an individual's tendency to choose more frequently rewarded stimuli, regardless of perceptual accuracy (Pizzagalli et al., 2008b). Studies generally show moderate (Henriques and Davidson, 2000; Vrieze et al., 2013) or high (Lawlor et al., 2019; Pizzagalli et al., 2008b) effect sizes, though a couple of studies show no effect (Liu et al., 2011; Reilly et al., 2020). Overall, depressed individuals appear to have moderate, significant differences in reward bias to HC such that their behaviour is less influenced by reward (Halahakoon et al., 2020).

**Reward Response Vigour:** Part of subcomponent 2 above (action and effort): *Reward response vigour* (RRV) reflects the speed with which an individual executes an action in order to obtain a reward. The difference between this and *option valuation* is that RRV relates to the actual *action taken*, not simply the choice to take it. RRV has been assessed overwhelmingly by the monetary incentive delay task (MID) (Knutson et al., 2000). Here, RRV can be assessed by comparing the difference between reaction times in rewarded and non-rewarded conditions. RRV has also once been assessed, in a similar manner, using the cued reinforcement reaction time task (CRRT) (Chase et al., 2010b). Overall, the MID (Admon et al., 2017; Arrondo et al., 2015; Carl et al., 2016; DelDonno et al., 2019b, 2019a, 2015; He et al., 2019; Pizzagalli et al., 2009; Sankar et al., 2019; Smoski et al., 2011; Takamura et al., 2017; Xie et al., 2014) and CRRT (Chase et al., 2010b) suggest that MDD participants do not have differences in RRV to HC (Halahakoon et al., 2020), though some individual studies do show that MDD participants have lower RRV than HC (Carl et al., 2016; DelDonno et al., 2019a; Pizzagalli et al., 2009; Sankar et al., 2019; Smoski et al., 2011; Takamura et al., 2017; Xie et al., 2014). Adjacent to RRV, some tasks require participants to expend (physical or mental) effort for

reward. Grip force tasks can measure the former and show mixed results (Cathomas et al., 2021; Cléry-Melin et al., 2011), while one study showed a large (negative) effect of MDD on the willingness to exert physical effort (Cléry-Melin et al., 2011) the other showed no effect (Cathomas et al., 2021). MDD also appears to have a moderate (negative) effect on the willingness to exert cognitive effort (Ang et al., 2022, p.).

**Reward Learning:** Subcomponent (6) from above, reward learning describes the process by which an individual uses feedback to change their behaviour over time, which is assumed to reflect the updating of value-expectations assigned to available behaviours (Husain and Roiser, 2018). This can be tested by a range of tasks. The most common is the Iowa gambling task (IGT). In this task, participants choose cards from four different decks, each with a different reward schedule. Some decks are advantageous while others are disadvantageous and learning can be approximated by subtracting disadvantageous from advantageous deck choices. Effect sizes tend to be small (Alexopoulos et al., 2015; Cella et al., 2010; Hegedűs et al., 2018; Jollant et al., 2016, 2005; Rinaldi et al., 2020; Saperia et al., 2019; Siqueira et al., 2021). One study yielded a medium (Gu et al., 2020) and another a high effect size (Must et al., 2006). Paradoxically, a couple of studies have reported better reward learning by MDD participants than HC (Deisenhammer et al., 2018; McGovern et al., 2014). Reward learning is also probed by various probabilistic learning tasks, in which participants choose between two stimuli with different reward contingencies and attempt to maximise earnings over time. These tasks yield a range of effect sizes, from negligible (Dezfouli et al., 2019; Moutoussis et al., 2018; Walsh et al., 2018) to small (Kumar et al., 2018; Liu et al., 2017; Reinen et al., 2021; Thoma et al., 2015) to medium (Hall et al., 2014; Nord et al., 2018) to large (Gradin et al., 2011; Herzallah et al., 2013) with one showing better learning by MDD participants than HC (Rothkirch et al., 2017). Overall, depressed individuals appear to have small but significant differences to HC in reward learning such that they tend to make less advantageous decisions (for the purpose of reward accumulation) in aggregate than HC (Halahakoon et al., 2020).

### **1.5 Behavioural Reward Processing in Depression: Summary**

To summarise, patients with MDD appear to have small differences, relative to HCs, in *option valuation* and *reward learning*, moderate differences in *reward bias* and no significant differences in *reward response vigour* (Halahakoon et al., 2020). Overall, MDD participants appear to have small-medium but significant differences to HCs in reward processing (Halahakoon et al., 2020). Some limitations of this review merit comment: I have grouped sometimes dissimilar measures in the same category. For example, option valuation contains studies that probe the effect of reward on the

willingness to exert effort (in 3 cases) and to take risks (in 6 cases). Additionally, it is not obvious what portion of the observed differences are due to medication effects. In the aforementioned meta-analysis (Halahakoon et al., 2020), colleagues and I divided depressed samples into ‘medicated’ or ‘non-medicated’ samples and conducted a moderation analysis to test whether observed group differences were due to medication. We found that medication status explained none of the variance either overall or within any reward processing subcategory. However the ‘medicated’ samples were often not entirely medicated, used a variety of medications (even within-study) and at different doses, making this result difficult to interpret. Finally, studies reviewed above had an average sample size of N=33/group. Given summary effect sizes in the aforementioned meta-analysis, studies would require a (considerably larger) sample size of N=136/group to achieve a power=0.8 at a significance=5% (two-tailed) (Halahakoon et al., 2020) suggesting that the vast majority of reviewed studies are under-powered.

### **1.6 The neural basis reward processing**

The field of neural reward processing is vast and a comprehensive account is beyond the scope of this thesis introduction. Instead I briefly outline the neural architecture of, and role of dopamine in, reward processing (section 1.7). I then describe in greater detail three important reward processing regions (used as regions of interest in chapter 3), referring both to anatomical and functional differences in MDD and to the effects of Pramipexole (the drug examined in chapters 2 and 3) in these regions (section 1.8).

### **1.7 The reward circuit in brief**

**The circuit:** The ventral tegmental area (VTA) and substantia nigra (SN) collectively contain the cell bodies of three large groups of dopamine secreting cells (Yamamoto and Vernier, 2011), which project primarily to the caudate, putamen and frontal cortex (Haber and Knutson, 2010). The frontal cortex, also outputs long range (non-dopaminergic) projections to the caudate and putamen, which in turn project to the substantia nigra pars reticulata (SNpr) and the globus pallidus (GP) (Haber and Knutson, 2010). The SNpr and GP then project to the thalamus which, in turn projects back to the frontal cortex, primarily the motor cortex, (Haber and Knutson, 2010) forming a feedback loop. The striatonigral (direct) pathway, is rich in excitatory D1-like receptors and is thought to reinforce actions. The striatopallidal (indirect) pathway is rich in inhibitory D2-like receptors (Missale et al., 1998) and is thought to suppress actions (Frank and Claus, 1996). The SNpr and GP also project back

to the VTA and SN bulbs, forming a second feedback loop (Haber and Knutson, 2010). This comprises the reward circuit, which integrates bottom up influences from the midbrain with top-down influences from the cortex, to guide reward-related action.

**The reward prediction error:** One of the signals transmitted by these three dopaminergic VTA/SN 'bulbs' is known as the 'reward prediction error' (RPE)(Schultz et al., 1997), which reports the difference between expected and received reward. Dopaminergic RPE activity is most accurately modelled by a class of reinforcement learning algorithms known as 'temporal difference' learning algorithms (Glimcher, 2011). In this class, the RPE signal reports the *continuous* updating of discounted future reward expectation, as opposed to the difference between expectation and outcome at the specific point of reward (non-/)delivery. This signal lies at the center of reward processing in the brain.

**Learning from the reward prediction error:** Co-occurring pre and post synaptic activity promotes long term potentiation (LTP) in the presence, and long term depression (LTD) in the absence, of dopamine (Wickens and Kötter, 1995). Neurons that produce movement remain active for a period of time after the production of the movement (Lau and Glimcher, 2007) and the subsequent presence or absence of dopamine (resulting from a positive or negative RPE) causes respectively LTP or LTD in the neuronal pathway that resulted in the +/-RPE (Glimcher, 2011). One prominent theory of basal ganglia function (Frank and Claus, 1996) proposes that this process of reinforcement occurs via two parallel and complementary pathways: phasic dopamine RPE signals stimulate D1-like receptors, which excites the direct pathway and reinforces the operations that preceded reward. Simultaneously they stimulate D2-like receptors, which inhibits the indirect pathway and again, effectively reinforces the preceding action (by electrochemically inhibiting this behaviorally inhibitory pathway) (Frank and Claus, 1996). Tonic dopamine dips do the opposite, facilitating no-go learning (Frank and Claus, 1996).

## 1.8 Regions of Interest

In this section I will describe three key neural regions involved in reward processing and summarise the role these regions are believed to play. These three regions are 1) the orbitofrontal cortex, 2) the medial prefrontal cortex and 3) the ventral striatum. These three regions serve as ROIs in chapter 3, in which I investigate the effects of Pramipexole, a novel anti-depressant, on neural reward processing.

The **orbitofrontal cortex (OFC)** receives inputs from a wide array of brain regions including parts of the sensory cortex involved in reward consumption (Elliott et al., 2000). The OFC is thought to process information about a diverse range of stimuli and convert information about 'what' stimuli are, into a common 'reward value' (Rolls et al., 2020) and maintain reward outcome expectations in working/representational memory (Schoenbaum and Roesch, 2005). Projections from the OFC provide one of the main inputs to the ventral striatum (Haber and Knutson, 2010), exerting a top-down effect on response selection (Frank and Claus, 1996). **The OFC in MDD:** OFC dysfunction is implicated in MDD (Rolls, 2021; Rolls et al., 2020). MDD participants appear to have lower OFC activity than HC during reward anticipation (Smoski et al., 2011) and higher OFC activity during reward feedback (Ng et al., 2019). There is also evidence of grey matter volume reduction in the OFC in MDD (Koolschijn et al., 2009; Lorenzetti et al., 2009). **The OFC and Pramipexole:** Reward-outcome related OFC activity is increased by Pramipexole in individuals with Parkinson's disease; an effect that is associated with greater (within task) risk taking (van Eimeren et al., 2009). In a group with (bipolar) depression, lower pre-treatment metabolic activity in the OFC predicted better symptom response to subsequent Pramipexole treatment, which in turn lowered metabolic activity within the OFC (Mah et al., 2011a)

The receipt of Rewards increases activity in the ventromedial prefrontal cortex (vmPFC), a region that overlaps considerably with the OFC (Haber and Knutson, 2010). Within the vmPFC, the **medial prefrontal cortex (mPFC)**, (which is distinct from the OFC), appears to respond to abstract (vs sensory) gain (Kim et al., 2006; Knutson et al., 2003; Kuhnen and Knutson, 2005). Direct electrophysiological recording during risky decision making (the IGT) demonstrates that mPFC neuronal activity correlates with positive RPEs (Oya et al., 2005). Likewise, blood-oxygen-level-dependant (BOLD) recordings suggest that mPFC activity fulfills the necessary and sufficient conditions of RPE signaling (Rutledge et al., 2010) and optogenetic evidence demonstrates that elevated mPFC activity inhibits reward-related behaviour (Ferenczi et al., 2016). **The mPFC in MDD:** mPFC activity during reward receipt is altered in MDD (Knutson et al., 2008; Kumar et al., 2015) and the magnitude of mPFC thinning is commensurate with the number of MDD episodes an individual has experienced (Treadway et al., 2015). **The mPFC and Pramipexole:** Pramipexole reduces metabolic activity in Brodmann area 10p (which overlaps with the mPFC) in (bipolar) depression (Mah et al., 2011a) and increases rostral mPFC activity during probabilistic reward feedback (Santesso et al., 2009).

The **ventral striatum (VS)** comprises loosely, the nucleus accumbens, olfactory tubercle and those regions of the caudate and putamen that lie rostral to the internal capsule (Haber and Knutson, 2010) and sits at the heart of neural reward processing. The VS receives input from all major

components of the reward circuit (the midbrain dopaminergic bulbs/frontal cortex/thalamus/hippocampus/amygdala) and projects to an equally diverse array of regions (midbrain/pons/pallidum/hypothalamus/periaqueductal gray/amygdala/nucleus basalis) (Haber and Knutson, 2010). The VS is reliably activated by reward, as opposed to arousal, and irrespective of reward-type (Blood and Zatorre, 2001; Elliott et al., 1997a; Martin-Sölch et al., 2001; Small et al., 2003). VS activity convincingly reflects the magnitude and valence of RPEs (McClure et al., 2003; Pessiglione et al., 2006). **The VS in MDD:** A number of studies show that MDD participants have lower ventral striatal activity than HC during reward anticipation across a range of paradigms (Arrondo et al., 2015; Hägele et al., 2015; Stoy et al., 2012). A large number of studies show the same for reward receipt (Admon et al., 2015; Foti et al., 2014; Hall et al., 2014; Johnston et al., 2015; Knutson et al., 2008; Redlich et al., 2015; Remijnse et al., 2009; Robinson et al., 2012; Satterthwaite et al., 2015; Smoski et al., 2009; Steele et al., 2007). Three meta-analyses report reduced VS during reward feedback in MDD participants vs HC (Keren et al., 2018; Ng et al., 2019; Zhang et al., 2013). **The VS and Pramipexole:** Pramipexole increases VS activity during reward anticipation and modulates prefrontal-striatal and insular-striatal connectivity (Ye et al., 2011). High baseline VS RPE signaling predicts greater clinical response to Pramipexole in MDD participants (Whitton et al., 2020).

## 1.9 Computational Modelling in Depression

Each data chapter in this thesis makes use of computational modelling. I first outline the potential use of theory-driven computational models before highlighting some of the limitations (or potential misuses) of these models and then finally summarising the behavioural computational literature on MDD.

### Why use computational modelling in psychiatric research?

Computational psychiatry encompasses a vast array of data-driven and theory-driven methods. In three of four data chapters, I use a popular class of computational model (reinforcement learning (RL)) to model participant behaviour and (in one case) the BOLD signal. In the remaining data chapter, I use a recursive Bayesian filter to infer the parameters that determine individuals' affective profile. Below, I briefly outline the case for using (theory-driven) computational models, particularly RL. I will describe specific models in the relevant data chapters.

As is apparent from the summary of the reward processing literature above, reward processing differences in depression are usually captured using group differences in observed

behaviour. However, behavioural differences can potentially result from a disparate range of latent processes that are difficult, if not impossible, to disentangle using descriptive models of behaviour alone. (Generative) computational models are formal hypotheses about these cognitive processes, implemented as algorithms that (often) contain free parameters. Models are tested against each other by assessing the accuracy and parsimony with which they reproduce participant behaviour or physiology. Participant-group differences are then typically assessed by comparing parameter values within the winning model. This approach is useful in two main ways:

- 1) Computational modelling ideally facilitates relatively precise investigation of underlying cognitive processes by a) allowing the formation and testing of complex hypotheses and b) replacing broad concepts with narrower (hopefully more tractable) ones. For example, instead of ‘anhedonia’, a participant may have ‘low reward sensitivity’, which is defined precisely as the magnitude of a reward sensitivity parameter that determines aspects of learning and, downstream of this, behaviour (Chen et al., 2015).
- 2) Computational modelling facilitates precise and biophysically plausible links between behaviour and neural activity (as reflected in e.g. the BOLD signal). RL models in particular closely model dopaminergic activity in the brain’s reward processing architecture (Schultz et al., 1997). Specifically, coordinated dopaminergic signals that originate in the ventral tegmental area reflect the RL belief update signal: the RPE. More generally, computational models facilitate links between different levels of description of brain function (Marr, 2010).

### **Over-interpretation and over-generalization of computational modelling in psychiatric research**

The power of computational modelling is easily overstated. Its promise of precision tempts over-generalisation and over-interpretation. Generalisability, in this context, is the idea that computational parameter values are specific to the person but independent of the task. This assumption demonstrably does not always hold (Eckstein et al., 2021). Over-interpretation in this context is the assumption that parameters represent specific cognitive/neural processes (Eckstein et al., 2021). There are cases in which this assumption is true (or true enough): the very study of RL in the brain was birthed by the remarkable confluence between VTA dopaminergic activity and the TD update signal (Schultz et al., 1997). But again, such strong assumptions do not always hold (Eckstein

et al., 2021). While one parameter can map to several mechanisms (in different contexts), several parameters can also map to the same mechanism. Huys et al highlighted the case of two very commonly used parameters which are sometimes interpreted as signifying quite different phenomena, but are mathematically identical (Huys et al., 2013). In the reward learning literature, rewards/outcomes are sometimes scaled by a ‘reward sensitivity’ parameter:

$$Q_{t+1(s)} = Q_{t(s)} + \alpha(\rho R_t - Q_{t(s)})$$

Where the expected value  $Q$  of stimulus  $s$  is updated by the reward-outcome  $R$  at each trial  $t$ . The influence of  $R$  on  $Q$  is determined by the learning rate parameter  $\alpha$  (i.e. the rate at which new information updates ones beliefs) and the reward sensitivity parameter  $\rho$ , which is often taken to reflect the subjective value of an objective unit of reward and has been linked to e.g. opioid signalling in the shell of the nucleus accumbens (Peciña and Berridge, 2005).

In the choice layer, the probability of choosing stimulus  $s$  is determined by the value of that stimulus and of all other stimuli. In the case of a 2-arm bandit this can be written as:

$$P_{t(s)} = \frac{1}{1 + \exp^{-\beta(Q_{t(s)} - Q_{t(s')})}}$$

Where  $P_{t(s)}$  is the probability of choosing stimulus  $s$  at trial  $t$  which is determined by the ( $Q$ )-values of stimuli  $s$  and  $s'$  as well as the ‘inverse temperature’ parameter  $\beta$ , which is often taken to reflect choice-determinism vs stochasticity or the tendency to explore vs exploit, and has been linked with e.g. noradrenergic neuromodulation (Aston-Jones and Cohen, 2005). However  $Q \propto \rho$  so that changes in  $\beta$  and  $\rho$  produce identical changes in behaviour. This means that identical between-group behavioural disparities can potentially be framed as arising from quite different underlying causes depending on how behaviour is modelled and the results interpreted. In chapter 3, I will explore the use of the BOLD signal in arbitrating between parameters (including  $\beta$  and  $\rho$ ) that produce similar, sometimes indistinguishable, behavioural results.

### 1.10 What does computational modelling tell us about depression?

The computational literature on depression does not clearly divide into conceptual categories like the reward processing literature in section 1.4. I have therefore instead categorised it by task-type: Those that require participants to **1)** learn stable contingencies, **2)** respond to varying contingencies, **3)** balance approach vs avoidance, **4)** make risk-based decisions and **5)** make timed reward-related responses.

**Learning stable contingencies:** In the depression literature, computational modelling is most commonly applied to probabilistic instrumental learning tasks (PILTs). In this type of paradigm, behaviour is modelled using Q-learning models (Chase et al., 2010a; Kumar et al., 2018; Liu et al., 2017) in all but one study, which uses a temporal difference (SARSA) model (Gradin et al., 2011). No studies found parameter differences between MDD participants and HC (Chase et al., 2010a; Gradin et al., 2011; Kumar et al., 2018; Liu et al., 2017). Ostensibly this suggests that there are no differences between MDD participants and HC in learning stable contingencies. However, in these tasks, asymptotic choice accuracy is largely equivalent to reward sensitivity ( $\rho$ )/decisions determinacy ( $\beta$ ). A couple of (non-computational) studies do show that MDD participants have lower asymptotic choice accuracy than HC in PILTs (Kumar et al., 2018; Walsh et al., 2018). In this sense, there is some evidence of reduced reward sensitivity/decision determinacy in MDD. Likewise, in a signal detection task, MDD participants were found to have lower reward sensitivity than HC (Huys et al., 2013). The same task accounts for the majority of the 'reward bias' subsection in section 1.4, above. Here, the 'response bias' (i.e. the propensity to report seeing a richly rewarded stimulus regardless of which stimulus is actually presented) describes the behaviour produced by high reward sensitivity/decision determinacy. If we assume that a low 'response bias' by MDD participants equates to low reward sensitivity/decision determinacy, there is considerable evidence for low reward sensitivity/decision determinacy in MDD (Henriques and Davidson, 2000; Lawlor et al., 2019; McGovern et al., 2014; Pizzagalli et al., 2008b; Vrieze et al., 2013).

**Responding to varying contingencies:** Two studies used reversal learning paradigms in MDD (Dombrowski et al., 2010; Mukherjee et al., 2020). One found that MDD participants had lower reward and punishment learning rates, lower value sensitivity and a higher bias term in the punishment condition, than HC. (Mukherjee et al., 2020). The other divided MDD participants into suicide attempters, ideators and non-ideators. MDD suicide attempters had a lower 'memory' parameter than controls i.e. suicide attempters' behaviour was more influenced by recent feedback relative to their reinforcement history (Dombrowski et al., 2010). So both reversal learning paradigms yield positive but different results. One study examined how depressed and anxious individuals modulate their rate of learning in response the frequency with which stimulus-outcome contingencies changed (Gagne et al., 2020). This study used a volatility paradigm (where reward contingencies are sometimes stable and sometimes unstable) with reward and punishment, as well as shock and no-shock conditions. Individuals with a high 'general factor' shared by depressed and anxious individuals had impaired learning rate modulation in response to volatility, regardless of condition.

**Approach-avoidance:** One study used an approach-avoidance paradigm with and without threat of shock. Here cases were heterogeneous, comprising participants with MDD or generalised anxiety disorder +/- co-morbid anxiety disorders. Cases had a higher avoidance bias parameters (that is, a greater tendency to inhibit action upon encountering punishing stimuli regardless of the optimal response to those stimuli) and a greater increase in this parameter from the 'no-threat' to 'threat' condition, than HC (Mkrtchian et al., 2017). Speculatively, this finding is consistent with the diathesis-stress model of depression.

**Risk-based decision-making:** Other frameworks that have sometimes been used are 1) prospect theory models and 2) drift diffusion models. Two studies that used a prospect theory framework found no differences in gambling behaviour between MDD participants and HC (Charpentier et al., 2017; Chung et al., 2017). However, dividing MDD participants into suicide attempters and non-attempters, a separate study found that, while non-attempters were not different from controls on any parameter measure, attempters had higher loss aversion as well as risk aversion, in both gain and loss conditions, than both non-attempters and controls (Baek et al., 2017). This is surprising as other (non-computational) studies suggest increased risk taking in suicide attempters (Ackerman et al., 2015, p.; Adams et al., 1973) which is more intuitive.

**Timed reward-related responses:** A drift diffusion model was used to model MDD participants' and HCs' responses in a flanker task. MDD participants had slower drift rates of both a reflexive mechanism that allowed flankers to bias responses, and of executive control, which overrode the reflexive mechanism to allow correct responses on incongruent trials (Dillon et al., 2015). MDD is known to be associated with executive dysfunction (Rock et al., 2014) and so the latter finding fits with the extant literature. It is difficult to know how to interpret the finding of slower reflexive drift rate. The authors suggest low tonic dopamine as a potential explanation for both the former and latter findings. Straying beyond current clinical MDD, one study applied drift diffusion modelling to those with current or past MDD and found that they had a lower drift rate than HC in a signal detection task (Vallesi et al., 2015). It would be interesting to investigate whether this result was driven by only those with current MDD or whether the remitted group also had slow drift rates, as the latter and former would respectively suggest that slow drift rates are trait/state related.

**Summary:** Computational findings on MDD are few, diverse and sometimes contradictory. A recent meta-analysis attempted (Pike and Robinson, 2022) to summarise at least part of the literature. Their initial 'conventional' meta-analysis suggested patients had lower inverse temperatures than HC. Their subsequent 'simulation' meta-analysis found that patients had larger punishment learning rates (and slightly lower reward learning rates) than HC, with no meaningful difference in inverse

temperature. This approach risks conflating nominally 'identical' parameters that could signify different things in different task contexts. That is to say, it risks overgeneralisation as outlined by Eckstein et al. (Eckstein et al., 2021). Computational modelling currently tells us little about depression. However, modelling can inform/frame our interpretation of more widely replicated findings. The main takeaway from the above collection of results is that low reward sensitivity/decision determinacy probably accounts for low asymptotic choice accuracy in PILTS and low response bias in the signal detection task referenced above, that MDD participants appear to achieve. As MDD participants appear to have consistent, significant and substantially low response biases in signal detection tasks (Halachakoon et al., 2020), this suggests that they probably have low reward sensitivity/decision determinacy.

### **1.11 The Effect of Pramipexole on behavioural and neural reward processing (Introduction to Chapters 2 and 3)**

Though the majority of antidepressants primarily target serotonin, others aim to remediate aberrant dopaminergic signaling, which is thought to be an important component of depressive pathophysiology (Belujon and Grace, 2017). Dopaminergic agents do not appear any more efficacious than serotonergic ones (Sinyor et al., 2010). However, the dopaminergic system can be targeted in a variety of ways. For example, monoamine oxidase inhibitors reduce the breakdown of monoamines (including dopamine but also neurotransmitters such as serotonin and noradrenaline) (Baker et al., 1992) while Bupropion prevents the reuptake of dopamine (as well as noradrenaline, serotonin and other neurotransmitters) from the synapse (Costa et al., 2019) and Quetiapine blocks D2-like receptors (Richelson and Souder, 2000). The anti-parkinsonian drug Pramipexole has garnered recent attention as a potential treatment for treatment resistant depression (Au-Yeung et al., 2022). Pramipexole agonises D2-like, primarily D3, receptors (Gerlach et al., 2003; Piercey, 1998). D3 receptors are found in the ventral (but not dorsal) striatum (Hall et al., 1996), which is thought to be central to the processing of rewards and implicated in the pathophysiology of depression (Höflich et al., 2019). Clinical evidence suggests a therapeutic effect of Pramipexole on depressive symptoms in Parkinson's disease (Lemke et al., 2006) (which is characterized by degeneration of dopaminergic neurons) (Kouli et al., 2018). Evidence for the efficacy of Pramipexole in major depression is mixed and summarized in a systematic review (Tundo et al., 2019). To briefly summarize the clinical evidence:

A randomized controlled trial found that 1.3mg/day of Pramipexole improves symptom measures in treatment resistant depression, but does not produce a clinical response over placebo (Cusin et al.,

2013). In another study, MDD participants that had not responded to one antidepressant were randomised to receive escitalopram with or without Pramipexole 2.25 mg (Franco-Chaves et al., 2013) and found no group differences. This study only had 13 participants in each group and only 4 of the +Pramipexole group completed the study. Another study found that different doses of Pramipexole monotherapy yielded different outcomes in MDD: 0.375mg was ineffective, 1 mg was effective and 5 mg was poorly tolerated (Corrigan et al., 2000). Two RCTs found a favourable effect of Pramipexole in bipolar depression (Goldberg et al., 2004; Zarate et al., 2004). A case-series (Fawcett et al., 2016) yielded compelling treatment effects of Pramipexole augmentation in 42 patients with unipolar or bipolar depression, 8 of whom had not benefitted from electroconvulsive therapy. In this case-series, the treatment dose ranged from 0.25-5mg with a mean of ~2.5 mg/day. Almost half i.e. 20 of the patients remitted while 12 responded. The majority of remaining patients did not tolerate the drug.

As outlined earlier in this chapter, there is some evidence that MDD participants have blunted responses to reward (Halakoon et al., 2020). Given its promising (if mixed) clinical effects, one might expect Pramipexole to enhance reward response. Previous experimental studies (Bodi et al., 2009; Gallant et al., 2016; Kanen et al., 2019; Murray et al., 2019; Pizzagalli et al., 2008a; Santesso et al., 2009; Whitton et al., 2020) generally indicate that Pramipexole blunts rather than enhances participants' behavioural responses to reward (Gallant et al., 2016; Murray et al., 2019; Pizzagalli et al., 2008a; Santesso et al., 2009). Studies that utilized a stimulus-response task (Gallant et al., 2016), a probabilistic instrumental learning task (Murray et al., 2019) and a signal detection task (Pizzagalli et al., 2008a) each suggested that Pramipexole blunted participants' behavioural responses to reward. Similarly, Pramipexole has been found to blunt the neural response to positive outcomes in reward sensitive brain regions such as the medial prefrontal cortex (mPFC), orbitofrontal cortex (OFC), striatal and dorsal anterior cingulate cortex (dACC) to a rewarding taste (McCabe et al., 2013) as well as ventral striatal (VS) and midbrain response to unexpected monetary gain (Riba et al., 2008).

One explanation for the seeming contradiction between the clinical and experimental evidence is that experimental studies have predominantly examined the effect of a single dose of Pramipexole while sustained treatment is required to improve symptoms (Fawcett et al., 2016; Lemke et al., 2006). From a pharmacological perspective, acute treatment with D2/3/4 agonists are believed to primarily influence inhibitory presynaptic auto-receptors, leading to reduced dopaminergic transmission, whereas sustained administration leads to auto-receptor down-regulation and enhanced transmission via agonism at post-synaptic D2-like receptors (Chen et al., 2005; Chernolet et al., 2008; Grace, 1991; Willner et al., 1994). This suggests that the clinically relevant effects of

Pramipexole on reward learning are likely to become apparent only after sustained administration of the drug. A couple of studies have examined the effect of a course of Pramipexole in clinical populations, one in euthymic bipolar individuals (Burdick et al., 2014), another in individuals with Parkinson's disease (Bodi et al., 2009). In both cases, a sustained course of Pramipexole enhanced reward sensitivity.

In Chapters 2 and 3, I will examine the effect of a sustained (2-week) course of Pramipexole on both behavioural (Chapter 2) and neural (Chapter 3) measures of reward learning in non-clinical participants. As registered in clinical trials.gov (<https://clinicaltrials.gov/ct2/show/NCT03681509>), I hypothesize that Pramipexole will induce the opposite pattern of reward-learning behaviour characteristic of depression and anhedonia, increasing asymptotic choice of stimuli associated with higher levels of reward by increasing subjective valuation of rewards (leading to increased BOLD to rewarding outcomes in reward sensitive areas of the brain such as the VS, OFC and mPFC) (Fouragnan et al., 2018).

#### **1.12 Linking affect instability in a reward learning task and in everyday life (Introduction to Chapter 4)**

Affect instability has a prevalence of 13.9% in the general population (Broome et al., 2015). While fluctuations of affect are central to conditions such as bipolar disorder (Harrison et al., 2018), affect instability is also prevalent (>49%) in a wide range of psychiatric diagnoses (Marwaha et al., 2014, 2013). Affect instability can predict the onset of ostensibly unrelated symptoms (e.g. auditory hallucinations) (Marwaha et al., 2014) and has been proposed as a candidate symptom dimension within the research domain criteria (RDoC) framework (Broome et al., 2015). There is reason to believe that affect instability is also relevant to depression. Data from the adult psychiatric morbidity survey 2007 suggests that 68.5% of men and 62.1% of women with depression experience affect instability (Marwaha et al., 2013).

The above statistics are from retrospective self-report. There is reason to also be interested in prospective data: retrospective reports risk recall bias and affective instability appears nuanced i.e. it is not a unidimensional phenomenon (Kuppens et al., 2010). Several studies suggest the utility of affect dynamics in predicting the onset of depressive episodes, beyond simply measuring mean affect levels. Emotional inertia (the autocorrelation of affect scores) prospectively predicts the onset of depression in adolescents (Kuppens et al., 2012). While at the opposite end of the age spectrum, older adults with higher baseline affective instability are more likely to experience a depressive

episode in the 6yr follow-up period (Eldesouky et al., 2018). Another study found that instability of negative affect predicts future depressive symptoms (Sperry et al., 2020). In contrast, a separate study found that reduced emotional reactivity at baseline predicted higher depressive symptom severity at one month and reduced likelihood of recovery over the follow-up period. Importantly, neither emotional reactivity nor MDD course in this study were accounted for by the severity/duration of initial symptoms (Peeters et al., 2010). Likewise, fluctuations in negative affect appear to predict treatment response to CBT (Husen et al., 2016). Tracking affect dynamics also appears important in predicting suicidality: affect instability appears to increase the odds of suicidal thinking five-fold (Marwaha et al., 2013) and independently of depression severity (Bowen et al., 2011). Additionally, the impact of impulsivity (an intuitively important factor) on suicidality is negated when affect instability is accounted for (Peters et al., 2016).

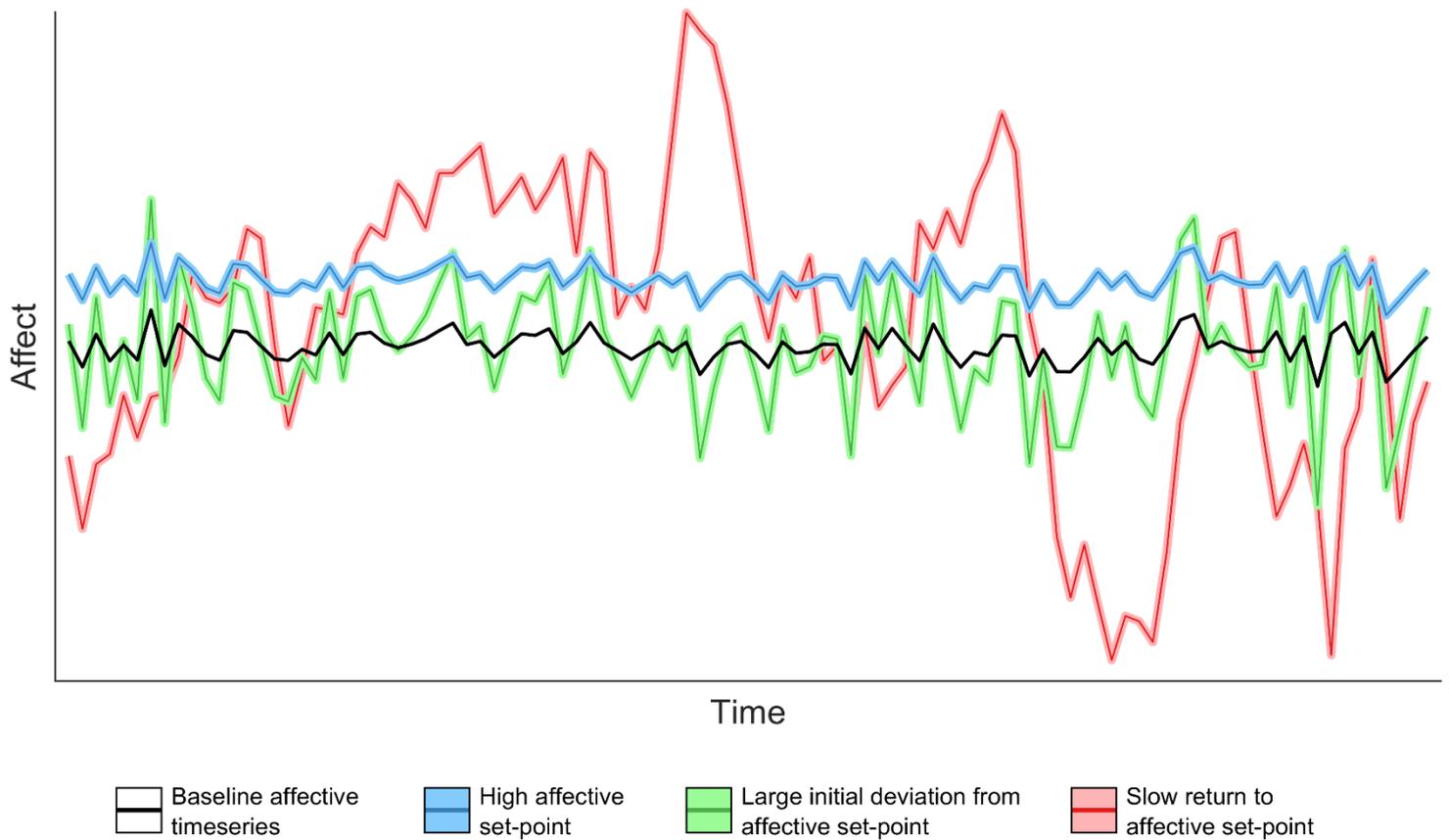
MDD is a heterogeneous condition, suggesting that treatments may best target specific components rather than the full syndrome. Affect instability may serve as one such component (or ‘dimension’) (Broome et al., 2015). As such, it would be useful to explore the nature of affect instability. The existing body of work on this topic, which largely uses a combination of ESM and descriptive (auto-correlation) models, presents some seemingly paradoxical results (Koval et al., 2013). For example, while some studies find that depression is associated with greater affective instability, others find that it is associated with greater affective inertia (Koval et al., 2013). Koval et al have proposed that some of these seeming paradoxes result from three factors (quote): *‘a) dependencies between indices of affect dynamics; b) measurement of affect at different timescales; and c) the influence of external events.’* (Koval et al., 2013). In their study, Koval et al tried to address some of these issues by collecting both lab based affective measures (emotional reactions to film clips), obtained during a short controlled schedule of events with experience sampling method (ESM) over a longer timescale (1 week) (Koval et al., 2013). Relating lab-based findings to ESM and using an autocorrelation model, this study found trend-level (i.e.  $p < 0.07$ ) positive correlations between ESM and lab-based measures for all but one dynamic measure. In chapter 4 I try to build on these findings. I assess the relationship of individuals’ affective dynamics at a short vs long timescale and describe these affective dynamics in terms of individuals’ affective responses to external events. I do so using an alternative model of affective dynamics (Pulcu and Browning, 2019).

**An alternative to auto-correlation:** In a prominent model of affect (Kuppens et al., 2010), individuals have **1)** an affective *‘set point’* i.e. their baseline affective state, from which internal and external events **2)** *shift affect*, causing (greater or lesser) affective variability, and to which **3)** affect *returns* over time. The formal description of this affect model is complex (Kuppens et al., 2010). To state and illustrate my hypotheses, I instead (more crudely) conceive of affect as a sum of exponential decays,

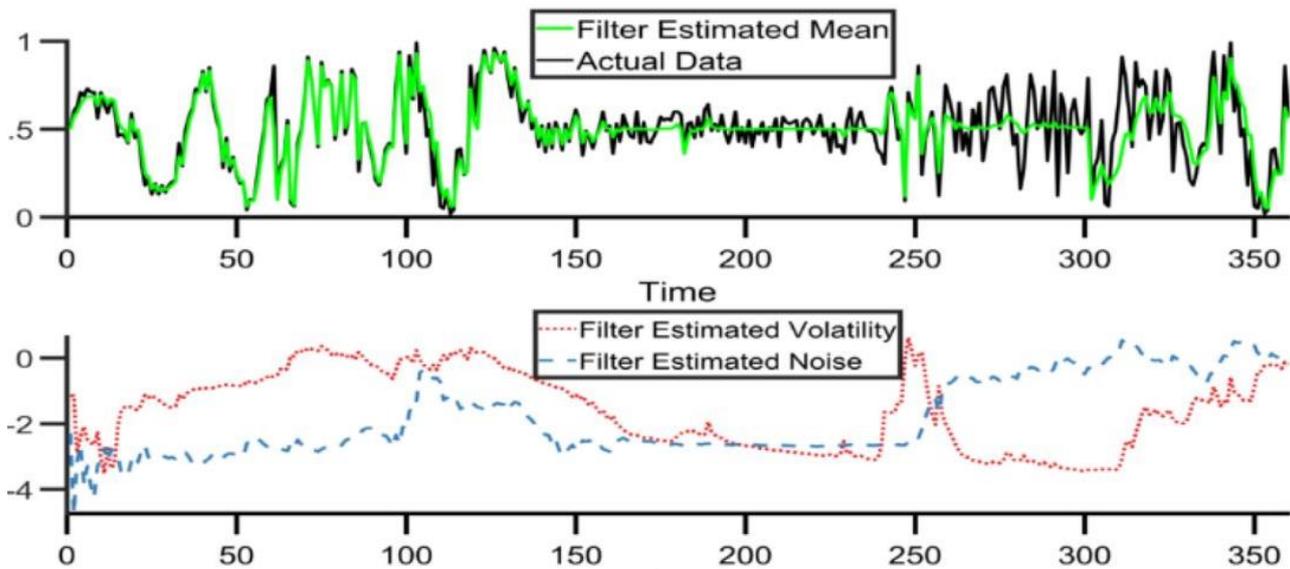
which still allows inter-individual variation in the 3 components outlined above i.e. the individuals' affective *set point*, the extent to which a given perturbation *deviates* affect from baseline and the rate at which affect *returns* to baseline.

$$affect_T = setpoint + \sum_1^T event_t(a \cdot e^{b \cdot (t-T)})$$

In this model, affect at time-point  $T$  is a weighted sum of affective deviations from baseline, caused by events at each time-point  $t$ , up to and including time-point  $t = T$ . The value of  $a$  determines the extent to which events deviate affect from baseline. The value of  $b$  determines the rate at which affective deviations return to baseline. This is somewhat similar to a prominent model that uses RL measures to reproduce participant affect (Rutledge et al., 2014). For illustrative purposes, I have plotted four synthetic affective time-series (figure 1). One 'baseline' affective time-series (black line), one with a high affective-'setpoint' (blue line), one in which events cause greater deviation in affect (green line) and one in which affect returns to baseline more slowly (pink line). The (black) 'baseline' and (blue) 'high' affective time-series appear to fluctuate around fairly consistent respective 'mean' affects (i.e. their affective set-points). These fluctuations could be thought of as 'noise'. The (green) 'reactive' line does the same, but with larger fluctuations. One might call the latter affective pattern (in which affect deviates substantially from, but quickly returns to, a consistent mean) 'noisy' affect. The (pink) 'slow affective return' line does not appear to fluctuate around a consistent mean. Instead, affect appears to fluctuate around a mean that itself changes over time. One might call this affective pattern (in which mean affect changes over time) 'volatile' affect.

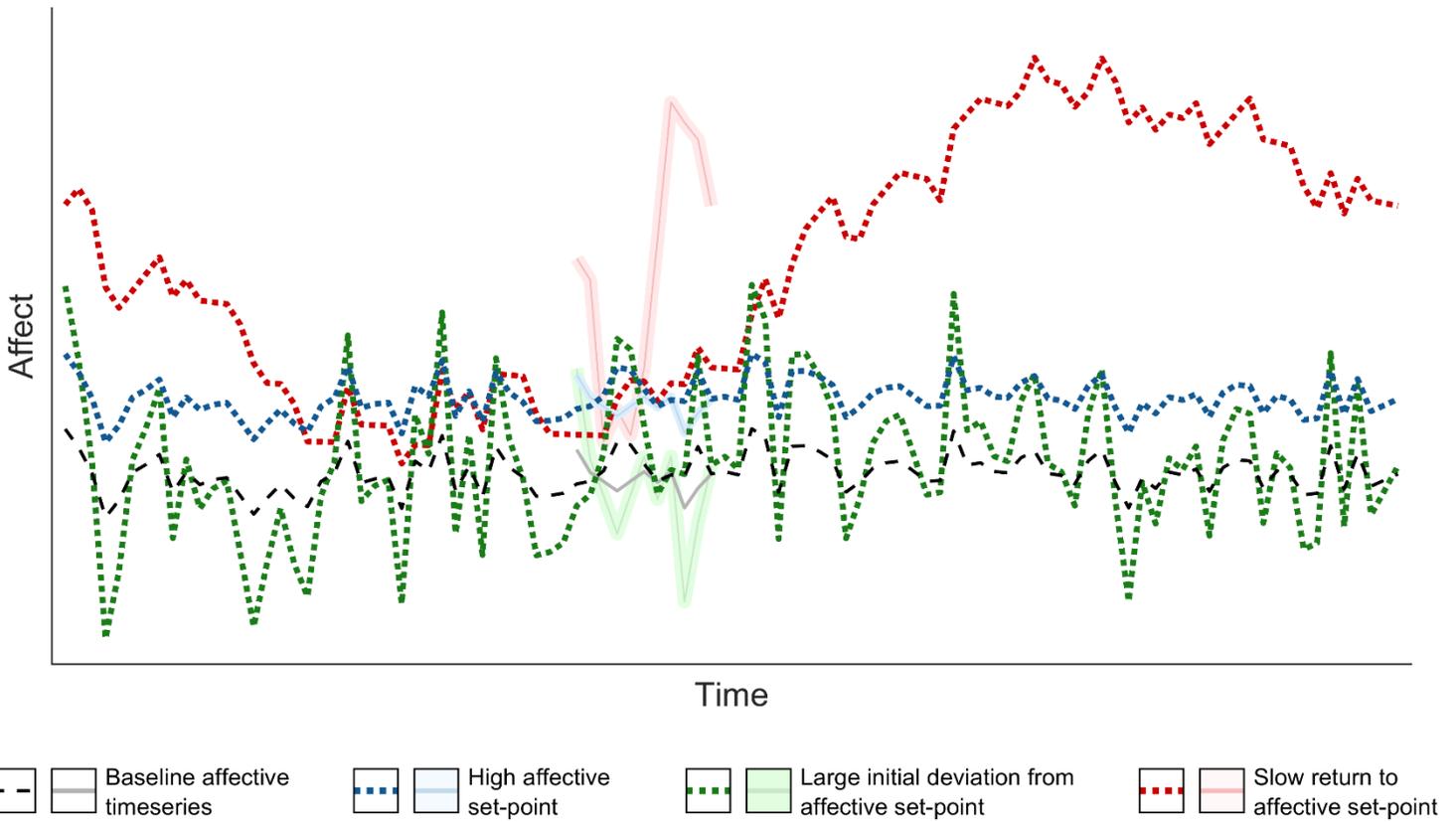


**Figure 1. Illustrative affective time-series.** External events perturb affect which then returns to baseline. The black line represents a ‘baseline’ affective time-series. The blue line illustrates how the baseline time-series changes if the affective set-point is higher but nothing else changes; ‘mean’ affect appears higher. The green line illustrates how the baseline time-series changes if each event perturbs affect to a greater extent, but nothing else changes; affect appears ‘noisy’. The pink line illustrates how the baseline time-series changes if affect returns more slowly to baseline after being perturbed, but nothing else changes; affect appears ‘volatile’. Pulcu et al have recently proposed a model of affect variability (a recursive Bayesian filter) that estimates affective ‘mean’, ‘volatility’ and ‘noise’ within a single framework (Pulcu et al., 2022) using participant affective reports.



**Figure 2. Classification of affective variability.** *Reproduced from Pulcu et al* (Pulcu et al., 2022). In the upper panel, the black line is a synthetic affective time-series. It is volatile at time 1-120 and 301-360, and noisy at time points 61-120 and 241-360. The green line in the upper panel is the Bayesian filter’s belief about the ‘mean’ affect at each time-point. The filter’s estimate of *volatility* (red line) and *noise* (blue line) are shown in the lower panel.

**Does affect at a smaller time-scale reflect affect at a larger time-scale?:** If we zoom in on a small portion of the set of time series illustrated in figure 1, we can see what it would like if we measured these 3 time series over a shorter timescale and at a higher frequency. This is illustrated in Figure 3. At this shorter timescale, the ‘high set point’ time-series continues to look like the baseline time-series with a higher ‘mean’ affect. The ‘reactive’ time-series continues to look more ‘noisy’ than the baseline time-series. The ‘slow reset’ time-series continues to look more ‘volatile’ than the baseline time-series.



**Figure 3. Illustrative affective time-series on a short time-scale.** Zooming in on a small proportion of affective time-series in figure 1 (illustrated by the faded solid lines), the ‘high set point’ time-series (dotted blue line) continues to look like the baseline time-series (dashed black line) with a higher ‘mean’ affect. The ‘reactive’ time-series (dotted green line) continues to look more ‘noisy’ than the baseline time-series. The ‘slow reset’ time-series (dotted red line) continues to look more ‘volatile’ than the baseline time-series.

In other words, measuring a person’s affective dynamics on a small time scale (such as during a reward learning task) may lend insight into their affective dynamics on a larger timescale (such as in ESM data), as suggested by Koval et al (Koval et al., 2013). If this were the case it would be potentially useful as it connects ESM data (which currently has no biological framework) to the reward learning literature (which does) and so may allow for e.g. translational insights into the biology of affect instability. The recursive Bayesian filter described by Pulcu et al (Pulcu et al., 2022) can be used on ESM data and equally on affective reports during a reward learning task allowing for links to be drawn between within-task and ESM affective dynamics.

In chapter 4, participants report ESM data 6 times a day, over 20 days. On two of those days, they also complete an online reward learning task during which they regularly report their affect. I estimate participants' affective mean, volatility and noise within a short timeframe with a known event schedule (i.e. during the reward learning task) and a long timeframe with unknown events (ESM data) using the recursive Bayesian filter described above (Pulcu et al., 2022).

I primarily assess whether affective parameters within the reward learning task reflect corresponding affective parameters in everyday life. I do so by regressing ESM affect parameters against within-task affect parameters. In a more exploratory analysis I test the association between affective parameters and the duration of affective responses to events in the relatively controlled setting of a reward learning task.

### **1.13 Does reinforcement learning underlie affect instability (Introduction to chapter 5)**

In chapter 4 I attempt to link affect, as captured through ESM, to affect within a reward learning paradigm. My motivation for doing so is that ESM data is highly proximate to the phenomena that we are trying to understand but is relatively unmoored from biology. Conversely, reward learning (and within this, the framework used in chapters 2 and 3 i.e. reinforcement learning (RL)) has a relatively well characterized biological substrate but does not tend to concern itself with subjective experience. RL emerged from the behaviourist tradition, which consciously rejected the study of subjective states, such as affect, in favour of observable objective data (e.g. behaviour). Nonetheless, reward-related behaviour (which is overwhelmingly what is modelled by RL in MDD research) is intuitively linked to subjective states, such as affect. A link implicit in the very study of reward-related behaviour in MDD (an affective disorder). Researchers have begun to investigate within-task affect through the lens of RL, typically by sampling participant affect (for example by using a visual analogue scale) while they complete a reward processing task. Participant behaviour is modelled using RL and the resultant trial-wise model metrics (such as value expectation and RPEs) are entered into separate models to produce affect scores. These studies demonstrate a close and consistent link between RL and reported affect (Blain and Rutledge, 2022; Chew et al., 2021; Keren et al., 2022; Rutledge et al., 2014).

If affect can be described in terms of RL, then so too should affective instability. Eldar and Niv have done just this (Eldar and Niv, 2015). Here, not only do RL measures produce affect, but affect in turn exerts an influence on the RL process. They conceive of affect as a (modified) weighted average of

RPEs and characterise affective instability as resulting from a positive feedback loop between affect and learning. Formally they describe affect  $m$  at time  $t$  as:

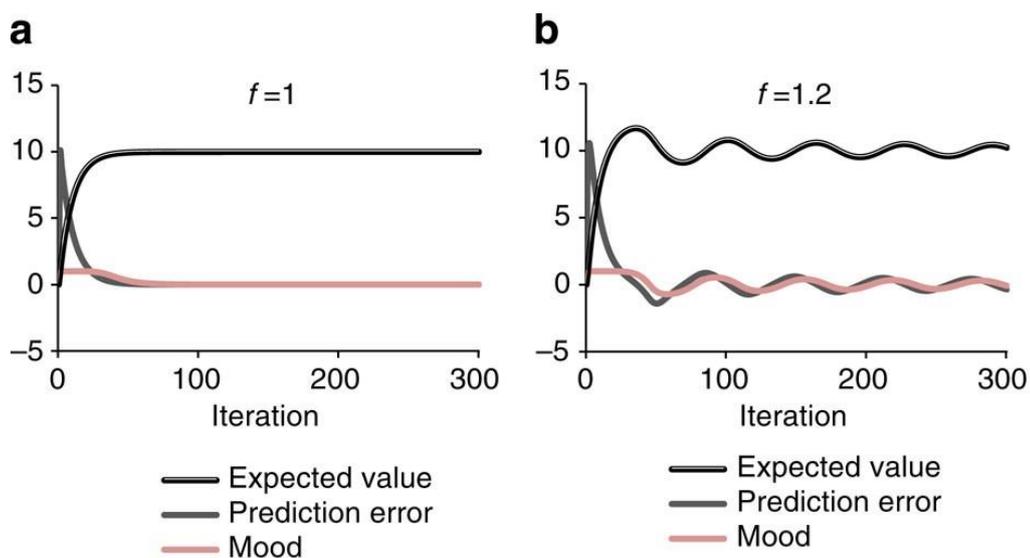
$$h_t = h_{t-1} + \eta(\delta_{t-1} - h_{t-1})$$

$$m_t = \tanh(h_t)$$

Where  $h_t$  is the weighted average of RPEs  $\delta$  and  $\eta$  is the step-size parameter that determines the extent to which RPEs are discounted as they recede into the past. Affect, in turn, modulates RPEs by altering the perception of subsequent rewards. Specifically, affect ( $m$ ) is a power placed on the bias term  $f$  which influences the perception of outcomes, and therefore RPEs.  $f^{m_t}$  effectively acts as a reward sensitivity term that changes according to current affect:

$$\delta_t = f^{m_t} \cdot R_t - Q_t$$

This creates a feedback loop so that affect depends on RPEs and RPEs depend on affect. The nature of this feedback loop depends on the magnitude of the bias term  $f$ . If  $f$  is larger than 1, the feedback loop is positive, resulting in affect instability.



**Figure 4. Simulations of the interaction between learning and affect.** *Reproduced from Eldar and Niv (Eldar and Niv, 2015).* The model was exposed to  $R = 10$  for 500 iterations with **a)**  $f$  set to '1'. The pink link represents affect ( $m$ ) and remains stable from quite early in the simulation. **b)**  $f$  set instead to '1.2'; affect becomes unstable.

In this study, participants that had stable affect in everyday life, as measured by the hypomanic personality questionnaire (HPS), were best described (within-task) by a model with  $f$  fixed at '1' (i.e.

no affect-learning interaction) while those who had unstable affect in everyday life were best described by the full model (i.e.  $f$  as a free parameter). Furthermore, HPS scores correlated (positively) with the  $f$  value (i.e. affective instability is associated with a positive, not negative affect-learning feedback loop) (Eldar and Niv, 2015). These findings suggest that affective instability is associated with positive feedback between affect and learning. This same modelling approach demonstrated that citalopram (a commonly used anti-depressant medication) enhanced the effect of positive affect induction on reward valuation in healthy volunteers (Michely et al., 2020). However, this is a nascent line of research and many questions remain unanswered. In this chapter I will use a novel task in which participants learn about and choose between rewarded options, before and after affect induction, while frequent measures of affect are collected. I use this data to explore Eldar and Niv's model in a number of ways: **1)** The model mechanism has so far been tested at a relatively coarse resolution. I test its ability to replicate affect at a resolution commensurate with another prominent RL model of affect (Rutledge et al., 2014)). **2)** The model posits that affect is a (modified) weighted average of RPEs. In the original study, this mechanism replicated participant affect, but the to-be-explained participant affect scores were actually included with the behavioural data to estimate the model parameters. I instead test whether the posited mechanism can replicate participant *affect* using *behavioural data* alone. **3)** The model has been successfully tested in one task. I test it in another. Finally I attempt to replicate the key original findings i.e. **a)** affective instability in everyday life (as measured by the HPS) is associated with a propensity toward an interaction between affect and learning and **b)** the extent of affective instability is commensurate with the extent of this interaction.

### 1.14 Summary

Depression is a prevalent and debilitating condition that warrants better treatments. Reward processing may facilitate better characterisation of MDD and its subtypes, which in turn should aid mechanistic MDD research. MDD participants appear to have reward processing differences to HC, overall and in several reward processing sub-components. Computational modelling is a promising, but as yet unproven approach to characterising these differences. In this thesis I use computational modelling and a reward processing framework to investigate areas relevant to MDD and its treatment. In **chapter 2** I characterise the effect of a 2-week course of Pramipexole, a novel adjunctive antidepressant, on participants' behaviour in a reward learning task. In **chapter 3**, I relate findings from chapter 2 to participants' task-based neural (as reflected in BOLD) activity. In **chapter 4** I fit a recursive Bayesian filter to participants' ESM and reward learning task data, in order to relate

the two. In **chapter 5**, I explore a recently described model of affect-production/affect-learning-interaction. I assess whether I can replicate the original finding (of an association between affect instability and an interaction between affect and learning) and whether the model can explain participant affect (without fitting the model to participant affect).

## Chapter 2: The effect of Pramipexole on behavioural reward processing

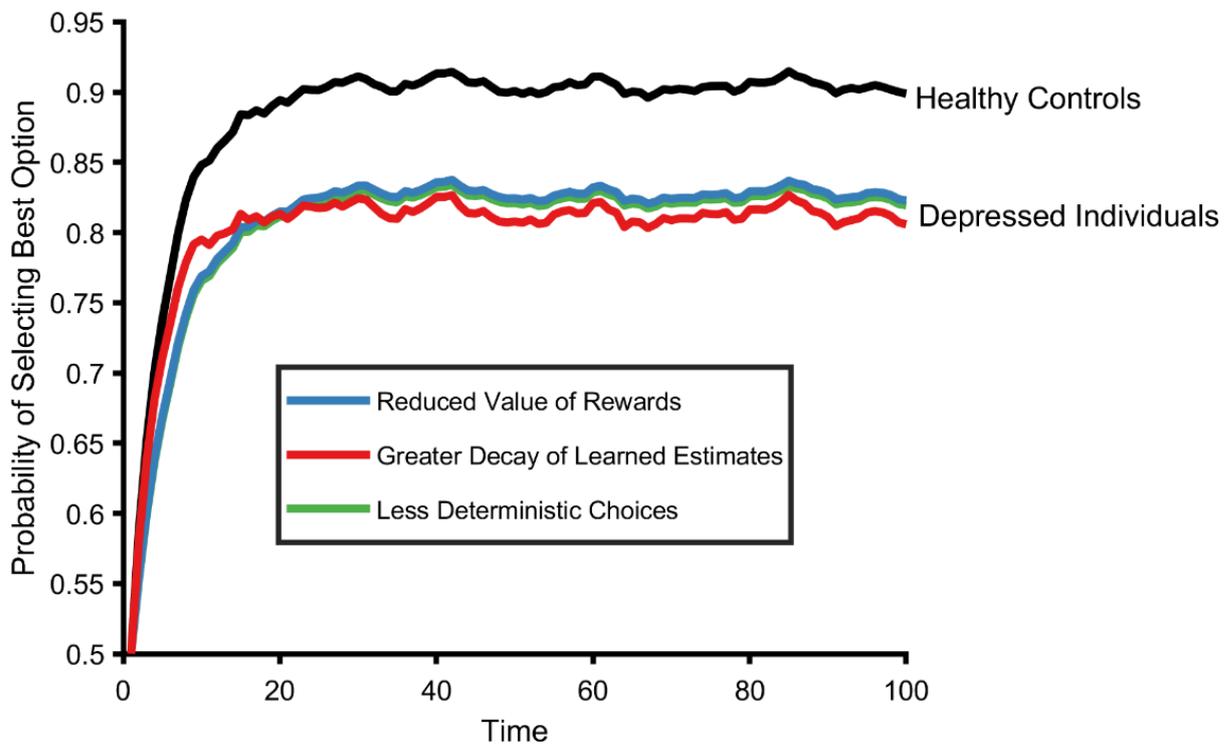
### 2.1 Introduction

As outlined in the thesis introduction, depressed individuals show reduced reward learning vs HC. Pramipexole (a D2-like receptor agonist) is a promising treatment for treatment resistant depression. Therefore one might expect Pramipexole to have the opposite effect of depression i.e. enhance reward learning. However, previous experimental studies (Bodi et al., 2009; Gallant et al., 2016; Kanen et al., 2019; Murray et al., 2019; Pizzagalli et al., 2008a; Santesso et al., 2009; Whitton et al., 2020) of Pramipexole generally indicate that it blunts rather than enhances participants' behavioural responses to reward (Gallant et al., 2016; Murray et al., 2019; Pizzagalli et al., 2008a; Santesso et al., 2009). One explanation for the seeming contradiction between the clinical and experimental evidence is that experimental studies have predominantly examined the effect of a single dose of Pramipexole while sustained treatment is required to improve symptoms (Fawcett et al., 2016; Lemke et al., 2006). From a pharmacological perspective, acute treatment with D2/3/4 agonists are believed to primarily influence inhibitory presynaptic auto-receptors, leading to reduced dopaminergic transmission, whereas sustained administration leads to auto-receptor down-regulation and enhanced transmission via agonism at post-synaptic D2-like receptors (Chen et al., 2005; Chernoloz et al., 2008; Grace, 1991; Willner et al., 1994). This suggests that the clinically relevant effects of Pramipexole on reward learning are likely to become apparent only after sustained administration of the drug.

In this Chapter, I will examine the effect of a sustained (2-week) course of Pramipexole on behavioural measures of reward learning in non-clinical participants. As registered in clinical trials.gov (<https://clinicaltrials.gov/ct2/show/NCT03681509>), I hypothesize that Pramipexole will induce the opposite pattern of reward-learning behaviour characteristic of depression and anhedonia, increasing asymptotic choice of stimuli associated with higher levels of reward.

I then model participant behaviour using variants of a simple Rescorla-Wagner learning model, to try and discover the mechanism under any observed behavioural reward-learning effect of Pramipexole. Computational characterisation using reinforcement learning (RL) models has identified three distinct alterations of learning and decision-making process that may produce the behaviour observed in depressed patients: they may make decisions less deterministically (Huys et al., 2013), they may treat rewards "as if" they were of reduced value (Huys et al., 2013), or their learned value estimates may decay to a greater degree over time (Collins and Frank, 2012). As registered in clinical

trials.gov (<https://clinicaltrials.gov/ct2/show/NCT03681509>), I hypothesize that Pramipexole will increase subjective valuation of rewards.



**Figure 1; Simulated choice accuracy in a reward learning task.** Depressed participants are observed to achieve lower asymptotic choice accuracy than healthy ones. This behaviour can be simulated using a Q-learning model. A reduced subjective value of rewards, a greater decay of learned value expectations or a propensity towards making less deterministic choices can each result in the model achieving lower asymptotic choice accuracy. In other words, (at least) three qualitatively distinct processes could lead to the observed behavioural effect of depression. The mechanics of the Q-learning model that simulated this data are described in the modelling methods below.

## 2.2 Methods: Participants

I conducted a randomized, placebo controlled experimental medicine study with a between-groups design. 42 non-clinical participants, between the age of 18 and 45, were randomized 1:1 to receive Pramipexole or placebo. Potential participants were excluded if they had ever been diagnosed with a psychiatric illness (determined using the SCID-5-CV) or had a first degree relative with a psychotic illness, were taking psychoactive medication, had any history of impulse control difficulties, had any

contraindication to Pramipexole, had taken any recreational drugs in the last three months, regularly drank more than 4 units of alcohol per day, smoked more than 5 cigarettes per day or drank more than 6 caffeinated drinks per day. Female participants who were pregnant, lactating, or not using a highly effective method of contraception were also excluded. Participants were young, highly educated and evenly split between females and males. Two participants (both in the placebo group) dropped out of the study due to side-effects (tremor in one case and racing thoughts in the other; in both cases, side-effects stopped following discontinuation of placebo).

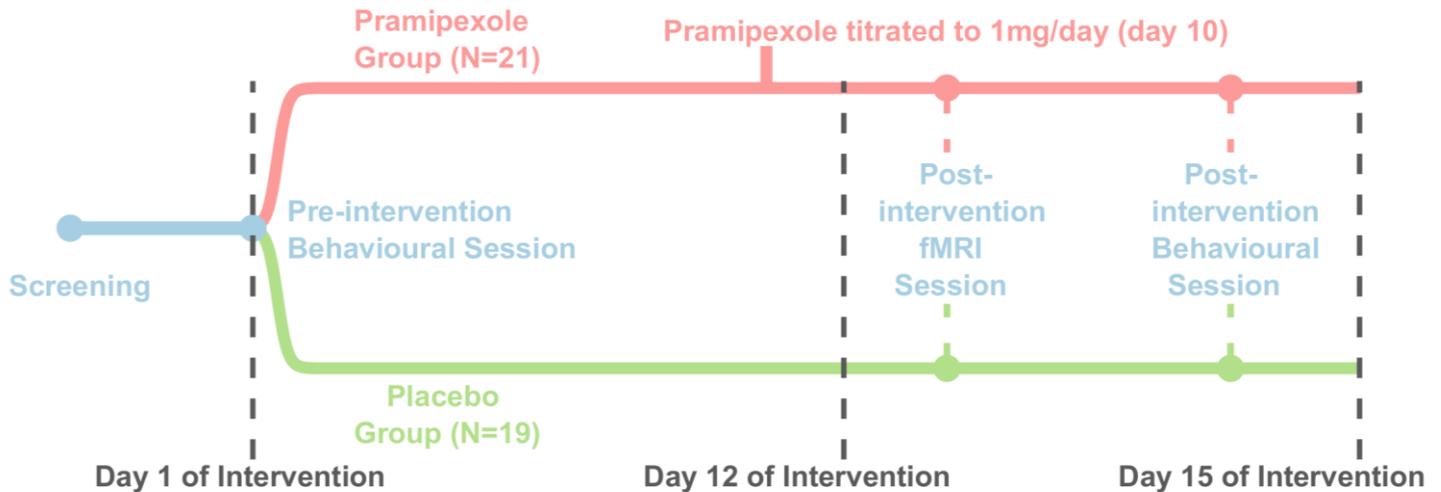
|                                                            | <b>Pramipexole<br/>(n = 21; 10 male)</b> | <b>Placebo<br/>(n = 19; 10 male)</b> |
|------------------------------------------------------------|------------------------------------------|--------------------------------------|
| <b>Age</b>                                                 | 22.5 (3.7)                               | 24.5 (6.9)                           |
| <b>Body mass index</b>                                     | 22.4 (2.6)                               | 24.0 (2.9)                           |
| <b>Years in full-time education</b>                        | 16.8 (2.9)                               | 17.5 (3.1)                           |
| <b>IQ estimate (Spot-the-word Test)</b>                    | 108.3 (8.1)                              | 111.9 (7.6)                          |
| <b>Neuroticism (Eysenck Personality Questionnaire)</b>     | 4.2 (3.7)                                | 4.3 (3.7)                            |
| <b>Psychoticism (Eysenck Personality Questionnaire)</b>    | 2.5 (2.1)                                | 2.8 (2.1)                            |
| <b>Extraversion (Eysenck Personality Questionnaire)</b>    | 14.7 (4.5)                               | 14.5 (3.7)                           |
| <b>Lie (Eysenck Personality Questionnaire)</b>             | 9.5 (4.6)                                | 7.5 (3.4)                            |
| <b>Trait Anxiety (State-Trait Anxiety Inventory)</b>       | 31.2 (9.1)                               | 32.1 (9.1)                           |
| <b>Depression at inclusion (Beck Depression Inventory)</b> | 1.6 (1.7)                                | 2.5 (4.0)                            |

**Table 1:** Demographic, physical, and psychological characteristics of the Pramipexole and Placebo groups, presented as ‘means (standard deviations)’

### 2.3 Methods: Study Design and Intervention

From a starting dose of 0.25mg of Pramipexole salt, the dose was increased in 0.25mg increments every 3 days, reaching a dose of 1mg/day by day 10. Participants continued to take 1mg/day for 3-5 days (until testing was completed). Following this, the dose was down-titrated over 3 days to avoid withdrawal effects. The apparent dose of the placebo was increased in the same manner.

Participants performed a probabilistic instrumental learning task (PILT; see below for details) before the intervention and then twice between days 12-15 of the intervention (one with fMRI data collection, one behavioural). In this chapter I will report on only the behavioural session results. I will then report on the fMRI session results in chapter 3.



**Figure 2; study design.** Following a screening session, participants underwent the pre-intervention behavioural testing session in which they performed the PILT task (described below). Participants received the first dose of Pramipexole/placebo at the end of this behavioural testing session. Between days 12-15 of the Pramipexole/placebo course, participants attended an fMRI session (in which they performed a computerized task whilst undergoing fMRI) and, on a separate day, a behavioural testing session which was identical to the pre-intervention behavioural testing session. In this chapter I will report on results from the two behavioural sessions.

## 2.4 Methods: Study Questionnaires

At the screening session, participants completed the Eysenck Personality Questionnaire (EPQ), Becks Depression Inventory (BDI) and Spot-the-word test (an estimate of IQ). At both behavioural testing sessions, participants completed the Befindlichkeitsskala (BFS), Positive and Negative Affect Schedule (PANAS), State-Trait Anxiety Inventory (STAI), Snaith-Hamilton Pleasure Scale (SHAPS), Temporal Experience of Pleasure Scale (TEPS), Oxford Happiness Questionnaire (OXH) and Impulsive-Compulsive Disorders in Parkinson's Disease-Rating Scale (QUIP). At the post-intervention behavioural testing session, participants additionally completed a side-effects questionnaire.

*Questionnaire scores:* No effect of drug treatment was found for any of the questionnaire measures, other than the anticipatory subscale of the TEPS, which was driven by a higher baseline score in the Pramipexole group (Table 2). Behavioural and neuroimaging analyses controlling for baseline TEPS anticipatory subscale scores are reported in the appendices 3 and 12.

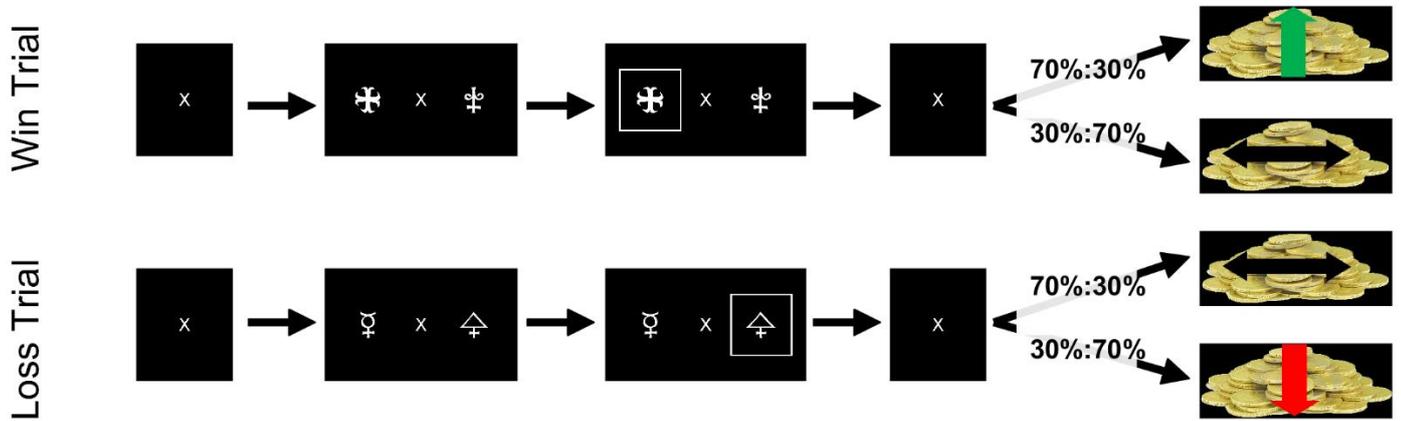
|                                 | Pramipexole  |             | Placebo      |             | ANCOVA  |
|---------------------------------|--------------|-------------|--------------|-------------|---------|
|                                 | Baseline     | On drug     | Baseline     | On drug     |         |
| <b>BFS</b>                      | 11.7 (12.4)  | 15.4 (14.7) | 13.8 (11.2)  | 14.8 (14.2) | p=0.51  |
| <b>STAI state</b>               | 28.9 (6.5)   | 27.8 (5.4)  | 28.6 (6.1)   | 29.1 (5.2)  | p=0.28  |
| <b>BDI</b>                      | 1.9 (2.5)    | 2.9 (3.4)   | 2.8 (3.3)    | 2.7 (3.3)   | p=0.32  |
| <b>PANAS positive present</b>   | 36.4 (7.7)   | 34.5 (8.3)  | 32.5 (8.6)   | 32.6 (7.9)  | p=0.5   |
| <b>PANAS negative present</b>   | 11.2 (1.5)   | 11.0 (1.5)  | 11.3 (1.3)   | 11.8 (2.7)  | p=0.15  |
| <b>PANAS positive today</b>     | 36.8 (8.0)   | 34.8 (9.1)  | 32.9 (8.2)   | 32.3 (8.1)  | p=0.73  |
| <b>PANAS negative today</b>     | 11.4 (2.0)   | 11.4 (1.7)  | 11.6 (1.7)   | 11.8 (2.4)  | p=0.55  |
| <b>PANAS positive last week</b> | 37.7 (9.2)   | 37.0 (7.7)  | 34.2 (7.6)   | 34.6 (9.3)  | p>0.99  |
| <b>PANAS negative last week</b> | 13.3 (2.8)   | 12.7 (3.2)  | 14.1 (4.1)   | 12.6 (2.7)  | p=0.71  |
| <b>SHAPS</b>                    | 0.5 (1.0)    | 0.7 (1.6)   | 0.2 (0.7)    | 0.3 (0.7)   | p=0.624 |
| <b>TEPS total</b>               | 85.3 (10.0)  | 82.1 (9.5)  | 79.3 (8.2)   | 79.2 (10.0) | p=0.61  |
| <b>TEPS anticipatory</b>        | 47.1 (6.1)*  | 44.6 (5.5)  | 42.1 (4.9)*  | 42.3 (5.4)  | P=0.38  |
| <b>TEPS consummatory</b>        | 38.1 (5.6)   | 37.5 (5.9)  | 37.3 (4.4)   | 36.8 (5.4)  | p=0.91  |
| <b>OXH</b>                      | 134.8 (19.2) | 138 (18.4)  | 132.4 (19.6) | 132 (19.7)  | p=0.08  |
| <b>QUIP</b>                     | 12.3 (8.3)   | 8.7 (8.5)   | 16.6 (11.2)  | 13.8 (10.4) | p=0.33  |

**Table 2:** Questionnaire scores before and after the intervention: Becks Depression Inventory (BDI), Befindlichkeitsskala (BFS), Positive and Negative Affect Schedule (PANAS), State-Trait Anxiety Inventory (STAI), Snaith-Hamilton Pleasure Scale (SHAPS), Temporal Experience of Pleasure Scale (TEPS), Oxford Happiness Questionnaire (OXH) and Impulsive-Compulsive Disorders in Parkinson's Disease-Rating Scale (QUIP). Scores are presented as 'means (standard deviations)'. Asterisks represent significant between-group, within-condition differences in questionnaire scores.

## 2.5 Methods: Task

Participants completed a modified version of a probabilistic instrumental learning task (PILT) described by Pessiglione et al (Pessiglione et al., 2006). The PILT is a 2-arm bandit task (Figure 3) with interleaved 'win' and 'loss' trials. In each trial, the participant is presented with two stimuli which have reciprocal probabilities (0.7 vs 0.3) of a 'win' outcome (+£0.20) vs a 'no win' outcome (£0.00) in reward condition trials, or a 'loss' outcome (-£0.20) vs a 'no loss' outcome (£0.00) in loss condition trials. Participants choose one of the two stimuli, following which they received visual feedback on the trial outcome and their current total earnings. Each block of the PILT consists of 30 reward trials

and 30 loss trials. Participants performed 3 blocks of the PILT in each behavioural testing session. Different task stimuli were used in each block. Participants started each session with £1.50 of funds. Participants received their winnings from this task (up to a maximum of £30).



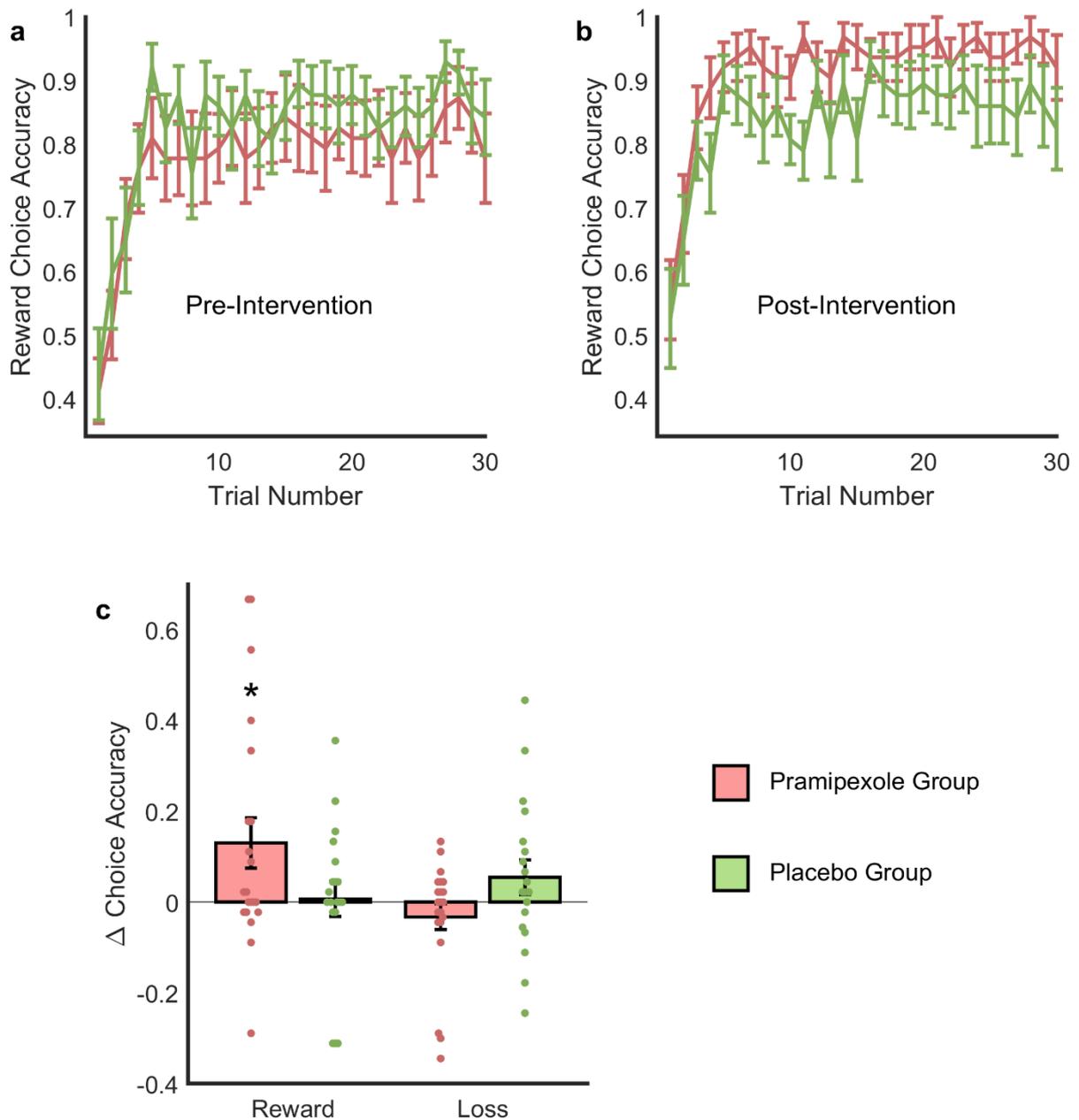
**Figure 3; Probabilistic Instrumental Learning Task (PILT).** In each trial, participants were presented with one of two possible pairs of shapes. For one of the shape pairs (top line), one shape was associated with winning money on 70% of trials and not winning on the remaining 30% (the other shape had reciprocal contingencies). For the other shape pair (bottom line), one shape was associated with losing money on 70% of trials and not losing on the remaining 30% (again, the other shape had reciprocal contingencies). Participants had to learn to select the shapes that were associated with the high probability of win/no-loss.

## 2.6 Methods: Model-Free Metric

The behavioural measure of interest was choice accuracy, defined as the proportion of advantageous choices made i.e. choosing the stimulus with 0.7 probability of 'win' in the reward condition or the stimulus with 0.7 probability of 'no-loss' in the loss condition. I used choice accuracy in the second half of each block (within-condition) as this provides a close estimate of asymptotic choice (Harrison et al., 2016; Walsh et al., 2018) found to be associated with depression. This metric was entered into a repeated measures ANOVA (Pramipexole vs placebo; win condition vs loss condition; pre vs post-intervention).

## 2.7 Results: Model Free Results

*Pramipexole Specifically Increases Asymptotic Choice of Rewarded Stimuli:* I found that there was a significant group\*valence\*time interaction for choice accuracy across behavioural sessions [Figure 2;  $F(1,38)=10.517$   $p=0.002$ ]. Win trial accuracy increased after treatment in the Pramipexole group [ $t(20)=2.347$   $p=0.029$ ], with no significant change in loss trial accuracy across sessions [ $t(20)=1.158$   $p=0.26$ ] and no change in either reward ( $p=0.86$ ) or loss trial accuracy ( $p=0.172$ ) in the placebo group. The Pramipexole and placebo groups did not differ significantly on reward or loss trial accuracy at baseline ( $p$ 's=0.435 and 0.395 respectively) or post-intervention ( $p$ 's=0.179 and 0.375 respectively). Using accuracy across all trials rather than those in the second half of blocks, the pattern of results remains the same (appendix 1).



**Figure 4; Reward accuracy in the pre and post intervention sessions.** (a-b) Learning curves depicting reward choice accuracy in the (a) pre-intervention and (b) post-intervention session. Green (pink) curves represent the placebo (Pramipexole) group. Error-bars represent SEM. The curves represent the proportion of participants who chose the advantageous shape (i.e. the shape associated with a 70% probability of receiving a 'win' outcome) in a given trial. (c) Mean (SEM) pre-vs-post intervention change in reward/loss condition choice accuracy.

## 2.8 Methods: Reinforcement Learning Models

I used a simple reinforcement learning model, which combined parameters from different, previously described models, to formalize the mechanistic question being addressed in this study. First a learning rule was used to update expectations about the association stimuli with the outcomes:

$$Q_{t+1(s)} = Q_{t(s)} + \alpha_j(\rho_j R_t - Q_{t(s)})$$

Here,  $Q_{t(s)}$  is the expectation about the value of shape  $s$  on trial  $t$ ,  $R_t$  is the observed outcome (1 for positive outcome, 0 for negative outcome),  $\alpha_j$  is the learning rate used for condition  $j$  (i.e. win or loss condition) and  $\rho_j$  is a reward sensitivity parameter for condition  $j$ . Expectations were initialized at  $Q_{0(s)} = 0.5$ , and the unchosen option,  $Q_{t(s')}$ , was updated with the reciprocal outcome. Following this, the model's expectations decayed back towards the initial value with the rate of decay controlled by a decay factor  $\phi$  (Collins and Frank, 2012):

$$Q \leftarrow Q + \phi(Q_0 - Q)$$

Finally, the  $Q$  values were fed into a softmax action selector to produce a choice:

$$P_{t(s)} = \frac{1}{1 + \exp^{-\beta(Q_{t(s)} - Q_{t(s')})}}$$

Here, the inverse temperature parameter,  $\beta$ , controls the degree to which the probability of the participant choosing shape  $s$ ,  $P_{t(s)}$ , is determined by the difference in  $Q$  values.

This model is over parameterized, the three parameters  $\rho$ ,  $\phi$  and  $\beta$  produce very similar effects on asymptotic choice (e.g. Figure 1) and therefore cannot be jointly estimated from participant behaviour. In order to account for changes in behaviour, two of the three parameters have to be fixed while the other (as well as the learning rate) remains free. Doing this is equivalent to making a statement about the presumed cause of the change in behaviour.

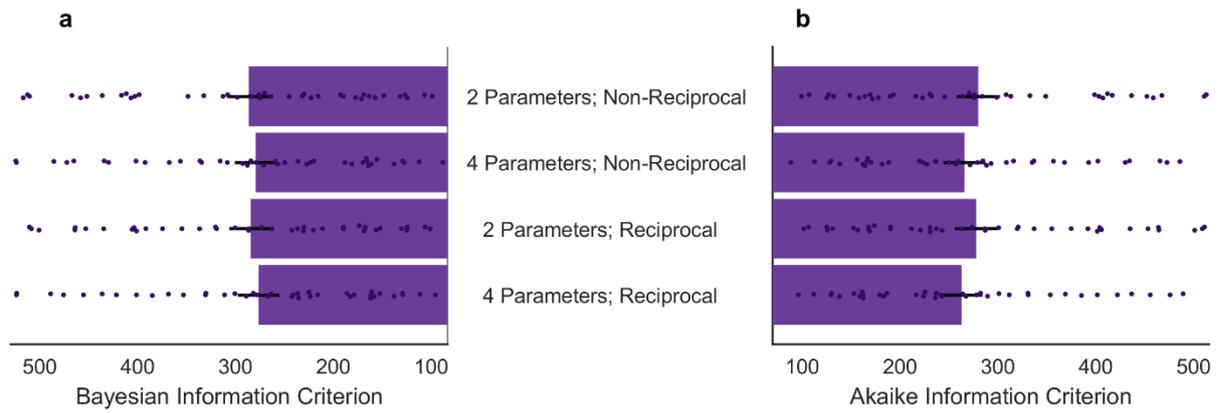
| Model-Type                        | Model | Parameters                                                             | Value-Updating          |
|-----------------------------------|-------|------------------------------------------------------------------------|-------------------------|
| Decision<br>Determinacy<br>Models | 1     | 2 parameters: $\alpha, \beta$                                          | Non-Reciprocal Updating |
|                                   | 2     | 4 parameters: $\alpha_{win}, \alpha_{loss}, \beta_{win}, \beta_{loss}$ | Non-Reciprocal Updating |
|                                   | 3     | 2 parameters: $\alpha, \beta$                                          | Reciprocal Updating     |
|                                   | 4     | 4 parameters: $\alpha_{win}, \alpha_{loss}, \beta_{win}, \beta_{loss}$ | Reciprocal Updating     |
| Reward Sensitivity<br>Models      | 5     | 2 parameters: $\alpha, \rho$                                           | Non-Reciprocal Updating |
|                                   | 6     | 4 parameters: $\alpha_{win}, \alpha_{loss}, \rho_{win}, \rho_{loss}$   | Non-Reciprocal Updating |
|                                   | 7     | 2 parameters: $\alpha, \rho$                                           | Reciprocal Updating     |
|                                   | 8     | 4 parameters: $\alpha_{win}, \alpha_{loss}, \rho_{win}, \rho_{loss}$   | Reciprocal Updating     |
| Belief Decay<br>Models            | 9     | 2 parameters: $\alpha, \phi$                                           | Non-Reciprocal Updating |
|                                   | 10    | 4 parameters: $\alpha_{win}, \alpha_{loss}, \phi_{win}, \phi_{loss}$   | Non-Reciprocal Updating |
|                                   | 11    | 2 parameters: $\alpha, \phi$                                           | Reciprocal Updating     |
|                                   | 12    | 4 parameters: $\alpha_{win}, \alpha_{loss}, \phi_{win}, \phi_{loss}$   | Reciprocal Updating     |

**Table 3. Models tested.** I tested three model-types ('belief decay', 'reward sensitivity' and 'decision determinacy'). They were Rescorla-Wagner models modified by the addition of, respectively, decay, reward sensitivity, and inverse temperature parameters. All models utilized learning rate parameters to update the valuation of task stimuli. For each model type, I tested whether participants utilized different parameter values for each of the two conditions or whether they used the same parameter values across both conditions. I also tested whether participants updated stimulus-values reciprocally, i.e. whether they updated the unchosen stimulus by the counterfactual outcome in a given trial. I expected that they would do so, as they were instructed that stimulus-outcome contingencies for individual trials would be reciprocal. I initialized stimulus values at 0.5 i.e. the midway point between the values of possible outcomes (0 and 1). I tested 4 variants of each model-type, totaling 12 models.

**2.9 Methods: Model Fitting:** In order to fit the three model variants described in Table 3 to participant choice, the joint posterior probability of the free parameters for each variant was calculated for each participant separately, given their choices. Each participant's parameter values were estimated as the expected value of the marginalised parameter distribution (Behrens et al., 2007; Browning et al., 2015).  $\rho$  and  $\beta$  parameters were sampled in log space while  $\alpha$  and  $\phi$  parameters were sampled in logit space. All statistical analyses were performed on transformed parameters.

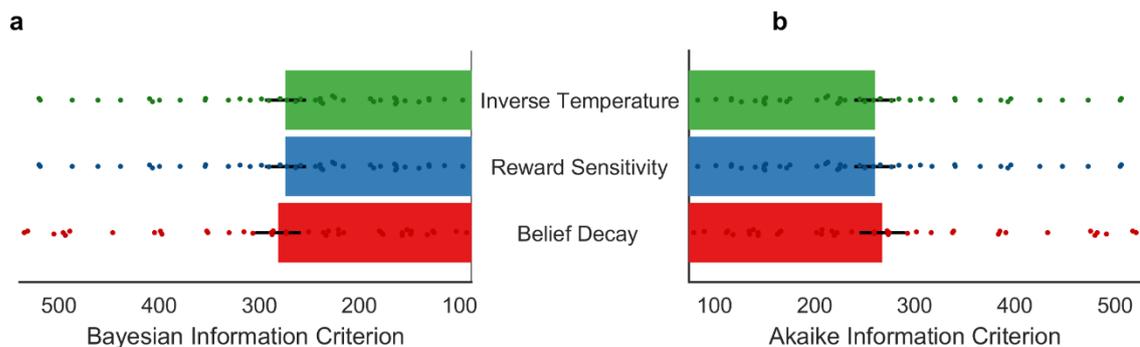
### **2.10 Methods: Model Selection**

I wanted, ultimately, to test which parameter (decision determinacy, reward sensitivity or decay) best accounted for the changes seen in the Pramipexole group from the pre-to-post intervention session. However, I first set out to determine which 'structure' of RL model to use (i.e. condition-specific vs independent parameters; reciprocal vs non-reciprocal updating). My a priori belief was that participants would utilize separate parameter values for the two conditions (as reflected in the hypothesis stated in the introduction to this chapter). I believed so because, clinically, Pramipexole appears to alter reward-related decision making seemingly independently of punishment-related decision making. I also believed a priori that participants update stimulus values reciprocally as they were instructed that in any given trial, one shape would yield a reward/punishment while the other would not. I first compared model structures by averaging BICs (Figure 5a) and AICs (Figure 5b) for each structure across parameter-categories (reward sensitivity/decay/inverse temperature). Those models that utilized condition-specific parameters performed better than those that didn't and reciprocally updating models performed better than the non-reciprocally updating models. The AIC/BIC results confirmed my a priori belief that participants used separate, condition-specific, parameters and that they updated values reciprocally.



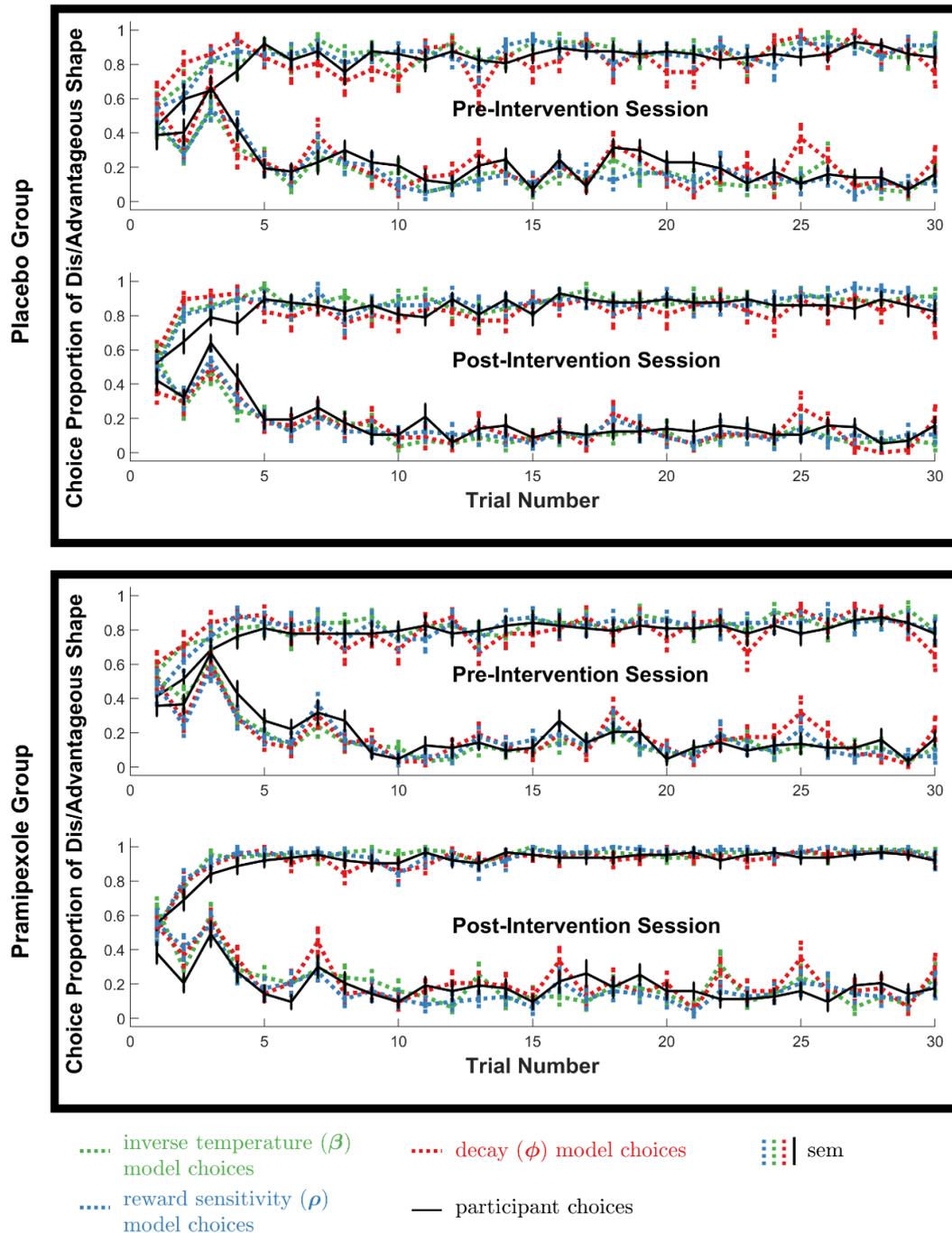
**Figure 5. Average (a) BIC and (b) AIC scores for each model structure:** 2 parameter models with non-reciprocal updating (models 1, 5 and 9), 4 parameter models with non-reciprocal updating (models 2, 6 and 10), 2 parameter models with reciprocal updating (models 3, 7 and 11), 4 parameter models with reciprocal updating (models 4, 8 and 12) (see table 3 for model descriptions).

I then compared BICs (Figure 6a) and AICs (Figure 6b) for each specific model (reward sensitivity i.e. model 8/decay i.e. model 12/inverse temperature i.e. model 4) within the winning structure-category. As can be seen, these latter three AICs/BICs are comparable, though the decay parameter model (12) does have a numerically larger AIC/BIC scores (mean AIC=276; mean BIC=281) than the inverse temperature and reward sensitivity models (mean AICs=260; mean BICs=274).



**Figure 6. Average (a) BIC and (b) AIC scores for each model within the winning structure-category** i.e. the reciprocally updating 4 parameter model that utilised inverse temperatures (model 4) reward sensitivity parameters (model 8) and decay parameters (model 12). Smaller AIC/BIC scores indicate a better model fit. AIC/BIC scores were calculated from log likelihoods that were summed across the 3 task blocks and across both sessions. Bars represent mean (SEM) AIC/BIC scores across all participants. Scatter plots overlaying bar graphs depict corresponding individual AIC/BIC values.

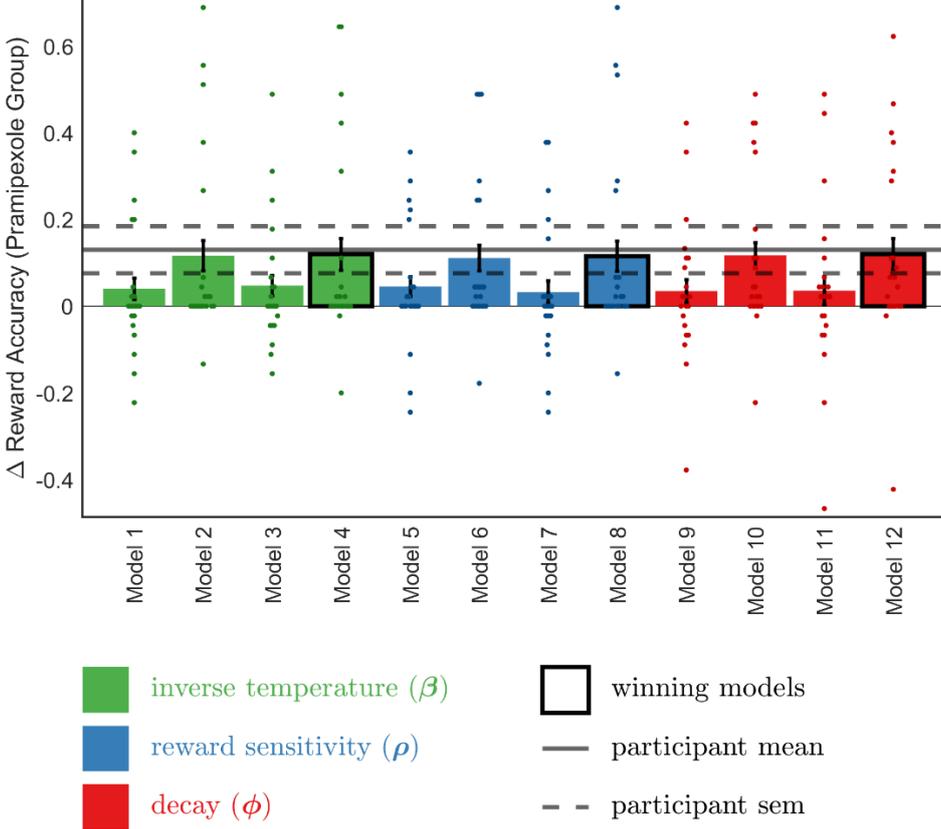
Taking trial-wise choices generated by each model using individual participants' parameters, I averaged these model-generated choices within group and condition to create learning curves for each model (equivalent to participants' learning curves in Figure 4a/b above). I plotted these learning curves against Pramipexole/placebo-group participants' actual learning curves and visually assessed the fidelity of the model-generated learning curves to participants' learning curves. The learning curves of those models that utilized condition-specific parameters appeared to have greater fidelity to participant learning curves than those that didn't. Little separated the apparent fidelity of learning curves generated using models that updated values reciprocally vs non-reciprocally. I created (2 groups \* 2 sessions \* 12 models =) 72 plots, which is impractical to display here. For illustrative purposes I have presented below 4 plots depicting participant learning curves vs those generated by the 3 winning models (Figures 5 and 6). As can be seen, the 3 models appear to have comparable fidelity to the participant behaviour though, arguably, the decision determinacy model (model 4) and reward sensitivity model (model 8) generated learning curves appear to have slightly higher fidelity to some of the participant learning curves than do the decay model (model 12) generated learning curves.



**Figure 7. Learning curves depicting model vs participant choice accuracy in each session.**

Approximately concave plots represent reward condition learning curves, that is, time series' depicting the proportion of participants/simulated-participants that chose the advantageous reward-condition stimulus in a given trial. Approximately convex plots represent loss condition learning curves, that is, time series' depicting the proportion of participants/simulated-participants that chose the disadvantageous loss-condition stimulus in a given trial. Solid black lines represent participant choices while dotted lines represent model-generated choices. Choices generated by model 4 are coloured green, choices generated by model 8 are blue and choices generated by model 12 are red. Error-bars represent SEM.

Finally, as the increase in reward choice accuracy within the Pramipexole group was the most pertinent change across sessions, I assessed how accurately the models replicated this change. I generated reward-condition choices using each model and individual participants' parameter values for each session, for each respective model. As I did for the model-free analysis, I used the proportion of 'accurate' choices in the latter half of each block. I plotted the change in this metric across sessions for each model (Figure 8). Models that utilized condition-specific parameters appear to perform better than those that utilize condition-independent parameters. There appears to be little difference between those models that update values reciprocally and those that update values non-reciprocally. This analysis therefore supports my a priori belief regarding condition-specific-vs-independent parameter values but does little to discriminate between reciprocally vs non-reciprocally updating models.



**Figure 8. Bar graphs depicting the change in model generated choice accuracy using pre-vs-post Pramipexole-group derived parameters.** Bars represent mean (SEM) change in the reward-accuracy of model-generated choices utilising parameters values inferred for participants that took Pramipexole. Inverse temperature model results are coloured green, reward sensitivity models blue and decay models red. 'Winning' models are framed in black. Scatter plots overlaying bar graphs depict corresponding individual values.

In summary, I used positive and negative metrics to assess 12 models comprising 4 versions (reciprocal vs non-reciprocal updating; condition-specific vs independent parameter values) of models that each used one of three different parameter-types (inverse temperature/reward sensitivity/decay) along with learning rate parameters. Models containing condition specific parameters performed better than those that didn't. Reciprocally updating models performed marginally better than non-reciprocally updating models. I therefore used models 4, 8 and 12 (described in Table 3) for the analyses below. Comparing models 4, 8 and 12, the inverse temperature (4) and reward sensitivity model (8) had slightly lower mean AIC and BIC scores than the decay model (12). All three models re-produce the effect of Pramipexole on reward accuracy to a comparable degree. All three models also reproduce participant learning curves to a comparable degree; though, arguably, Models 4 and 8 do so with marginally higher fidelity than model 12 with respect to some of the learning curves.

### **2.11 Results: Model Based Results**

*The Behavioural Effects of Pramipexole May be Attributed to Increased Reward Sensitivity, Reduced Choice Stochasticity or Reduced Reward Value Decay:*

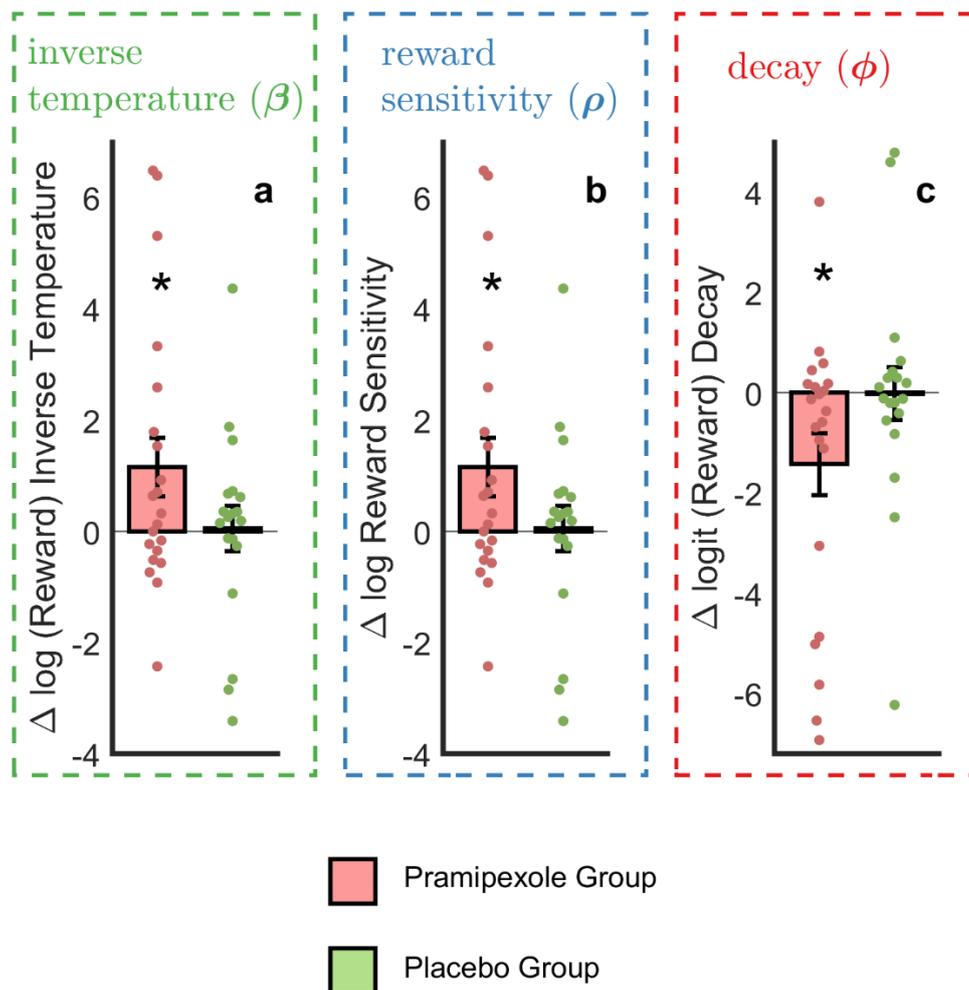
As outlined above, I fitted participants' behaviour to the winning versions of the reinforcement learning model described in the methods section; all three models were able to account for participant behaviour (see Figures 6-8).

Entering (separately) reward sensitivity, inverse temperature and decay parameter values into respective repeated measures ANOVAs in the same manner as for the observed effect of Pramipexole (Pramipexole vs placebo; win condition vs loss condition; pre vs post-intervention), I find a significant group \* valence \* time interaction in each case: **1**) Reward sensitivity: [F(1,38)=5.81 p=0.021], **2**) decay: [F(1,38)=7.96 p=0.008] and **3**) inverse temperature: [F(1,38)=5.81 p=0.021].

Underlying these results, win condition reward sensitivity/inverse temperatures increased, and decay decreased, in the Pramipexole group after treatment ( $p_s \leq 0.04$ ; Figure 9 a-c), with no significant change in loss trial reward sensitivity/inverse temperature/decay across sessions ( $p_s \geq 0.31$ ). Reward sensitivity/inverse temperature/decay did not change in either reward ( $p_s \geq 0.90$ ; Figure 9 a-c) or loss condition ( $p_s \geq 0.38$ ) in the placebo group. The Pramipexole and placebo groups did not differ significantly on reward or loss sensitivity/inverse temperature/decay at baseline ( $p's \geq 0.563$  and  $0.533$  respectively) or post-intervention ( $p's \geq 0.090$  and  $0.398$  respectively). Therefore

the observed effect of Pramipexole on choice behaviour may result from increased reward sensitivity, decreased reward value decay, or increased inverse temperature.

There were no significant changes in learning rate in any of the three models, in either condition, in either group, across sessions ( $p \geq 0.20$ ). Likewise, repeated measures ANOVAs of learning rate parameters in each model showed no main effects or interactions that included group differences ( $p \geq 0.10$ ). The Pramipexole and placebo groups did not differ significantly on reward or loss learning rates at baseline ( $p \geq 0.7626$  and  $0.6981$  respectively) or post-intervention ( $p \geq 0.2568$  and  $0.1054$  respectively).



**Figure 9; Change in model parameters in the reward condition.** The bar graphs show the pre-post intervention change when the (a) reward sensitivity, (b) inverse temperature, and (c) decay parameter were allowed to vary. The parameter changes illustrated here produced the pre-vs-post intervention model behaviour (in the reward condition) illustrated in figures 7 and 8. Green (pink) bars represent the placebo (Pramipexole) group. Error-bars represent SEM. Scatter plots overlaying bar graphs depict corresponding individual values. Asterisks represent significant ( $p < 0.05$ ) pre-vs-post intervention change.

## 2.12 Discussion

I investigated the effect of the dopamine D2-like receptor agonist Pramipexole on reward-related decision making in healthy volunteers. As predicted, a sub-acute course of Pramipexole (titrating to 1mg/day) increased choice accuracy in the reward (but not loss) condition. Fitting three behavioural RL models (RS, IT and Decay), I found that all three recapitulated participant behaviour to a comparable degree.

Pramipexole increased asymptotic choice of highly rewarded stimuli. This is in contrast to the majority of previous experimental studies of Pramipexole which have found that it impairs reward learning (Gallant et al., 2016; Murray et al., 2019; Pizzagalli et al., 2008a; Santesso et al., 2009). However, past studies have generally administered a single low dose of Pramipexole. Translational studies suggest that acute administration reduces, whereas sustained administration enhances, dopamine transmission (Chen et al., 2005; Chernoloz et al., 2008; Grace, 1991; Willner et al., 1994). Human imaging studies also suggest dichotomous effects of acute vs sustained administration. Arterial spin labelling demonstrated an *increased* metabolic rate in depression-relevant brain regions, in healthy volunteers, following acute pramipexole administration (Michels et al., 2016) (approximated by local cerebral blood flow (Hoge et al., 1999)). In contrast, a positron emission tomography study demonstrated *reduced* metabolic rate in depression-relevant brain regions following sustained pramipexole administration in bipolar depression (Mah et al., 2011b) (approximated by uptake of a glucose analogue (Sokoloff et al., 1977)). This latter effect is diametric to that of unipolar and bipolar depression (Drevets, 1999; Mah et al., 2007). The current results are consistent with the few reward processing studies of patient groups in which longer treatment durations were used, and which also found an increase in reward choice following Pramipexole treatment (Bodi et al., 2009; Burdick et al., 2014). Together, these studies indicate that putatively post-synaptically acting, sustained dosing of Pramipexole acts to enhance reward learning. Depression is associated with reduced reward learning (Blanco et al., 2013; Halahakoon et al., 2020; Henriques et al., 1994; Huys et al., 2013; Huys and Browning, in press; Kunisato et al., 2012; Liu et al., 2011; Must et al., 2006; Pizzagalli et al., 2005; Robinson and Chase, 2017; Steele et al., 2007; Vrieze et al., 2013). Overall, therefore, the impact of Pramipexole on learning is opposite to that associated with depression and is consistent with the antidepressant effects of the drug (Cusin et al., 2013; Fawcett et al., 2016; Tundo et al., 2019).

Pramipexole had no effect on loss condition accuracy. This finding is consistent with previously reported effects of dopaminergic agents on the performance of healthy volunteers in probabilistic learning tasks (Eisenegger et al., 2014; Pessiglione et al., 2006). It has been speculated that, while the

win condition engages only appetitive processing, the loss condition may engage both appetitive and aversive processing (Pessiglione et al., 2006). This distinction is supported by the observation that punishing stimuli appear to evoke different neural responses than ‘non-rewarding’ stimuli. That is, they evoke (phasic) aversive prediction errors which are different from positive or negative RPEs, and are produced by a population of VTA neurons that are distinct from those that produce RPEs (Brischoux et al., 2009; Coizet et al., 2006; Guarraci and Kapp, 1999; Matsumoto et al., 2016; Matsumoto and Hikosaka, 2009). Interestingly, the relationship between reward vs punishment processing in depression appears similarly dichotomous to Pramipexole’s effect on reward vs punishment learning: depressed individuals are clearly *hyposensitive* to reward (Halakoon et al., 2020), whereas the relationship between depression and punishment processing appears more complicated (Elliott et al., 1997; Eshel and Roiser, 2010).

I fitted three models to the behavioural data: RS, IT and decay. The RS model tested the hypothesis that Pramipexole increases the subjective value of outcomes, thereby amplifying the learned values of the stimuli. The decay model tested the hypothesis that Pramipexole reduces the forgetting/decaying of beliefs about stimulus values, consequently facilitating more advantageous choices. The IT model tested the hypothesis that Pramipexole does not affect learning at all and instead shifts participants towards more deterministic (vs stochastic) choices. Comparing the models to the behavioural data did not differentiate between them as all of them fitted and recapitulated participants’ choice behaviour comparably well. Previous computational analyses of Pramipexole’s effect have suggested an effect on learning rate (Huys et al., 2013) or no effect on any computational parameter (Murray et al., 2019). However, as stated above, a single low dose of Pramipexole is thought to inhibit phasic dopaminergic signaling (Piercey et al., 1996). In contrast, 14-day administration of Pramipexole increases tonic activity (Chernoloz et al., 2012). Consistent with my findings, high dose Sulpride (which is thought to *reduce* tonic dopamine activity) reduced accuracy and ITs in the reward, but not loss, condition in a task similar to the current one (Eisenegger et al., 2014). This effect was driven by a subset of participants who had a genetically determined lower density of striatal D2 receptors (Eisenegger et al., 2014). Genetic determinants of D2 receptor density may also be relevant to the effect of Pramipexole in depression.

The current study has a number of limitations. Most obviously, the population recruited were non-clinical healthy participants. A non-clinical population was selected to reduce phenotypic variation among participants and thus enhance the sensitivity of this experimental medicine study to detect the pharmacological effects of Pramipexole. However, this design is not able to assess the degree to which change in reward learning mediates clinical response in patients. Answering this question requires a clinical trial of Pramipexole in which patients complete the PILT task before and after

initiating treatment with Pramipexole or placebo. Such a trial is currently underway (Au-Yeung et al., 2022).

### **2.13 Conclusion**

Using a 2-week course of Pramipexole titrated to 1mg/day, I find that Pramipexole enhances learning from rewards. Fitting three models (RS, IT and decay) to the behavioural data, I find little to discriminate between them behaviorally. Though an increase in RS, increase in IT or decrease in decay have comparable effects on choice behaviour, each has a distinct effect on reward prediction errors. In chapter three, I will attempt to arbitrate between the three models by using functional magnetic resonance imaging to examine Pramipexole's effect on neural reward prediction errors.

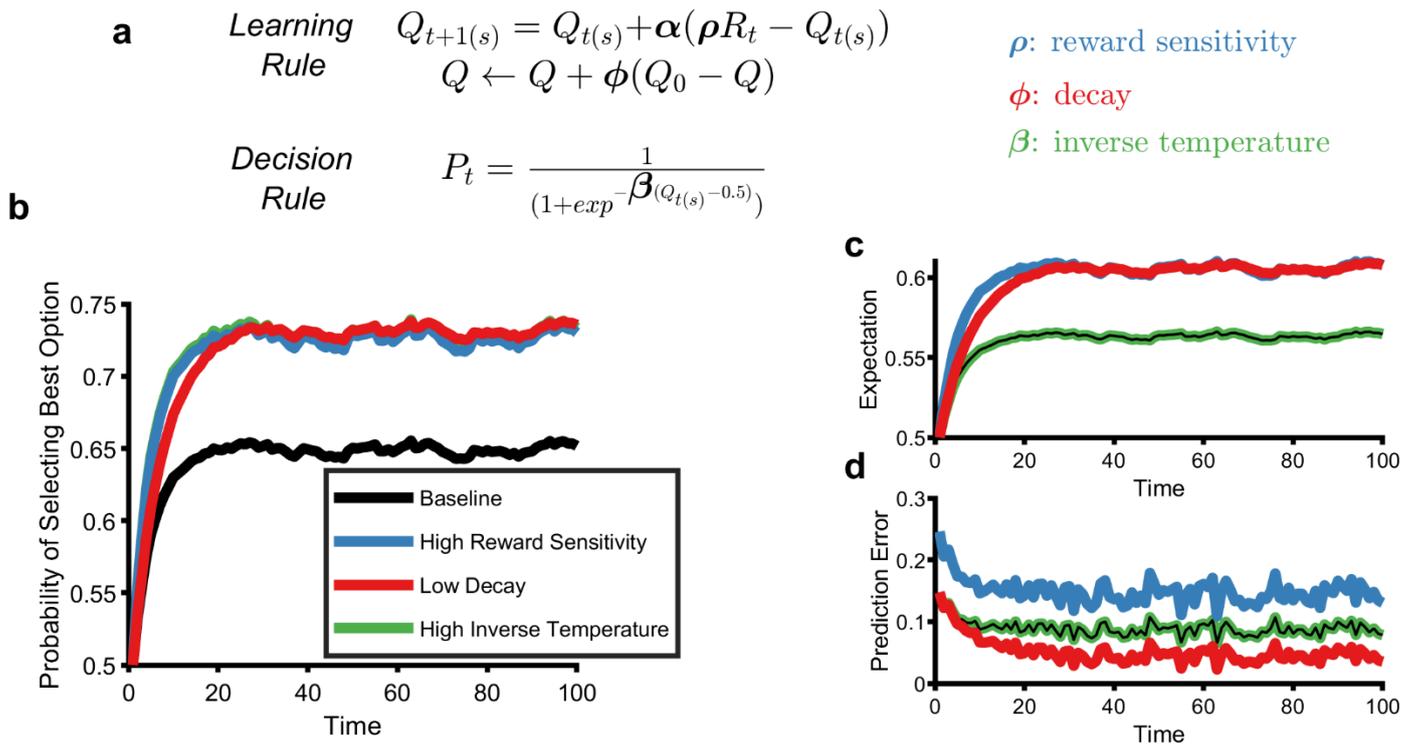
## Chapter 3: The effect of Pramipexole on neural reward processing

### 3.1 Introduction

In chapter 2 I found that a 2 week course of Pramipexole increased accuracy in the reward but not loss condition, in a PILT. I fitted three models to the behavioural data and found that each explained the behavioural observations to a comparable degree. Though these models can produce very similar behavior, they make different predictions about how Pramipexole affects RPEs in the win condition. The RS model predicts that Pramipexole amplifies the value of a reward and therefore amplifies neural responses to rewards (e.g. win outcomes). The decay model predicts that Pramipexole attenuates neural responses to rewards as they are more expected (as a consequence of Pramipexole preventing the Q-value from decaying to baseline). The IT model predicts that Pramipexole has no effect on neural responses to rewards. These predictions are outlined in table 1 and illustrated in Figure 1.

**Table 1:** The three causal proposals of changes in asymptotic choice associated with the reward sensitivity, decay and inverse temperature parameters of the reinforcement learning model.

| Presumed Cause                                                         | Free Parameter                         | Effect on Asymptotic Choice                  | Effect on Expectation                       | Effect on Prediction Error                  |
|------------------------------------------------------------------------|----------------------------------------|----------------------------------------------|---------------------------------------------|---------------------------------------------|
| <i>A differential response to experienced reward</i>                   | Reward Sensitivity, $\rho$             | Increased with increased reward sensitivity  | Increased with increased reward sensitivity | Increased with increased reward sensitivity |
| <i>A differential rate of decay of learned expectations</i>            | Decay parameter, $\phi$                | Increased with decreased decay               | Increased with decreased decay              | Decreased with decreased decay              |
| <i>A differential effect of learned values on decision probability</i> | Inverse temperature parameter, $\beta$ | Increased with increased inverse temperature | No effect                                   | No effect                                   |



**Figure 1 a** Recap of the model described in the previous chapter: Rewards,  $R$ , are combined with expectations,  $Q$ , in a learning rule and then fed into a decision rule. Distinct parameters modify each component of this process: a reward sensitivity parameter,  $\rho$ , influences the effective size of experienced rewards, a decay parameter,  $\phi$ , influences the degree to which expectations are maintained between trials, a learning rate parameter,  $\alpha$ , influences the rate at which rewards alter expectations and an inverse temperature parameter,  $\beta$ , influences the degree to which expectations are used to determine choices. **b** As demonstrated in the previous chapter (and simulated here) a high reward sensitivity, high inverse temperature or a low decay produce effectively indistinguishable learning curves. The learning curves displayed here were generated by the learning and decision-making rules described in ‘a’ using a  $\rho$  of 0.6, a  $\phi$  0.12, a  $\beta$  of 10 and an  $\alpha$  of 0.1. Increases in either  $\rho$  (blue line) or  $\beta$  (green line) and decreases in  $\phi$  (red line) produce equivalent changes in asymptotic choice. In other words, three qualitatively distinct processes lead to the same behavioural effect. As a result, choice data on its own cannot be used to distinguish between these processes. However, the internal model variables do differ, and thus can discriminate, between these processes. Panel **c** illustrates model expectations,  $Q$ . As can be seen, either increasing  $\rho$  or decreasing  $\phi$  causes an increase in expectations, whereas  $\beta$ , which influences decision-making rather than learning, does not change expectations (i.e. the green and black lines are identical). Panel **d** illustrates the prediction errors (PE) of the models, which are able to discriminate the three parameters. Again, changes in  $\beta$  have no effect, whereas an increase in  $\rho$  leads to increased PEs and a reduction in  $\phi$  leads to decreased PEs.

In this chapter I attempt to arbitrate between these three hypotheses regarding the latent effect of Pramipexole driving the change in choice behaviour across sessions. In order to do so I need to measure participants' internal model variables during the task. These internal variables (reward expectations and RPEs) appear to be coded by neural activity in various parts of the brains reward circuit (Elliott et al., 2000; Fouragnan et al., 2018; Haber and Knutson, 2010), as outlined in the overall thesis introduction. I will assess neural activity in three regions of interest (ROIs) within this circuit:

- 1) The **orbitofrontal cortex (OFC)** is thought to maintain outcome expectations in working/representational memory (Schoenbaum and Roesch, 2005) and exert a top-down effect on response selection (Frank and Claus, 1996).
- 2) The **medial prefrontal cortex (mPFC)** is thought to signal RPEs (Oya et al., 2005; Rutledge et al., 2010).
- 3) The **ventral striatum (VS)** is central to the reward circuit and activity within it convincingly reflects RPEs (McClure et al., 2003; Pessiglione et al., 2006; Rutledge et al., 2010).

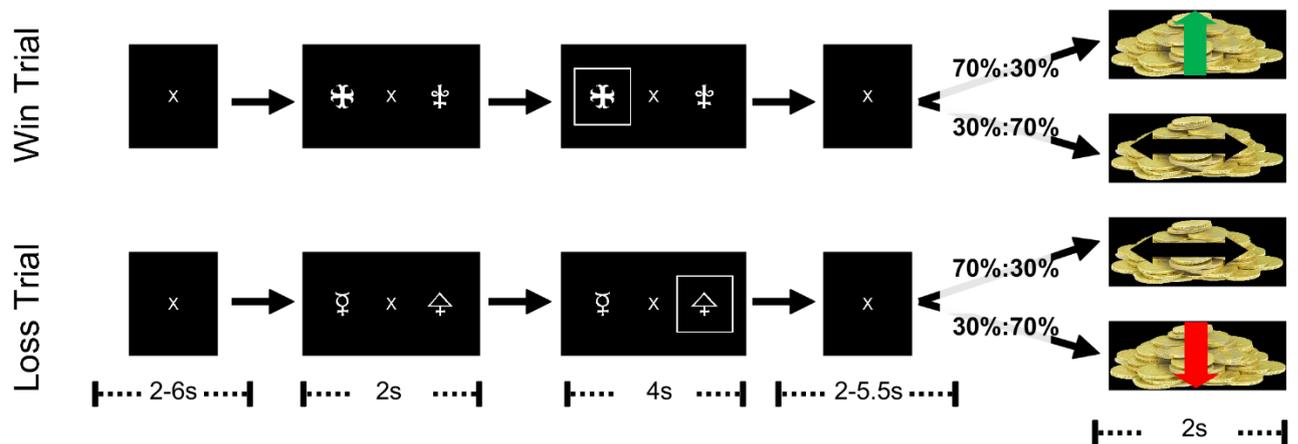
I used functional magnetic resonance imaging (fMRI) to measure the strength of the blood oxygen level dependent (BOLD) signal, which is a reasonable measure of neuronal activity (Rees et al., 2000). I investigated participants' BOLD activity during task engagement to attempt to discover which (if any) of the 3 models reflects Pramipexole's modulation of reward processing within the context of the PILT task. As registered in clinical trials.gov (<https://clinicaltrials.gov/ct2/show/NCT03681509>), I hypothesized that Pramipexole would increase subjective valuation of rewards (leading to increased BOLD to rewarding outcomes in reward sensitive areas of the brain such as the VS, OFC and mPFC) (Fouragnan et al., 2018).

### 3.2 Study Recap

I outlined the study design in more detail in chapter 2. To recap, 42 participants were randomized to receive either Pramipexole or placebo. The Pramipexole group was titrated up to a dose of 1mg/day by day 10, which they continued to take for 3-5 days until testing was completed, and then down-titrated till discontinuation. The apparent dose of the placebo was altered in the same manner. Participants performed the PILT before the intervention and then twice between days 12-15 of the intervention (one with fMRI data collection, one behavioural).

### 3.3 Imaging Task

In the fMRI session, participants played the PILT as described in chapter 2 with two differences. Instead of three runs, they played two runs. To allow for the timing constraints of the haemodynamic response, the task was slowed considerably. To increase design efficiency, the inter-trial period and the period between choice and outcome-presentation were jittered, as illustrated in figure 3.



**Figure 3. Probabilistic Instrumental Learning Task (PILT).** In each trial, participants were presented with one of two possible pairs of shapes. For one of the shape pairs (top line), one shape was associated with winning money on 70% of trials and not winning on the remaining 30% (the other shape had reciprocal contingencies). For the other shape pair (bottom line), one shape was associated with losing money on 70% of trials and not losing on the remaining 30% (again, the other shape had reciprocal contingencies). Participants had to learn to select the shapes that were associated with the high probability of win/no-loss. In the fMRI scanner, participants completed 2 runs of 60 trials each. The inter-trial interval (during which time an 'x' appears on the screen) was jittered using a flat distribution of 2-6 seconds. Likewise, an intra-trial interval, which occurred between the time that trial stimuli disappeared and the trial outcome was revealed, was jittered using a flat distribution of 2-5.5 seconds.

### 3.4 MRI Image Acquisition

MRI images were acquired at the Oxford Centre for Human Brain Activity (OHBA), University of Oxford, using a 3T Siemens Prisma scanner with a 32-channel head-coil. T1-weighted structural images had a voxel resolution of 1mm isotropic. The echo time was 3.97 msec, repetition time was 1900msec, flip angel was 8° and field of view was 192mm. BOLD Images were T2-weighted echo-planar images. Sixty slices were acquired with a voxel resolution of 2.4mm isotropic, repetition time

of 800msec, echo time of 30msec, flip angle of 52°, field of view of 216mm and the multiband acceleration factor was 6 interleaved. Fieldmaps were collected with echo times of 4.92 and 7.38ms, repetition time of 590msec and flip angle of 46°. Cardiac and Respiratory data were collected using a pulse oximeter and respiratory belt respectively.

### 3.5 MRI Data analysis

MRI data were analysed using FSL (FMRIB Software Library v6.0) tools. Data were pre-processed to reduce unwanted variability in the data. Pre-processing involved 1) removal of non-brain structures using BET (Smith, 2002), 2) motion correction using MCFLIRT (Jenkinson et al., 2002), 3) spatial smoothing with a Gaussian kernel of FWHM 5mm, 4) un-warping using fieldmaps and 5) high-pass temporal filtering with a cut-off of 60 seconds. Functional images were registered non-linearly to corresponding structural images via a high contrast functional image and BBR (Jenkinson et al., 2002; Jenkinson and Smith, 2001). Cardiac and Respiratory data were used to create 33 denoising regressors using PNM (Brooks et al., 2008).

Task events were represented using separate explanatory variables for the presentation of stimuli (2s period during which stimuli were first presented) in win and loss trials, and separate variables representing the four possible outcomes (2s period during presentation of win, no-win, loss or no-loss outcomes). Additional EVs were included to account for respiratory and cardiac noise. Activity associated with expectation during learning was captured as the relative difference between signal strength during the stimuli presentation period for win vs. loss condition trials. Post-hoc analyses then compared expectation associated activity between the 1<sup>st</sup> vs 2<sup>nd</sup> half of trials in a block (i.e. when expectation should be low relative to when it should be high, Figure 1c). The contrast between 'win' vs 'no-win' outcomes, and 'no-loss' vs 'loss outcomes' were used as simple non-model-based measures of prediction errors. I supplemented this analysis with three model-based analyses in which the estimated prediction errors from each of the three models (decay, beta and rho) respectively, were used as parametric regressors in place of the binary outcome regressors. In the model-based analyses, explanatory variables were: the presentation of stimuli (2s period during which stimuli were first presented) in win and loss trials, the presentation of outcomes (2s period during which outcomes were presented) in win-condition and loss-condition trials and win/loss condition RPEs, these were the same as the win/loss-condition outcome EVs but with demeaned RPEs as parametric modulators. First-level analyses were run for each participant and both blocks of the task. The outputs of these analyses were then averaged, within subject, across the blocks and

entered into a higher-level random-effects analysis which assessed the difference between the two groups. See appendices 6-8 for further details.

The higher-level analyses were restricted to three anatomical ROIs:

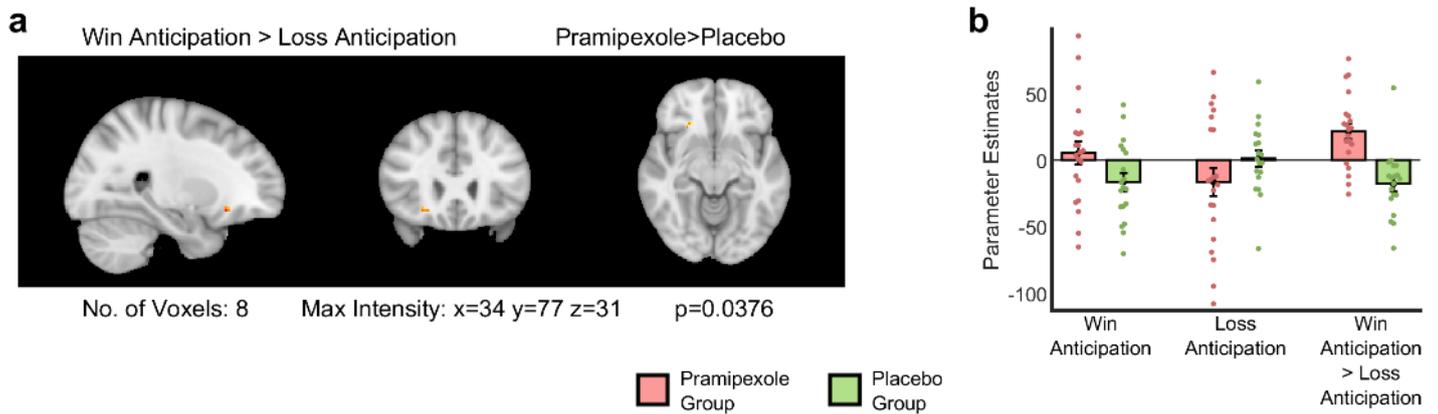
- 1) mPFC (defined using the Harvard-Oxford Structural Atlas library (Desikan et al., 2006))
- 2) OFC (defined using the Harvard-Oxford Structural Atlas library (Desikan et al., 2006))
- 3) Ventral Striatum (defined using the Oxford-GSK-Imanova Structural–anatomical Striatum Atlas (Tziortzi et al., 2014)).

Group level inference was performed using the FSL nonparametric permutation tool (Randomise) with 5000 permutations, threshold free cluster enhancement method and family-wise error correction ( $p < 0.05$ ).

### **3.6 Results: Model Free Analysis; Anticipation**

#### *Pramipexole Increases BOLD Signal during Anticipation of Rewards vs. Losses*

Anticipation of reward stimuli, as measured using the activity during presentation of stimuli in reward relative to loss trials, was increased in participants receiving Pramipexole relative to placebo in the OFC ROI (Figure 4). This is consistent with the RS or Decay models, but not the IT model. There were no significant clusters for this contrast in the mPFC or VS ROIs.

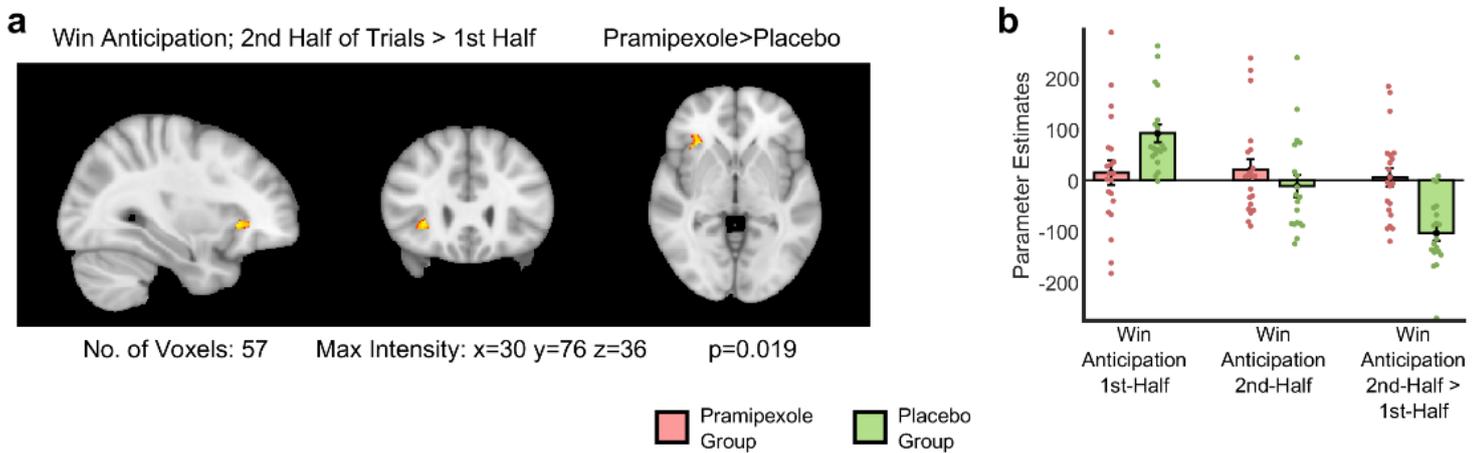


**Fig 4. Model Free Analysis: Anticipation** The red-yellow coloured area represents the cluster of significantly increased activity in the Pramipexole vs placebo group in the OFC ROI during win-anticipation > loss-anticipation. Areas of significantly increased activity are threshold free cluster enhancement corrected with a family-wise error cluster significance level of  $p \leq 0.05$ . **(b)** Parameter estimates extracted from the area of significantly increased activity in Fig 4a associated with win-anticipation, loss-anticipation and win-anticipation > loss-anticipation. Green (pink) bars represent the placebo (Pramipexole) group. Error-bars represent SEM. Scatter plots overlaying bar graphs depict corresponding individual parameter estimates.

#### *Pramipexole Increases BOLD Signal during Anticipation of Rewards across trials*

To probe this result, I examined the development of win and loss expectations during the task blocks. As illustrated in Figure 1c, expectations develop as learning proceeds. I captured this process by subtracting the response to stimuli in the first half of trials (trial 1-15) from the latter half (trials 16-30) separately for reward and loss trials. There was no significant group (Pramipexole vs placebo)\*trial( 1-15 vs 16-30)\*condition (reward anticipation vs loss anticipation) interaction. Within the reward condition, participants receiving Pramipexole had greater increase in activity across the block than those receiving placebo in the OFC ROI (Peak voxel  $x=30, y=76, z=36$ ; voxel size:57;  $p=0.019$ ) (Figure 5a-b). There were no clusters for this contrast within mPFC or VS ROIs, nor for the OFC/mPFC/VS ROIs during loss stimulus presentation. On the surface, this result is consistent with Pramipexole causing an increase of reward sensitivity or reduction in reward expectation decay, as both processes lead to increased reward expectation (Figure 1d). It is not consistent with an increase in inverse temperature, which does not require a change in expectations. However examining activity within this area in the 1<sup>st</sup> and 2<sup>nd</sup> half of trials separately, the placebo group had relatively high activity in the 1<sup>st</sup> half of trials, which dropped considerably in the 2<sup>nd</sup> half of trials, while the Pramipexole group had comparable activity in the two halves. While it can be difficult to

straightforwardly interpret extracted signal effects from functionally defined clusters, it is not obvious that the overall interaction signifies an increase in expectation by the Pramipexole group over time.

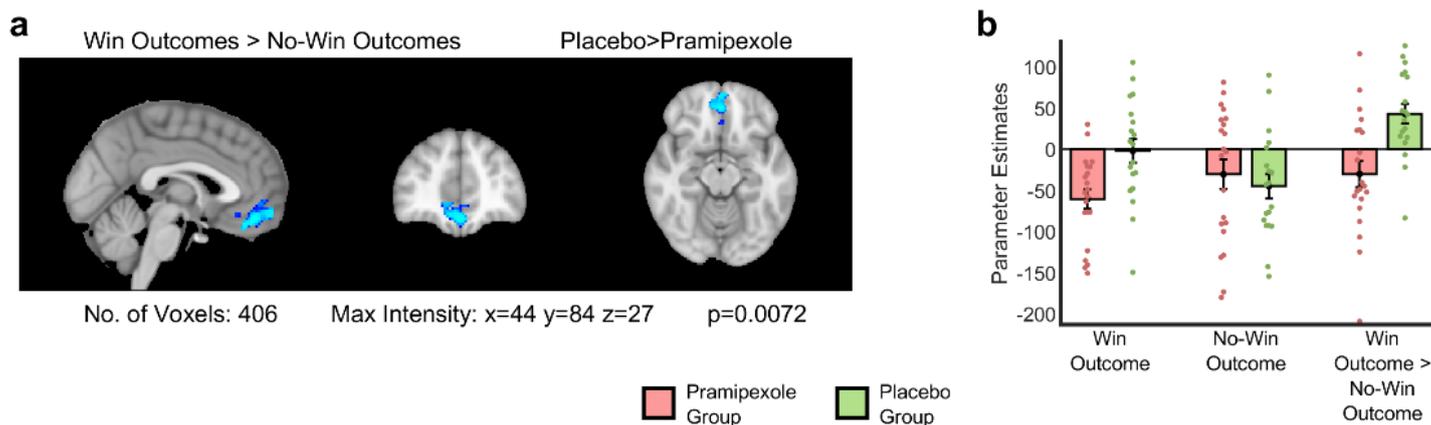


**Fig 5. Post-Hoc Analysis: Win-Anticipation.** (a) The (red-yellow) coloured area represents the cluster of significantly increased activity in the Pramipexole vs placebo group in the OFC ROI during win-anticipation in the 2<sup>nd</sup> half > 1<sup>st</sup> half of win-condition trials. Areas of significantly increased activity are threshold free cluster enhancement corrected with a family-wise error cluster significance level of  $p \leq 0.05$ . (b) parameter estimates extracted from the area of significantly increased activity in 5a associated with win-anticipation in the 1<sup>st</sup> half, 2<sup>nd</sup> half, and 2<sup>nd</sup> half > 1<sup>st</sup> half of win-condition trials. green (pink) bars represent the placebo (Pramipexole) group. Error-bars represent SEM. Scatter plots overlaying bar graphs depict corresponding individual parameter estimates.

### 3.7 Results: Model Free Analysis; Outcome

#### *Pramipexole Decreases the BOLD Signal Associated with Reward Receipt vs Reward Omission*

The response to rewarded outcomes, as measured using activity associated with win relative to no-win outcomes, was reduced in the Pramipexole group relative to the placebo group in the mPFC ROI (Figure 6a-b). Pramipexole did not influence activity in loss relative to no-loss trials, or in the OFC or VS ROIs.



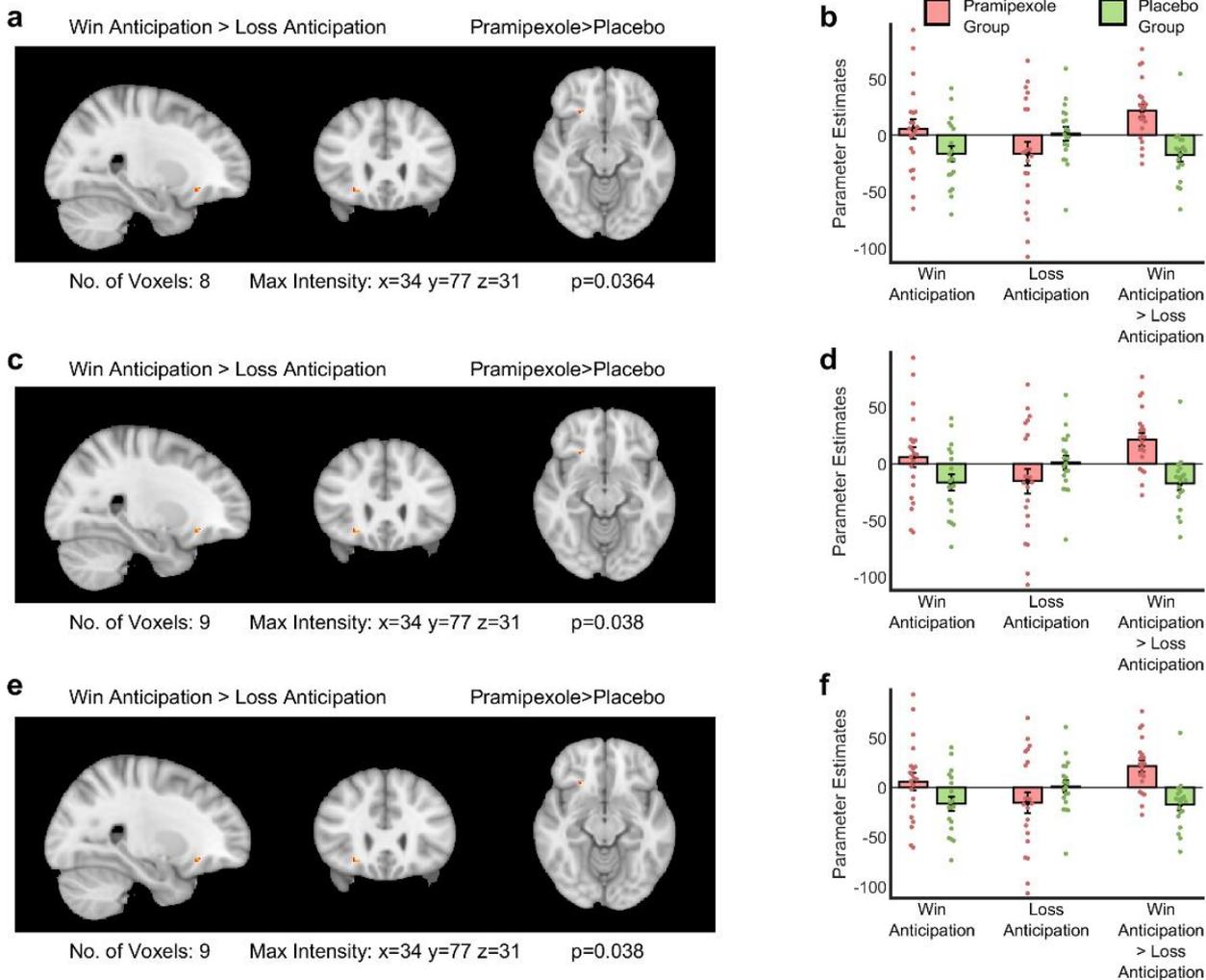
**Fig 6. Model Free Analysis: Outcome (a)** The (blue) coloured area represents the cluster of significantly decreased activity in the Pramipexole vs placebo group in the mPFC ROI associated with win-outcomes > no-win-outcomes. Areas of significantly decreased activity are threshold free cluster enhancement corrected with a family-wise error cluster significance level of  $p \leq 0.05$ . **(b)** Parameter estimates extracted from the area of significantly decreased activity in Fig 6a associated with win-outcomes, no-win-outcomes, and win-outcomes > no-win-outcomes. Green (pink) bars represent the placebo (Pramipexole) group. Error-bars represent SEM. Scatter plots overlaying bar graphs depict corresponding individual parameter estimates.

### 3.8 Results: Model Based Analysis; Comparing Reward Sensitivity, Inverse Temperature and Decay Models

To confirm the above result, I fitted each model (RS, IT, Decay) to each person's imaging session behavioural data. I then extracted RPEs for each model, for each person and used them as parametric regressors for 3 separate model-based analyses (RS, IT, Decay).

#### *Pramipexole Increases BOLD Signal during Anticipation of Rewards vs. Losses*

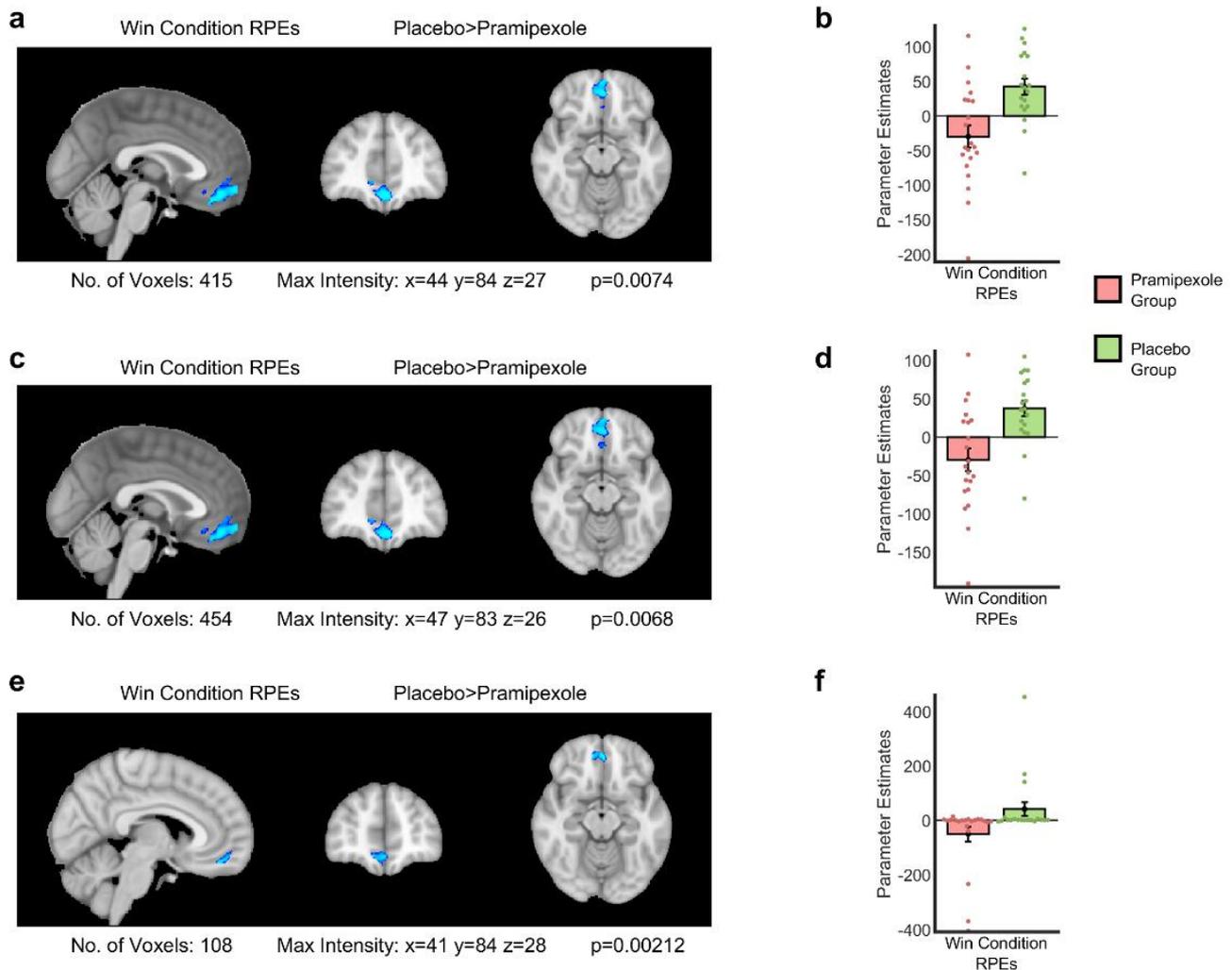
As with the model free analysis, anticipation of reward stimuli (vs loss stimuli) was increased in participants receiving Pramipexole relative to placebo in the OFC ROI (Figure 7). This is consistent with the RS or Decay models, both of which posit increased expectation in the reward vs loss condition, but not the IT model, which does not. There were no significant clusters for this contrast in the mPFC or VS ROIs.



**Fig 7. Model based fMRI (expectation) results: (a,c,e)** Win-anticipation > loss-anticipation contrast from the **(a)** decay-model based analysis **(c)** beta-model based analysis and **(e)** rho model based analysis. The red-yellow coloured areas represent the cluster of significantly increased activity in the Pramipexole vs placebo group in the OFC ROI. Areas of significantly increased activity are threshold free cluster enhancement corrected with a family-wise error cluster significance level of  $p \leq 0.05$ . **(b,d,f)** Parameter estimates extracted from the area of significantly increased activity in *Fig 7 a, c* and *e* respectively, associated with win-anticipation, loss-anticipation and win-anticipation > loss-anticipation. Green (pink) bars represent the placebo (Pramipexole) group. Error-bars represent SEM. Scatter plots overlaying bar graphs depict corresponding individual parameter estimates.

## *Pramipexole Decreases the BOLD Signal Associated with RPEs*

Win-condition RPEs were reduced in the Pramipexole group relative to the placebo in the mPFC ROI (Figure 8) in all three model based results. This is consistent with internal decay-model mechanisms, but not internal RS or IT-model mechanisms, as illustrated in figure 1.



**Fig 8. Model based fMRI (outcome) results: (a,c,e)** Win-condition RPE contrasts from the (a) decay-model based analysis (c) beta-model based analysis and (e) rho-model based analysis. The (blue) coloured area represents the cluster of significantly decreased activity in the Pramipexole vs placebo group for win-condition RPEs in the mPFC ROI. Areas of significantly decreased activity are threshold free cluster enhancement corrected with a family-wise error cluster significance level of  $p \leq 0.05$ . (b,d,f) Parameter estimates extracted from the areas of significantly decreased activity in Fig 8 a, c and e respectively, associated with win-condition RPEs. Green (pink) bars represent the placebo (Pramipexole) group. Error-bars represent SEM. Scatter plots overlaying bar graphs depict corresponding individual parameter estimates.

### 3.9 Discussion

As demonstrated in Chapter 2, Pramipexole treatment may increase asymptotic reward choice by a number of distinct cognitive mechanisms. It is not possible to arbitrate between these mechanistic hypotheses using choice behaviour from the PILT task alone. I used fMRI to gauge activity in areas of the brain associated with reward processing, during the presentation of task stimuli and the receipt of trial outcomes. Using this activity I inferred the internal processes that might underlie the observed behaviour. Pramipexole enhances (OFC) BOLD activity during the anticipation of rewarded trials and suppresses the (mPFC) BOLD response to win outcomes. This suggests that it enhances reward learning by reducing the decay of value estimates and suggests a cognitive mechanism by which it may ameliorate the reduced reward learning characteristic of depression and anhedonia (Chen et al., 2015; Halahakoon et al., 2020; Husain and Roiser, 2018).

During reward-anticipation (vs loss-anticipation), Pramipexole increased activity in the OFC ROI. I infer from this contrast that Pramipexole amplifies expectation in the win condition. Consistent with this inference, the contrast is significant within the OFC, activity in which has been found to track reward expectation (Moorman and Aston-Jones, 2014; Schoenbaum et al., 1998; Tremblay and Schultz, 1999) and is altered in depression (Forbes et al., 2006; Xie et al., 2021). The Decay and RS models attribute the behavioural effect of Pramipexole to its effect on the Q value, which represents the 'expected value' of a stimulus. By contrast, Pramipexole has no effect on the Q value in the IT model. A significant between-group contrast to reward (vs loss) anticipation, suggests an effect on reward-condition Q values and is therefore consistent with the Decay and RS models. This contrast is arguably as much in the insula as in the OFC. Activity in the insula has also been associated with anticipation of rewards (Oldham et al., 2018).

I found that the Pramipexole group had higher OFC activity in the second half of trials vs first half of trials, relative to the placebo group. This contrast could equally be interpreted as the placebo group having higher activity in the first half of trials vs the second half of trials, relative to the Pramipexole group. Probing the within trial-half contrasts, the placebo group activity was far higher than the Pramipexole group in first half of trials (within condition/run) and fell to below the level of activity in the Pramipexole group in the second half. Although extracted signal effects from functionally defined clusters should be interpreted cautiously, it is not obvious that the overall interaction signifies an increase in expectation by the Pramipexole group over time. It is also not clear that activity during this phase purely represents expectation. It is in this phase of the task that participants likely (cognitively) select which option they will chose, so activity during this phase may also include

decision effects (e.g. an inverse temperature effects). Therefore there is some ambiguity about what this result signifies.

Pramipexole reduced the response to win (vs no-win) outcomes and reward prediction errors in the mPFC. Activity in the mPFC reflects RPEs (Fouragnan et al., 2018; Rutledge et al., 2010) and anterior vmPFC activity correlates with experienced value in simple choice tasks (Haber and Knutson, 2010; Smith et al., 2010). In a previous EEG study utilizing a signal detection task, Pramipexole amplified mPFC activity during (positive) outcomes (Santesso et al., 2009). Though we find the opposite, the two findings may be consistent with each other, as we use a regime of Pramipexole that is purported to have a diametric effect on dopaminergic transmission and we find a diametric effect on (correlates of) neural activity. Depression itself is associated with a reduced BOLD response to rewarding outcomes (Greenberg et al., 2015; Huys and Browning, in press; Kumar et al., 2008; Pizzagalli et al., 2009; Ruppel et al., 2020), which suggests that the reduced learning in patients (Blanco et al., 2013; Halahakoon et al., 2020; Henriques et al., 1994; Huys et al., 2013; Huys and Browning, in press; Kunisato et al., 2012; Liu et al., 2011; Must et al., 2006; Pizzagalli et al., 2005; Robinson and Chase, 2017; Steele et al., 2007; Vrieze et al., 2013) is the result of a lower effective value of rewards rather than a difference in decay of value estimates. The current findings indicate that Pramipexole does not act directly to reverse the cognitive profile of depressed patients, but rather improves reward learning via a separate mechanism. This result may go some way to explaining why the clinical response to Pramipexole in depression seems to be higher in patients with intact, rather than impaired, baseline reward learning (Whitton et al., 2020). Specifically, as Pramipexole does not increase reward sensitivity, the impact of the drug on reward learning, and presumably on symptoms of depression, will depend on an intact response to rewarding outcomes, and will be reduced in those patients with an impaired response.

Pramipexole's effect on decay may potentially explain Pramipexole-induced compulsivity (Kolla et al., 2010) due to (mal-adaptively) persistent salience of previously rewarding stimuli/activities. This is compatible with reports, by patients suffering from Pramipexole-induced compulsive disorders, of persistent pre-occupation with rewarding activities even in the absence of obvious cues (Murphy et al., 2021). It is not obvious how Pramipexole might implement the reduction of value decay. As stated in the thesis introduction, dopamine RPEs affect future behaviour by altering the rate of LTP/D. The rate of early phase LTP/D decay is determined by the synaptic availability of  $\alpha$ -amino-3-hydroxy-5-methyl-isoxazole-4-propionic acid receptors (AMPAr)(Dong et al., 2015). Exposure to the D-2 like receptor agonist Quinpirole alters surface availability of AMPAr in the striatum (Ferrario et al., 2011) and increases perseverative errors when learned contingencies need to be unlearned (Pezze et al., 2007). Speculatively, it is possible that Pramipexole might act via a similar mechanism.

It could be argued that the 'decay' is similar to working memory (Arnsten et al., 2015). Previous modelling work has demonstrated that simple learning tasks are often solved using a mixture of working memory and reinforcement learning based processes, with working memory acting to reduce prediction error responses by maintaining representations of current value (Collins and Frank, 2018). The observed effect of Pramipexole may therefore reflect an increase in the degree to which participants rely on working memory when completing the PILT. However, a general enhancement of working memory should also influence loss learning, rather than produce a reward-specific effect as found here. It is therefore necessary to evoke some form of valence-specific working memory effect to explain the current findings. Ultimately, the potential role of working memory in the effect of Pramipexole would best be tested by manipulating memory load during learning (Collins and Frank, 2012).

### **3.10 Limitations**

There were no between-group behavioural differences in the imaging session (reported in appendix 4). It is unclear why we see no behavioural effect in the imaging session while we observe a marked effect across behavioural sessions. One possibility is that the alien experience of performing the task in an MRI scanner influenced the behavioural results in this session. Another is that the lower number of task blocks in the imaging session vs the behavioural sessions may have resulted in a lower signal-to-noise ratio in the latter relative to the former. Lastly, as the imaging session took place before the post-intervention behavioural session, it is possible that the participants were not exposed to the maximum dose of Pramipexole for long enough by the time of the imaging session, to produce a significant behavioural result.

I have not accounted for the possibility that Pramipexole may have an effect on the BOLD signal independent of neural activity. Arterial spin labelling was carried out during the same fMRI session and demonstrated that Pramipexole did not influence resting brain perfusion or global perfusion; these finds are reported elsewhere (Martens et al., 2021).

### **3.11 Conclusion**

A 2-week course of Pramipexole enhanced asymptotic choice of highly rewarded stimuli, while reducing the neural response to rewarding outcomes. These results indicate that Pramipexole enhances reward learning by reducing the decay of learned value estimates and suggests a potential cognitive mechanism by which it may act to ameliorate symptoms of depression.

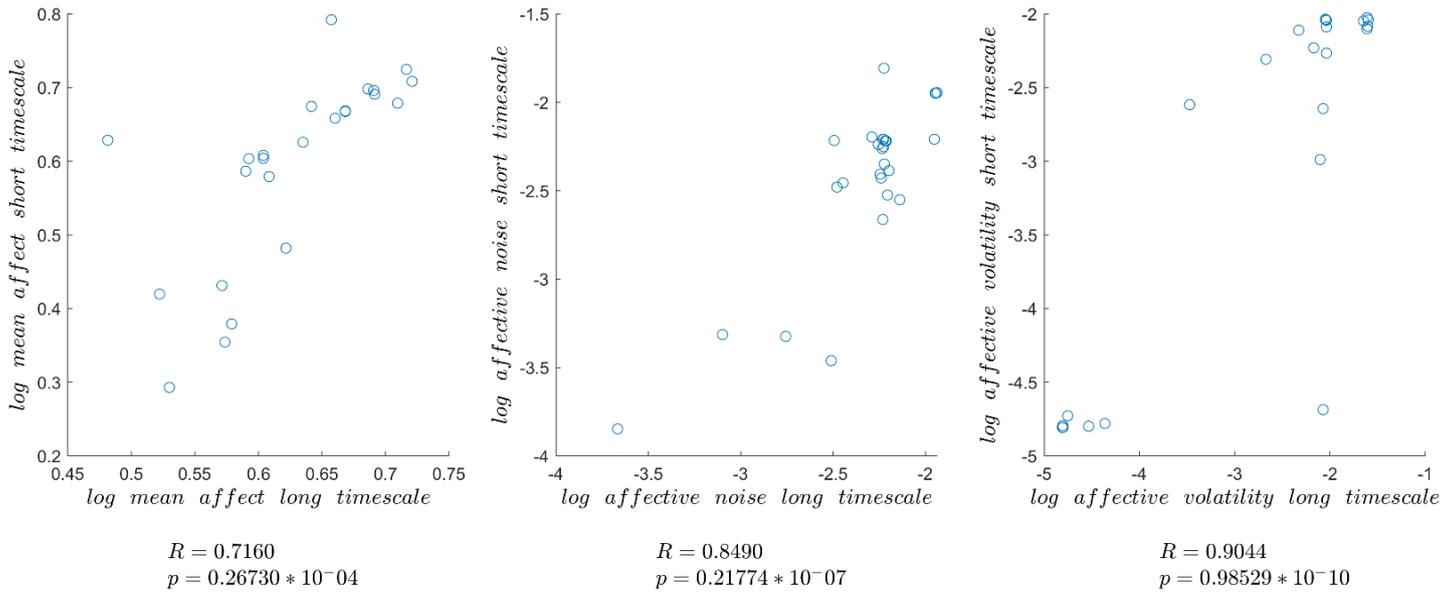
## Chapter 4: Linking mood variability at different time scales

### 4.1 Introduction

Mood variability is an important aspect of major depressive disorder (MDD), playing roles in both aetiology and prognosis (Bowen et al., 2013; Marwaha et al., 2013; Peters et al., 2016).

Understanding the mechanisms underlying mood instability is therefore important for understanding depression. While some studies have probed the mechanisms underlying mood variability, many of these studies have approached mood variability as a unidimensional (or even binary) phenomenon. Conversely, experience sampling method (ESM) studies suggest that understanding the role of mood variability in MDD necessitates understanding mood variability in a more nuanced manner (Koval et al., 2013). If such a nuanced analysis of ESM data could be linked to affective dynamics during reward learning (which has a stronger biological framework than ESM) it may allow for e.g. translational insights into the biology of affect instability. Koval et al have previously tried to link lab based affective measures (emotional reactions to film clips), obtained during a short controlled schedule of events with experience sampling method (ESM) over a longer timescale (1 week) (Koval et al., 2013). Using an autocorrelation model, this study found trend-level (i.e.  $p < 0.07$ ) positive correlations between ESM and lab-based measures for all but one dynamic measure.

In a recent paper, Pulcu et al (Pulcu et al., 2022) have proposed a model (a Bayesian filter) that parses ESM data into 'mean affect', 'volatility' and 'noise'. The term affective 'noise' ( $SD$ ) describes affective fluctuations around a stable mean ( $\mu$ ) while the term affective 'volatility' ( $\nu\mu$ ) describes apparent changes in the mean affect ( $\mu$ ) over time. This model can be applied to affective time-series measured at a large time-scale (such as ESM data) and, equally, at a small time-scale (such as a series of affective reports during a reward learning task). As with the study by Koval et al. (Koval et al., 2013), individuals' Bayesian affective parameters ( $\mu$ ,  $\nu\mu$ ,  $SD$ ) at a short time-frame (e.g. during a reward learning task) may lend insight into their affective parameters at a longer time frame (e.g. ESM). This idea is illustrated below in figure 1 which shows that affective parameters acquired from a small portion (3%) of affective time-series, generally appear to correlate with equivalent parameters acquired from the full time-series.



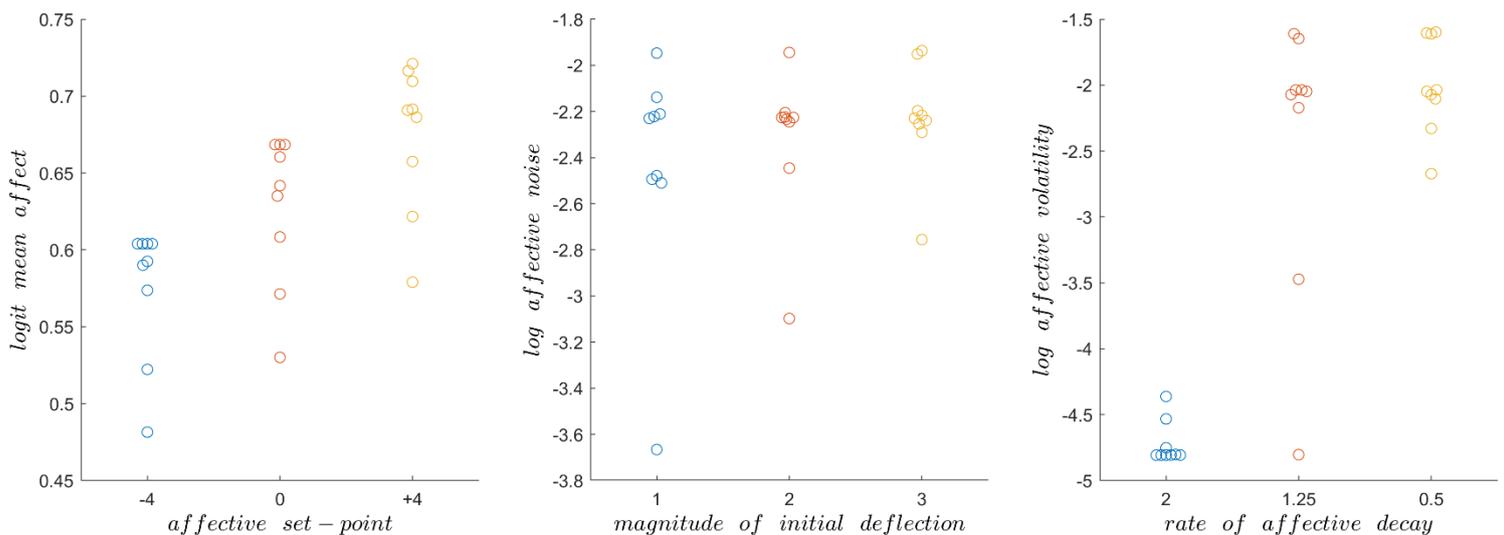
**Figure 1. Bayesian filter parameters; Short time-frame vs Long time-frame.** I used a randomly generated reward schedule (100,000 time-points consisting of 99%=0, 1%=1, 1%=-1) to generate 27 ‘affective’ time-series using the formula described in the thesis introduction\* with each combination of parameters from 3 affective set-points (-4, 0, 4), 3 deflection magnitudes (1, 2, 3) and 3 decay rates (0.5, 1.25, 2). I then sampled 100 evenly spaced data-points from each time series and passed these through the Bayesian filter described by Pulcu et al (Pulcu et al., 2022) (and described below in the methods section). To create ‘short time-frame’ affective time-series, I extracted 100 evenly spaced data-points from a small randomly and independently selected (3%) portion of each of the 27 affective time-series described above. These ‘short-timeframe’ affective time-series were then passed through the Bayesian filter and the resulting parameters were correlated against corresponding parameters from the respective ‘full time-series’. As illustrated in the three panels, parameters acquired at a relatively high (~33 times) sampling frequency during a short time period (3% of the full time-series), correlated with those acquired at a relatively low sampling frequency during a long time-period. All other correlations between ‘short-timeframe’ and ‘full time-series’ parameters were non-significant ( $p \geq 0.0571$ ) bar the correlation between ‘short-timeframe’  $\nu mu$  and ‘full time-series’  $SD$  ( $p=0.0201$ ).

\*This formula is:

$$affect_T = setpoint + \sum_1^T event_t(a \cdot e^{b \cdot (t-T)})$$

Affect at time-point  $T$  is a weighted sum of affective deviations from baseline, caused by events at each time-point  $t$ , up to and including time-point  $t = T$ . The value of  $a$  determines the extent to which events deviate affect from baseline. The value of  $b$  determines the rate at which affective deviations return to

baseline. This is an attempt to crudely approximate a theory of affect put forward by Kuppens et al (Kuppens et al., 2010) which suggests that individuals **1)** have a ‘baseline’ affective set-point from which **2)** affect gets *deflected* by internal or external events and to which affect **3)** *returns* over time. The three components in this affective model (i.e. affective *set-point*, initial *deflection* and *return* to baseline) plausibly loosely differentially relate to the parameters  $\mu$ ,  $SD$  and  $\nu\mu$  respectively such that: **1)** A high/low baseline affective set-point is somewhat analogous to high/low  $\mu$ . **2)** Large/small initial affective deflections from baseline are somewhat analogous to more/less ‘noisy’ affect (i.e. a larger or smaller  $SD$ ). **3)** Least intuitively, the rate at which affect returns to baseline might relate to ‘volatility’. Fast/slow return to baseline is somewhat analogous to less/more ‘volatile’ affect (i.e. lower/higher  $\nu\mu$ ). This is loosely illustrated in figure 2. The biggest caveat here is that in the exponential decay model, the affective ‘set-point’ is (by definition) set, while in the Bayesian analysis,  $\mu$  can change over time.



**Figure 2. Illustration of the effect of exponential decay components on Bayesian filter parameters.**

As illustrated in the 3 panels above, the filter’s estimated  $\mu$  appears to generally increase as the affective set point increases, the filter’s estimated  $SD$  appears to generally increase as the initial deflection increases and the filter’s estimated  $\nu\mu$  appears to generally increase as the decay rate decreases.

In this chapter I primarily assess whether affective parameters within the reward learning task reflect corresponding affective parameters in everyday life. I do so by regressing ESM affect parameters against within-task affect parameters.

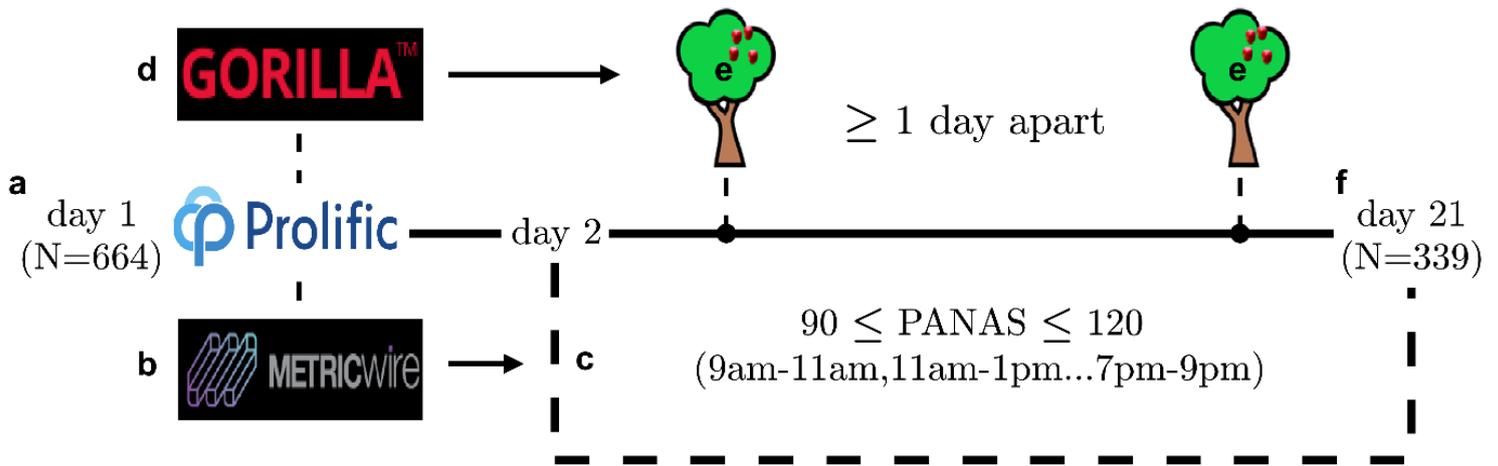
In a more exploratory analysis I test the association between affective parameters and the duration of affective responses to events in the relatively controlled setting of a reward learning task. Specifically, I assess whether **a**) affective noise ( $SD$ ) depends on the immediate affective impact of events (which is somewhat analogous to the *initial deflection*), whether **b**) affective volatility ( $\nu\mu$ ) depends on the extent to which the affective responses persist (which is somewhat analogous the *rate of affective decay*), and whether **c**) affective mean ( $\mu$ ) has no such relationship with the immediate or persistent impact of events (as it is roughly the *affective set-point*). Finally I test whether affective parameters in everyday life directly reflect these proxies to components of the hypothetical exponential decay model. I do so by repeating analyses a, b and c, substituting ESM parameters for within-task parameters.

**My hypothesis is:**

-Affective mean, volatility and noise within a short timeframe (the reward learning task) reflect corresponding parameters on a longer timeframe (ESM data).

In the exploratory analysis I explore whether the immediate affective impact of events (a proxy for the 'initial affective deflection') underlies affective 'noise'; the persistent affective impact of events (a proxy for the 'rate of affective decay') underlies affective volatility; mean affect does not depend on the affective impact of events (as it would not, if it reflected the affective set-point).

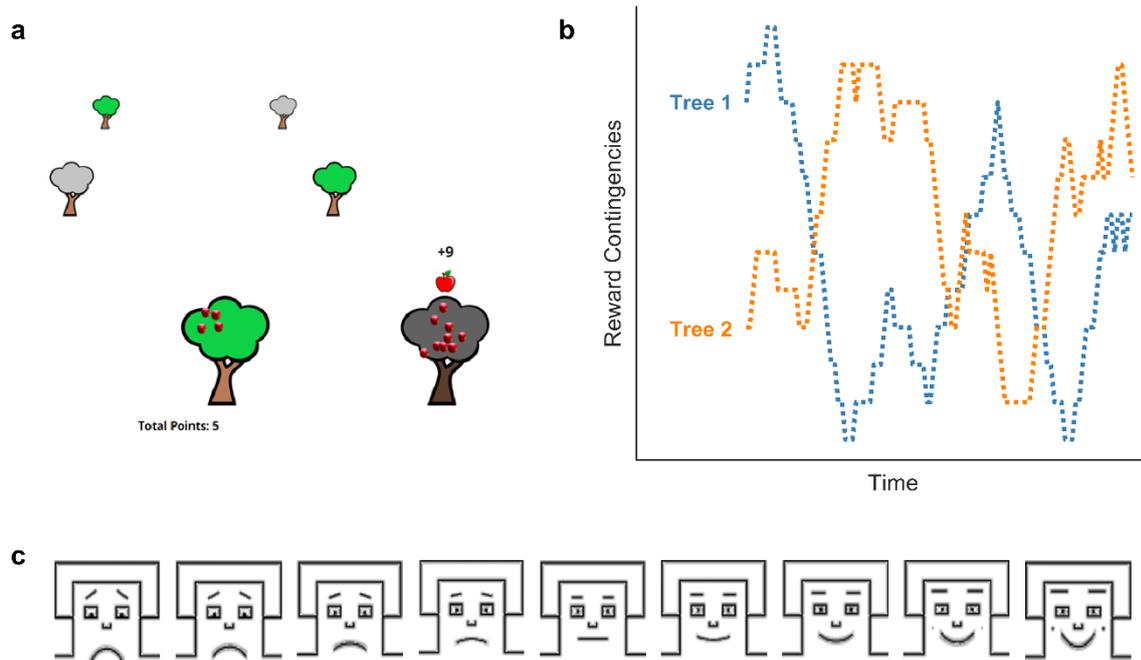
**4.2 Methods; Recruitment and Study Design:** 664 Participants were recruited through the online platform Prolific (prolific.co), of whom 339 completed the study. I asked participants to join the study only if they had no psychiatric history, were aged  $\geq 18$  yrs and spoke fluent English. Upon clicking the study link, participants were directed to MetricWire (metricwire.com), a digital health research platform, where they were instructed on how to download and install the MetricWire app onto their smartphone. On the app, participants completed the consent process and answered a number of initial questionnaires. These were the Centre for Epidemiological Studies-Depression (Radloff, 1977), the Hypomanic Personality Scale (Eckblad and Chapman, 1986) the State and Trait Anxiety Inventory (Spielberger et al., 1983) and the Temporal Experience of Pleasure Scale (Gard et al., 2006). Starting at 9am on the following day, participants received 6 Positive and Negative Affect Scale (PANAS) questionnaires (Watson et al., 1988) per day for 20 days (120 questionnaires in total) through the MetricWire app. In the PANAS, participants are asked to rate the extent to which 20 adjectives matched their current affective state on a 5-point Likert scale (from 'very slightly, or not at all' to 'extremely'). Ten of these adjectives had positive valence (e.g. 'enthusiastic') and 10 had negative valence (e.g. afraid). Participants had a 2hr window in which to complete each of the 6 daily PANAS questionnaires (9-11 am,.....7-9pm). Within each time-window, participants received phone-notifications every 20 minutes until they completed the current questionnaire. Participants were required to complete at least 90 PANAS questionnaires in order to complete the study. Participants were additionally directed from Prolific to the Gorilla online task platform where they completed a reward learning task on two separate days within this 20-day period.



**Figure 4** **a)** 664 Participants were recruited through the online platform, Prolific. **b)** Participants used the MetricWire app to answer a number of initial questionnaires and then **c)** 90-120 PANAS questionnaires, which they received 6 times a day, for 20 days. **d)** Participants used the Gorilla online task platform to complete **e)** a reward processing task on two separate days within the 20-day period. This task is described below. **f)** 339 Participants completed the study.

**4.3 Methods; Online Reward Processing Task:** Participants completed the online reward processing task (illustrated in figure 5) on two separate days. They completed two runs of the task on each day. In this chapter I will restrict my analysis to only the first run of each day. I will analyse the second run in the following chapter. Each run of the task consisted of 120 trials. In each trial, participants chose one of two coloured trees, each of which contained a number of (visible) apples. The number of apples on each tree differed between 0 and 10 from trial to trial (and between the two trees in the vast majority of trials). The two trees could appear in two of three positions (left, middle, right). Each tree (identifiable by its colour) was associated with a probability of ‘winning’ apples. These probabilities were unknown to the participants and varied across the run. Participants were instructed to learn these changing underlying probabilities in order to maximise the number of apples they won. Participants chose one tree per trial by using either the mouse or an arrow-key. A ‘win’ outcome triggered a cash register sound and the appearance of a number, signifying the number of apples won (within-trial). A ‘no-win’ outcome triggered an ‘error’ sound and the appearance of a cross over the chosen tree. A ‘running total’ was kept at the bottom of the screen and updated by 0.25 points for each apple won. Stimulus-outcome contingencies (illustrated in Figure 5b) varied over time, jumping reciprocally  $\pm 0.5$  on three occasions and, additionally,  $\pm 0.15$  non-reciprocally every 19-29 trials; underlying contingencies remained between 0.25-0.75 throughout the task; stimulus contingencies were anti-correlated. Participants were required to

achieve a minimum score on a demo run of the task to ensure that they understood the task before proceeding to the actual task. After every third trial of the task, participants answered the question “how are you feeling right now?” by selecting one of 9 faces (extremely unhappy to extremely happy) from the self-assessment mannequin (Bradley and Lang, 1994).



**Figure 5:** a) In each trial, participants chose one of two coloured trees to try and win apples. b) Each tree was associated with a probability of winning apples, which were unknown to participants, varied within-run and were anti-correlated (as described in the main text) c) After every third trial of the task, participants answered the question “how are you feeling right now?” by selecting a face from the self-assessment mannequin (Bradley and Lang, 1994).

**4.4 Methods; Reimbursement:** Participants received £20 reimbursement if they completed the study. That is, if they answered  $\geq 90$  PANAS questionnaires and completed the online task on both days, plus an additional £5 if they answered  $\geq 110$  PANAS questionnaires. Participants did not receive monetary reimbursement for the points won in the task.

**4.5 Methods; Bayesian Filter:** I utilized a previously described Bayesian filter (Pulcu et al., 2022) to estimate the causal process that generated reported affect. The code for this filter can be found at <https://osf.io/j7md3/>

The filter assumes that reported affect  $y$ , at time-point  $t$ , is generated from a Gaussian distribution with mean  $\mu_t$  and standard deviation  $\exp(SD_t)$ .

$$y_t \sim \mathcal{N}(\mu_t, \exp(SD_t))$$

Differences between current and prior affect might be due to the width of this distribution (affective noise;  $SD$ ) or due to a change in the mean ( $\mu$ ) of the distribution.  $\mu_{t+1}$  is sampled from a Gaussian distribution with mean  $\mu_t$  and standard deviation  $\exp(\nu\mu_t)$ . Therefore, the extent to which the mean changes over time is determined by the volatility parameter,  $\nu\mu_t$ .

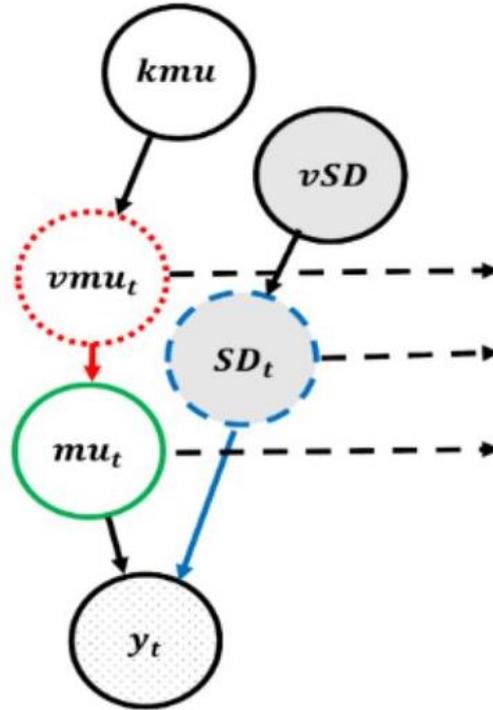
$$p(\mu_{t+1}) \sim \mathcal{N}(\mu_t, \exp(\nu\mu_t))$$

$\nu\mu$  and  $SD$  can themselves change over time, in the same manner as  $\mu$ .  $\nu\mu_{t+1}$  is itself sampled from a Gaussian with mean  $\nu\mu_t$  and standard deviation  $\exp(k\nu\mu)$ .

$$p(\nu\mu_{t+1}) \sim \mathcal{N}(\nu\mu_t, \exp(k\nu\mu))$$

$SD_{t+1}$  is also sampled from a Gaussian, with mean  $SD_t$  and standard deviation  $\exp(\nu SD)$ .

$$p(SD_{t+1}) \sim \mathcal{N}(SD_t, \exp(\nu SD))$$



**Figure 6. An illustration of the generative model that produces affect ratings ( $y_t$ ).** *Reproduced from Pulcu et al (Pulcu et al., 2022).* Affect ratings  $y$  at time-point  $t$  are sampled from a gaussian distribution with mean(/median/mode)  $\mu_t$  and standard deviation  $SD_t$ .  $\mu_t$  is itself sampled from a gaussian distribution with standard deviation  $vmu_t$ , which, in turn, is sampled from a gaussian distribution with standard deviation  $kmu$ .  $SD_t$  is sampled from a gaussian distribution with standard deviation  $vSD$ .

The filter estimates the joint probability of the five parameters  $\mu$ ,  $SD$ ,  $vmu$ ,  $vSD$ ,  $kmu$ , at time point  $t$ : given the affect ratings up to that time-point ( $y_1, y_2, \dots, y_{t-1}$ ). The filter assumes that all affect ratings from a given time-point onwards are completely described by the joint distribution of parameters at that time-point. This is equivalent to assuming that the joint probability of parameters at a given time-point,  $p(joint_t)$ , depends only on the joint probability of parameters and observed affect rating at the previous time-point; respectively  $p(joint_{t-1})$  and  $y_{t-1}$ .

$$p(joint_t) = p(\mu_t, vmu_t, kmu, SD_t, vSD | y_{t-1}, joint_{t-1})$$

$p(joint)$  is initialised as a flat distribution and updated between each affective observation. At each time-point the model uses the affect score at that time-point to update its belief about the joint distribution of parameters. Once it receives a data point, it updates as follows.

The model first updates the joint probability distribution of parameters, given the observation at time-point  $t - 1$  (using Bayes' rule):

$$p(\text{joint}_{t-1} | y_{t-1}) = \frac{p(y_{t-1} | \text{joint}_{t-1}) \cdot p(\text{joint}_{t-1})}{p(y_{t-1})}$$

The likelihood of the observation at  $t - 1$  given the joint distribution at  $t - 1$  i.e.  $p(y_{t-1} | \text{joint}_{t-1})$  is a function only of  $mu_t$  and  $SD_t$  and so the partially updated probability of the joint distribution given the observation i.e.  $p(\text{joint}_{t-1} | y_{t-1})$  assumes that  $vmu$ ,  $vSD$  and  $kmu$  remain static, but they don't. The probability distributions of  $mu$ ,  $SD$  and  $vmu$  at time-point  $t$  depend on the observation at  $t - 1$ , the probability distribution of the respective parameter after time  $t - 1$  and the probability distributions of  $vmu$ ,  $vSD$  and  $kmu$  (respectively) after  $t - 1$ . The next step of the update is therefore to adjust the joint probability distribution to account for these dependencies. For example, the probability distribution of  $mu$  before the observation at time-point  $t$  and after the observation at  $t - 1$  depends on the observation at  $t - 1$  (already accounted for above), the probability distribution of  $mu$  just before the observation at time-point  $t - 1$  and also on the probability distribution of  $vmu$  just before the observation at time-point  $t$ . The probability distribution for  $mu_t$  given the probability distribution for  $vmu_t$  is calculated by integrating over  $mu_{t-1}$ .

$$\int p(mu_t | mu_{t-1}, vmu_t) d(mu_{t-1}) = p(mu_t | vmu_t)$$

The model performs the same operation for parameters  $SD$ , and  $vmu$ , each time 'integrating out' the influence of the prior magnitude-probabilities of the respective parameter on the current magnitude-probabilities of that same parameter, in the same manner as for  $mu$  above.

$$\int p(SD_t | SD_{t-1}, vSD_t) d(SD_{t-1}) = p(SD_t | vSD_t)$$

$$\int p(vmu_t | vmu_{t-1}, kmu_t) d(vmu_{t-1}) = p(vmu_t | kmu_t)$$

The resulting probability distributions are combined with the updates that resulted from the affective observation described above to give the joint probability distribution of parameters, given the most recent observation, accounting for the influence of each higher level parameter on the rate of change of each respective lower level parameter:

$$p(\text{joint}_t | y_{t-1}) = p(\text{joint}_{t-1} | y_t) \cdot p(vmu_t | kmu_t) \cdot p(mu_t | vmu_t) \cdot p(SD_t | vSD_t)$$

This joint distribution is used to calculate the marginal probability distribution for each parameter. In turn, these marginal distributions are used to calculate a point estimate for each parameter, for each trial.

The filter sampled  $\mu$  in logit space and so, before I entered  $\mu$  in any statistical analyses, I first logit transformed the  $\mu$  values. The filter sampled all other parameters in log space and so I log transformed them before I entered them into statistical analyses. The filter calculates a value for each parameter for each time-point. For the purpose of extracting data to enter into statistical analyses, I took  $\mu$  to be the mean  $\mu$  across the entire time-series, I took other parameters to be the mean of the last 5 values of that parameter in the time-series. Regressions and correlations were implemented in MatLab. Though the Bayesian filter uses 5 parameters ( $\mu, SD, \nu\mu, \nu SD, k\mu$ ) I restricted my analysis to only 3 parameters ( $\mu, SD, \nu\mu$ ). This was because 1) My hypotheses do not include  $\nu SD$  or  $k\mu$  and 2) the PANAS time-series' contained at most 120 affective reports and the task time-series contained 41 affective reports per run. With this number of data points,  $\nu SD$  and  $k\mu$  cannot be precisely estimated.

**4.6 Methods; Regressing affective parameters against affective responses to task events:** In the introduction to this chapter I stated that  $SD$  could be thought to reflect the immediate affective impact of events, that  $\nu\mu$  could be thought to reflect the extent to which the affective impact of events persists and that  $\mu$  could be thought to reflect the affective set-point (with the caveat that  $\mu$  can change over time whereas the affective set-point can't), and so is unlikely to have a relationship with the affective impact of events. In order to test this, I calculated the association between within-task mood variability parameters and affective responses to trial outcomes. I first regressed affect-report scores against the preceding 9 trial outcomes, within-subject. The resulting regression coefficients represent the dependency of reported affect on each of the 9 preceding trial outcomes. I then averaged each group of 3 of these regression coefficients (participants reported affect every 3 trials). The resulting 3 'averaged regression coefficients' represent the dependence of each reported affect score on the events that took place between it and the previous affect reports. These averaged regression coefficients' were then used collectively as the independent variables in each of 3 subsequent between-subject regressions in which the dependent variables were, respectively, within-task  $\mu$ ,  $\nu\mu$  and  $SD$  as this should illustrate the relationship between each parameter and the affective consequence of task events.

#### 4.7 Results: Demographics and questionnaire scores

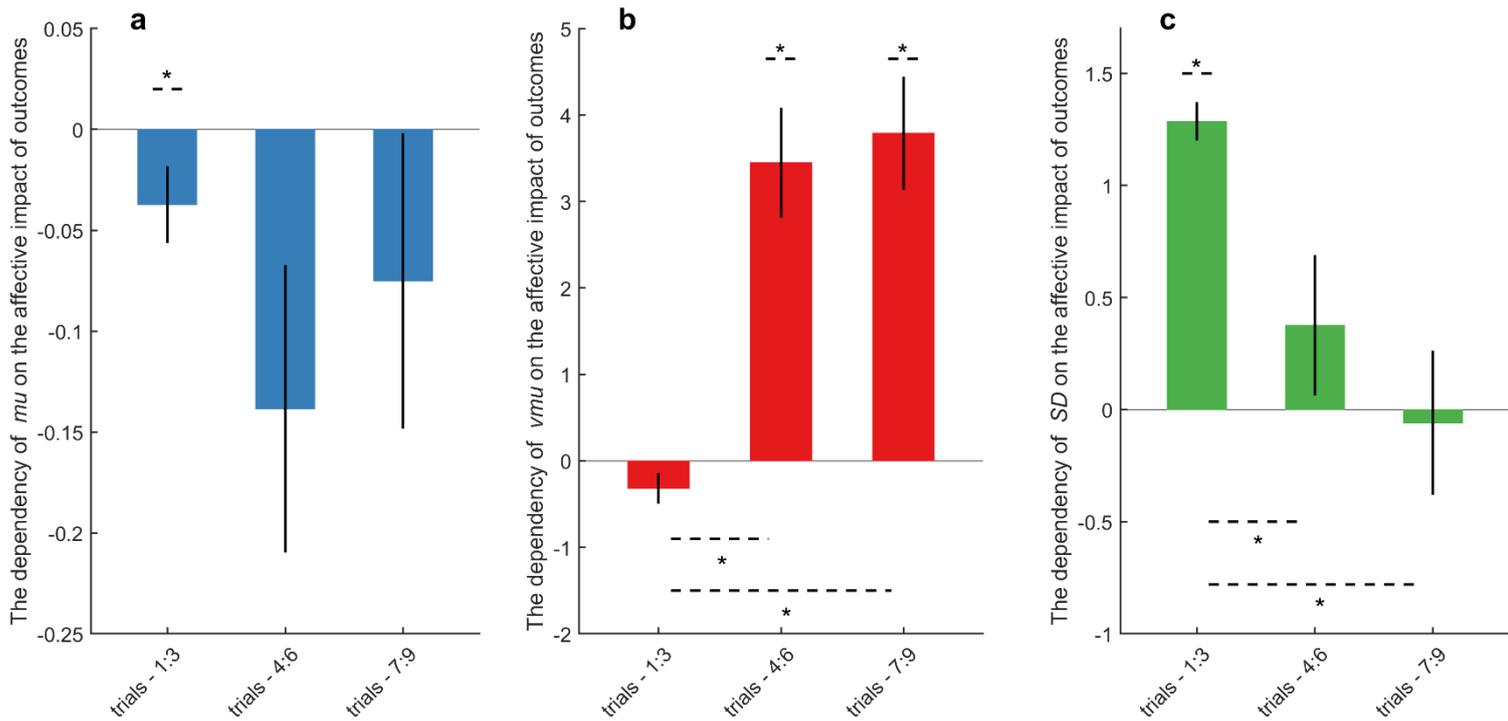
664 participants clicked on the prolific study link, of whom 339 completed the study.

|                       |                                                                     |
|-----------------------|---------------------------------------------------------------------|
| <b>Age</b>            | Not known: 13 ; of the rest (326): 27.9yrs std: 9.216yrs            |
| <b>Sex</b>            | Male: 193 (56.9%); Female: 132 (38.9%); Not known: 14 (4.1%)        |
| <b>First Language</b> | English: 65 (19.2%); Not English: 253 (74.6%); Not known: 21 (6.2%) |
| <b>CESD</b>           | Mean: 16.3; std: 9.294                                              |
| <b>HPS</b>            | Mean: 64.0; std: 11.125                                             |
| <b>STAI (state)</b>   | Mean: 38.1; std: 11.483                                             |
| <b>STAI (trait)</b>   | Mean: 44.5; std: 11.561                                             |
| <b>TEPS</b>           | Mean: 82.8; std: 10.614                                             |

**Table 1.** Participant demographics and initial questionnaire scores.

#### 4.8 Results: The relationships between task affective parameters and affective responses to task events

I regressed affective parameters against affective responses to task events (as described in section 4.6). I found that task  $SD$  is, as expected, associated with the immediate affective impact of task events and that this association decreases monotonically from more recent to more distant events. This is consistent with the idea that  $SD$  reflects the magnitude of the initial affective perturbation. I found the opposite (and, again, expected) pattern for  $vmu$ :  $vmu$  is not associated with the immediate affective impact of task events and is, instead, associated with the extent to which the affective impact of task events persist. This association increases monotonically from more recent to more distant events. This is consistent with the idea that  $vmu$  reflects the rate at which affect returns to baseline following an initial perturbation. I found that task  $mu$  is negatively associated with the immediate response to task outcomes. This outcome is unexpected and it is unclear what it signifies. Perhaps that a high  $mu$  is associated with non-reactivity to external events. However, this result reaches significance by a hair's breadth ( $p=0.0436$ ) and so should be interpreted cautiously. This latter result is not consistent with the conceptual ('exponential decay') model, which suggests that  $mu$  should be independent of affective responses to events. Expectedly, there is no obvious relationship between  $mu$  and affective responses of different durations, as there was for  $SD$  and  $vmu$ .



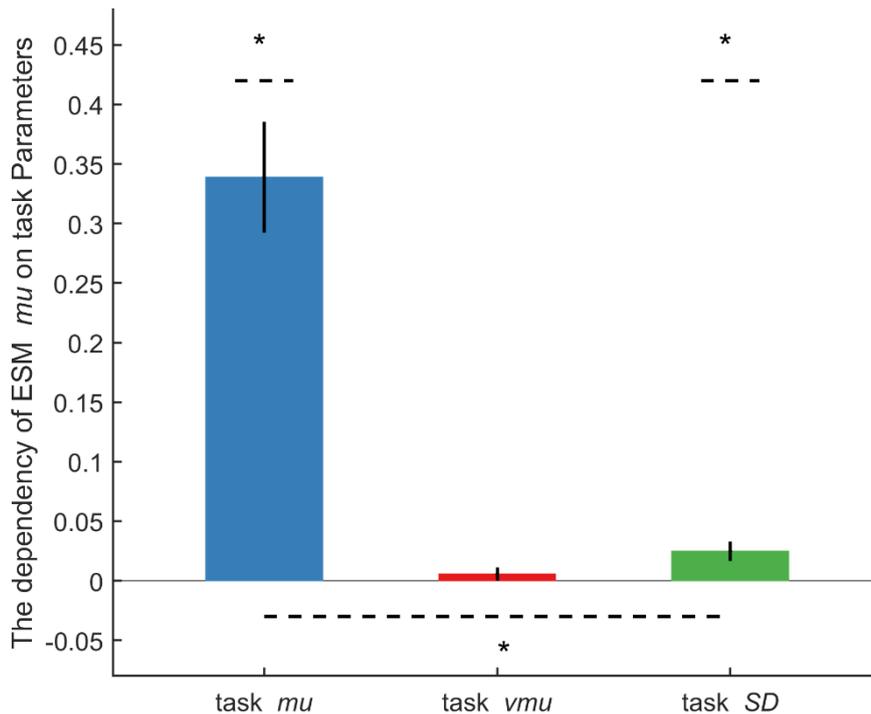
**Figure 7. The relationship between within-task affective parameters and affective responses to within-task events.** Bars represent regression coefficients for the dependency of task **a)**  $\mu$  (blue bars), **b)**  $vmu$  (red bars) and **c)**  $SD$  (green bars) on the affective impact of trial outcomes. In each panel, the three bars represent (from left to right) the dependency of the relevant parameter on the mean affective impact of sequentially preceding groups of 3 trials (the left-most bar/regression coefficient: trials -1:3; the middle bar: trials -4:6; the right-most bar: 'trials -7:9'). Bars are regression coefficients and error bars are the standard error of the respective regression coefficient. Dashed lines and asterisks *above* bars signify that the regression coefficient is significantly different from zero (as measured by a t-test). Dashed lines and asterisks *below* bars signify that the respective coefficients are significantly different to each other (as measured by non-overlapping 95% confidence intervals).

**Conclusion/Summary:** Task  $SD$  is associated with the immediate impact of task events, which is consistent with the idea that it reflects the magnitude of initial affective deflection from baseline. Task  $vmu$  is associated with the persistent impact of task events, which is consistent with the idea that it reflects the rate at which affect returns to baseline. Unexpectedly, task  $\mu$  is negatively associated with the immediate impact of task events, it is unclear what this signifies.

#### 4.9 The relationship between corresponding task-based and ESM affective parameters

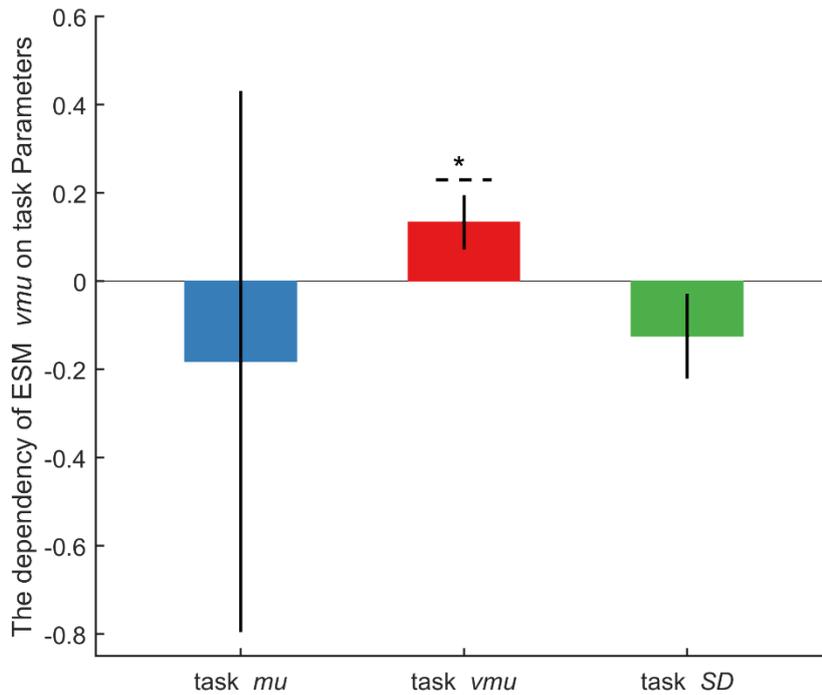
My primary hypothesis is that Bayesian affective parameters within a short timeframe may reflect corresponding parameters on a longer timeframe (which, if true, would contribute to the ecological validity of the task based measures of affective variability using the Bayesian filter). I tested this idea by regressing ESM parameters against task parameters. If my hypothesis is correct, within-task parameters should specifically reflect corresponding parameters in the ESM dataset (i.e. ESM  $\mu$  should depend specifically on task  $\mu$ ; ESM  $\nu\mu$  should depend specifically on task  $\nu\mu$ ; ESM  $SD$  should depend specifically on task  $SD$ ).

**$\mu$** : I averaged  $\mu$ ,  $\nu\mu$  and  $SD$  each across (the first runs of) the two task sessions. I then regressed ESM  $\mu$  against task  $\mu$ ,  $\nu\mu$  and  $SD$ . I found that ESM  $\mu$  was associated with task  $\mu$  and task  $SD$ , but that the regression coefficient for task  $\mu$  was significantly higher than for task  $SD$ , as illustrated in Figure 8. This is not exactly as predicted (as ESM  $\mu$  was associated with task  $SD$ ) but the results suggest that **ESM  $\mu$  is most associated with task  $\mu$** . This is consistent with my hypothesis.



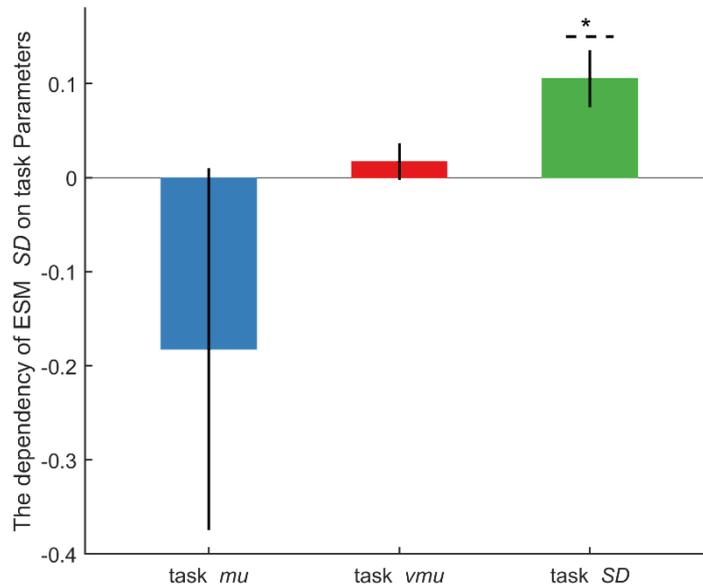
**Figure 8. Dependency of ESM  $\mu$  on task parameters.** Bars are regression coefficients from regressing ESM  $\mu$  against task  $\mu$ ,  $\nu\mu$  and  $SD$ . ESM  $\mu$  is dependent on both task  $\mu$  and task  $SD$ . Error bars represent the standard error of the respective regression coefficient. Dashed lines and asterisks above bars signify that the co-efficient is significantly different from zero (as measured by a t-test). Dashed lines and asterisks below bars signify that the respective coefficients are significantly different to each other: The lower 95% CI for the task  $\mu$  regression coefficient is 0.2497 while the upper 95% CI for the  $SD$  regression coefficient is 0.0384 so that the regression coefficient for task  $\mu$  is significantly larger than for task  $SD$ .

**$\nu\mu$ :** I averaged  $\mu$ ,  $\nu\mu$  and  $SD$  each across the two task sessions. I then regressed ESM  $\nu\mu$  against task  $\mu$ ,  $\nu\mu$  and  $SD$  as illustrated in Figure 9. **I found that ESM  $\nu\mu$  was specifically associated with task  $\nu\mu$ .** This is consistent with my hypothesis.



**Figure 9. Dependency of ESM  $vmu$  on task parameters.** Bars are regression coefficients from regressing ESM  $vmu$  against task  $\mu$ ,  $vmu$  and  $SD$ . ESM  $vmu$  is dependent specifically on task  $vmu$ . Error bars represent the standard error of the respective regression coefficient. Dashed lines and asterisks above bars represent that the co-efficient is significantly different from zero (as measured by a t-test).

**$SD$ :** I averaged  $\mu$ ,  $vmu$  and  $SD$  each across the two task sessions. I then regressed ESM  $SD$  against task  $\mu$ ,  $vmu$  and  $SD$  as illustrated in Figure 10. **I found that ESM  $SD$  was specifically associated with task  $SD$ .** This is consistent with my hypothesis.



**Figure 10. Dependency of ESM *SD* on task parameters.** Bars are regression coefficients from regressing ESM *SD* against task *mu*, *vmu* and *SD*. ESM *SD* is dependent specifically on task *SD*. Error bars represent the standard error of the respective regression coefficient. Dashed lines and asterisks above bars represent that the co-efficient is significantly different from zero (as measured by a t-test).

**Conclusion/Summary:** Task affect variables are specifically associated with their ESM counterparts when controlling for other task variables. The only exception is ESM *mu* which was additionally associated with task *SD*.

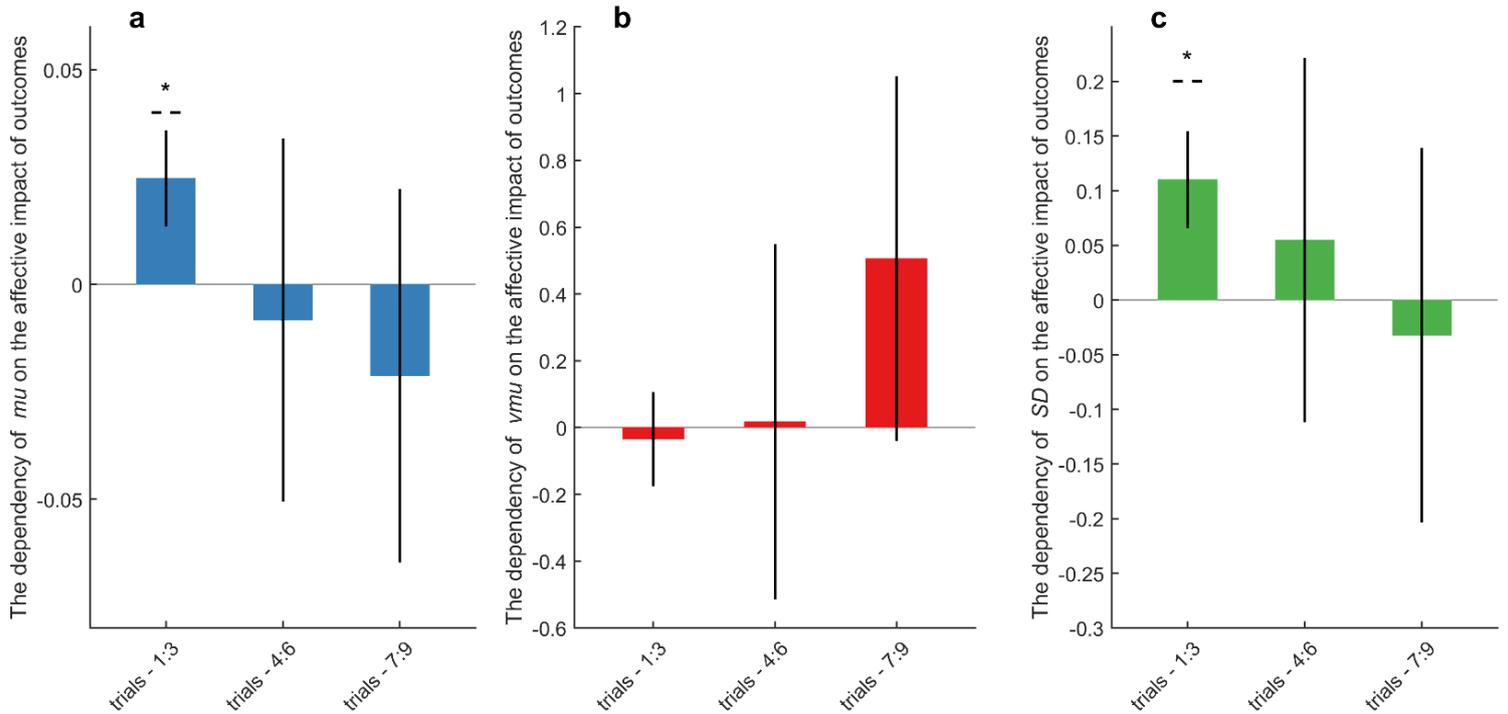
#### 4.10 The relationships between ESM affective parameters and affective responses to task events

Having found a (mostly) specific association between task and ESM affective parameters I next examined the direct relationship between ESM affective parameters and affective responses to trial outcomes. In order to do so I repeated the analyses at the start of the results section, but replacing within-task affective parameters with equivalent ESM parameters. I hypothesized that the results would follow the same basic pattern as for the within-task affective parameters:

- ESM *SD* would be associated with the immediate affective impact of task events and this association would decrease monotonically from more recent to more distant events.

- ESM *vmu* would be associated with the persistent affective impact of task events and this association would increase monotonically from more recent to more distant events.

-ESM  $\mu$  would not show any clear pattern of association with either the immediate or persistent impact of task events.



**Figure 11. The relationship between ESM affective parameters and affective responses to within-task events.** Bars represent regression coefficients for the dependency of ESM **a)**  $\mu$  (blue bars), **b)**  $\nu\mu$  (red bars) and **c)**  $SD$  (green bars) on the affective impact of trial outcomes. In each panel, the three bars represent (from left to right) the dependency of the relevant parameter on the mean affective impact of sequentially preceding groups of 3 trials (the left-most bar: trials -1:3; the middle bar: trials -4:6; the right-most bar: 'trials -7:9'). Bars are regression coefficients and error bars are the standard error of the respective regression coefficient. Dashed lines and asterisks signify that the regression co-efficient is significantly different from zero (as measured by a t-test).

I found that the pattern of results for PANAS  $SD$  is similar to the pattern of results for task  $SD$ . First, PANAS  $SD$  is significantly (and specifically) associated with the affective impact of recent trial events. Second, the association of PANAS  $SD$  on task events decreases (numerically) monotonically from more to less recent task events. Unlike with task  $SD$ , this decrease is not statistically significant.

I found no significant association of PANAS  $\nu\mu$  with the affective impact of task events.

Interestingly, the regression coefficients increase monotonically (numerically) from more to less

recent events, which may **hint** that PANAS *vmu* may be associated with the affective response to more distant events (i.e. it reflects the persistence of affective responses to task events).

The pattern of results for PANAS *mu* is strikingly similar to that of PANAS *SD* (and task *SD*). This was unexpected and may signify that mean affect may be positively associated with emotional reactivity. This is the opposite of what the within-task results suggested. Scatter plots of ESM *mu*, *SD* and *vmu* versus each of task *mu*, *SD* and *vmu* (as well as accompanying correlation coefficients) are shown in appendix 14.

**Conclusion/Summary:** As with task *SD*, ESM *SD* is associated with the immediate impact of task events, which is consistent with the idea that *SD* in everyday life reflects the magnitude of initial affective deflections from baseline. ESM *vmu* does not bear any significant relationship to the duration of responses to task events. ESM *mu* shows the same pattern of associations as ESM *SD*.

#### 4.11 Discussion

In this chapter I first tested the association between Bayesian affective parameters and affective responses to task events. Affective noise (*SD*) was associated with the immediate affective impact of events. Affective volatility (*vmu*) was associated with the persistence of affective responses. These findings were consistent with the idea that *SD* reflects an initial affective deflection from one's baseline and that *vmu* reflects the rate at which affect returns to baseline. Affective *mu* was (just about) significantly negatively associated with the immediate impact of task events. This finding was *not* consistent with *mu* reflecting an affective set-point. I then attempted to link within-task affective parameters to corresponding affective parameters in everyday life. Consistent with my hypothesis, ESM ('real-life') *mu*, *vmu* and *SD* were respectively associated with task *mu*, *vmu* and *SD*, even controlling for the other within-task parameters. Finally I tested the direct relationship between ESM parameters and within-task affective responses. I found that affective noise (*SD*) was associated with the immediate affective impact of events, which was consistent with my hypotheses. I found no significant associations between affective volatility (*vmu*) and the persistence of affective responses. ESM affective mean (*mu*) was associated with the immediate impact of task events. These two findings are not consistent with my hypotheses.

The core finding in this chapter is that each task parameter is associated with its corresponding ESM parameter. This is the case even controlling for other task parameters. In the cases of affective noise (*SD*) and volatility (*vmu*), this is a *specific* association. Though ESM mean affect (*mu*) is additionally associated with task noise (*SD*), the regression coefficient for task mean affect (*mu*) is significantly

larger than for noise ( $SD$ ). Building upon previous work in the area (Koval et al., 2013), I have drawn a link between within task and corresponding ESM affective variability parameters. ESM does not currently have a strong biological framework whereas the biology of reward processing is comparatively better characterized. Linking the two may facilitate translational channels between reward learning and mood instability research. There are currently two strands of reward-processing based explorations of affect (Eldar et al., 2016; Eldar and Niv, 2015; Rutledge et al., 2014). Both lines of research have investigated the relationship between striatal activity and reported (within-task) affective variability (as well as summary reports of 'real-world' affective instability in the case of Eldar et al (Eldar and Niv, 2015)). Rutledge et al find that ventral striatal BOLD responses are significantly correlated with trial-by-trial changes in several internal model metrics (Rutledge et al., 2014). Eldar finds that the magnitude of striatal responses are more influenced by mood in those with high HPS scores (which are indicative of affective instability) (Eldar and Niv, 2015). The Rutledge model of affect (Rutledge et al., 2014) is essentially a sum of exponentially decaying affective perturbations with separate parameters determining the magnitude of initial deflection and the subsequent rate of decay. As this model is already known to reflect striatal activity (Rutledge et al., 2014), one could e.g. explore the effect of the rate of decay on striatal activity, and use the Bayesian filter described here and in Pulcu et al (Pulcu et al., 2022) to relate this to participants' ESM  $\nu mu$ .

One interpretation of ESM affective noise ( $SD$ ) is that it represents the immediate and transient affective impact of events while ESM volatility ( $\nu mu$ ) represents the persistent affective impact of events. We cannot be sure of this from the ESM data. I assessed whether (within-task) affective noise ( $SD$ ) and volatility ( $\nu mu$ ) are differentially dependant on the duration of affective responses to events.  $SD$  is dependent on the immediate affective impact of task events, and not on the more persistent affective impact of task events (decreasing monotonically as trials recede into the past). Also as hypothesised,  $\nu mu$  is not dependent on the immediate affective impact of task events, and instead on the more persistent affective impact of task events (increasing monotonically as trials recede into the past). These results are consistent with the interpretation that ESM affective noise ( $SD$ ) represents the immediate/transient, and ESM volatility ( $\nu mu$ ) the persistent affective impact of events. They are also loosely consistent with the idea that affect can be described as a sum of exponential affective decays with  $SD$  reflecting initial affective deflections and  $\nu mu$  reflecting the rate of return to affective baseline. Not as hypothesized, within-task  $\mu$  is (just about) negatively associated with the immediate impact of task events.

I also assessed the association between ESM parameters and the duration of within-task affective responses. As hypothesised, ESM  $SD$  is directly associated with the immediate affective impact of task events, and not on the more persistent affective impact of task events. Affective noise ( $SD$ ) is

somewhat analogous to affective instability and loosely the opposite of emotional inertia, both of which are relevant to depression (Eldesouky et al., 2018; Kuppens et al., 2012; Sperry et al., 2020). The current finding suggests that the underpinnings of ESM affective noise ( $SD$ ) could be probed by reward-learning tasks. I did not find a significant association between ESM affective volatility ( $\nu mu$ ) and the persistence of affective responses. Unexpectedly, ESM  $mu$  showed the same pattern as ESM  $SD$ : ESM  $mu$  is directly associated with the immediate affective impact of task events, and not on the more persistent affective impact of task events. The associations decrease monotonically as task events recede into the past. This finding is surprising, difficult to explain and warrants exploration.

#### **4.12 Limitations**

This chapter aimed to relate reward-processing task based measures to ESM mood dynamics. The underlying motivation was to ground depression-related mood dynamics (as measured using ESM) in the relatively well established field of reward processing, which has a better characterized neurobiological basis. In order to do so, I used a relatively new model of ESM affective variability (Pulcu et al., 2022). While this model has shown some utility in characterizing mood dynamics in BPAD and EUPD, it remains untested for unipolar depression. Therefore, it is unclear how relevant my results are for depression research specifically, particularly as I did not recruit a depressed population.

Previous studies find utility in decomposing affect. For example, compared to healthy adults, depressed adults appear to have greater instability of specifically *negative* affect following specifically *positive* events (Thompson et al., 2012). I have treated ‘affect’ as a unidimensional measure. In doing so I have possibly brushed over useful and interesting aspects of the data. For example, Leemput et al found that episodes of depression were preceded by increased mutual reinforcement of within-valence emotions and increased mutual suppression of between-valence emotions (van de Leemput et al., 2014). Future studies that assess the underpinnings of affective variability using the PANAS may wish to examine the dynamics of the various subcomponents within the PANAS score.

#### **4.13 Conclusion**

Affective parameters can differentially capture the immediate vs persistent affective impact of task events in a reward learning task. These parameters are specifically associated with corresponding ESM parameters from everyday life. This suggests that combining reward learning and ESM using

Bayesian affect parameters provides a plausible method for studying the reward processing underpinnings of affective variability.

## Chapter 5: Does reinforcement learning underlie affect instability

### 5.1 Introduction

As outlined in the thesis introduction, there has been recent interest in the connection between affect and RL. Several studies have demonstrated a link between RL measures of learning and moment-to-moment changes in reported affect (Blain and Rutledge, 2022; Chew et al., 2021; Keren et al., 2022; Rutledge et al., 2014). In these studies participants complete reward processing tasks, during which time they additionally report affect using a visual analogue scale. Following behavioural RL modelling, model metrics (such as RPEs) are inputted to separate models that generate a series of affect scores, which are shown to correlate with participant-reported affect scores (Rutledge et al., 2014). Given that affect appears to be linked to RL, it seems plausible that affective instability can be expressed in terms of RL. Eldar and Niv have done so (Eldar and Niv, 2015), proposing that affective instability arises from a feedback loop between affect and learning such that affect modulates the perception of outcomes, which in-turn modulate affect, and so on (Eldar and Niv, 2015). Eldar and Niv found that affective instability in everyday life (as measured using the hypomanic personality inventory) was associated with a within-task interaction between affect and learning. This is an exciting result. However, this is a nascent line of research and many questions remain unanswered. Among these:

#### 1. Testing the affect-generating mechanism

-The model mechanism has so far been tested at a relatively coarse resolution (replicating one affect report every 17 trials). Its **ability to accurately replicate affect at a higher resolution is not yet known** (for example once every 3 trials, as is the case for another prominent RL model of affect (Rutledge et al., 2014)).

-The model posits that affect is a (modified) weighted average of RPEs. In the original study, this mechanism replicated participant affect, but the to-be-explained participant affect scores were actually included with the choice data to estimate the model parameters. **It is not yet known if the posited mechanism can replicate participant affect using behavioural choice data alone** (i.e. without including the affect scores in the fitting process).

- The model has been successfully tested in one task, however its **generalisability to other contexts (e.g. other tasks) is not yet known**.

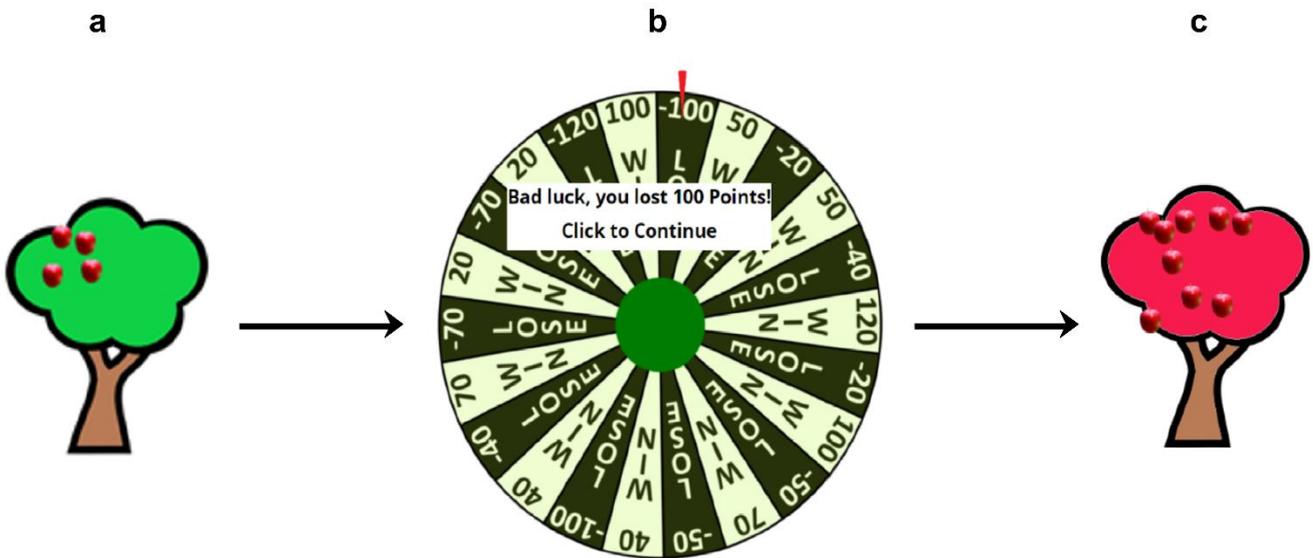
## 2. Replicating the original key findings

-The key findings in the original study were that affective instability in everyday life (as measured by the HPS) was associated with a within-task interaction between affect and learning, and that the extent of affective instability was commensurate with the extent of this interaction. **The association between within-task affect-learning interaction and affective instability in everyday life has not yet been replicated.**

*In this Chapter I attempt to address these questions. I first test Eldar and Niv's basic affect-generating mechanism on a separate reward learning task, using more frequent affect sampling and without fitting to affect report data. I then attempt to replicate the original key finding (the association between affect-learning interaction within-task and affective instability in everyday life).*

## 5.2 Methods and Results

Recruitment and demographics are as described in Chapter 4. The task used in this chapter is also the same as that described in chapter 4, with some additions:



**Figure 1. Task Description.** The task description and analysis in chapter 4 restricted itself to the first task-run of each day. However, immediately following the (a) initial run, participants completed (b) a wheel of fortune (WoF) task as affect induction. The WoF contained 22 balanced and ostensibly equally probable outcomes (up to  $\pm 120$  points). Unknown to participants the WoF outcomes were deterministic: on one day, the participant won 100 point (that is, 100 points were added to participants' apple-task score); this was intended as a positive affect manipulation. On the other day, the participant lost 100 point (that is, 100 points were subtracted from participants apple-task scores); this was intended as a negative affect manipulation. This was counterbalanced across participants. Following this, participants completed a (c) second run of the reward processing task. This second run was implemented in the same manner as the first run. Each run contained a unique pair of tree colours within-participant.

This task is different from the original task on which the model was tested (Eldar and Niv, 2015). Two differences are: 1) the original task had fixed stimulus-outcome contingencies while the present task has wandering contingencies (as described in chapter 4, Figure 5). 2) The original task had fixed reward magnitudes, while in the present task, reward magnitudes vary between trial and between stimuli intra-trial. Reward magnitudes here should not affect RPEs as reward magnitudes vary from

trial to trial and are explicitly shown (and so are of no value in updating the expected value of selecting a given stimulus). However they are likely to affect choice. Given this, I tested 3 variants of the original model, altering the choice layer variously to take/not take account of the varying reward magnitudes.

**Model Descriptions:** The learning rate ( $\alpha$ ) updates the expected value for each shape ( $Q$ ), after each trial ( $t$ ) using the reward prediction error ( $\delta$ ) for that trial, which is dependent on the perceived value of the trial outcome ( $R_{percieved}$ ).

$$Q_{t+1} = Q_t + \alpha(\delta_t)$$

$$\delta_t = R_{percieved(t)} - Q_t$$

( $R_{percieved}$ ) is the trial outcome ( $R$ ), as modified by the participant's affect ( $m$ ). The influence of affect on perception is mediated by a 'affect-learning interaction' parameter ( $f$ ) so that if  $f > 1$  the participant experiences a positive feedback loop between affect and positive outcomes, while  $1 > f > 0$  leads to negative feedback between affect and positive outcomes (and  $f = 1$  results in no interaction).

$$R_{percieved(t)} = R_t \cdot f^{m_t}$$

Affect ( $m$ ) is determined by the history of reward prediction errors ( $h$ ), modified to constrain  $m$  to  $0 \leq m \leq 1$

$$m_t = \tanh(h_t)$$

$h$  is updated in the same manner as the  $Q$  values (above) but using a different update parameter ( $\eta$ ).

$$h_{t+1} = h_t + \eta(\delta_t)$$

$Q$  values (i.e. model beliefs regarding the value of stimuli) are then entered into a softmax function to generate trial-wise probabilities of selecting stimulus  $s$ . Here  $\beta$  is an inverse temperature parameter and  $\rho_t$  is the number of points/apples associated with a given stimulus on a given trial ( $\rho$  is therefore not a free parameter).

$$P(s) = \frac{1}{1 + e^{-\beta(\rho_{t(s)} \cdot Q_{t(s)} - \rho_{t(s')} \cdot Q_{t(s')})}}$$

I compared the model as originally described to two variants which attempt to account for the variability in available reward.

**Model 1a** is the model described above with  $\rho$  set to 1. This is the model as originally described (Eldar and Niv, 2015).

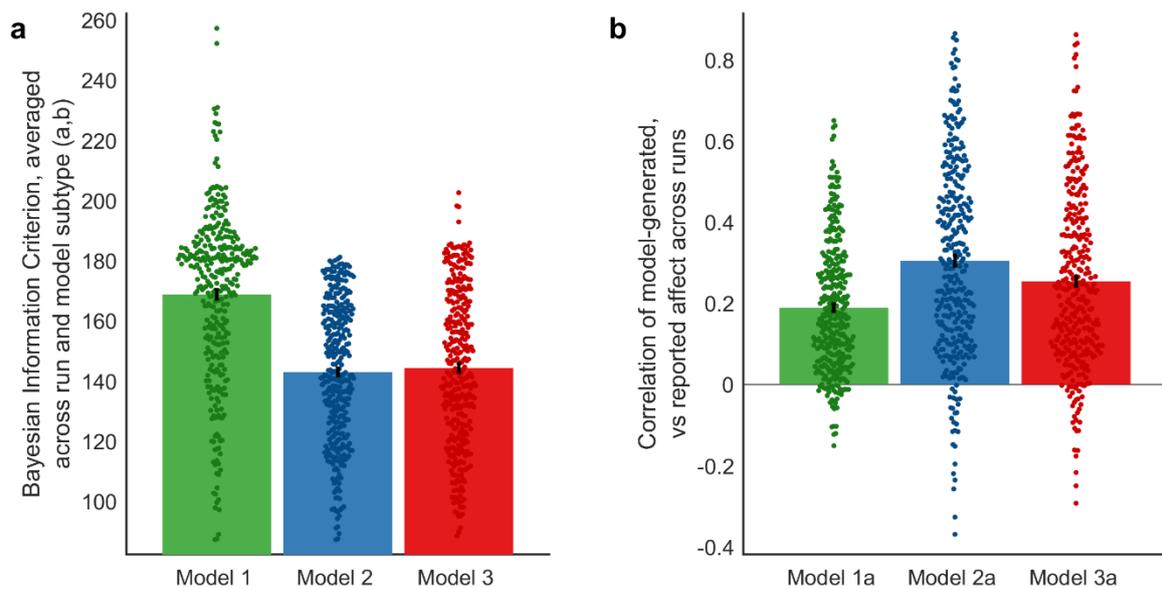
**Model 2a** is the model described above with  $\beta$  set to 1.

**Model 3a** is the full model as described above.

These three models each assume a bidirectional dependence of affect and learning. In order to test versions of each model in which affect depends on learning but not vice versa, I additionally fitted each model with  $f$  set to 1: respectively **model 1b, 2b** and **3b**. These 3 models (i.e. the 'b' models) cannot test the affect generating mechanism as there is no principled way to constrain the affect-update parameter ( $\eta$ ) without fitting the model to affect data.

**Model Fitting:** I calculated the full joint posterior probability of parameters for each model, for each participant. Model choice probabilities were combined with participant choices to generate trial-by-trial likelihoods for the latter. The likelihoods were logged, summed across trials, exponentiated and normalized to generate posterior probabilities for each tested parameter combination, (given that the participant is using the model in question). These probabilities were marginalized and the resulting probability distributions were weighted by the corresponding values of the parameter of interest and summed to calculate the expected value of the parameter in question (again, given that the participant used the model in question). I fitted models only to behaviour (and not also affect).

**Model Comparison:** I used 2 criteria to compare the models: **1)** Model fit to behavioural data (using the BIC), averaged across runs and model subtypes i.e. the (no/)affect-interaction version. **2)** In the case of the affect-interaction subtypes, it was possible to additionally determine the affect generated by each model. I correlated model-generated affect to participant-reported affect across all runs (i.e. taking into account inter as well as intra-run changes in affect). It was not possible to produce model-generated affect for the no-affect-interaction version of each model without fitting model affect to participant affect reports.

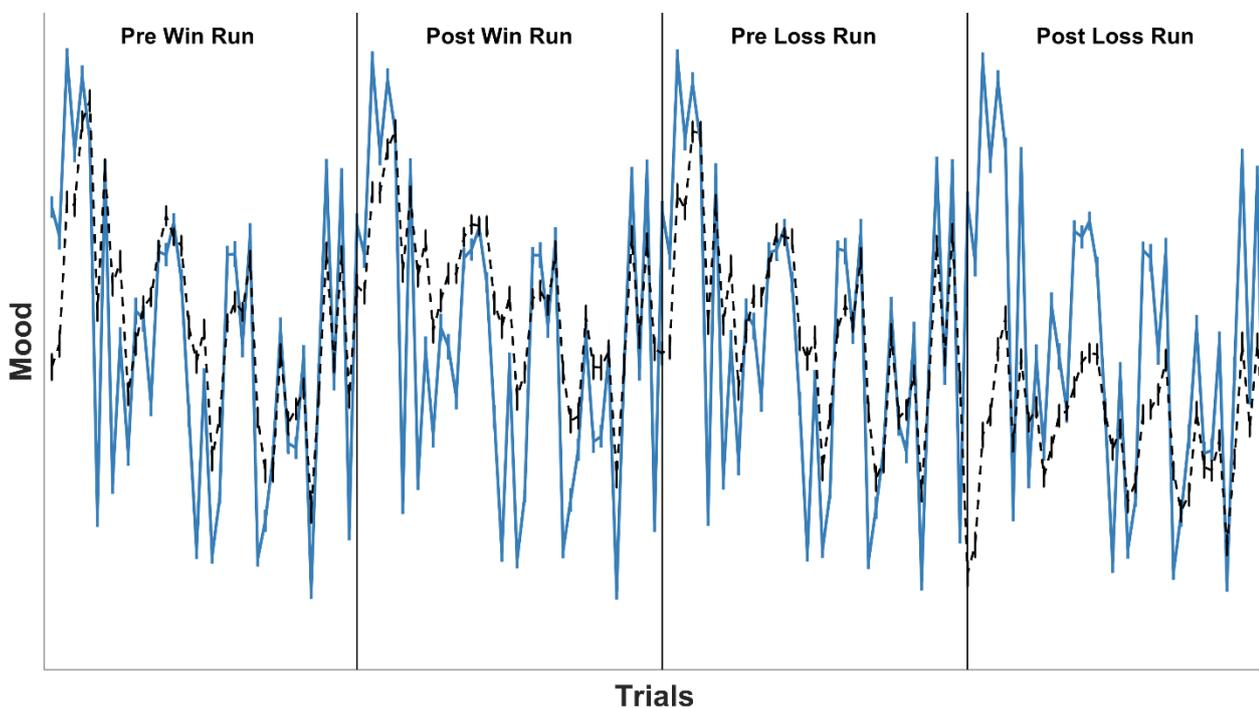


**Figure 2 (a)** Bar chart illustrating the mean (SEM) model BICs (averaged across participants, model sub-type and run), accounting only for participant behaviour. Overlying scatter plots illustrate individual participants' BICs. Model 2 had the (marginally) lowest 'behavioural' BIC. **(b)** Bar chart illustrating the mean (SEM) correlation between model-generated and participant-reported affect, across all runs (averaged across participants). Model 2a generated affect had the highest correlation to participant reported affect across runs. Overlying scatter plots illustrate individual participants' BICs/affect correlations.

Figure 2a illustrates model fit to behavioural data (BICs). Model 2 and 3 (in which available points/apples influence choice) fit participant behaviour considerably better than model 1 (in which they do not). As illustrated in Figure 2a, model 2 (mean BIC: 142.95) performs *marginally* better than model 3 (mean BIC: 144.36). As it was not possible to produce model-generated affect for the no-affect-interaction version of each model, I compared affect only for 'affect-interaction' versions of each model (Figure 2b). Here too, Model 1 performed worst. However, here, model 2 more convincingly distinguished itself from model 3 (mean affect correlations were respectively  $R=0.31$  and  $R=0.25$ ). It could be argued that this difference in mean participant-vs-model affect correlation is made less convincing by the large range of correlations in both cases. Performing a paired samples t-test for the two sets of correlations, I find that model 2-generated affect has a significantly higher correlation to participant affect than model 3-generated affect, ( $p=0.22 \times 10^{-11}$ ). Overall, model 2 fitted behaviour and affect most closely.

### 5.3 Testing the affect-generating mechanism

Figure 2 illustrates how closely model 2 replicated reported affect. Generally the model replicated reported affect well. Assessing affect correlation *within*-runs, I find that in the pre-WoF runs, mean model-generated affect correlates to participant affect at  $R=0.44$  (SEM: 0.020) in the pre-Wof-win, and  $R=0.46$  (SEM: 0.015) in the pre-Wof-loss runs. These correlations are commensurate with a prominent model of affect (Rutledge et al., 2014) which achieves a mean correlation of  $R=0.47$  with 5 (vs 3) model parameters and after first optimising parameters by fitting the model to reported to affect. As illustrated in figure 3, model-2a-generated affect generally fits affect quite closely, but deviates considerably from reported affect in the post WoF-loss run. In the latter case, mean model-generated vs participant-reported affect correlation is  $R=0.27$ .



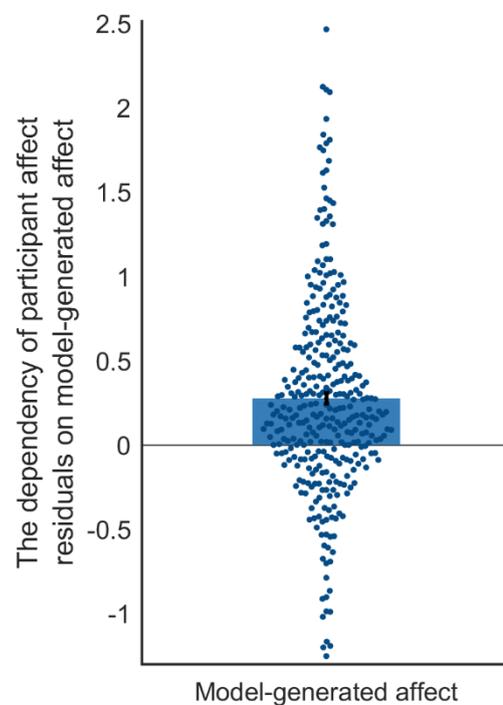
**Figure 3. Plot of model-generated affect (fitting only to behaviour) vs participant affect-reports.**

The solid blue line is mean model generated affect and the dashed black line is mean participant affect reports. Vertical lines are SEM. All affect scores are z-scored across runs.

### 5.4 How accurate is the model, accounting for other contributors to affect?

Though figure 3 shows a relatively convincing correlation between model generated vs participant reported affect, it could still be argued that affect depends instead on other factors and that, once

those factors are accounted for, the model does little to explain affect. In order to test this, I first regressed participant reported affect against three measures on which affect could also depend: **1)** trial outcomes (win/no-win), **2)** the number of apples won in win trials, and **3)** (-)the number of apples on the chosen tree in no-win trials. Here the effect of trial outcomes was significant ( $p=0.0465$ ) while the effects of apples won/not-won were not ( $p \geq 0.1796$ ). I took the resulting residuals to signify the portion of participant-reported affect scores that were not accounted for by these independent variables. I regressed these residuals against model-generated affect. The residuals were significantly dependent on model generated affect (mean beta: 0.2764, sem: 0.0321; single-sample t-test  $p=0.2742 \times [10^{-15}]$ ). Model generated affect therefore recapitulates participant reported affect to an extent that is not accounted for by immediate the impact of task outcomes. This supports the validity of the posited affect-generating mechanism.



**Figure 4** Bar chart illustrating the dependency of participant-reported affect on model-generated affect, accounting for trial outcomes (win/no-win), apples won and apples ‘not won’ in no-win trials. Black vertical lines are SEMs. Overlying scatter plot illustrates individual participant betas.

**Summary:** The affect generating mechanism in the Eldar/Niv model is able to accurately replicate affect at a *considerably* higher resolution than that at which it was originally tested (every 3<sup>rd</sup> vs ~17<sup>th</sup> trial), in a separate task to that in which it was originally tested. The mechanism appears valid to the extent that it is able to recapitulate the influence of RPEs on participant affect using *only* behavioural RPEs, without also being fitted affect data. Even under these relatively stringent conditions, it is able

to capture participant affect even after accounting for trial outcomes. The model therefore seems to be able to capture the influence of learning on affect. However it is not able to account for the effects of prior affect induction, suggesting that it cannot capture the effect of (or lack of effect of) affect on learning.

### 5.5 Replicating the original result: preliminary model fitting

In the original study, Eldar and Niv compared model fit for their full model ( $f$  as a free parameter) vs one without affect interaction ( $f = 1$ ), showing that those with high HPS scores were better described by the full model while those with low HPS scores were better described by the model without affect-learning interaction. It is not possible to produce affect with the latter model without fitting the model to affective data (or without choosing an arbitrary affect update term). I therefore first fitted affect to participant affect-reports in the manner described by Eldar and Niv (Eldar and Niv, 2015).

The likelihood ( $L$ ) of the participant reported affect scores ( $PtA$ ) given each combination of model parameter values ( $\theta$ ) is calculated for each trial ( $t$ ) as:

$$L(PtA_t | \theta) = B_t | C_t$$

Where

$$B_t = e^{-(PtA_t - (ModA_t | \theta))^2}$$

In which ( $ModA | \theta$ ) is the model-generated affect-score given a particular set of model parameter values ( $\theta$ ), z-scored across runs and

$$C_t = \sum_{p=1}^9 e^{-(Z(p) - ModA_t)^2}$$

where ( $Z(p)$ ) is the z-score of each affect-rating ( $p$ ) that the participant can make, given the mean and standard deviation of the participant's affect-reports across runs.

I then fitted models to participant data as described above, but instead of summing over log likelihoods of only choice behaviour, I additionally summed over reported affect (as Eldar and Niv did).

**Models tested:** Despite model fitting, the previous winning models (2a/b) were still unable to reproduce the post-WoF-loss affective timeseries (see Figure 5). I therefore additionally considered

variants of the model. All variants were identical to the winning model above with respect to learning and choice, varying only with respect to affect generation. As before, each variant was fitted both with and without affect interaction.

**Model 2.1:** The winning model from the model comparison above.

**Model 2.2:** Model 2.1, plus an additional parameter ( $\tau$ ) representing participants' initial affect, represented as an initial 'RPE history' value ( $h_0$ ), (affect  $m$  is determined entirely by the 'RPE-history' value  $h$ ).

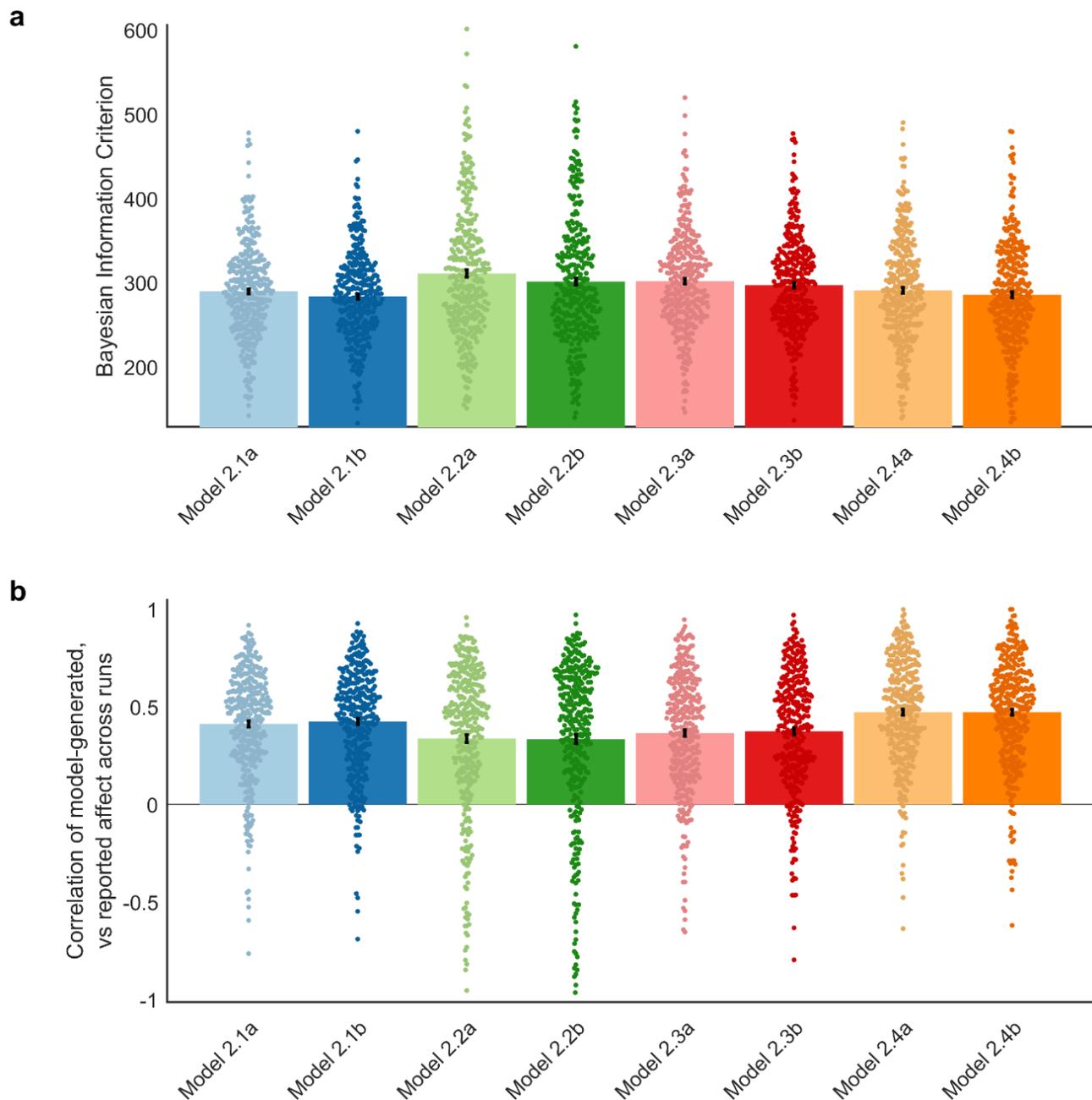
$$h_0 = \tau$$

**Model 2.3:** Model 2.1 with separate affect updates rates for positive and negative trial outcomes;  $x$  denotes outcome type (*win/no-win*).

$$h_{t+1} = h_t + \eta_x(\delta_t)$$

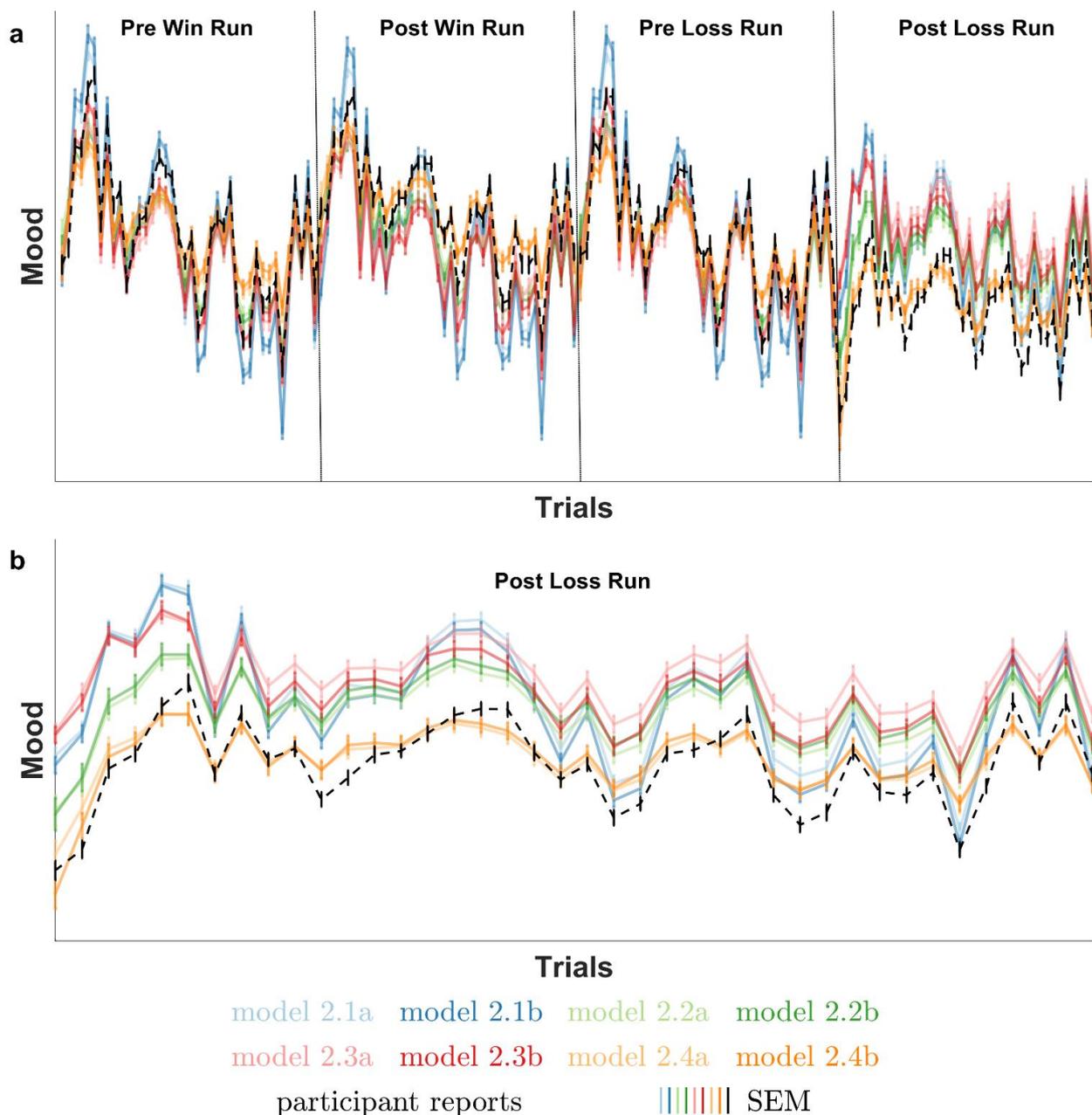
**Model 2.4:** Model 2.1 with the additional parameters from Model 2.2 and 2.3.

Model fitting results are illustrated in Figure 5.



**Figure 5 (a)** Bar chart illustrating mean (SEM) model BICs (averaged across participants, model subtype and run). Model 2.1b had the (marginally) lowest ‘behavioural’ BIC (283.94) followed closely by Model 2.4b (285.90). **(b)** Bar chart illustrating the mean (SEM) correlations between model-generated and participant-reported affect, across all runs (averaged across participants). Affect generated by Model 2.4a/b had the highest correlation to participant reported affect across runs. Overlying scatter plots illustrate individual participants’ BICs/affect correlations.

Importantly Models 2.4(a/b) reproduce post-WoF-loss affect more accurately than the other models, as illustrated in Figure 7.



**Figure 7 (a).** Plot of model-generated affect vs participant affect-reports. Solid coloured lines are the mean affect generated by each model; the dashed black line is mean reported participant affect score. Vertical lines are SEM. All affect scores are z-scored across run. **(b)** This is the same as 7a but displaying only the post-WoF-loss run. Models 2.4a/b reproduce post-WoF-loss affect more accurately than the other models.

Model 2.1 (a and b) had the lowest BICs of the no-affect-interaction and affect-interaction models respectively. However they were only marginally lower than the Model 2.4 equivalents (~2 units in each case). Model 2.4 had the best affect correlation across runs (0.472). Crucially, visual inspection of the affect plots showed that model 2.4 (a and b) were considerably better at reproducing the effect of the Wof-loss on subsequent affect reports than the other models. I therefore took the

model 2.4a/b pair to be the winning pair of models. Model 2.4b had 4 parameters: A behavioural learning rate parameter ( $\alpha$ ), an initial affect parameter ( $\tau$ ) and two affect updates rates ( $\eta_{win/no-win}$ ). Model 2.4a additionally had an affect-learning interaction' parameter ( $f$ ) to total 5 parameters.

### **Replicating the original result: Does an interaction between affect and learning underlie affective instability?**

In their study, Eldar and Niv demonstrated that an interaction between affect and learning was associated with affective instability in everyday life. They did so by comparing the full model (with  $f$  as a free parameter, which allows for affect-learning interaction) vs a restricted model (with  $f = 1$ , which does not allow for affect-learning interaction) using  $\log_{10}$  Bayes factors. They showed that participants with high HPS scores (which signify relative affective instability) were better described by the full model, while those with low HPS scores (which signify relative affective stability) were better described by  $f = 1$ .

I calculated  $\log_{10}$  Bayes factors for the winning pair of models: Model 2.4a (in which  $f$  is a free parameter) vs Model 2.4b (in which  $f = 1$ ) for each person and run. I then correlated the Bayes factors with HPS scores. I found no significant correlations between HPS and Bayes factors overall or in any condition/run. It is possible that those with affective instability might show more of a affect-learning interaction following affect induction and that there may be a differential effect of positive vs negative affect induction. I therefore conducted a repeated measures ANCOVA (Pre-vs-Post WoF, Win-vs-Loss) with HPS score as a co-variate. I found no significant main effects or interactions.

Eldar and Niv also demonstrated that the magnitude of the  $f$  parameter in the full model correlated significantly with participants' HPS scores. That is to say, the extent of affect-learning interaction within-task correlated with the extent of affective instability in everyday life. I found no significant correlations between HPS and  $f$  values, overall or in any condition/run. As with the  $\log_{10}$  Bayes factors, I conducted a repeated measures ANCOVA (Pre-vs-Post WoF, Win-vs-Loss) adding HPS score as a co-variate, in order to test whether affectively unstable individuals might reveal a propensity for affect instability following affect induction. This analysis revealed a main effect of Run  $F(337)= 4.632$   $p=0.032$  and a Run\*HPS interaction  $F(337)= 4.514$   $p=0.034$ . This latter findings appears to result from very small (non-significant;  $ps \geq 0.424$ ) positive correlations between HPS and  $f$  values pre-WoF, and very small (non-significant;  $ps \geq 0.154$ ) negative correlations between HPS and  $f$  values post-WoF.

This result does not support the idea that affectively unstable individuals have greater affect instability following affect induction.

**Summary:** I was unable to replicate the original findings regarding affect-learning interaction and affective instability in a separate reward learning task.

## 5.6 Discussion

In this Chapter I have explored an existing RL model of affect (Eldar and Niv, 2015). Overall, the model was able to accurately replicate participant reported affect in a task on which it had not previously been tested. This is the first time that the model has replicated frequent affect reports (once every 3 vs ~17 trials). In runs without prior affect induction, model-generated affect correlated with participant affect to an extent commensurate with another prominent model of affect (Rutledge et al., 2014). The model was able to do so without fitting to participant affect. The model was unable to replicate affect in the post-WoF-loss run without fitting to participant affect. I was not able to replicate the original key finding: affective instability in real life was not associated with affect-learning interaction in this sample.

The affect-learning interaction model used in this study has previously been used to study the impact of affect on learning (Eldar and Niv, 2015; Michely et al., 2020). As they were not focused on giving a detailed account of learning on affect, these studies recorded participant affect scores every ~16-18 trials. Studies that instead focus on the moment-by-moment impact of learning on affect (Blain and Rutledge, 2022; Chew et al., 2021; Keren et al., 2022; Rutledge et al., 2014) record participant affect scores more frequently (every ~2-4 trials). In the present study, using frequent affect reporting (every 3 trials) I have tested the ability of the affect-learning interaction model to account for moment-by-moment variation of participant affect. In runs without prior affect induction, the model produced intra-run affect correlations that were commensurate with those produced by a more commonly used model of the impact of learning on affect (Blain and Rutledge, 2022; Chew et al., 2021; Keren et al., 2022; Rutledge et al., 2014). While the latter model was fitted to affect, the present model was not, and was instead fitted only to behaviour. Furthermore, the present model used only 3 (vs 5) model parameters. As such the Eldar/Niv model appears a promising means of linking RL, which has a relatively well characterized neural implementation (Glimcher, 2011), and affect, which is highly (self-evidently) relevant to affective disorder research.

I tested the model's capacity to accurately replicate affect under relatively stringent conditions. First, model fitting was restricted to only behaviour and not also affect. As far as I am aware, this is the

first time any model has been shown to produce an accurate affective time-series, without first fitting to affect. This is important as it removes the circularity inherent in reproducing a time-series by fitting to that time-series. Second, the model's ability to capture participant affect was tested *after* accounting for the immediate impact of factors that may alter affect, and to which model generated affect likely correlates. Again, to my knowledge, this is the first time the fidelity of model-generated affect has been tested to this extent. Third, this was done on the entire affective time-series (across all runs as a whole, two of which were post affect induction) as opposed to intra-run (within which correlations were generally higher). All of this serves to validate the basic affect generating mechanism: affect can be described as a weighted sum of RPEs.

Without fitting to participant affect, the model was unable to replicate affect in the post-WoF-loss run. This appears to be because the WoF-Loss had a significant and persistent impact on affect, but very little (arguably no) impact on behaviour/performance (see Appendix 20). This suggests that affect does not impact the perception of rewards (in a manner that impacts reward-learning behaviour). I subsequently tried to formally test whether an interaction between affect and reward-perception underlies affective instability (i.e. I tried to replicate the original finding by Eldar and Niv). I did not find affective instability in everyday life to be associated with an interaction between affect and learning. Rather, I did not find affective instability in everyday life to be associated with an interaction between affect and learning *in the current task*. This task may not have been ideal for asking this particular question, as discussed below in the limitations section.

## 5.7 Limitations

Reward magnitudes (i.e. the number of available apples) varied inter and intra trial and impacted behaviour (as reflected in the poor fit of Models 1a/b). I chose to account for the effect of apples in the choice (softmax) layer (by replacing the inverse temperature with the number of apples on each tree). Arguably, I could equally have accounted for the effect of the apples in the learning layer. I assumed that the number of available apples would not impact the Q-value as the number of available apples were explicitly shown and not related to previous trials. I therefore decided to use them in the choice layer. This means that the number of apples does not directly impact affect in my model. This choice is supported by one of the regressions in this chapter, which demonstrates that participant affect depended on whether they won/did not win the trial, but not on the number of apples won/not won.

In the current task, reward contingencies wander across trials (making it harder to rely on learned stimulus values) and can differ greatly between stimuli within trials. This combination of task characteristics appears to incentivize participants to weigh reward magnitude heavily in their decision making (as evidenced by the success of the model that replaces the inverse temperature parameter with reward magnitudes). Reward magnitudes here play an identical role to  $f^m$  in determining choice probability i.e. in the models,  $f^m$  is effectively reward sensitivity and reward magnitudes are inverse temperatures, two measures that have identical effects on choice probability. Therefore, variable reward magnitudes may obscure any present  $f^m$  signal. This task is therefore not ideal for answering this question. A future study may wish to repeat this investigation using a task better suited to elicit this effect, if present (Eldar and Niv, 2015).

## **5.8 Conclusion**

Affect is a dimension of experience that is (like any dimension of subjective experience) not directly amenable to scientific study, and yet is the crux of affective disorder research. RPEs, by contrast, have a well characterized neural implementation. This chapter demonstrates that affect reports (presumably a reasonable proxy for subjective affect) can be modelled as a weighted sum of RPEs, by a model that is blind to the affect reports themselves. This chapter does not find evidence for an interaction between affect and learning, however, the present task had a limited ability to test this idea. Future studies should further explore the link between RPEs and affect, and continue to test the interaction between affect and learning.

## Chapter 6: General Discussion

### 6.1 Assessment of the results reported in this thesis

In this thesis I used computational modelling and, in particular, reinforcement learning, to study aspects of cognition and affect that are relevant to MDD.

In **chapter 2** I investigated the effect of a (clinically relevant) ~2 week course of Pramipexole on reward-related decision making in healthy volunteers. As hypothesized, a sub-acute course of Pramipexole increased choice accuracy in the reward (but not loss) condition. Expectedly, behavioural computational modelling did not clearly arbitrate between the three proposed mechanisms underlying Pramipexole's effect on reward choice accuracy: 1) increased reward sensitivity, 2) increased choice determinacy and 3) decreased belief decay. Each of the three models were able to recapitulate participant behaviour with comparable fidelity.

In **chapter 3** I attempted to arbitrate between the models in chapter 2, using fMRI. I gauged activity in areas of the brain associated with reward processing during the presentation of task stimuli and the receipt of trial outcomes. Using this activity I inferred the internal processes that might underlie the observed behaviour. Pramipexole enhances BOLD activity during the anticipation of rewarded trials and suppresses BOLD RPE activity. This suggests that it enhances reward learning by reducing the decay of value estimates.

In **chapter 4** I tested the association between within-task and 'real-life' (ESM) affective parameters. I found that each of the three 'real-life' affective parameters (ESM  $\mu$ , ESM  $\nu\mu$  and ESM  $SD$ ) were associated with the corresponding within-task parameter, even controlling for other within-task parameters. This result suggests that affective dynamics in everyday life are reflected in affective dynamics during reward learning tasks. This requires replication but if it is found to be robust, it might suggest that reward learning tasks are an ecologically valid tool for studying affective instability.

**Chapter 5** explored a link between two seemingly disparate areas: 1) behavioural RL and 2) subjective affect. Specifically, I explored an existing RL model of affective instability (Eldar and Niv, 2015). The model was able to replicate participant reported affect in a task on which it had not previously been tested, at a finer granularity than it had previously been tested (once every 3 vs ~17 trials), to an extent commensurate with another prominent model of affect (Rutledge et al., 2014), without fitting to participant affect. However, I was unable to replicate Eldar and Niv's original key

finding (i.e. that affective instability in real life was associated with affect-learning interaction within-task).

## 6.2 General conclusions from the thesis

In **chapter 2**, Pramipexole raises the choice probability asymptote i.e. the impact of Pramipexole on learning is opposite to that of MDD (Admon et al., 2015; Foti et al., 2014; Hall et al., 2014; Johnston et al., 2015; Knutson et al., 2008; Redlich et al., 2015; Remijnse et al., 2009; Robinson et al., 2012; Satterthwaite et al., 2015; Smoski et al., 2009; Steele et al., 2007). This is consistent with Pramipexole's antidepressant effect (Cusin et al., 2013; Fawcett et al., 2016; Tundo et al., 2019). However, MDD participants have blunted reward prediction errors (Admon et al., 2015; Foti et al., 2014; Hall et al., 2014; Johnston et al., 2015; Knutson et al., 2008; Redlich et al., 2015; Remijnse et al., 2009; Robinson et al., 2012; Satterthwaite et al., 2015; Smoski et al., 2009; Steele et al., 2007). This suggests that blunted behavioural response to reward in MDD are probably due to reduced reward sensitivity. The neuroimaging results in **chapter 3** suggest that Pramipexole decreases value decay (rather than increasing reward sensitivity). This implies that Pramipexole does not remediate reward processing impairment in depression, rather it compensates for it. It also implies that Pramipexole's treatment effect may require intact reward sensitivity (as the effect of increased value preservation presumably requires an adequate pre-existing capacity to represent value). This is consistent with findings by Whitton and colleagues: a 6 week course of Pramipexole benefitted those depressed individuals with intact pre-intervention reward sensitivity (Whitton et al., 2020).

In **chapter 4** Bayesian affective dynamic parameters in everyday life are reflected in Bayesian affective dynamic parameters during a reward learning task. This suggests that reward learning tasks may be an ecologically valid way of studying affective instability. While **chapter 5** suggests that affect might be accurately modelled using *only* participant behaviour, strengthening the tie between subjective affect and objective behavior that was implied by Eldar and Niv's original findings (Eldar and Niv, 2015). If this pair of findings is accurate/robust, they may *suggest* a translational channel between animal behavioural research, which studies the biological implementation of reward learning in detail, and ESM affect data, which is more proximate to human experience, but not as amenable to detailed biological investigation.

### **6.3 Are computational models useful in understanding depression and its treatment?**

Computational modelling most commonly purports to help us understand mental disorder by modelling the algorithmic layer which is thought to mediate the effect of implementation (e.g. neural activity) on output (e.g. task stimulus choice) (Marr and Poggio, 1976). More broadly, computational models allow us to explore aspects of data that are not amenable to more traditional statistical analysis (e.g. distinguishing affective volatility from affective noise (Pulcu et al., 2022)). In this thesis, computational models have shown utility in helping to couple different levels of analysis. This allows us to reach conclusions that any one level of analysis on its own might not.

In Chapters 2 and 3, computational models allowed me to couple the effect of Pramipexole on PILT performance with its effect on neural activity, in a relatively rigorous manner (by applying Q-learning algorithms to the former, and using the resulting RPEs to model BOLD activity the latter.) In Chapter 4, applying a Bayesian filter to affect reports allowed me to parse affect into 1) mean affect, 2) short affective fluctuations (noise) and 3) sustained affective perturbations (volatility). This allowed me to discover specific associations between mean affect/volatility/noise during a reward-task and the respective constructs in real life (using ESM data). In Chapter 5, a slightly modified application of an existing model allowed me to recapitulate participant affect using only participant behaviour and task outcomes, lending further weight to the already purported connection between RL and affect (Eldar et al., 2016).

In summary, the chapters in this thesis draw links from neural activity to reward learning behaviour (in chapters 2 and 3), from reward learning behaviour to within-task affect (in chapter 5) and from within task affect to affect in everyday life (in chapter 4). This hints at the intriguing possibility that all of these different levels may potentially one day be linked together, from neural activity to affect in everyday life.

### **6.4 Outstanding questions**

The present results suggest several potential follow-up lines of enquiry. Following on from Chapters 2 and 3: does value preservation mediate Pramipexole's effect on depression? This is best answered in the context of a randomized controlled trial (RCT). Such an RCT is currently underway (Au-Yeung et al., 2022). Another follow-up question is: If a dopaminergic agent, such as Pramipexole, does not remediate reward sensitivity, what might? Reward sensitivity (or consummatory hedonic capacity) is linked to opioid activity (Treadway and Zald, 2011). Perhaps novel opioid agents will have some utility in directly remediating low reward sensitivity in MDD (Browne et al., 2022; Elias et al., 2022).

Speculatively, framing Pramipexole as preserving value estimates could reframe some previous seemingly paradoxical findings. For example, Pramipexole caused riskier choice-making *but (seemingly) paradoxically*, decreased rostral basal ganglia/midbrain responses to high gains (Riba et al., 2008). Framed as decay reduction: Pramipexole both increased risky choice-making and reduced rostral basal ganglia/midbrain responses by preserving value estimates. Future studies may also find that reframing Pramipexole's effect in this manner may resolve seemingly paradoxical results. The Bayesian filter used in Chapter 4 has previously shown that Lithium therapy specifically reduces volatility of positive affect (Pulcu et al., 2022). An RCT is currently underway testing the effect of Pramipexole on Bipolar depression (PAX-BD) (Azim et al., 2021). If participants' ESM data were obtained, applying the same filter could help answer the question: does Pramipexole reduce volatility of *negative* affect in bipolar depression? Or does it, instead, have a countervailing effect on some aspect of positive affect. Incorporating the methods from Chapter 5, and asking participants in PAX-BD (Azim et al., 2021) to perform a reward learning task, we could ask: is any observed effect of Pramipexole on volatility of *negative* affect associated with a reduction of the magnitude of the RL negative affective reactivity parameter ( $\eta_{loss}$ ). We could likewise ask (in a separate RCT): is the observed effect of Lithium on volatility of *positive* affect (Pulcu et al., 2022) associated with a reduction of the magnitude of the RL positive affective reactivity parameter ( $\eta_{win}$ )?

## 6.5 Limitations

This thesis has several limitations. First, the thesis tries to answer questions relevant to depression using entirely non-clinical populations. One question that arises from this choice is whether the results are generalizable to clinical populations. However, and in keeping with the RDOC framework (Cuthbert, 2014), mental health symptoms are arguably a continuum. We can therefore gather potentially valuable insights into mental health symptoms from healthy populations. Non-clinical studies are therefore a reasonable first line in the investigation of mental health phenomena. These studies will need to be followed up by clinical work. Second, this thesis assumes that artificial (trivial) cognitive tasks can probe patients' real world cognitive functioning (e.g. PILT performance can give insights into how people actually learn about important events). This is a limitation in the field generally. There is evidence that such tasks can evidence the *presence* of differences between depressed and healthy individuals (Halahakoon et al., 2020); it is arguably reasonable to think that such tasks can give us insight into the *nature* of these differences also (Halahakoon et al., 2020). Chapter 4, which compares mood parameters in ESM with task data, has attempted to address this

question, but it is also important to assess the relationship between task-based measures and real world (e.g. functional) outcomes.

## **6.6 Conclusion**

In this thesis I studied a number of areas relevant to depression using computational methods, particularly RL. In Chapters 2 and 3, I investigated the effect of a promising antidepressant on neural and behavioural reward learning, finding that Pramipexole increases reward accuracy by preserving value estimates. In Chapter 4 I investigated the relationship between affective dynamics in a reward learning task and those in everyday life, finding that within-task affective mean, volatility and noise are specifically associated with corresponding parameters in everyday life. In Chapter 5 I explored an existing model of affect, finding that the model was able to replicate affect at a finer granularity than previously tested and with reasonable fidelity, even when fitted only to behaviour. I did not find an association between affective instability and mood-learning interaction in the current sample. This thesis adds to the growing body of work that suggests the potential utility of computational methods in psychiatric research.

# Appendices

## Chapter 2

### Appendix 1: Accounting for all trials in the behavioural sessions

In chapter 2, I calculated model-free behavioural results using participants' choice accuracy in the second half of trials. This was because the focus of that analysis was the effect of Pramipexole on asymptotic choice accuracy. However, it could be argued that cutting off half the trials is an arbitrary choice. Including all trials in the analysis, the pattern of results remains the same:

The group\*valence\*session interaction remains significant [ $F(1,38)=7.572$   $p=0.009$ ]. The increase in Pramipexole group reward condition accuracy across sessions remains significant [ $t(20)=2.705$   $p=0.014$ ]. The change in Pramipexole group loss condition accuracy across sessions, as well as the change in placebo group reward and loss accuracy remain non-significant (all  $ps \geq 0.172$ ).

### Appendix 2: Win-Stay Lose Shift Analysis

To check for the possibility that Pramipexole simply changed a tendency to stay/shift after a win/loss, I performed a win-stay lose-shift analysis. In other words, I investigated the effect of the previous (relevant) outcome on choices in each trial. The measure of interest was the probability of choosing the same stimulus as in the previous trial (within condition) binned by the outcome of the (previous) trial.

A repeated measures ANOVA for Condition (Reward vs Punishment), Outcome in the preceding trial (Positive vs Negative) by Session (Pre-Intervention vs Post-Intervention) by Group (Pramipexole vs Placebo) yielded no significant main effects or interactions that included group, either in the second half of trials ( $ps > 0.161$ ) or for all trials ( $ps > 0.191$ ).

### **Appendix 3: The effect of adding baseline TEPS anticipatory subscale scores as a covariate**

By chance, at the initial testing session, the two groups differed significantly in the anticipatory subscale of the TEPS questionnaire. I added individuals' baseline TEPS anticipatory subscale scores to the main behavioural analyses:

After adding baseline anticipatory TEPS sub-scores as a covariate, the group\*condition\*behavioural-session interaction remained significant:  $F(1,37)=8.59$   $p=0.006$ .

Likewise, the group\*condition\*behavioural-session interaction for computational parameter values remained significant:

Inverse temperature parameter [ $F(1,37)=5.521$   $p=0.024$ ]

Reward sensitivity parameter [ $F(1,37)=5.521$   $p=0.024$ ]

Decay parameter [ $F(1,37)=6.332$   $p=0.016$ ].

## **Chapter 3**

### **Appendix 4: Behavioural results, fMRI session**

There are no significant reward/loss accuracy differences between the pre-intervention behavioural session and imaging session:

The main behavioural finding presented in Chapter 2 was a significant group\*valence\*pre-vs-post intervention-behavioural-session interaction for choice accuracy. I did the equivalent analysis for the pre-intervention behavioural session vs the post-intervention imaging session and found no such group\*valence\*pre-intervention-vs-imaging-session interaction for choice accuracy [ $F(1,38)=10.517$   $p=0.096$ ].

Nor did I find any significant changes in reward/loss condition accuracy across sessions in either group (all  $ps \geq 0.086$ .)

I found no significant group\*valence\*(pre-intervention-vs-imaging-session interactions for reward sensitivity, decay or decision determinacy parameters (all  $ps \geq 0.309$ ), nor any significant changes in win/loss trial parameter values across sessions in either group (all  $ps \geq 0.087$ .)

Interestingly, adding individuals' baseline TEPS anticipatory scores as a co-variate yielded a significant group\*condition\* pre-intervention-vs-fMRI-session interaction for choice accuracy [ $F(1,38)=4.457$   $p=0.042$ ]. The group\*condition\*pre-intervention-vs-fMRI-session interactions for computational parameter values (inverse temperature parameter, reward sensitivity parameter and decay parameters) remained non-significant ( $p \geq 0.167$ ).

#### **Appendix 5: Model Free Analysis; using the default high-pass filter**

In fMRI pre-processing I used a high-pass filter cut-off of 60 seconds. This is within the recommendations in the FSL user guide (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT/UserGuide>): *'for event-related designs.....the cutoff can typically be reduced at least to 50s'*. However, the default Feat high-pass filter cut-off is 100s (and I am told this is a more typically used cut-off). As a sanity check, I have repeated the main (model-free, all trials) analysis using a high-pass filter cut-off of 100s. My findings are as follows:

- 1) Using a high-pass filter cutoff of 60s, BOLD activity during reward anticipation in the R OFC was increased in participants receiving Pramipexole relative to placebo. **Re-analyzing with a cut-off of 100s, this result is no longer significant.**
- 2) Using a high-pass filter cutoff of 60s, BOLD activity in the mPFC for 'win' outcomes (vs 'no-win' outcomes), was increased in participants receiving placebo relative to those receiving Pramipexole. **Re-analyzing with a cut-off of 100s, this result remains significant (447 voxels). Additionally the same contrast is now significant in a small area in the R OFC (13 voxels).**
- 3) **The re-analysis yielded no additional significant between-group contrasts.**

## Appendix 6: Model free analysis, individual level design

I used 6 explanatory variables in the main model-free analysis:

**1, 2) Win, Loss Anticipation:** Two-second periods in which stimuli were presented (before choices could be made) in win/loss-condition trials.

**3, 4, 5, 6) Win, No-Win, Loss, No-Loss Outcome:** Two-second periods in which win/win-condition neutral/loss-condition neutral/loss outcomes were presented.

There were 14 Individual-level contrasts in the main model free analysis:

| Contrast                             | Win Anticipation | Loss Anticipation | Win Outcome | No-Win Outcome | Loss Outcome | No-Loss Outcome |
|--------------------------------------|------------------|-------------------|-------------|----------------|--------------|-----------------|
| Win Anticipation                     | 1                | 0                 | 0           | 0              | 0            | 0               |
| Loss Anticipation                    | 0                | 1                 | 0           | 0              | 0            | 0               |
| Win Outcome                          | 0                | 0                 | 1           | 0              | 0            | 0               |
| No-Win Outcome                       | 0                | 0                 | 0           | 1              | 0            | 0               |
| Loss Outcome                         | 0                | 0                 | 0           | 0              | 1            | 0               |
| No-Loss Outcome                      | 0                | 0                 | 0           | 0              | 0            | 1               |
| Win Anticipation > Loss Anticipation | 1                | -1                | 0           | 0              | 0            | 0               |
| Loss Anticipation > Win Anticipation | -1               | 1                 | 0           | 0              | 0            | 0               |
| Win Outcome > No-Win Outcome         | 0                | 0                 | 1           | -1             | 0            | 0               |
| No-Win Outcome > Win Outcome         | 0                | 0                 | -1          | 1              | 0            | 0               |
| Loss Outcome > No-Loss Outcome       | 0                | 0                 | 0           | 0              | 1            | -1              |
| No-Loss Outcome > Loss Outcome       | 0                | 0                 | 0           | 0              | -1           | 1               |
| Win Outcome > Loss Outcome           | 0                | 0                 | 1           | 0              | -1           | 0               |
| Loss Outcome > Win Outcome           | 0                | 0                 | -1          | 0              | 1            | 0               |

## Appendix 7: Model based Analyses, individual level design

I used 6 explanatory variables in the model-based analysis:

**1, 2) Win, Loss Anticipation:** Two-second periods in which stimuli were presented (before choices could be made) in win/loss-condition trials.

**3, 4) Win, Loss Condition Outcome:** Two-second periods in which win condition/loss condition outcomes were presented.

**5, 6) Win, Loss-Condition RPEs:** The timings are the same as for EVs 3/4, but with within-participant, trial-wise RPEs as parametric modulators.

There were 12 Individual-level contrasts in the main model based analysis:

| Contrast                                       | Win Anticipation | Loss Anticipation | Win-Condition Outcome | Loss-Condition Outcome | Win-Condition RPEs | Loss-Condition RPEs |
|------------------------------------------------|------------------|-------------------|-----------------------|------------------------|--------------------|---------------------|
| Win Anticipation                               | 1                | 0                 | 0                     | 0                      | 0                  | 0                   |
| Loss Anticipation                              | 0                | 1                 | 0                     | 0                      | 0                  | 0                   |
| Win-Condition Outcome                          | 0                | 0                 | 1                     | 0                      | 0                  | 0                   |
| Loss-Condition Outcome                         | 0                | 0                 | 0                     | 1                      | 0                  | 0                   |
| Win-Condition RPEs                             | 0                | 0                 | 0                     | 0                      | 1                  | 0                   |
| Loss-Condition RPEs                            | 0                | 0                 | 0                     | 0                      | 0                  | 1                   |
| Win Anticipation > Loss Anticipation           | 1                | -1                | 0                     | 0                      | 0                  | 0                   |
| Loss Anticipation > Win Anticipation           | -1               | 1                 | 0                     | 0                      | 0                  | 0                   |
| Win-Condition Outcome > Loss-Condition Outcome | 0                | 0                 | 1                     | -1                     | 0                  | 0                   |
| Loss-Condition Outcome > Win-Condition Outcome | 0                | 0                 | -1                    | 1                      | 0                  | 0                   |
| Win-Condition RPEs > Loss-Condition RPEs       | 0                | 0                 | 0                     | 0                      | 1                  | -1                  |
| Loss-Condition RPEs > Win-Condition RPEs       | 0                | 0                 | 0                     | 0                      | -1                 | 1                   |

## Appendix 8: Group-level Design

In both the model-free/based analyses, Individual-level outputs were averaged across runs and entered into group analyses. Model free/based group analyses were identical in design.

There were 2 group level explanatory variables:

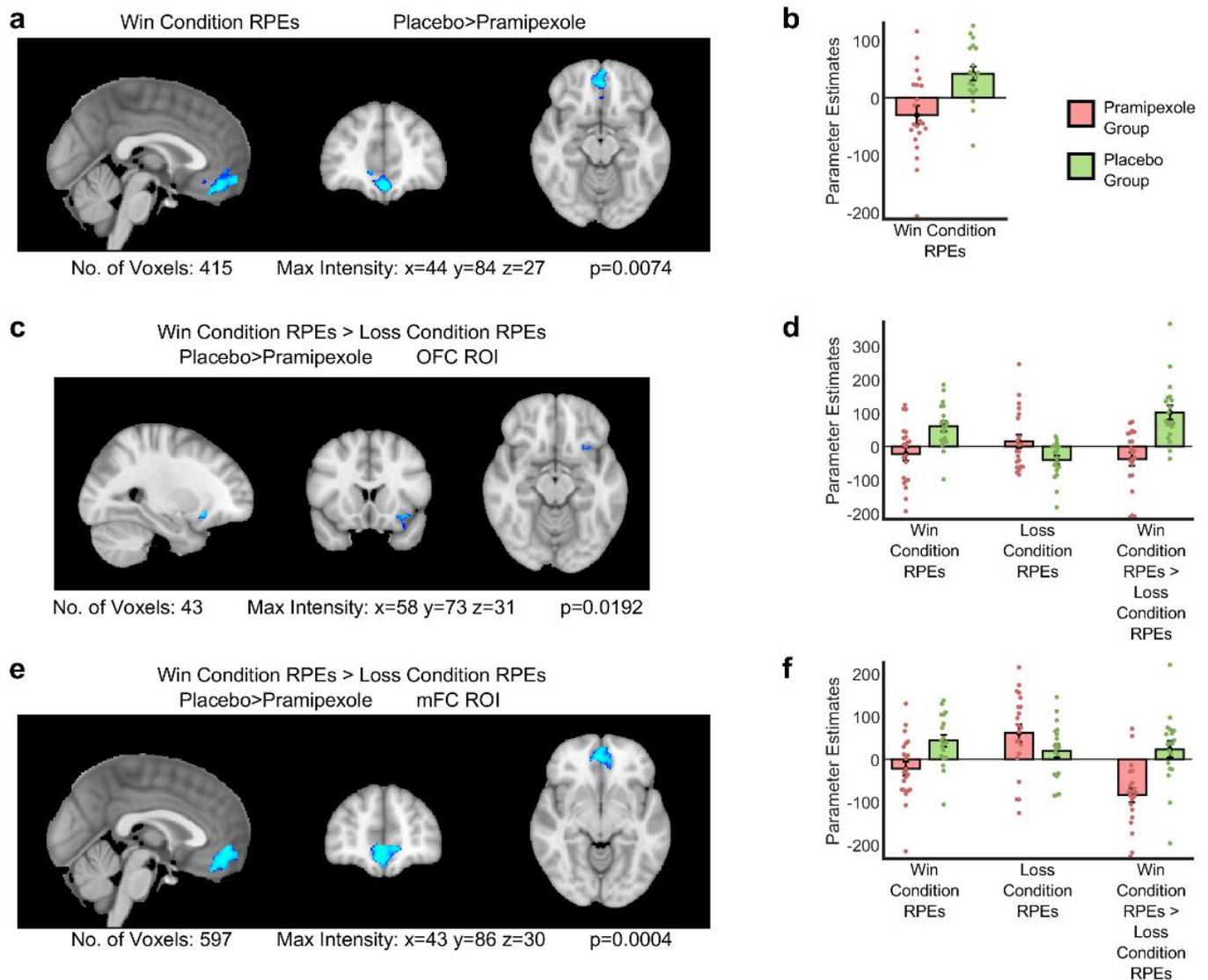
**1) Pramipexole Group:** The participant was randomized to receive Pramipexole.

**2) Placebo Group:** The participant was randomized to receive placebo.

There were 6 group level contrasts:

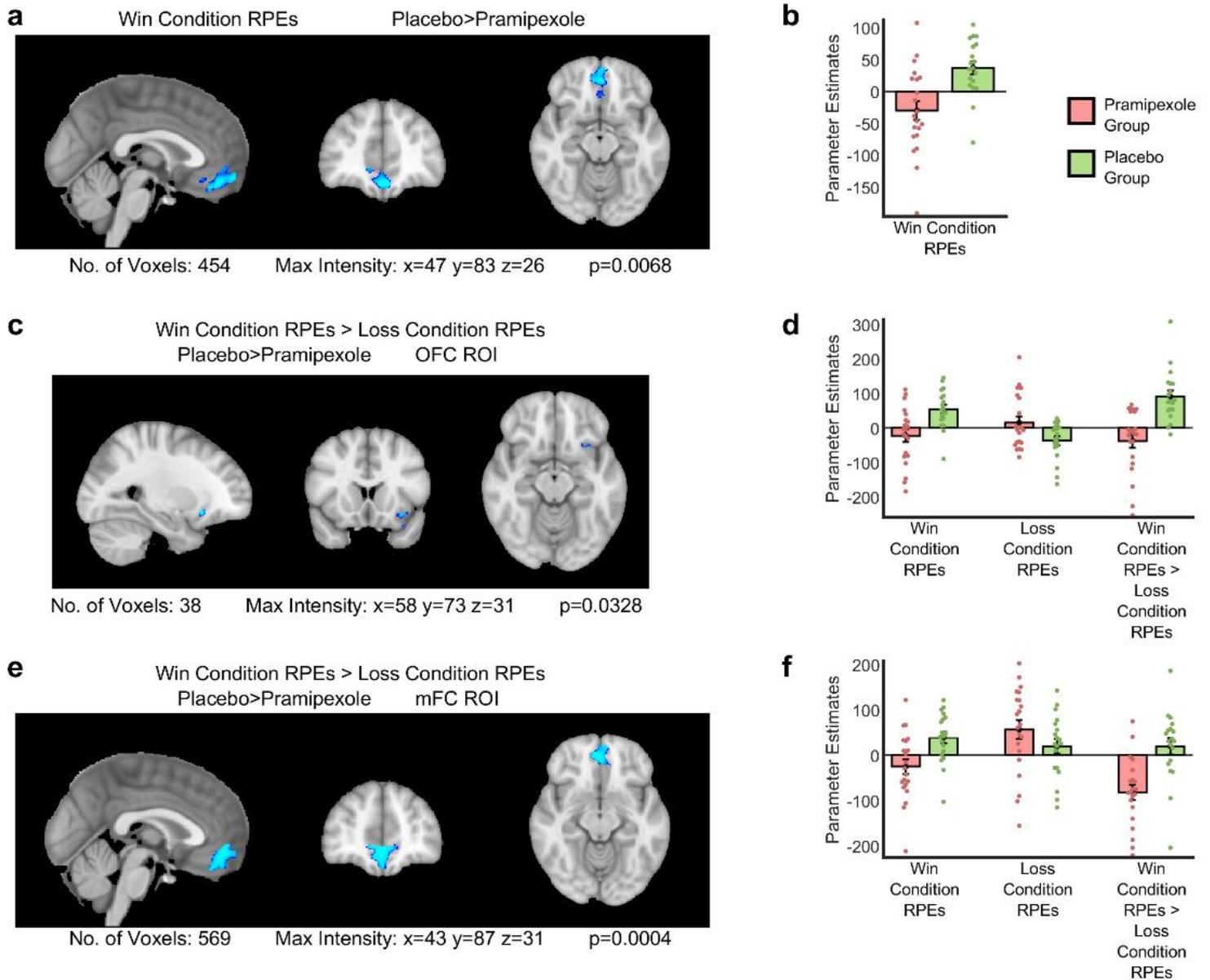
| Contrast                          | Pramipexole Group | Placebo Group |
|-----------------------------------|-------------------|---------------|
| Group Mean                        | 1                 | 1             |
| Pramipexole Group                 | 1                 | 0             |
| Placebo Group                     | 0                 | 1             |
| Pramipexole Group > Placebo Group | 1                 | -1            |
| Placebo Group > Pramipexole Group | -1                | 1             |
| Negative Group Mean               | -1                | -1            |

## Appendix 9: fMRI results, decay model



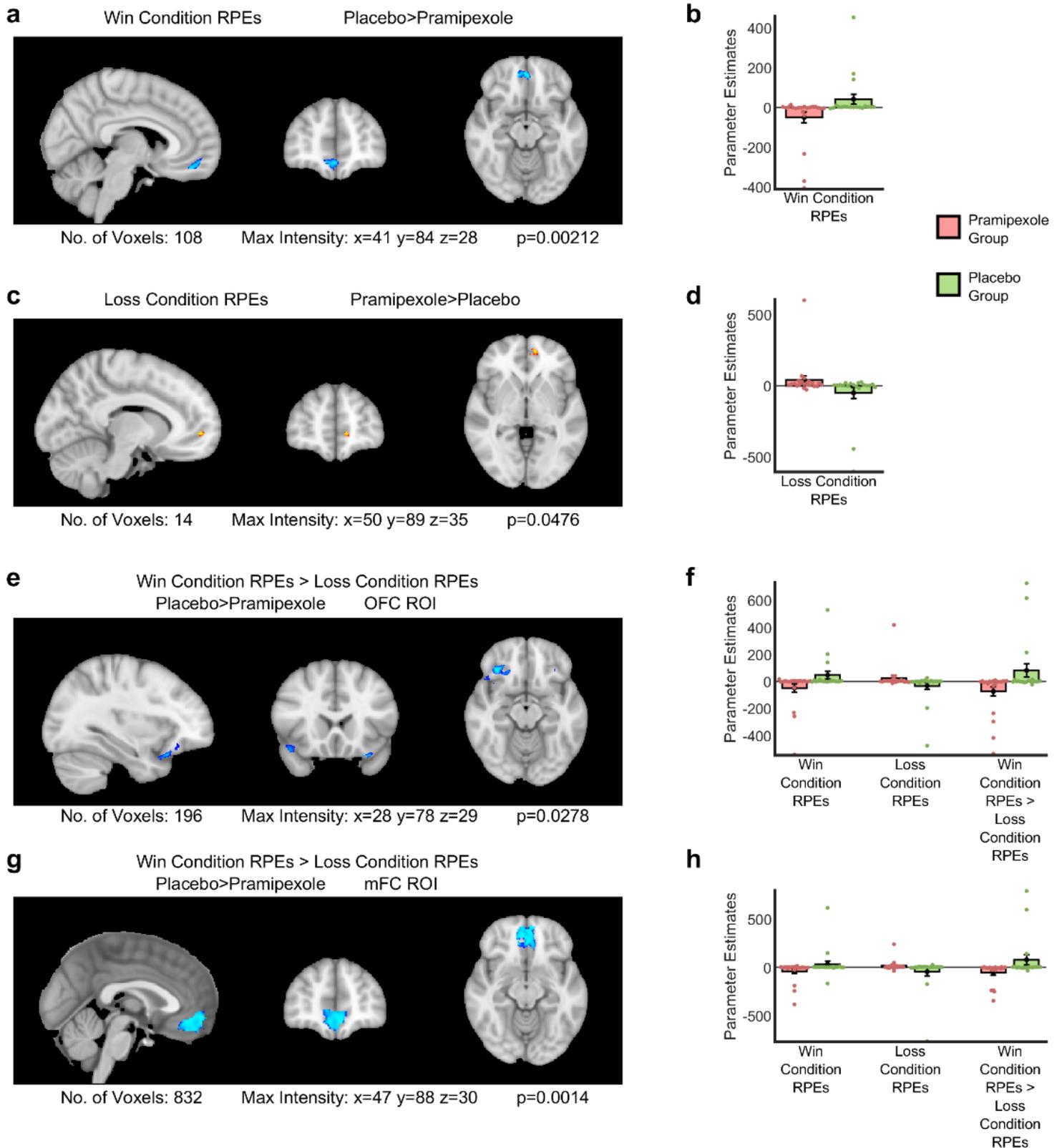
**Fig A1. Model based fMRI (outcome) results:** Decay model contrasts for **(a)** Win-condition RPEs; mPFC ROI **(c)** Win-condition RPEs > Loss-condition RPEs; OFC ROI and **(e)** Win-condition RPEs > Loss-condition RPEs; mPFC ROI. The (blue) coloured area represents clusters of significantly decreased activity in the Pramipexole vs placebo group. Areas of significantly decreased activity are threshold free cluster enhancement corrected with a family-wise error cluster significance level of  $p \leq 0.05$ . **(b,d,f)** Parameter estimates extracted from the areas of significantly decreased activity in *Fig A1 a, c* and *e* respectively. Green (pink) bars represent the placebo (Pramipexole) group. Error-bars represent SEM. Scatter plots overlaying bar graphs depict corresponding individual parameter estimates.

## Appendix 10: fMRI results, inverse temperature model



**Fig A2. Model based fMRI (outcome) results:** IT model contrasts for **(a)** Win-condition RPEs; mPFC ROI **(c)** Win-condition RPEs > Loss-condition RPEs; OFC ROI and **(e)** Win-condition RPEs > Loss-condition RPEs; mPFC ROI. The (blue) coloured area represents clusters of significantly decreased activity in the Pramipexole vs placebo group. Areas of significantly decreased activity are threshold free cluster enhancement corrected with a family-wise error cluster significance level of  $p \leq 0.05$ . **(b,d,f)** Parameter estimates extracted from the areas of significantly decreased activity in *Fig A2 a, c* and *e* respectively. Green (pink) bars represent the placebo (Pramipexole) group. Error-bars represent SEM. Scatter plots overlaying bar graphs depict corresponding individual parameter estimates.

## Appendix 11: fMRI results, reward sensitivity model



The figure description is on the next page.

**Fig A3. Model based fMRI (outcome) results:** RS model contrasts for **(a)** Win-condition RPEs; mPFC ROI **(c)** Loss-condition RPEs; mPFC ROI **(e)** Win-condition RPEs > Loss-condition RPEs; OFC ROI and **(g)** Win-condition RPEs > Loss-condition RPEs; mPFC ROI. The red-yellow coloured area represents a cluster of significantly increased activity in the Pramipexole vs placebo group. The blue coloured areas represent clusters of significantly decreased activity in the Pramipexole vs placebo group. Areas of significantly increased/decreased activity are threshold free cluster enhancement corrected with a family-wise error cluster significance level of  $p \leq 0.05$ . **(b,d,f,h)** Parameter estimates extracted from the areas of significantly decreased (increased in the case of A3c) activity in *Fig A3 a, c, e and g* respectively. Green (pink) bars represent the placebo (Pramipexole) group. Error-bars represent SEM. Scatter plots overlaying bar graphs depict corresponding individual parameter estimates.

### **Appendix 12: Adding Baseline TEPS anticipation scores to fMRI analyses**

As mentioned above, by chance, the baseline anticipatory TEPS subscale was different between the two groups.

I repeated the analyses that yielded the key imaging results in Chapter 3, adding baseline TEPS anticipatory sub-scale scores as a co-variate. I found that:

-The Win Anticipation > Loss Anticipation; Pramipexole > Placebo and Win Anticipation 2<sup>nd</sup> Half > 1<sup>st</sup> Half; Pramipexole > Placebo contrasts are no longer significant.

-The Win Out > No-Win Out; Placebo > Pramipexole contrast remains significant but the area of significant activity reduces to 5 voxels (Peak voxel  $x=46, y=81, z=25$ ; voxel size:5;  $p=0.0434$ ).

-The (Decay) Win Condition RPE; Placebo > Pramipexole contrast remained significant but area of significant activity reduces to 3 voxels (Peak voxel  $x=46, y=81, z=25$ ; voxel size:3;  $p=0.0476$ ).

## Chapter 4

### Appendix 13: Do within-task parameters represent consistent traits?

The main analysis in this chapter assesses the association between task-based parameters and corresponding ESM parameters. These analyses consisted of regressing each of ESM  $\mu$ ,  $\nu\mu$  and  $SD$  against, collectively, task  $\mu$ ,  $\nu\mu$  and  $SD$  averaged across sessions. Implicit in 'averaging across sessions' is the assumption that these values are relatively consistent across sessions. I tested this by correlating each parameter across the two task sessions:

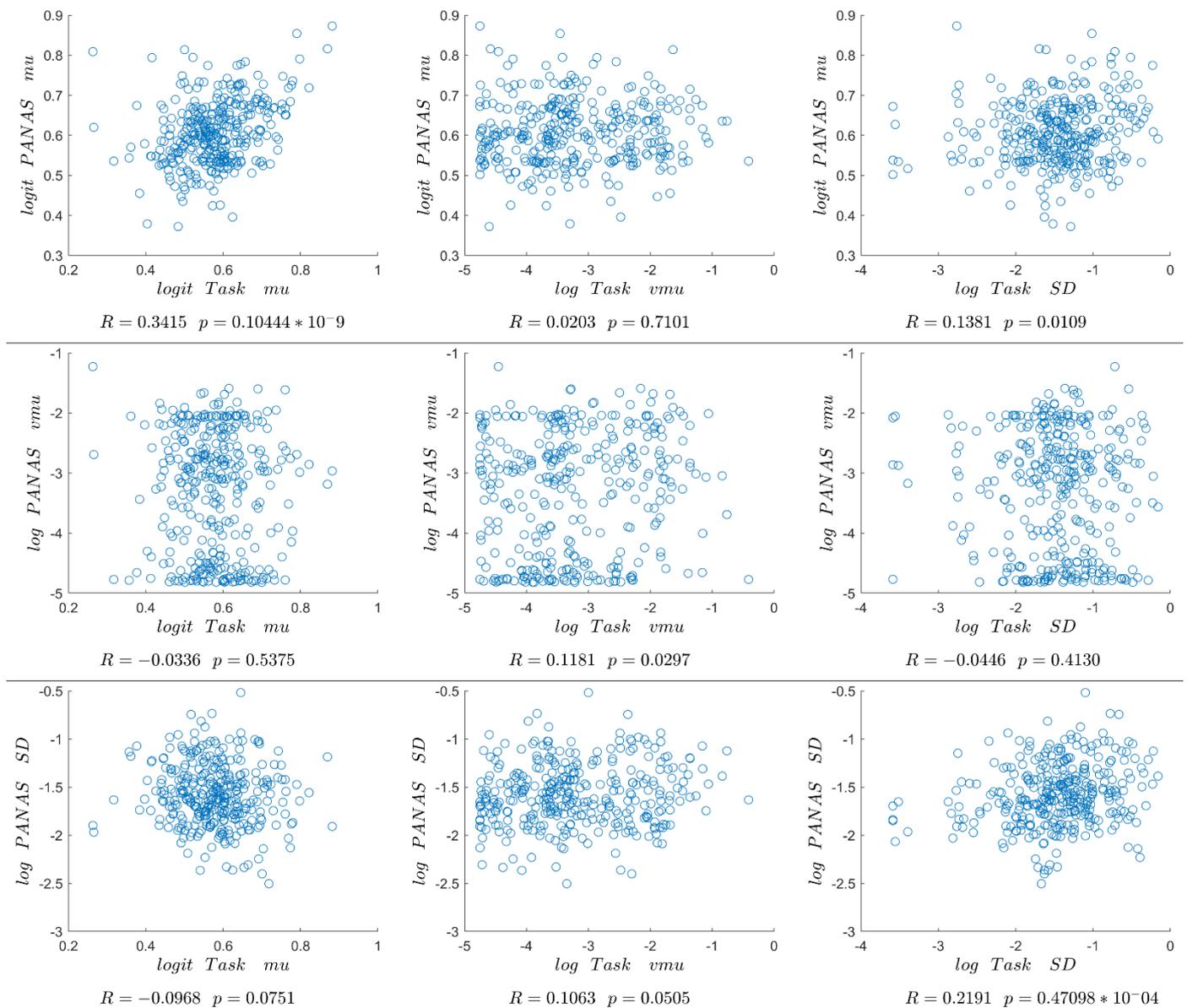
**Task  $\mu$ :**  $R\ 0.686\ p\ 0.18228*10^{-47}$

**Task  $\nu\mu$ :**  $R\ 0.2765\ p\ 0.22973*10^{-6}$

**Task  $SD$ :**  $R\ 0.662\ p\ 0.40717*10^{-43}$

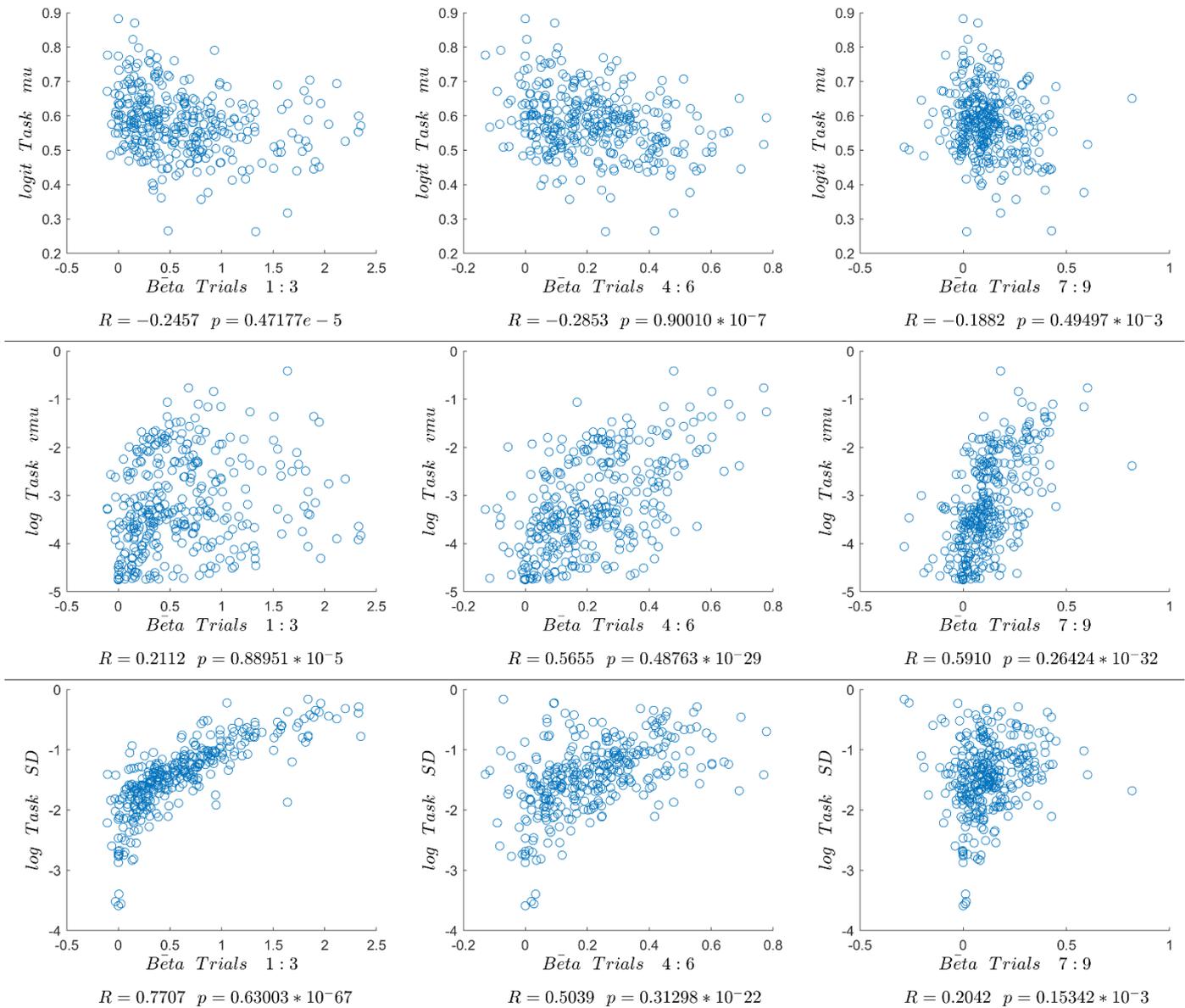
It therefore appears from task data that measures of affective variability represent consistent traits, though not that consistent in the case of volatility.

## Appendix 14: Scatter plots accompanying main analysis



**Figure A4.** Scatter plots of logit PANAS  $\mu$  (top row), log PANAS  $\nu\mu$  (middle row) and log PANAS SD (bottom row), each vs logit task  $\mu$  (left column), log task  $\nu\mu$  (middle column) and log task SD (right column). Each with the accompanying correlation coefficient. As with the regressions in Chapter 4, task parameters are specifically correlated with the equivalent PANAS parameter (the only exception is logit PANAS  $\mu$ , which is additionally significantly correlated with log task SD).

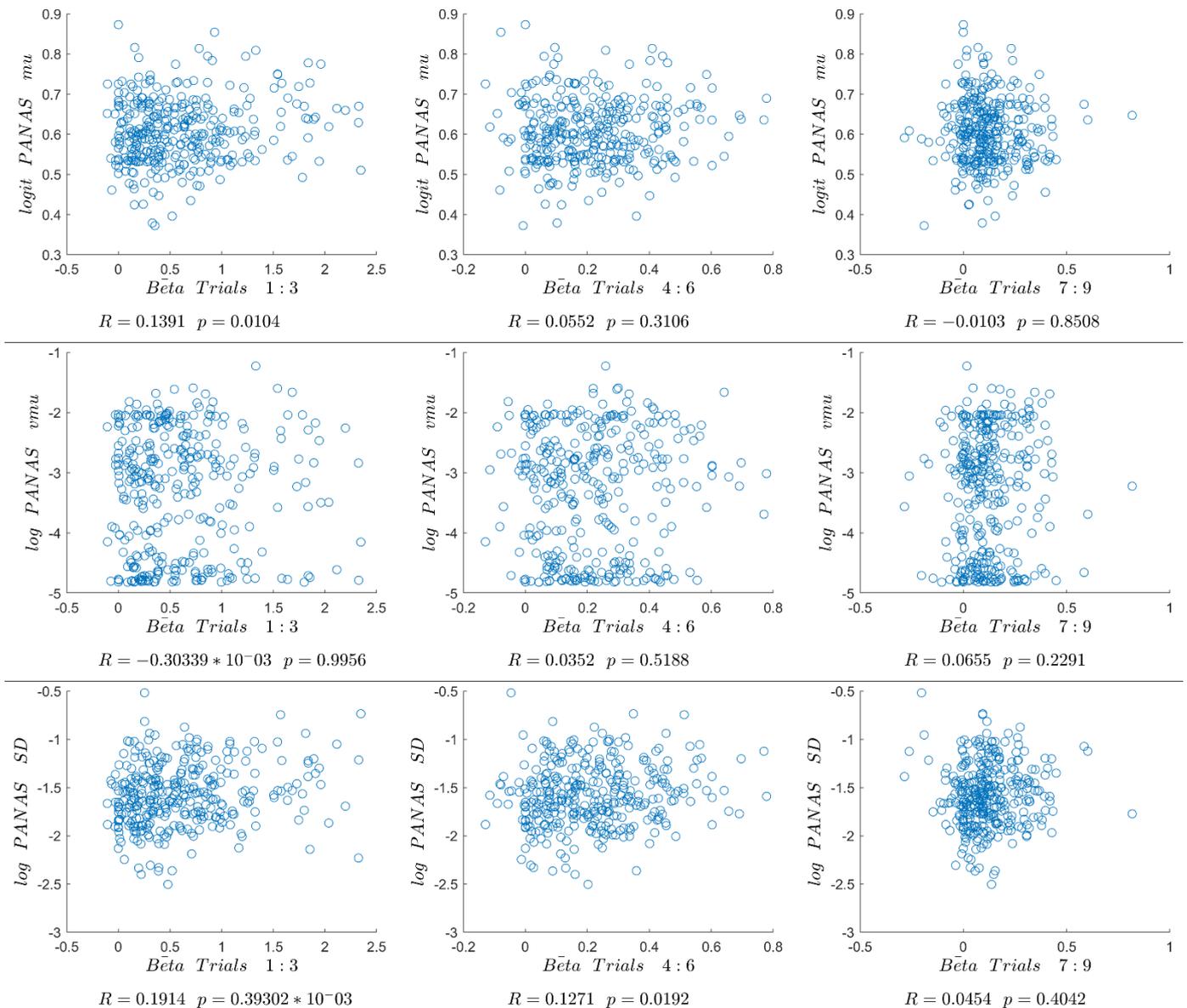
## Appendix 15: Scatter plots accompanying task parameters vs task betas analysis



**Figure A5. Scatter plots** Affect-report scores were regressed against the preceding 9 trial outcomes, within-subject, to generate betas that represented the dependency of reported affect on each of the preceding 9 trial outcomes. Each group of 3 betas was then averaged (representing, the dependence of each affect score on the 3 events that took place immediately before it, the 3 events that took place before that and the 3 events that took place before that; these are labelled, respectively,  $\bar{B}\bar{e}\bar{t}\bar{a}$  Trials 1:3,  $\bar{B}\bar{e}\bar{t}\bar{a}$  Trials 4:6 and  $\bar{B}\bar{e}\bar{t}\bar{a}$  Trials 7:9. These panels depict the correlation between logit Task  $\mu$ , (top row) log Task  $\nu\mu$  (middle row) and log Task SD (bottom row), each vs  $\bar{B}\bar{e}\bar{t}\bar{a}$  Trials 1:3 (left column),  $\bar{B}\bar{e}\bar{t}\bar{a}$  Trials 4:6 (middle column) and  $\bar{B}\bar{e}\bar{t}\bar{a}$  Trials 7:9 (right column). Each with the accompanying correlation coefficient. Consistent with the original regression

analysis, log Task *vmu* is most correlated with the persistent affective impact of task events while log Task *SD* is most correlated with the immediate affective impact of task events. logit Task *mu* is negatively correlated with the affective impact of task events.

## Appendix 16: Scatter plots accompanying PANAS parameters vs task betas analysis



**Figure A6. Scatter plots** Affect-report scores were regressed against the preceding 9 trial outcomes, within-subject, to generate betas that represented the dependency of reported affect on each of the preceding 9 trial outcomes. Each group of 3 betas was then averaged (representing, the dependence of each affect score on the 3 events that took place immediately before it, the 3 events that took place before that and the 3 events that took place before that; these are labelled, respectively,  $\bar{\beta}$  Trials 1:3,  $\bar{\beta}$  Trials 4:6 and  $\bar{\beta}$  Trials 7:9). These panels depict the correlation between logit PANAS  $\mu$ , (top row) log PANAS  $\nu\mu$  (middle row) and log PANAS SD (bottom row), each vs  $\bar{\beta}$  Trials 1:3 (left column),  $\bar{\beta}$  Trials 4:6 (middle column) and  $\bar{\beta}$  Trials 7:9 (right column). Each with the accompanying correlation coefficient. Consistent with the original

regression analysis, PANAS *SD* is most correlated with the immediate impact of trail events (as is PANAS *mu*, though less so; PANAS *vmu* appears to have no correlation to the impact of trail events.)

#### **Appendix 17: Are Bayesian parameters correlated with each other?**

One of the criticism levelled at auto-correlation models of affect are that their metrics are dependent on one-another (Koval et al., 2013). Bayesian affective parameters are not mathematically dependent in the same way as the autocorrelation measures. I additionally assessed whether they traded off against each other in practice. I did so by correlating the parameters against each other.

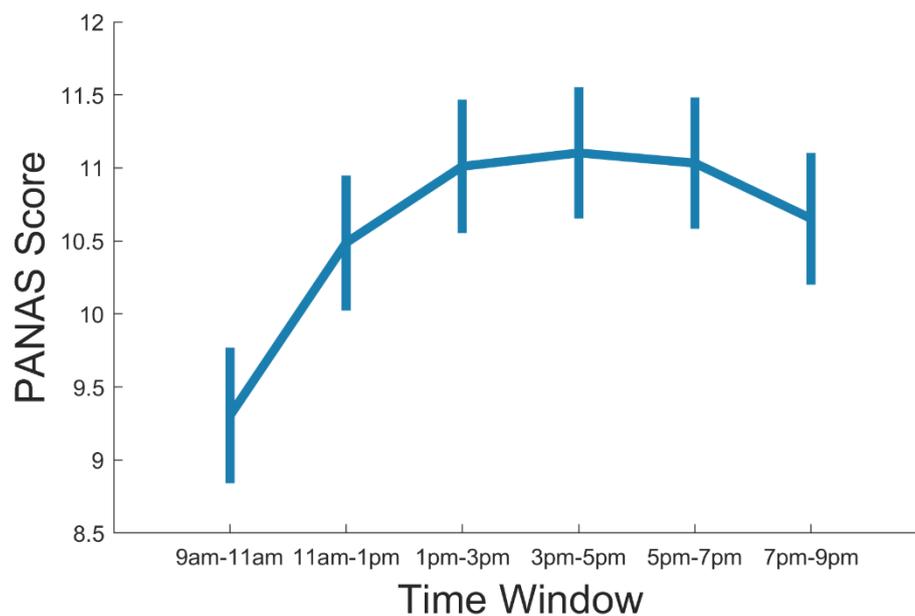
**Task Parameters:** *mu* was negatively correlated with *vmu* ( $R=-0.2240$   $p=0.31545 \cdot 10^{-4}$ ) and with *SD* ( $R=-0.1635$   $p=0.0025$ ). *vmu* was positively correlated with *SD* ( $R=0.2124$   $p=0.0001$ ).

**PANAS Parameters:** *mu*, *vmu* and *SD* were not significantly correlated with each other ( $p \geq 0.1024$ ).

From the analysis of task parameters, there does seem to be some correlation between parameters. However these correlations are relatively low. It appears that the Bayesian parameters measure at least partially independent phenomena.

## Appendix 18: Do ESM parameters reflect inter-day, or solely intra-day, affective dynamics?

PANAS scores appear to have a diurnal variation. A repeated measures ANOVA with one factor and 6 levels, (for each of the 6 daily time-windows {9-11am, 11am-1pm,....7-9pm}) revealed a significant effect of time-window  $F=40.4$   $p \leq 0.001$ . The diurnal affective variation is illustrated in Figure A7 below, mood is lowest in the morning, peaks in the mid-afternoon and then falls slightly in the evening.



**Figure A7. PANAS scores across the day.** PANAS scores for each time-window, averaged across days and participants. Vertical bars represent SEM. Mood is lowest in the morning, peaks in the mid-afternoon and then falls slightly in the evening.

Given this apparent diurnal variation, I explored whether the variance in participants' affective variability parameters was entirely due to differences in diurnal variation or whether they reflected affective dynamics across a longer timescale. In order to do so, I divided the original PANAS time series into 6 separate PANAS time series consisting of responses within each of 6 daily time-windows (9-11am, 11am-1pm,....7-9pm). I fitted the Bayesian filter to each of these time-series to produce a separate set of mood variability parameters for each. I then regressed each mood variability parameter for each window against all three mood variability parameters (collectively) for the full PANAS time-series. If parameters for the overall time-series reflected longer term affective dynamics, and not just the diurnal variation, there should be a dependence of the 'time-window'-series parameters on the equivalent parameter from the overall time-series.

I first regressed PANAS  $\mu$  (for each of the 6 'time-window' times-series) against the three mood variability parameters (collectively) from the overall PANAS time-series. PANAS  $\mu$  for each time-window is significantly dependent on the overall PANAS  $\mu$  ( $p \leq 0.45 * 10^{-165}$ ). I then did the same for PANAS  $\nu\mu$  for each of the 6 'time-window' times-series. PANAS  $\nu\mu$  for each time-window is significantly dependent on the overall PANAS  $\nu\mu$  ( $p \leq 0.17 * 10^{-04}$ ). Finally I did the same for PANAS  $SD$  for each of the 6 'time-window' times-series. PANAS  $SD$  for each time-window is significantly dependent on the overall PANAS  $SD$  ( $p \leq 0.21 * 10^{-35}$ ).

**Conclusion:** PANAS  $\mu$ ,  $\nu\mu$  and  $SD$  for each time-window were each dependant on overall PANAS  $\mu$ ,  $\nu\mu$  and  $SD$ . Therefore, overall PANAS parameters reflect variance in inter-day, and not only intra-day, affective dynamics.

## Chapter 5

### Appendix 19: Are model parameters consistent across sessions?

Eldar and Niv's analysis suggests that model parameters reflect consistent traits. In the present study, participants performed the task on two separate days. I correlated parameters across sessions and runs to assess whether they are consistent across sessions.

I first did so for the winning model without fitting to affect:

**Affect-learning interaction parameter ( $f$ ):** Pre-WoF interaction parameters were moderately correlated ( $R=0.31$   $p=0.97 * 10^{-8}$ ), as were Pre-vs-post WoF interaction parameters (Win:  $R=0.46$   $p=0.18 * 10^{-18}$ ; Loss:  $R=0.44$   $p=0.81 * 10^{-17}$ ) and post Win-vs-Loss WoF interaction parameters ( $R=0.36$   $p=0.16 * 10^{-10}$ ). The interaction parameter is therefore a stable trait.

**Behavioural learning rate ( $\alpha$ ):** Pre-WoF behavioural learning rates were moderately correlated ( $R=0.66$   $p=0.12 * 10^{-42}$ ), as were Pre-vs-post WoF behavioural learning rates (Win:  $R=0.68$   $p=0.10 * 10^{-46}$ ; Loss:  $R=0.63$   $p=0.29 * 10^{-38}$ ) and post Win-vs-Loss WoF behavioural learning rates ( $R=0.61$   $p=0.17 * 10^{-35}$ ). The behavioural learning rate is therefore a stable trait.

**Affect update parameter ( $\gamma$ ):** Pre-WoF affect update parameters were not correlated ( $R=0.078$   $p=0.15$ ), nor were Pre-vs-post WoF affect update parameters (Win:  $R=-0.075$   $p=0.17$ ; Loss:  $R=0.063$   $p=0.25$ ) or the post Win-vs-Loss WoF affect update parameters ( $R=0.034$   $p=0.536$ ).

Here I do the same for the winning model when fitted to affective reports:

**Initial mood parameter ( $\tau$ ):** Pre-WoF initial mood parameters were (expectedly) not significantly correlated ( $R=-0.0180$   $p=0.7407$ ). Pre-vs-post WoF Initial mood was not significantly correlated ( $p \geq 0.1769$ ) while post Win-vs-Loss WoF initial mood parameters were weakly negatively correlated ( $R=-0.1769$   $p=0.0086$ ). Surprisingly, the two post-mood-induction runs were (weakly) negatively correlated. This suggests that the effect of positive and negative events on a person's affect are negatively correlated.

**Behavioural learning rate ( $\alpha$ ):** Pre-WoF behavioural learning rates were moderately correlated ( $R=0.5374$   $p=0.91987 \cdot 10^{-26}$ ), as were Pre-vs-post WoF behavioural learning rates (Win:  $R=0.5491$   $p=0.438 \cdot 10^{-27}$ ; Loss:  $R=0.6616$   $p=0.4734 \cdot 10^{-43}$ ) and post Win-vs-Loss WoF behavioral learning rates ( $R=0.5809$   $p=0.5572 \cdot 10^{-31}$ ). The behavioural learning rate is therefore a stable trait.

**Mood reactivity parameter for positive outcomes ( $\gamma_{win}$ ):** Pre-WoF mood reactivity parameters for positive outcomes were weakly correlated across session ( $R=0.2861$   $p=0.8283 \cdot 10^{-7}$ ). Pre-vs-post WoF mood reactivity parameter for positive outcomes were weakly correlated (Pre-vs-post Win:  $R=0.2317$   $p=0.16349 \cdot 10^{-4}$ ; Pre-vs-post Loss:  $R=0.2819$   $p=0.12983 \cdot 10^{-6}$ ). Post Win-vs-Loss WoF mood reactivity parameter for positive outcomes were also weakly correlated ( $R=0.3519$   $p=0.25755 \cdot 10^{-10}$ ). Reactivity to positive outcomes is therefore a somewhat stable trait, though correlations are low.

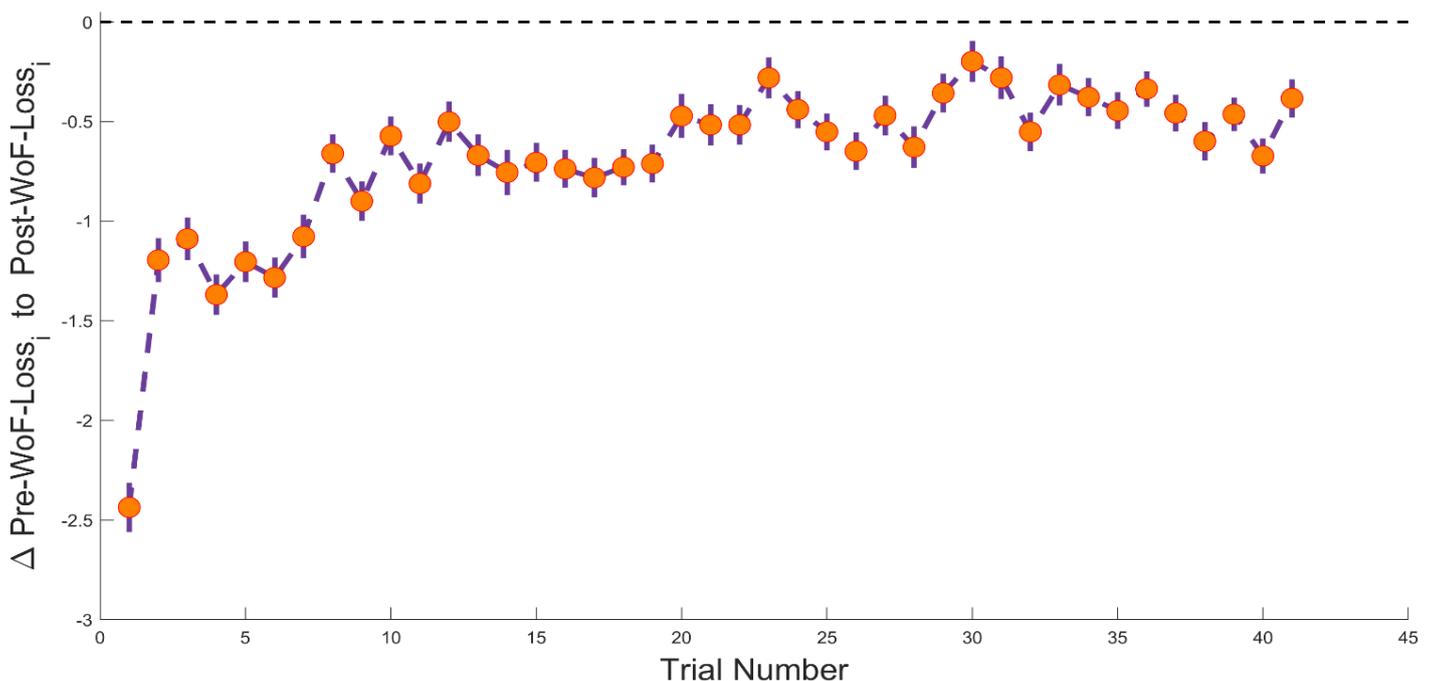
**Mood reactivity parameter for negative outcomes ( $\gamma_{loss}$ ):** Pre-WoF mood reactivity parameters for negative outcomes were weakly correlated across session ( $R=0.3208$   $p=0.14976 \cdot 10^{-8}$ ). Pre-vs-post WoF mood reactivity parameter for negative outcomes were weakly correlated (Win:  $R=0.2758$   $p=0.2462 \cdot 10^{-6}$ ; Loss:  $R=0.3673$   $p=0.28964 \cdot 10^{-11}$ ). Post Win-vs-Loss WoF mood reactivity parameter for negative outcomes were also weakly correlated ( $R=0.3107$   $p=0.50835 \cdot 10^{-8}$ ). Reactivity to negative outcomes is therefore a somewhat stable trait, though correlations are low.

*Summary:* Apart from initial mood, the value each parameter was positively correlated within-parameter, across participants, between a) the pre WoF run b) pre-vs-post WoF runs (of each condition) and c) post-win-vs-loss runs (in which case initial mood parameters were anticorrelated). This suggests that the behavioral learning rate parameter and two mood reactivity parameters each represent consistent traits, though correlations for mood reactivity parameters were low.

**Appendix 20: Why does the model not capture affect in the post-WoF-Loss run when fit only to behaviour?**

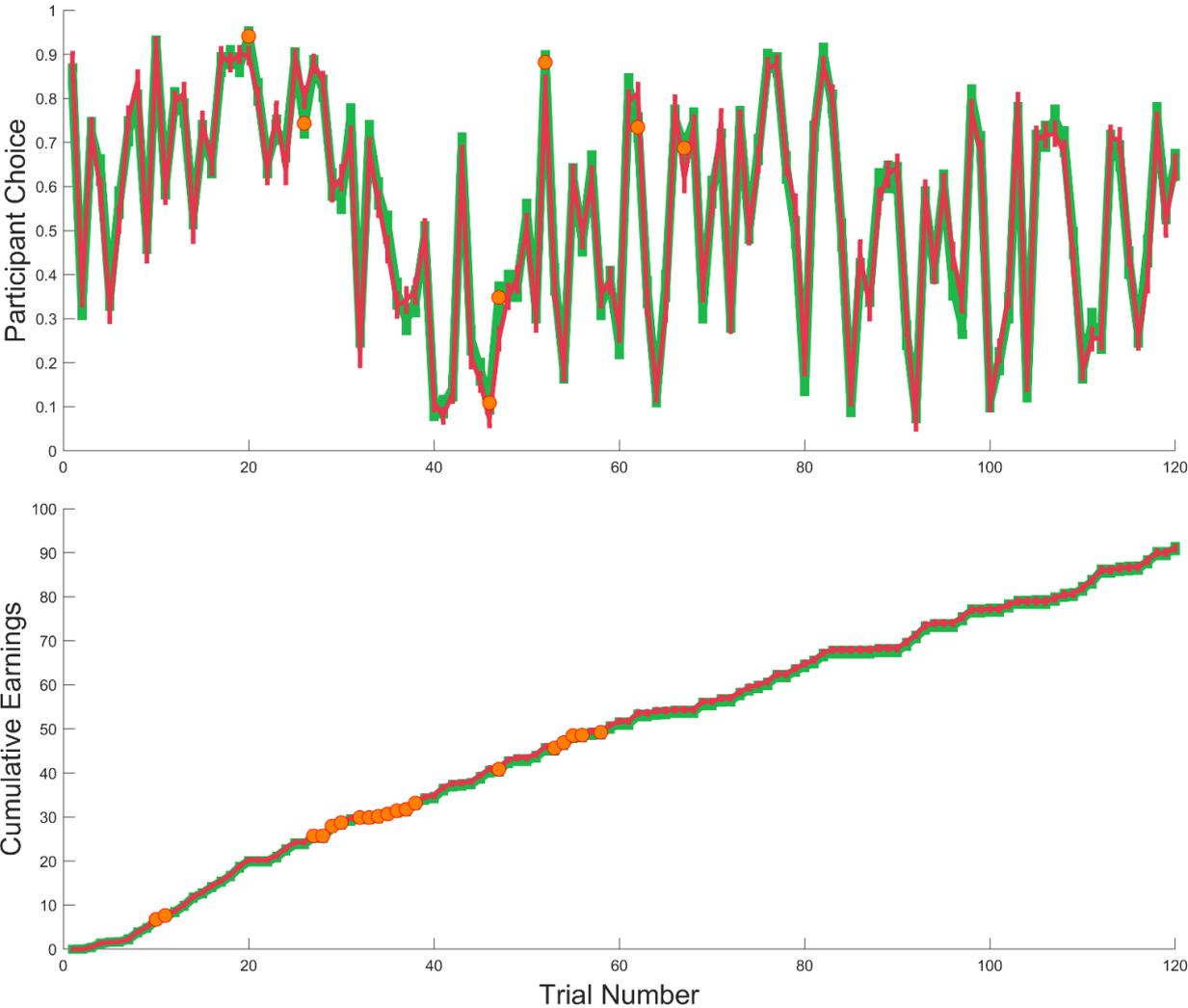
As is apparent from chapter 5 figure 3, and from the within-run model-vs-participant affect correlations, the model recapitulates post-WoF-loss affect relatively poorly. In this appendix, I examine why.

**The impact of the WoF-loss on affect:** To examine why the model does not perform as well in the post-WoF-loss run, I first examined the impact of the WoF-loss on affect. I compared affect at equivalent points along pre-vs-post WoF-loss runs. I used paired sample t-tests to compare pre-vs-post WoF-loss mood reports at equivalent time-points within-run. The results of this analysis are presented in figure A8. As is apparent from the figure, every post WoF-loss affect rating was significantly lower than the equivalent pre-WoF-loss rating. The WoF-loss therefore had significant and persistent impact on affect in the subsequent run.



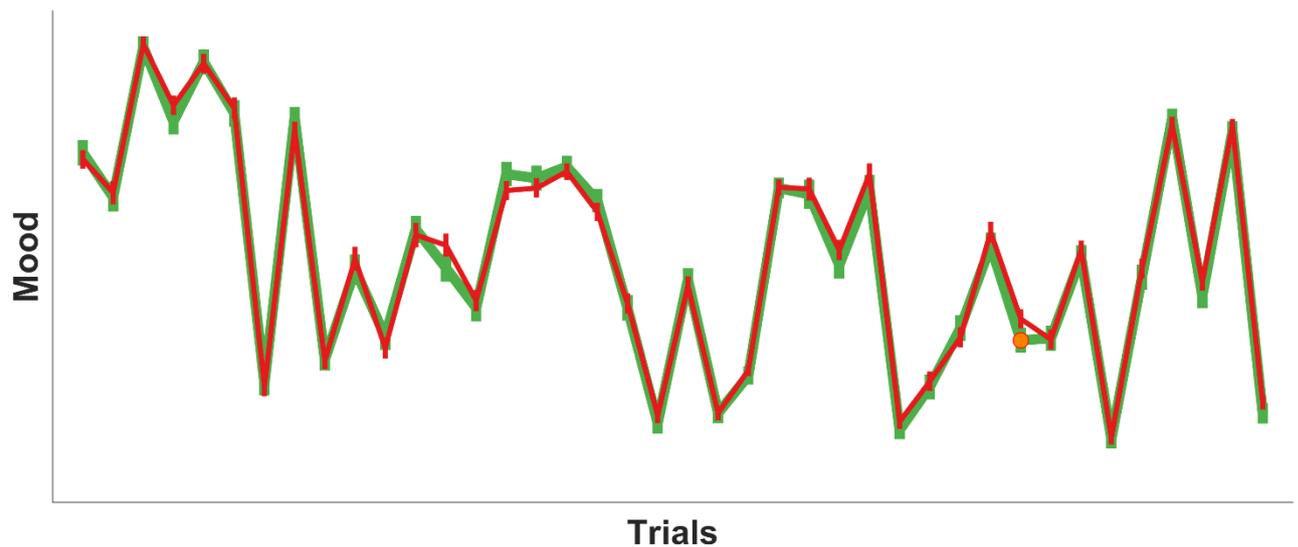
**Figure A8. The effect of the WoF-loss on affect:** The dashed purple line represent the mean difference between pre-vs-post WoF-loss affect reports at each time-point ( $\Delta_{WoF_{loss_i}} = Post_{WoF_{loss_i}} - Pre_{WoF_{loss_i}}$ ). The dashed purple line represents  $\bar{\Delta}_{WoF_{loss_i}}$  across participants. The error bars represent the respective SEMs. The dashed black line represents  $\bar{\Delta}_{WoF_{loss_i}} = 0$ . Separate paired sample t-tests were performed for mood reports at each time-point (comparing pre-vs-post WoF-loss mood reports, within trial). Orange circles signify that the t-test returned a  $p \leq 0.001$ .

**The effect of the WoF-loss on choice/performance:** I then compared trial-wise stimulus choice and earnings in the pre-vs-post WoF-loss in the same manner as for affect. Participants behaved almost identically in the pre-vs-post WoF-loss runs. Participant behaviour never deviated to the point that the a p-value reached a threshold of 0.001 (the threshold used for affect deviation above). The p value was  $\leq 0.05$  in only 7 of 120 trials. Likewise, at no point did pre-vs-post WoF-loss run performance (as measured by earnings) deviate to the point that the a p-value reached a threshold of 0.001. The p value was  $\leq 0.05$  in only 19 of 120 trials. This is in contrast to affect (Fig A5) which deviated sufficiently between runs that affect-report-wise paired t-tests comparing pre-vs-post affect reached a p threshold of 0.001 in all cases.



**Figure A9 a) Participant Choice:** The proportion of participants who chose stimulus 1 in each pre-WoF-loss trial (red) and post-WoF-loss trial (green). **b) Participant Performance:** Mean points accumulated at each pre-WoF-loss trial (red) and post-WoF-loss trial (green). Orange circles signify that within-trial, between-runs paired t-test returned a  $p \leq 0.05$ . Vertical bars represent SEM.

Given that behaviour and outcomes were so similar in the pre-vs-post WoF-Loss runs, and given that the model is fit only to behaviour, model parameters for pre-vs-post WoF-Loss runs were very similar: between-runs paired t-tests for pre-WoF-Loss vs post-WoF-Loss  $\alpha$ ,  $\eta$  and  $f$  were all non-significant ( $p_s \geq 0.08$ ). Given the similarity of the pre-vs-post WoF-Loss parameter values, the model produced very similar affective time series, with only one affective rating that differed significantly between the two runs, as illustrated in Figure A7.



**Figure A10 a) Model Generated Mood** for the pre-WoF-loss run (red) overlaid on post-WoF-loss run (green). Orange circles signify that the t-test returned a  $p \leq 0.05$ . Vertical bars represent SEM.

As can be seen from the participant affect, choice and performance plots (Figures A5,6), the WoF-loss has a significant and persistent impact on participant affect (Figure A5) but very little (arguably no) impact on choice or trial outcomes, (A6) which generate the RPEs that the model uses to generate affect. It is therefore unsurprising that the model was unable to replicate the impact of the WoF-Loss on affect.

#### **Appendix 21: Does an interaction between affect and learning underlie ESM measures of affective instability?**

In chapter 5 I calculated  $\log_{10}$  Bayes factors for the winning pair of models: Model 2.4a (in which  $f$  is a free parameter) vs Model 2.4b (in which  $f = 1$ ) for each person and run. I then correlated the Bayes factors with HPS scores. HPS served as a measure of affective instability. However, from

chapter 4 I have separate (ESM) measures of affective instability i.e. ESM *vmu* and ESM *SD*. Here I repeated for ESM *vmu* and *SD* the analyses I performed for HPS scores in chapter 5.

I found no significant correlations between ESM *vmu* and Bayes factors overall or in any condition/run. I conducted a repeated measures ANCOVA (Pre-vs-Post WoF, Win-vs-Loss) with ESM *vmu* score as a co-variate. I found no significant effects/interactions that included the co-variate *vmu*.

I found no significant correlation between ESM *SD* and Bayes factors overall. There was a weak correlation between ESM *SD* and Bayes factors in the pre-win run ( $R=0.11$   $p=0.04$ ). There was a weak correlation between ESM *SD* and Bayes factors in any other run. I conducted a repeated measures ANCOVA (Pre-vs-Post WoF, Win-vs-Loss) with ESM *SD* score as a co-variate. I found no significant effects/interactions that included the co-variate *SD*.

- Ackerman, J.P., McBee-Strayer, S.M., Mendoza, K., Stevens, J., Sheftall, A.H., Campo, J.V., Bridge, J.A., 2015. Risk-Sensitive Decision-Making Deficit in Adolescent Suicide Attempters. *Journal of Child and Adolescent Psychopharmacology* 25, 109–113. <https://doi.org/10.1089/cap.2014.0041>
- Adams, R.L., Giffen, M.B., Garfield, F., 1973. Risk taking among suicide attempters. *Journal of Abnormal Psychology* 82, 262–267. <https://doi.org/10.1037/h0035153>
- Admon, R., Kaiser, R.H., Dillon, D.G., Beltzer, M., Goer, F., Olson, D.P., Vitaliano, G., Pizzagalli, D.A., 2017. Dopaminergic Enhancement of Striatal Response to Reward in Major Depression. *American Journal of Psychiatry* 174, 378–386. <https://doi.org/10.1176/appi.ajp.2016.16010111>
- Admon, R., Nickerson, L.D., Dillon, D.G., Holmes, A.J., Bogdan, R., Kumar, P., Dougherty, D.D., Iosifescu, D.V., Mischoulon, D., Fava, M., Pizzagalli, D.A., 2015. Dissociable cortico-striatal connectivity abnormalities in major depression in response to monetary gains and penalties. *Psychological Medicine* 45, 121–131. <https://doi.org/10.1017/S0033291714001123>
- Alexopoulos, G.S., Manning, K., Kanellopoulos, D., McGovern, A., Seirup, J.K., Banerjee, S., Gunning, F., 2015. Cognitive control, reward-related decision making and outcomes of late-life depression treated with an antidepressant. *Psychological Medicine* 45, 3111–3120. <https://doi.org/10.1017/S0033291715001075>
- Ang, Y.-S., Gelda, S.E., Pizzagalli, D.A., 2022. Cognitive effort-based decision-making in major depressive disorder. *Psychological Medicine* 1–8. <https://doi.org/10.1017/S0033291722000964>
- Arnsten, A.F.T., Wang, M., Paspalas, C.D., 2015. Dopamine's Actions in Primate Prefrontal Cortex: Challenges for Treating Cognitive Disorders. *Pharmacol Rev* 67, 681–696. <https://doi.org/10.1124/pr.115.010512>
- Arrondo, G., Segarra, N., Metastasio, A., Ziauddeen, H., Spencer, J., Reinders, N.R., Dudas, R.B., Robbins, T.W., Fletcher, P.C., Murray, G.K., 2015. Reduction in ventral striatal activity when anticipating a reward in depression and schizophrenia: a replicated cross-diagnostic finding. *Frontiers in Psychology* 6. <https://doi.org/10.3389/fpsyg.2015.01280>
- Aston-Jones, G., Cohen, J.D., 2005. Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. *J Comp Neurol* 493, 99–110. <https://doi.org/10.1002/cne.20723>
- Au-Yeung, S.K., Griffiths, J., Roberts, S., Edwards, C., Yu, L.-M., Bogacz, R., Rendell, J., Attenburrow, M.-J., Watson, S., Chan, F., Cipriani, A., Cleare, A., Harmer, C.J., Kessler, D., Evans, J., Lewis, G., Singh, I., Simon, J., Harrison, P.J., Cowen, P., Shanyinde, M., Geddes, J., Browning, M., 2022. PAX-D: study protocol for a randomised placebo-controlled trial evaluating the efficacy and mechanism of pramipexole as add-on treatment for people with treatment resistant depression. *Evidence-Based Mental Health* 25, 77–83. <https://doi.org/10.1136/ebmental-2021-300282>
- Azim, L., Hindmarch, P., Browne, G., Chadwick, T., Clare, E., Courtney, P., Dixon, L., Duffelen, N., Fouweather, T., Geddes, J.R., Goudie, N., Harvey, S., Helder, T., Holstein, E.-M., Martin, G., Mawson, P., McCaffery, J., Morriss, R., Simon, J., Smith, D., Stokes, P.R.A., Walker, J., Weetman, C., Wolstenhulme, F., Young, A.H., Watson, S., McAllister-Williams, R.H., 2021. Study protocol for a randomised placebo-controlled trial of pramipexole in addition to mood stabilisers for patients with treatment resistant bipolar depression (the PAX-BD study). *BMC Psychiatry* 21, 334. <https://doi.org/10.1186/s12888-021-03322-y>
- Baek, K., Kwon, J., Chae, J.-H., Chung, Y.A., Kralik, J.D., Min, J.-A., Huh, H., Choi, K.M., Jang, K.-I., Lee, N.-B., Kim, S., Peterson, B.S., Jeong, J., 2017. Heightened aversion to risk and loss in depressed patients with a suicide attempt history. *Scientific Reports* 7, 11228. <https://doi.org/10.1038/s41598-017-10541-5>

- Baker, G.B., Coutts, R.T., McKenna, K.F., Sherry-McKenna, R.L., 1992. Insights into the mechanisms of action of the MAO inhibitors phenelzine and tranylcypromine: a review. *J Psychiatry Neurosci* 17, 206–214.
- Behrens, T.E.J., Woolrich, M.W., Walton, M.E., Rushworth, M.F.S., 2007. Learning the value of information in an uncertain world. *Nature Neuroscience* 10, 1214–1221. <https://doi.org/10.1038/nn1954>
- Belujon, P., Grace, A.A., 2017. Dopamine System Dysregulation in Major Depressive Disorders. *Int J Neuropsychopharmacol* 20, 1036–1046. <https://doi.org/10.1093/ijnp/pyx056>
- Blain, B., Rutledge, R.B., 2022. Momentary subjective well-being depends on learning and not reward. *eLife* 9, e57977. <https://doi.org/10.7554/eLife.57977>
- Blanco, N.J., Otto, A.R., Maddox, W.T., Beevers, C.G., Love, B.C., 2013. The influence of depression symptoms on exploratory decision-making. *Cognition* 129, 563–568. <https://doi.org/10.1016/j.cognition.2013.08.018>
- Blood, A.J., Zatorre, R.J., 2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences* 98, 11818–11823. <https://doi.org/10.1073/pnas.191355898>
- Bodi, N., Keri, S., Nagy, H., Moustafa, A., Myers, C.E., Daw, N., Dibo, G., Takats, A., Bereczki, D., Gluck, M.A., 2009. Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. *Brain* 132, 2385–2395. <https://doi.org/10.1093/brain/awp094>
- Bowen, R., Baetz, M., Leuschen, C., Kalynchuk, L.E., 2011. Predictors of suicidal thoughts: Mood instability versus neuroticism. *Personality and Individual Differences* 51, 1034–1038. <https://doi.org/10.1016/j.paid.2011.08.015>
- Bowen, R.C., Wang, Y., Balbuena, L., Houmphan, A., Baetz, M., 2013. The relationship between mood instability and depression: implications for studying and treating depression. *Med Hypotheses* 81, 459–462. <https://doi.org/10.1016/j.mehy.2013.06.010>
- Bradley, M.M., Lang, P.J., 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 49–59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- Brischoux, F., Chakraborty, S., Brierley, D.I., Ungless, M.A., 2009. Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proceedings of the National Academy of Sciences of the United States of America* 106, 4894–4899. <https://doi.org/10.1073/pnas.0811507106>
- Brooks, J.C.W., Beckmann, C.F., Miller, K.L., Wise, R.G., Porro, C.A., Tracey, I., Jenkinson, M., 2008. Physiological noise modelling for spinal functional magnetic resonance imaging studies. *Neuroimage* 39, 680–692. <https://doi.org/10.1016/j.neuroimage.2007.09.018>
- Broome, M.R., Saunders, K.E.A., Harrison, P.J., Marwaha, S., 2015. Mood instability: significance, definition and measurement. *Br J Psychiatry* 207, 283–285. <https://doi.org/10.1192/bjp.bp.114.158543>
- Browne, C.A., Wulf, H., Lucki, I., 2022. Kappa Opioid Receptors in the Pathology and Treatment of Major Depressive Disorder, in: Liu-Chen, L.-Y., Inan, S. (Eds.), *The Kappa Opioid Receptor, Handbook of Experimental Pharmacology*. Springer International Publishing, Cham, pp. 493–524. [https://doi.org/10.1007/164\\_2020\\_432](https://doi.org/10.1007/164_2020_432)
- Browning, M., Behrens, T.E., Jocham, G., O'Reilly, J.X., Bishop, S.J., 2015. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience* 18, 590–596. <https://doi.org/10.1038/nn.3961>
- Burdick, K.E., Braga, R.J., Gopin, C.B., Malhotra, A.K., 2014. Dopaminergic Influences on Emotional Decision Making in Euthymic Bipolar Patients. *Neuropsychopharmacology* 39, 274–282. <https://doi.org/10.1038/npp.2013.177>
- Carl, H., Walsh, E., Eisenlohr-Moul, T., Minkel, J., Crowther, A., Moore, T., Gibbs, D., Petty, C., Bizzell, J., Dichter, G.S., Smoski, M.J., 2016. Sustained anterior cingulate cortex activation during

- reward processing predicts response to psychotherapy in major depressive disorder. *Journal of Affective Disorders* 203, 204–212. <https://doi.org/10.1016/j.jad.2016.06.005>
- Casacalenda, N., Perry, J.C., Looper, K., 2002. Remission in Major Depressive Disorder: A Comparison of Pharmacotherapy, Psychotherapy, and Control Conditions. *American Journal of Psychiatry* 159, 1354–1360. <https://doi.org/10.1176/appi.ajp.159.8.1354>
- Cathomas, F., Klaus, F., Guetter, K., Seifritz, E., Hartmann-Riemer, M.N., Tobler, P.N., Kaiser, S., Kaliuzhna, M., 2021. . *Schizophrenia Bulletin Open* 2, sgab022. <https://doi.org/10.1093/schizbullopen/sgab022>
- Cella, M., Dymond, S., Cooper, A., 2010. Impaired flexible decision-making in major depressive disorder. *Journal of Affective Disorders* 124, 207–210. <https://doi.org/10.1016/j.jad.2009.11.013>
- Charpentier, C.J., Aylward, J., Roiser, J.P., Robinson, O.J., 2017. Enhanced Risk Aversion, But Not Loss Aversion, in Unmedicated Pathological Anxiety. *Biological Psychiatry* 81, 1014–1022. <https://doi.org/10.1016/j.biopsych.2016.12.010>
- Chase, Frank, M.J., Michael, A., Bullmore, E.T., Sahakian, B.J., Robbins, T.W., 2010a. Approach and avoidance learning in patients with major depression and healthy controls: Relation to anhedonia. *Psychological Medicine* 40, 433–440. <https://doi.org/10.1017/S0033291709990468>
- Chase, Michael, A., Bullmore, E., Sahakian, B., Robbins, T., 2010b. Paradoxical enhancement of choice reaction time performance in patients with major depression. *Journal of Psychopharmacology* 24, 471–479. <https://doi.org/10.1177/0269881109104883>
- Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., Kusumi, I., 2015. Reinforcement learning in depression: A review of computational research. *Neuroscience & Biobehavioral Reviews* 55, 247–267. <https://doi.org/10.1016/j.neubiorev.2015.05.005>
- Chen, Y.C.I., Choi, J.K., Andersen, S.L., Rosen, B.R., Jenkins, B.G., 2005. Mapping dopamine D2/D3 receptor function using pharmacological magnetic resonance imaging. *Psychopharmacology* 180, 705–715. <https://doi.org/10.1007/s00213-004-2034-0>
- Chernoloz, O., El Mansari, M., Blier, P., 2012. Long-term administration of the dopamine D3/2 receptor agonist pramipexole increases dopamine and serotonin neurotransmission in the male rat forebrain. *Journal of Psychiatry and Neuroscience* 37, 113–121. <https://doi.org/10.1503/jpn.110038>
- Chernoloz, O., Mansari, M. El, Blier, P., 2008. Sustained Administration of Pramipexole Modifies the Spontaneous Firing of Dopamine, Norepinephrine, and Serotonin Neurons in the Rat Brain. *Neuropsychopharmacology* 2009 34:3 34, 651–661. <https://doi.org/10.1038/npp.2008.114>
- Chew, B., Blain, B., Dolan, R.J., Rutledge, R.B., 2021. A Neurocomputational Model for Intrinsic Reward. *J Neurosci* 41, 8963–8971. <https://doi.org/10.1523/JNEUROSCI.0858-20.2021>
- Chung, D., Kadlec, K., Aimone, J.A., McCurry, K., King-Casas, B., Chiu, P.H., 2017. Valuation in major depression is intact and stable in a non-learning environment. *Scientific reports* 7, 44374. <https://doi.org/10.1038/srep44374>
- Clark, L., Dombrowski, A.Y., Siegle, G.J., Butters, M.A., Shollenberger, C.L., Sahakian, B.J., Szanto, K., 2011. Impairment in risk-sensitive decision-making in older suicide attempters with depression. *Psychology and Aging* 26, 321–330. <https://doi.org/10.1037/a0021646>
- Cléry-Melin, M.-L., Schmidt, L., Lafargue, G., Baup, N., Fossati, P., Pessiglione, M., 2011. Why Don't You Try Harder? An Investigation of Effort Production in Major Depression. *PLoS ONE* 6, e23178. <https://doi.org/10.1371/journal.pone.0023178>
- Coizet, V., Dommert, E.J., Redgrave, P., Overton, P.G., 2006. Nociceptive responses of midbrain dopaminergic neurones are modulated by the superior colliculus in the rat. *Neuroscience* 139, 1479–1493. <https://doi.org/10.1016/j.neuroscience.2006.01.030>
- Collins, A.G.E., Frank, M.J., 2018. Within- and across-trial dynamics of human EEG reveal cooperative interplay between reinforcement learning and working memory. *Proceedings of the National Academy of Sciences* 115, 2502–2507. <https://doi.org/10.1073/PNAS.1720963115>

- Collins, A.G.E., Frank, M.J., 2012. How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience* 35, 1024–1035. <https://doi.org/10.1111/j.1460-9568.2011.07980.x>
- Corrigan, M.H., Denahan, A.Q., Wright, C.E., Ragual, R.J., Evans, D.L., 2000. Comparison of pramipexole, fluoxetine, and placebo in patients with major depression. *Depress Anxiety* 11, 58–65. [https://doi.org/10.1002/\(sici\)1520-6394\(2000\)11:2<58::aid-da2>3.0.co;2-h](https://doi.org/10.1002/(sici)1520-6394(2000)11:2<58::aid-da2>3.0.co;2-h)
- Costa, R., Oliveira, N.G., Dinis-Oliveira, R.J., 2019. Pharmacokinetic and pharmacodynamic of bupropion: integrative overview of relevant clinical and forensic aspects. *Drug Metabolism Reviews* 51, 293–313. <https://doi.org/10.1080/03602532.2019.1620763>
- Cusin, N, I., DV, I., AA, N., M, F., AJ, R., RH, P., 2013. A randomized, double-blind, placebo-controlled trial of pramipexole augmentation in treatment-resistant major depressive disorder. *The Journal of clinical psychiatry* 74, e636–e641. <https://doi.org/10.4088/JCP.12M08093>
- Cuthbert, B.N., 2014. The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World psychiatry : official journal of the World Psychiatric Association (WPA)* 13, 28–35. <https://doi.org/10.1002/wps.20087>
- Deisenhammer, E.A., Schmid, S.K., Kemmler, G., Moser, B., Delazer, M., 2018. Decision making under risk and under ambiguity in depressed suicide attempters, depressed non-attempters and healthy controls. *Journal of Affective Disorders* 226, 261–266. <https://doi.org/10.1016/j.jad.2017.10.012>
- DelDonno, S.R., Karstens, A.J., Cerny, B., Kling, L.R., Jenkins, L.M., Stange, J.P., Nusslock, R., Shankman, S.A., Langenecker, S.A., 2019a. The Titrated Monetary Incentive Delay Task: Sensitivity, convergent and divergent validity, and neural correlates in an RDoC sample. *Journal of Clinical and Experimental Neuropsychology* 41, 512–529. <https://doi.org/10.1080/13803395.2019.1585519>
- DelDonno, S.R., Mickey, B.J., Pruitt, P.J., Stange, J.P., Hsu, D.T., Weldon, A.L., Zubieta, J.K., Langenecker, S.A., 2019b. Influence of childhood adversity, approach motivation traits, and depression on individual differences in brain activation during reward anticipation. *Biological Psychology* 146. <https://doi.org/10.1016/j.biopsycho.2019.05.009>
- DelDonno, S.R., Weldon, A.L., Crane, N.A., Passarotti, A.M., Pruitt, P.J., Gabriel, L.B., Yau, W., Meyers, K.K., Hsu, D.T., Taylor, S.F., Heitzeg, M.M., Herbener, E., Shankman, S.A., Mickey, B.J., Zubieta, J.-K., Langenecker, S.A., 2015. Affective personality predictors of disrupted reward learning and pursuit in major depressive disorder. *Psychiatry Research* 230, 56–64. <https://doi.org/10.1016/j.psychres.2015.08.011>
- Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021>
- Dezfouli, A., Griffiths, K., Ramos, F., Dayan, P., Balleine, B.W., 2019. Models that learn how humans learn: The case of decision-making and its disorders. *PLoS Computational Biology* 15. <https://doi.org/10.1371/journal.pcbi.1006903>
- Dillon, D.G., Wiecki, T., Pechtel, P., Webb, C., Goer, F., Murray, L., Trivedi, M., Fava, M., McGrath, P.J., Weissman, M., Parsey, R., Kurian, B., Adams, P., Carmody, T., Weyandt, S., Shores-Wilson, K., Toups, M., McInnis, M., Oquendo, M.A., Cusin, C., Deldin, P., Bruder, G., Pizzagalli, D.A., 2015. A Computational Analysis of Flanker Interference in Depression. *Psychol Med* 45, 2333–2344. <https://doi.org/10.1017/S0033291715000276>
- Dombrovski, A.Y., Clark, L., Siegle, G.J., Butters, M.A., Ichikawa, N., Sahakian, B., Szanto, K., 2010. REWARD/PUNISHMENT REVERSAL LEARNING IN OLDER SUICIDE ATTEMPTERS. *The American journal of psychiatry* 167, 699. <https://doi.org/10.1176/appi.ajp.2009.09030407>
- Dombrovski, A.Y., Siegle, G.J., Szanto, K., Clark, L., Reynolds, C.F., Aizenstein, H., 2012. The temptation of suicide: striatal gray matter, discounting of delayed rewards, and suicide

- attempts in late-life depression. *Psychological Medicine* 42, 1203–1215.  
<https://doi.org/10.1017/S0033291711002133>
- Dong, Z., Han, H., Li, H., Bai, Y., Wang, W., Tu, M., Peng, Y., Zhou, L., He, W., Wu, Xiaobin, Tan, T., Liu, M., Wu, Xiaoyan, Zhou, W., Jin, W., Zhang, S., Sacktor, T.C., Li, T., Song, W., Wang, Y.T., 2015. Long-term potentiation decay and memory loss are mediated by AMPAR endocytosis. *J Clin Invest* 125, 234–247. <https://doi.org/10.1172/JCI77888>
- Drevets, W.C., 1999. Prefrontal Cortical-Amygdalar Metabolism in Major Depression. *Annals of the New York Academy of Sciences* 877, 614–637. <https://doi.org/10.1111/j.1749-6632.1999.tb09292.x>
- Eckblad, M., Chapman, L.J., 1986. Development and validation of a scale for hypomanic personality. *Journal of abnormal psychology* 95 3, 214–22.
- Eckstein, M.K., Master, S.L., Xia, L., Dahl, R.E., Wilbrecht, L., Collins, A.G.E., 2021. Learning Rates Are Not All the Same: The Interpretation of Computational Model Parameters Depends on the Context. <https://doi.org/10.1101/2021.05.28.446162>
- Eisenegger, C., Naef, M., Linssen, A., Clark, L., Gandamaneni, P.K., Müller, U., Robbins, T.W., 2014. Role of dopamine D2 receptors in human reinforcement learning. *Neuropsychopharmacology* 39, 2366–2375. <https://doi.org/10.1038/npp.2014.84>
- Eldar, E., Niv, Y., 2015. Interaction between emotional state and learning underlies mood instability. *Nat Commun* 6, 6149. <https://doi.org/10.1038/ncomms7149>
- Eldar, E., Rutledge, R.B., Dolan, R.J., Niv, Y., 2016. Mood as Representation of Momentum. *Trends in Cognitive Sciences* 20, 15–24. <https://doi.org/10.1016/J.TICS.2015.07.010>
- Eldesouky, L., Thompson, R.J., Oltmanns, T.F., English, T., 2018. Affective Instability Predicts the Course of Depression in Late Middle-Age and Older Adulthood. *J Affect Disord* 239, 72–78. <https://doi.org/10.1016/j.jad.2018.06.038>
- Elias, E., Zhang, A.Y., Manners, M.T., 2022. Novel Pharmacological Approaches to the Treatment of Depression. *Life* 12, 196. <https://doi.org/10.3390/life12020196>
- Elliott, R., Dolan, R.J., Frith, C.D., 2000. Dissociable Functions in the Medial and Lateral Orbitofrontal Cortex: Evidence from Human Neuroimaging Studies. *Cerebral Cortex* 10, 308–317. <https://doi.org/10.1093/cercor/10.3.308>
- Elliott, R., Frith, C.D., Dolan, R.J., 1997a. Differential neural response to positive and negative feedback in planning and guessing tasks. *Neuropsychologia* 35, 1395–1404. [https://doi.org/10.1016/S0028-3932\(97\)00055-9](https://doi.org/10.1016/S0028-3932(97)00055-9)
- Elliott, R., Sahakian, B.J., Herrod, J.J., Robbins, T.W., Paykel, E.S., 1997b. Abnormal response to negative feedback in unipolar depression: Evidence for a diagnosis specific impairment. *Journal of Neurology Neurosurgery and Psychiatry* 63, 74–82. <https://doi.org/10.1136/jnnp.63.1.74>
- Eshel, N., Roiser, J.P., 2010. Reward and Punishment Processing in Depression. *BIOL PSYCHIATRY* 68, 118–124. <https://doi.org/10.1016/j.biopsych.2010.01.027>
- Fawcett, J., Rush, A.J., Vukelich, J., Diaz, S.H., Dunklee, L., Romo, P., Yarns, B.C., Escalona, R., 2016. Clinical experience with high-dosage pramipexole in patients with treatment-resistant depressive episodes in unipolar and bipolar depression. *American Journal of Psychiatry* 173, 107–111. <https://doi.org/10.1176/appi.ajp.2015.15060788>
- Ferenczi, E.A., Zalocusky, K.A., Liston, C., Grosenick, L., Warden, M.R., Amatya, D., Katovich, K., Mehta, H., Patenaude, B., Ramakrishnan, C., Kalanithi, P., Etkin, A., Knutson, B., Glover, G.H., Deisseroth, K., 2016. Prefrontal cortical regulation of brainwide circuit dynamics and reward-related behavior. *Science* 351. <https://doi.org/10.1126/science.aac9698>
- Ferrari, A.J., Charlson, F.J., Norman, R.E., Patten, S.B., Freedman, G., Murray, C.J.L., Vos, T., Whiteford, H.A., 2013. Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. *PLoS Medicine* 10, e1001547. <https://doi.org/10.1371/journal.pmed.1001547>

- Ferrario, C.R., Li, X., Wolf, M.E., 2011. Effects of acute cocaine or dopamine receptor agonists on AMPA receptor distribution in the rat nucleus accumbens. *Synapse (New York, N.Y.)* 65, 54. <https://doi.org/10.1002/SYN.20823>
- Forbes, E.E., May, J.C., Siegle, G.J., Ladouceur, C.D., Ryan, N.D., Carter, C.S., Birmaher, B., Axelson, D.A., Dahl, R.E., 2006. Reward-Related Decision-Making in Pediatric Major Depressive Disorder: An fMRI Study. *J Child Psychol Psychiatry* 47, 1031–1040. <https://doi.org/10.1111/j.1469-7610.2006.01673.x>
- Foti, D., Carlson, J.M., Sauder, C.L., Proudfit, G.H., 2014. Reward dysfunction in major depression: Multimodal neuroimaging evidence for refining the melancholic phenotype. *Neuroimage* 101, 50–58. <https://doi.org/10.1016/j.neuroimage.2014.06.058>
- Fouragnan, E., Retzler, C., Philiastides, M.G., 2018. Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. *Hum Brain Mapp* 39, 2887–2906. <https://doi.org/10.1002/hbm.24047>
- Franco-Chaves, J.A., Mateus, C.F., Luckenbaugh, D.A., Martinez, P.E., Mallinger, A.G., Zarate, C.A., 2013. Combining a dopamine agonist and selective serotonin reuptake inhibitor for the treatment of depression: a double-blind, randomized pilot study. *J Affect Disord* 149, 319–325. <https://doi.org/10.1016/j.jad.2013.02.003>
- Frank, M.J., Claus, E.D., 1996. Anatomy of a Decision: Striato-Orbitofrontal Interactions in Reinforcement Learning, Decision Making and Reversal. Beiser & Houk.
- Gagne, C., Zika, O., Dayan, P., Bishop, S.J., 2020. Impaired adaptation of learning to contingency volatility in internalizing psychopathology. *eLife* 9, e61387. <https://doi.org/10.7554/eLife.61387>
- Gallant, H., Vo, A., Seergobin, K.N., MacDonald, P.A., 2016. Pramipexole impairs stimulus-response learning in healthy young adults. *Frontiers in Neuroscience* 10. <https://doi.org/10.3389/fnins.2016.00374>
- Gard, D.E., Gard, M.G., Kring, A.M., John, O.P., 2006. Anticipatory and consummatory components of the experience of pleasure: A scale development study. *Journal of Research in Personality* 40, 1086–1102. <https://doi.org/10.1016/j.jrp.2005.11.001>
- Gerlach, M., Double, K., Arzberger, T., Leblhuber, F., Tatschner, T., Riederer, P., 2003. Dopamine receptor agonists in current clinical use: comparative dopamine receptor binding profiles defined in the human striatum. *J Neural Transm* 110, 1119–1127. <https://doi.org/10.1007/s00702-003-0027-5>
- Glimcher, P.W., 2011. Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 108, 15647–15654. <https://doi.org/10.1073/pnas.1014269108>
- Goldberg, J.F., Burdick, K.E., Endick, C.J., 2004. Preliminary randomized, double-blind, placebo-controlled trial of pramipexole added to mood stabilizers for treatment-resistant bipolar depression. *Am J Psychiatry* 161, 564–566. <https://doi.org/10.1176/appi.ajp.161.3.564>
- Grace, A.A., 1991. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia. *Neuroscience* 41, 1–24. [https://doi.org/10.1016/0306-4522\(91\)90196-U](https://doi.org/10.1016/0306-4522(91)90196-U)
- Gradin, V.B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., Steele, J.D., 2011. Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* 134, 1751–1764. <https://doi.org/10.1093/brain/awr059>
- Greenberg, T., Chase, H.W., Almeida, J.R., Stiffler, R., Zevallos, C.R., Aslam, H.A., Deckersbach, T., Weyandt, S., Cooper, C., Touns, M., Carmody, T., Kurian, B., Peltier, S., Adams, P., McInnis, M.G., Oquendo, M.A., McGrath, P.J., Fava, M., Weissman, M., Parsey, R., Trivedi, M.H., Phillips, M.L., 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.

- <https://doi.org/10.1176/appi.ajp.2015.14050594> 172, 881–891.  
<https://doi.org/10.1176/APPI.AJP.2015.14050594>
- Gu, Y.-T., Zhou, C., Yang, J., Zhang, Q., Zhu, G.-H., Sun, L., Ge, M.-H., Wang, Y.-Y., 2020. A transdiagnostic comparison of affective decision-making in patients with schizophrenia, major depressive disorder, or bipolar disorder. *Psych J* 9, 199–209.  
<https://doi.org/10.1002/pchj.351>
- Guarraci, F.A., Kapp, B.S., 1999. An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. *Behavioural Brain Research* 99, 169–179. [https://doi.org/10.1016/S0166-4328\(98\)00102-8](https://doi.org/10.1016/S0166-4328(98)00102-8)
- Haber, S.N., Knutson, B., 2010. The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology*. <https://doi.org/10.1038/npp.2009.129>
- Hägele, C., Schlagenhaut, F., Rapp, M., Sterzer, P., Beck, A., Bermpohl, F., Stoy, M., Ströhle, A., Wittchen, H.-U., Dolan, R.J., Heinz, A., 2015. Dimensional psychiatry: reward dysfunction and depressive mood across psychiatric disorders. *Psychopharmacology (Berl)* 232, 331–341.  
<https://doi.org/10.1007/s00213-014-3662-7>
- Halahakoon, D.C., Kieslich, K., O’Driscoll, C., Nair, A., Lewis, G., Roiser, J.P., 2020. Reward-Processing Behavior in Depressed Participants Relative to Healthy Volunteers: A Systematic Review and Meta-analysis. *JAMA Psychiatry*. <https://doi.org/10.1001/jamapsychiatry.2020.2139>
- Hall, G.B.C., Milne, A.M.B., MacQueen, G.M., 2014. An fMRI study of reward circuitry in patients with minimal or extensive history of major depression. *European Archives of Psychiatry and Clinical Neuroscience* 264, 187–198. <https://doi.org/10.1007/s00406-013-0437-9>
- Hall, H., Halldin, C., Dijkstra, D., Wikström, H., Wise, L.D., Pugsley, T.A., Sokoloff, P., Pauli, S., Farde, L., Sedvall, G., 1996. Autoradiographic localisation of D3-dopamine receptors in the human brain using the selective D3-dopamine receptor agonist (+)-[3H]PD 128907. *Psychopharmacology* 128, 240–247. <https://doi.org/10.1007/s002130050131>
- Hall, Richard C.W., Platt, D.E., Hall, Ryan C.W., 1999. Suicide Risk Assessment: A Review of Risk Factors for Suicide in 100 Patients Who Made Severe Suicide Attempts. *Psychosomatics* 40, 18–27. [https://doi.org/10.1016/S0033-3182\(99\)71267-3](https://doi.org/10.1016/S0033-3182(99)71267-3)
- Harrison, N.A., Voon, V., Cercignani, M., Cooper, E.A., Pessiglione, M., Critchley, H.D., 2016. A Neurocomputational Account of How Inflammation Enhances Sensitivity to Punishments Versus Rewards. *Biological Psychiatry* 80, 73.  
<https://doi.org/10.1016/J.BIOPSYCH.2015.07.018>
- Harrison, P.J., Geddes, J.R., Tunbridge, E.M., 2018. The Emerging Neurobiology of Bipolar Disorder. *Trends in Neurosciences* 41, 18–30. <https://doi.org/10.1016/j.tins.2017.10.006>
- He, Z., Zhang, D., Muhlert, N., Elliott, R., 2019. Neural substrates for anticipation and consumption of social and monetary incentives in depression. *Soc Cogn Affect Neurosci* 14, 815–826.  
<https://doi.org/10.1093/scan/nsz061>
- Hegedűs, K.M., Szkaliczki, A., Gál, B.I., Andó, B., Janka, Z., Álmos, P.Z., 2018. Decision-making performance of depressed patients within 72 h following a suicide attempt. *Journal of Affective Disorders* 235, 583–588. <https://doi.org/10.1016/j.jad.2018.04.082>
- Henriques, J.B., Davidson, R.J., 2000. Decreased responsiveness to reward in depression. *Cognition & Emotion* 14, 711–724. <https://doi.org/10.1080/02699930050117684>
- Henriques, J.B., Glowacki, J.M., Davidson, R.J., 1994. Reward fails to alter response bias in depression. *J Abnorm Psychol* 103, 460–466. <https://doi.org/10.1037//0021-843x.103.3.460>
- Herzallah, M.M., Moustafa, A.A., Natsheh, J.Y., Abdellatif, S.M., Taha, M.B., Tayem, Y.I., Sehwal, M.A., Amleh, I., Petrides, G., Myers, C.E., Gluck, M.A., 2013. Learning from negative feedback in patients with major depressive disorder is attenuated by SSRI antidepressants. *Frontiers in Integrative Neuroscience* 7, 67. <https://doi.org/10.3389/fnint.2013.00067>
- Höflich, A., Michenthaler, P., Kasper, S., Lanzenberger, R., 2019. Circuit Mechanisms of Reward, Anhedonia, and Depression. *Int J Neuropsychopharmacol* 22, 105–118.  
<https://doi.org/10.1093/ijnp/pyy081>

- Hoge, R.D., Atkinson, J., Gill, B., Crelier, G.R., Marrett, S., Pike, G.B., 1999. Linear coupling between cerebral blood flow and oxygen consumption in activated human cortex. *Proceedings of the National Academy of Sciences* 96, 9403–9408. <https://doi.org/10.1073/pnas.96.16.9403>
- Husain, M., Roiser, J.P., 2018. Neuroscience of apathy and anhedonia: a transdiagnostic approach. *Nature Reviews Neuroscience* 19, 470–484. <https://doi.org/10.1038/s41583-018-0029-9>
- Husen, K., Rafaeli, E., Rubel, J.A., Bar-Kalifa, E., Lutz, W., 2016. Daily affect dynamics predict early response in CBT: Feasibility and predictive validity of EMA for outpatient psychotherapy. *J Affect Disord* 206, 305–314. <https://doi.org/10.1016/j.jad.2016.08.025>
- Huys, Q.J., Pizzagalli, D.A., Bogdan, R., Dayan, P., 2013. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of Mood & Anxiety Disorders* 3, 12. <https://doi.org/10.1186/2045-5380-3-12>
- Huys, Q.J.M., Browning, M., in press. A Computational View on the Nature of Reward and Value in Anhedonia, in: *Anhedonia: Preclinical, Translational, and Clinical Integration, Current Topics in Behavioral Neurosciences*.
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* 17, 825–841. [https://doi.org/10.1016/S1053-8119\(02\)91132-8](https://doi.org/10.1016/S1053-8119(02)91132-8)
- Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* 5, 143–156. [https://doi.org/10.1016/S1361-8415\(01\)00036-6](https://doi.org/10.1016/S1361-8415(01)00036-6)
- Johnston, B.A., Tolomeo, S., Gradin, V., Christmas, D., Matthews, K., Douglas Steele, J., 2015. Failure of hippocampal deactivation during loss events in treatment-resistant depression. *Brain* 138, 2766–2776. <https://doi.org/10.1093/brain/awv177>
- Jollant, F., Bellivier, F., Leboyer, M., Astruc, B., Torres, S., Verdier, R., Castelnau, D., Malafosse, A., Courtet, P., 2005. Impaired Decision Making in Suicide Attempters. *American Journal of Psychiatry* 162, 304–310. <https://doi.org/10.1176/appi.ajp.162.2.304>
- Jollant, F., Richard-Devantoy, S., Ding, Y., Turecki, G., Bechara, A., Near, J., 2016. Prefrontal inositol levels and implicit decision-making in healthy individuals and depressed patients. *European Neuropsychopharmacology* 26, 1255–1263. <https://doi.org/10.1016/j.euroneuro.2016.06.005>
- Kanen, J.W., Ersche, K.D., Fineberg, N.A., Robbins, T.W., Cardinal, R.N., 2019. Computational modelling reveals contrasting effects on reinforcement learning and cognitive flexibility in stimulant use disorder and obsessive-compulsive disorder: remediating effects of dopaminergic D2/3 receptor agents. *Psychopharmacology* 236, 2337–2358. <https://doi.org/10.1007/s00213-019-05325-w>
- Keren, H., O’Callaghan, G., Vidal-Ribas, P., Buzzell, G.A., Brotman, M.A., Leibenluft, E., Pan, P.M., Meffert, L., Kaiser, A., Wolke, S., Pine, D.S., Stringaris, A., 2018. Reward Processing in Depression: A Conceptual and Meta-Analytic Review Across fMRI and EEG Studies. *American Journal of Psychiatry* appi.ajp.2018.1. <https://doi.org/10.1176/appi.ajp.2018.17101124>
- Keren, H., Zheng, C., Jangraw, D.C., Chang, K., Vitale, A., Rutledge, R.B., Pereira, F., Nielson, D.M., Stringaris, A., 2022. The temporal representation of experience in subjective mood. *eLife* 10, e62051. <https://doi.org/10.7554/eLife.62051>
- Kim, H., Shimojo, S., O’Doherty, J.P., 2006. Is Avoiding an Aversive Outcome Rewarding? Neural Substrates of Avoidance Learning in the Human Brain. *PLoS Biol* 4, e233. <https://doi.org/10.1371/journal.pbio.0040233>
- Knutson, B., Bhanji, J.P., Cooney, R.E., Atlas, L.Y., Gotlib, I.H., 2008. Neural Responses to Monetary Incentives in Major Depression. *Biological Psychiatry* 63, 686–692. <https://doi.org/10.1016/j.biopsych.2007.07.023>
- Knutson, B., Fong, G.W., Bennett, S.M., Adams, C.M., Hommer, D., 2003. A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *NeuroImage* 18, 263–272. [https://doi.org/10.1016/S1053-8119\(02\)00057-5](https://doi.org/10.1016/S1053-8119(02)00057-5)

- Knutson, B., Westdorp, A., Kaiser, E., Hommer, D., 2000. FMRI Visualization of Brain Activity during a Monetary Incentive Delay Task. *NeuroImage* 12, 20–27.  
<https://doi.org/10.1006/NIMG.2000.0593>
- Kolla, B., Mansukhani, M., R, B., JM, B., 2010. Impact of dopamine agonists on compulsive behaviors: a case series of pramipexole-induced pathological gambling. *Psychosomatics* 51, 271–273.  
<https://doi.org/10.1176/APPI.PSY.51.3.271>
- Koolschijn, P.C.M.P., van Haren, N.E.M., Lensvelt-Mulders, G.J.L.M., Hulshoff Pol, H.E., Kahn, R.S., 2009. Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Hum Brain Mapp* 30, 3719–3735.  
<https://doi.org/10.1002/hbm.20801>
- Kouli, A., Torsney, K.M., Kuan, W.-L., 2018. Parkinson’s Disease: Etiology, Neuropathology, and Pathogenesis, *Parkinson’s Disease: Pathogenesis and Clinical Aspects* [Internet]. Codon Publications. <https://doi.org/10.15586/codonpublications.parkinsonsdisease.2018.ch1>
- Koval, P., Pe, M.L., Meers, K., Kuppens, P., 2013. Affect dynamics in relation to depressive symptoms: variable, unstable or inert? *Emotion* 13, 1132–1141. <https://doi.org/10.1037/a0033579>
- Kuhnen, C.M., Knutson, B., 2005. The Neural Basis of Financial Risk Taking. *Neuron* 47, 763–770.  
<https://doi.org/10.1016/j.neuron.2005.08.008>
- Kumar, P., Goer, F., Murray, L., Dillon, D.G., Beltzer, M.L., Cohen, A.L., Brooks, N.H., Pizzagalli, D.A., 2018. Impaired reward prediction error encoding and striatal-midbrain connectivity in depression. *Neuropsychopharmacology* 43, 1581–1588. <https://doi.org/10.1038/s41386-018-0032-x>
- Kumar, P., Slavich, G.M., Berghorst, L.H., Treadway, M.T., Brooks, N.H., Dutra, S.J., Greve, D.N., O’Donovan, A., Bleil, M.E., Maninger, N., Pizzagalli, D.A., 2015. Perceived Chronic Stress Exposure Modulates Reward-Related Medial Prefrontal Cortex Responses to Acute Stress in Depression. *J Affect Disord* 180, 104–111. <https://doi.org/10.1016/j.jad.2015.03.035>
- Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., Steele, J.D., 2008. Abnormal temporal difference reward-learning signals in major depression. *Brain* 131, 2084–2093.  
<https://doi.org/10.1093/brain/awn136>
- Kunisato, Y., Okamoto, Y., Ueda, K., Onoda, K., Okada, G., Yoshimura, S., Suzuki, S., Samejima, K., Yamawaki, S., 2012. Effects of depression on reward-based decision making and variability of action in probabilistic learning. *J Behav Ther Exp Psychiatry* 43, 1088–1094.  
<https://doi.org/10.1016/j.jbtep.2012.05.007>
- Kuppens, P., Oravecz, Z., Tuerlinckx, F., 2010. Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology* 99, 1042–1060. <https://doi.org/10.1037/a0020962>
- Kuppens, P., Sheeber, L.B., Yap, M.B.H., Whittle, S., Simmons, J.G., Allen, N.B., 2012. Emotional inertia prospectively predicts the onset of depressive disorder in adolescence. *Emotion* 12, 283–289. <https://doi.org/10.1037/a0025046>
- Lau, B., Glimcher, P.W., 2007. Action and outcome encoding in the primate caudate nucleus. *J Neurosci* 27, 14502–14514. <https://doi.org/10.1523/JNEUROSCI.3060-07.2007>
- Lawlor, V.M., Webb, C.A., Wiecki, T.V., Frank, M.J., Trivedi, M., Pizzagalli, D.A., Dillon, D.G., 2019. Dissecting the impact of depression on decision-making. *Psychological Medicine*.  
<https://doi.org/10.1017/S0033291719001570>
- Lemke, M.R., Brecht, H.M., Koester, J., Reichmann, H., 2006. Effects of the dopamine agonist pramipexole on depression, anhedonia and motor functioning in Parkinson’s disease. *Journal of the Neurological Sciences* 248, 266–270. <https://doi.org/10.1016/j.jns.2006.05.024>
- Liu, Chan, R.C.K., zhi Wang, L., Huang, J., Cheung, E.F.C., yong Gong, Q., Gollan, J.K., 2011. Deficits in sustaining reward responses in subsyndromal and syndromal major depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 35, 1045–1052.  
<https://doi.org/10.1016/j.pnpbp.2011.02.018>

- Liu, Valton, V., Wang, L.-Z., Zhu, Y.-H., Roiser, J.P., 2017. Association between habenula dysfunction and motivational symptoms in unmedicated major depressive disorder. *Social Cognitive and Affective Neuroscience* 12, 1520–1533. <https://doi.org/10.1093/scan/nsx074>
- Lorenzetti, V., Allen, N.B., Fornito, A., Yücel, M., 2009. Structural brain abnormalities in major depressive disorder: A selective review of recent MRI studies. *Journal of Affective Disorders* 117, 1–17. <https://doi.org/10.1016/j.jad.2008.11.021>
- Mah, L., Zarate, C.A., Jr., Nugent, A.C., Singh, J.B., Manji, H.K., Drevets, W.C., 2011a. Neural mechanisms of antidepressant efficacy of the dopamine receptor agonist pramipexole in treatment of bipolar depression. *International Journal of Neuropsychopharmacology* 14, 545–551. <https://doi.org/10.1017/S1461145710001203>
- Mah, L., Zarate, C.A., Nugent, A.C., Singh, J.B., Manji, H.K., Drevets, W.C., 2011b. Neural mechanisms of antidepressant efficacy of the dopamine receptor agonist pramipexole in treatment of bipolar depression. *Int J Neuropsychopharmacol* 14, 545–551. <https://doi.org/10.1017/S1461145710001203>
- Mah, L., Zarate, C.A., Singh, J., Duan, Y.-F., Luckenbaugh, D.A., Manji, H.K., Drevets, W.C., 2007. Regional cerebral glucose metabolic abnormalities in bipolar II depression. *Biol Psychiatry* 61, 765–775. <https://doi.org/10.1016/j.biopsych.2006.06.009>
- Marr, D., 2010. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. The MIT Press. <https://doi.org/10.7551/mitpress/9780262514620.001.0001>
- Marr, D., Poggio, T., 1976. From Understanding Computation to Understanding Neural Circuitry.
- Martens, M.A.G., Kaltenboeck, A., Halahakoon, D.C., Browning, M., Cowen, P.J., Harmer, C.J., 2021. An Experimental Medicine Investigation of the Effects of Subacute Pramipexole Treatment on Emotional Information Processing in Healthy Volunteers. *Pharmaceuticals* 2021, Vol. 14, Page 800 14, 800. <https://doi.org/10.3390/PH14080800>
- Martin-Sölch, C., Magyar, S., König, G., Missimer, J., Schultz, W., Leenders, K., 2001. Changes in brain activation associated with reward processing in smokers and nonsmokers. *Exp Brain Res* 139, 278–286. <https://doi.org/10.1007/s002210100751>
- Marwaha, S., Broome, M.R., Bebbington, P.E., Kuipers, E., Freeman, D., 2014. Mood instability and psychosis: analyses of British national survey data. *Schizophr Bull* 40, 269–277. <https://doi.org/10.1093/schbul/sbt149>
- Marwaha, S., Parsons, N.R., Broome, M.R., 2013. Mood instability, mental illness and suicidal ideas : results from a household survey. *Social Psychiatry and Psychiatric Epidemiology* Volume 48, 1431–1437.
- Matsumoto, H., Tian, J., Uchida, N., Watabe-Uchida, M., 2016. Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *eLife* 5. <https://doi.org/10.7554/eLife.17328>
- Matsumoto, M., Hikosaka, O., 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459, 837–841. <https://doi.org/10.1038/nature08028>
- McCabe, C., Harwood, J., Brouwer, S., Harmer, C.J., Cowen, P.J., 2013. Effects of pramipexole on the processing of rewarding and aversive taste stimuli. *Psychopharmacology* 228, 283–290. <https://doi.org/10.1007/s00213-013-3033-9>
- McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal Prediction Errors in a Passive Learning Task Activate Human Striatum. *Neuron* 38, 339–346. [https://doi.org/10.1016/S0896-6273\(03\)00154-5](https://doi.org/10.1016/S0896-6273(03)00154-5)
- McClure, S.M., York, M.K., Montague, P.R., 2004. The neural substrates of reward processing in humans: the modern role of fMRI. *Neuroscientist* 10, 260–268. <https://doi.org/10.1177/1073858404263526>
- McGovern, A.R., Alexopoulos, G.S., Yuen, G.S., Morimoto, S.S., Gunning-Dixon, F.M., 2014. Reward-related decision making in older adults: relationship to clinical presentation of depression. *International journal of geriatric psychiatry* 29, 1125–31. <https://doi.org/10.1002/gps.4200>

- Michels, L., Scherpiet, S., Stämpfli, P., Herwig, U., Brühl, A.B., 2016. Baseline Perfusion Alterations Due to Acute Application of Quetiapine and Pramipexole in Healthy Adults. *Int J Neuropsychopharmacol* 19, pyw067. <https://doi.org/10.1093/ijnp/pyw067>
- Michely, J., Eldar, E., Martin, I.M., Dolan, R.J., 2020. A mechanistic account of serotonin's impact on mood. *Nat Commun* 11, 2335. <https://doi.org/10.1038/s41467-020-16090-2>
- Missale, C., Russel Nash, S., Robinson, S.W., Jaber, M., Caron, M.G., 1998. Dopamine receptors: From structure to function. *Physiological Reviews*. <https://doi.org/10.1152/physrev.1998.78.1.189>
- Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J.P., Robinson, O.J., 2017. Modeling Avoidance in Mood and Anxiety Disorders Using Reinforcement Learning. *Biol Psychiatry* 82, 532–539. <https://doi.org/10.1016/j.biopsych.2017.01.017>
- Moorman, D.E., Aston-Jones, G., 2014. Orbitofrontal Cortical Neurons Encode Expectation-Driven Initiation of Reward-Seeking. *The Journal of Neuroscience* 34, 10234. <https://doi.org/10.1523/JNEUROSCI.3216-13.2014>
- Moutoussis, M., Rutledge, R.B., Prabhu, G., Hryniewicz, L., Lam, J., Ousdal, O.T., Guitart-Masip, M., Fonagy, P., Dolan, R.J., 2018. Neural activity and fundamental learning, motivated by monetary loss and reward, are intact in mild to moderate major depressive disorder. *PLoS ONE* 13. <https://doi.org/10.1371/journal.pone.0201451>
- Mukherjee, D., Filipowicz, A.L.S., Vo, K., Satterthwaite, T.D., Kable, J.W., 2020. Reward and punishment reversal-learning in major depressive disorder. *J Abnorm Psychol* 129, 810–823. <https://doi.org/10.1037/abn0000641>
- Murphy, F.C., Rubinsztein, J.S., Michael, A., Rogers, R.D., Robbins, T.W., Paykel, E.S., Sahakian, B.J., 2001. Decision-making cognition in mania and depression. *Psychological medicine* 31, 679–93.
- Murphy, L., Ly, T., DiMario, K., Mihowich, T., Fedoroff, J.P., 2021. Treatment of Pramipexole-Induced Problematic Sexual Behaviors. *Prim Care Companion CNS Disord* 23, 0–0. <https://doi.org/10.4088/PCC.20br02663>
- Murray, G.K., Knolle, F., Ersche, K.D., Craig, K.J., Abbott, S., Shabbir, S.S., Fineberg, N.A., Suckling, J., Sahakian, B.J., Bullmore, E.T., Robbins, T.W., 2019. Dopaminergic drug treatment remediates exaggerated cingulate prediction error responses in obsessive-compulsive disorder. *Psychopharmacology* 236, 2325–2336. <https://doi.org/10.1007/s00213-019-05292-2>
- Must, A., Szabó, Z., Bódi, N., Szász, A., Janka, Z., Kéri, S., 2006. Sensitivity to reward and punishment and the prefrontal cortex in major depression. *Journal of Affective Disorders* 90, 209–215. <https://doi.org/10.1016/j.jad.2005.12.005>
- Ng, T.H., Alloy, L.B., Smith, D.V., 2019. Meta-analysis of reward processing in major depressive disorder reveals distinct abnormalities within the reward circuit. *Translational Psychiatry* 9, 293. <https://doi.org/10.1038/s41398-019-0644-x>
- Nord, C.L., Lawson, R.P., Huys, Q.J.M., Pilling, S., Roiser, J.P., 2018. Depression is associated with enhanced aversive Pavlovian control over instrumental behaviour. *Scientific Reports* 8. <https://doi.org/10.1038/s41598-018-30828-5>
- Oldham, S., Murawski, C., Fornito, A., Youssef, G., Yücel, M., Lorenzetti, V., 2018. The anticipation and outcome phases of reward and loss processing: A neuroimaging meta-analysis of the monetary incentive delay task. *Human Brain Mapping* 39, 3398–3418. <https://doi.org/10.1002/hbm.24184>
- O'Reardon, J.P., Amsterdam, J.D., 1998. Treatment-Resistant Depression: Progress and Limitations. *Psychiatric Annals* 28, 633–640. <https://doi.org/10.3928/0048-5713-19981101-08>
- Oya, H., Adolphs, R., Kawasaki, H., Bechara, A., Damasio, A., Howard, M.A., 2005. Electrophysiological correlates of reward prediction error recorded in the human prefrontal cortex. *Proc Natl Acad Sci U S A* 102, 8351–8356. <https://doi.org/10.1073/pnas.0500899102>
- Peciña, S., Berridge, K.C., 2005. Hedonic hot spot in nucleus accumbens shell: where do mu-opioids cause increased hedonic impact of sweetness? *J Neurosci* 25, 11777–11786. <https://doi.org/10.1523/JNEUROSCI.2329-05.2005>

- Peeters, F., Berkhof, J., Rottenberg, J., Nicolson, N.A., 2010. Ambulatory emotional reactivity to negative daily life events predicts remission from major depressive disorder. *Behav Res Ther* 48, 754–760. <https://doi.org/10.1016/j.brat.2010.04.008>
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045. <https://doi.org/10.1038/nature05051>
- Peters, E.M., Balbuena, L., Marwaha, S., Baetz, M., Bowen, R., 2016. Mood instability and impulsivity as trait predictors of suicidal thoughts. *Psychol Psychother* 89, 435–444. <https://doi.org/10.1111/papt.12088>
- Pezze, M.A., Dalley, J.W., Robbins, T.W., 2007. Differential roles of dopamine D1 and D2 receptors in the nucleus accumbens in attentional performance on the five-choice serial reaction time task. *Neuropsychopharmacology* 32, 273–283. <https://doi.org/10.1038/sj.npp.1301073>
- Piercey, M.F., 1998. Pharmacology of pramipexole, a dopamine D3-preferring agonist useful in treating Parkinson's disease. *Clin Neuropharmacol* 21, 141–151.
- Piercey, M.F., Hoffmann, W.E., Smith, M.W., Hyslop, D.K., 1996. Inhibition of dopamine neuron firing by pramipexole, a dopamine D3 receptor-preferring agonist: Comparison to other dopamine receptor agonists. *European Journal of Pharmacology* 312, 35–44. [https://doi.org/10.1016/0014-2999\(96\)00454-2](https://doi.org/10.1016/0014-2999(96)00454-2)
- Pike, A.C., Robinson, O.J., 2022. Reinforcement Learning in Patients With Mood and Anxiety Disorders vs Control Individuals: A Systematic Review and Meta-analysis. *JAMA Psychiatry* 79, 313–322. <https://doi.org/10.1001/jamapsychiatry.2022.0051>
- Pizzagalli, D.A., Jahn, A.L., O'Shea, J.P., 2005. Toward an objective characterization of an anhedonic phenotype: a signal-detection approach. *Biol Psychiatry* 57, 319–327. <https://doi.org/10.1016/j.biopsych.2004.11.026>
- Pizzagalli, Evins, A.E., Schetter, E.C., Frank, M.J., Pajtas, P.E., Santesso, D.L., Culhane, M., 2008a. Single dose of a dopamine agonist impairs reinforcement learning in humans: Behavioral evidence from a laboratory-based measure of reward responsiveness. *Psychopharmacology* 196, 221–232. <https://doi.org/10.1007/s00213-007-0957-y>
- Pizzagalli, Holmes, A.J., Dillon, D.G., Goetz, E.L., Birk, J.L., Bogdan, R., Dougherty, D.D., Iosifescu, D.V., Rauch, S.L., Fava, M., 2009. Reduced Caudate and Nucleus Accumbens Response to Rewards in Unmedicated Individuals With Major Depressive Disorder. *American Journal of Psychiatry* 166, 702–710. <https://doi.org/10.1176/appi.ajp.2008.08081201>
- Pizzagalli, Iosifescu, D., Hallett, L.A., Ratner, K.G., Fava, M., 2008b. Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task. *Journal of psychiatric research* 43, 76–87. <https://doi.org/10.1016/j.jpsychires.2008.03.001>
- Pulcu, E., Browning, M., 2019. The Misestimation of Uncertainty in Affective Disorders. *Trends in Cognitive Sciences* 23, 865–875. <https://doi.org/10.1016/j.tics.2019.07.007>
- Pulcu, E., Saunders, K.E.A., Harmer, C.J., Harrison, P.J., Goodwin, G.M., Geddes, J.R., Browning, M., 2022. Characterizing Affective Variability in Bipolar Disorder and Borderline Personality Disorder, and the Effects of Lithium, Using a Generative Model of Affect. <https://doi.org/10.1101/2022.02.18.22271166>
- Radloff, L.S., 1977. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement* 1, 385–401. <https://doi.org/10.1177/014662167700100306>
- Redlich, R., Dohm, K., Grotegerd, D., Opel, N., Zwieterlood, P., Heindel, W., Arolt, V., Kugel, H., Dannlowski, U., 2015. Reward Processing in Unipolar and Bipolar Depression: A Functional MRI Study. *Neuropsychopharmacology* 40, 2623–2631. <https://doi.org/10.1038/npp.2015.110>
- Rees, G., Friston, K., Koch, C., 2000. A direct quantitative relationship between the functional properties of human and macaque V5. *Nat Neurosci* 3, 716–723. <https://doi.org/10.1038/76673>

- Reilly, E.E., Whitton, A.E., Pizzagalli, D.A., Rutherford, A.V., Stein, M.B., Paulus, M.P., Taylor, C.T., 2020. Diagnostic and dimensional evaluation of implicit reward learning in social anxiety disorder and major depression. *Depress Anxiety* 37, 1221–1230. <https://doi.org/10.1002/da.23081>
- Reinen, J.M., Whitton, A.E., Pizzagalli, D.A., Slifstein, M., Abi-Dargham, A., McGrath, P.J., Iosifescu, D.V., Schneier, F.R., 2021. Differential reinforcement learning responses to positive and negative information in unmedicated individuals with depression. *Eur Neuropsychopharmacol* 53, 89–100. <https://doi.org/10.1016/j.euroneuro.2021.08.002>
- Remijnse, P.L., Nielen, M.M.A., Balkom, A.J.L.M. van, Hendriks, G.-J., Hoogendijk, W.J., Uylings, H.B.M., Veltman, D.J., 2009. Differential frontal–striatal and paralimbic activity during reversal learning in major depressive disorder and obsessive–compulsive disorder. *Psychological Medicine* 39, 1503–1518. <https://doi.org/10.1017/S0033291708005072>
- Riba, J., Krämer, U.M., Heldmann, M., Richter, S., Münte, T.F., 2008. Dopamine Agonist Increases Risk Taking but Blunts Reward-Related Brain Activity. *PLoS ONE* 3, e2479. <https://doi.org/10.1371/journal.pone.0002479>
- Richelson, E., Souder, T., 2000. Binding of antipsychotic drugs to human brain receptors: Focus on newer generation compounds. *Life Sciences* 68, 29–39. [https://doi.org/10.1016/S0024-3205\(00\)00911-5](https://doi.org/10.1016/S0024-3205(00)00911-5)
- Rinaldi, R., Lefebvre, L., Joachim, A., Rossignol, M., 2020. Decision-making of patients with major depressive disorder in the framework of action control. *Cogn Neuropsychiatry* 25, 71–83. <https://doi.org/10.1080/13546805.2019.1685481>
- Robinson, O.J., Chase, H.W., 2017. Learning and Choice in Mood Disorders: Searching for the Computational Parameters of Anhedonia. *Comput Psychiatr* 1, 208–233. [https://doi.org/10.1162/CPSY\\_a\\_00009](https://doi.org/10.1162/CPSY_a_00009)
- Robinson, O.J., Cools, R., Carlisi, C.O., Sahakian, B.J., Drevets, W.C., 2012. Ventral Striatum Response During Reward and Punishment Reversal Learning in Unmedicated Major Depressive Disorder. *Am J Psychiatry* 169, 152–159. <https://doi.org/10.1176/appi.ajp.2011.11010137>
- Rock, P.L., Roiser, J.P., Riedel, W.J., Blackwell, A.D., 2014. Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological Medicine* 44, 2029–2040. <https://doi.org/10.1017/S0033291713002535>
- Rolls, E.T., 2021. Attractor cortical neurodynamics, schizophrenia, and depression. *Transl Psychiatry* 11, 215. <https://doi.org/10.1038/s41398-021-01333-7>
- Rolls, E.T., Cheng, W., Feng, J., 2020. The orbitofrontal cortex: reward, emotion and depression. *Brain Communications* 2, fcaa196. <https://doi.org/10.1093/braincomms/fcaa196>
- Rothkirch, M., Tonn, J., Köhler, S., Sterzer, P., 2017. Neural mechanisms of reinforcement learning in unmedicated patients with major depressive disorder. *Brain* 140, 1147–1157. <https://doi.org/10.1093/brain/awx025>
- Ruff, C.C., Fehr, E., 2014. The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience* 15, 549–562. <https://doi.org/10.1038/nrn3776>
- Rupprechter, S., Romaniuk, L., Series, P., Hirose, Y., Hawkins, E., Sandu, A.-L., Waiter, G.D., McNeil, C.J., Shen, X., Harris, M.A., Campbell, A., Porteous, D., Macfarlane, J.A., Lawrie, S.M., Murray, A.D., Delgado, M.R., McIntosh, A.M., Whalley, H.C., Steele, J.D., 2020. Blunted medial prefrontal cortico-limbic reward-related effective connectivity and depression. *Brain* 143, 1946–1956. <https://doi.org/10.1093/BRAIN/AWAA106>
- Rutledge, R.B., Dean, M., Caplin, A., Glimcher, P.W., 2010. Testing the Reward Prediction Error Hypothesis with an Axiomatic Model. *J Neurosci* 30, 13525–13536. <https://doi.org/10.1523/JNEUROSCI.1747-10.2010>
- Rutledge, R.B., Skandali, N., Dayan, P., Dolan, R.J., 2014. A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences* 111, 12252–12257. <https://doi.org/10.1073/pnas.1407535111>

- Sankar, A., Yttredahl, A.A., Fourcade, E.W., Mickey, B.J., Love, T.M., Langenecker, S.A., Hsu, D.T., 2019. Dissociable neural responses to monetary and social gain and loss in women with major depressive disorder. *Frontiers in Behavioral Neuroscience* 13. <https://doi.org/10.3389/fnbeh.2019.00149>
- Santesso, D.L., Evins, A.E., Frank, M.J., Schetter, E.C., Bogdan, R., Pizzagalli, D.A., 2009. Single dose of a dopamine agonist impairs reinforcement learning in humans: Evidence from event-related potentials and computational modeling of striatal-cortical function. *Human Brain Mapping* 30, 1963. <https://doi.org/10.1002/HBM.20642>
- Saperia, S., Da Silva, S., Siddiqui, I., Agid, O., Daskalakis, Z.J., Ravindran, A., Voineskos, A.N., Zakzanis, K.K., Remington, G., Foussias, G., 2019. Reward-driven decision-making impairments in schizophrenia. *Schizophrenia Research* 206, 277–283. <https://doi.org/10.1016/j.schres.2018.11.004>
- Satterthwaite, T.D., Kable, J.W., Vandekar, L., Katchmar, N., Bassett, D.S., Baldassano, C.F., Ruparel, K., Elliott, M.A., Sheline, Y.I., Gur, R.C., Gur, R.E., Davatzikos, C., Leibenluft, E., Thase, M.E., Wolf, D.H., 2015. Common and Dissociable Dysfunction of the Reward System in Bipolar and Unipolar Depression. *Neuropsychopharmacol* 40, 2258–2268. <https://doi.org/10.1038/npp.2015.75>
- Schoenbaum, G., Chiba, A.A., Gallagher, M., 1998. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neuroscience* 1:2 1, 155–159. <https://doi.org/10.1038/407>
- Schoenbaum, G., Roesch, M., 2005. Orbitofrontal Cortex, Associative Learning, and Expectancies. *Neuron* 47, 633–636. <https://doi.org/10.1016/j.neuron.2005.07.018>
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275, 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>
- Sinyor, M., Schaffer, A., Levitt, A., 2010. The Sequenced Treatment Alternatives to Relieve Depression (STAR\*D) Trial: A Review. *Can J Psychiatry* 55, 126–135. <https://doi.org/10.1177/070674371005500303>
- Siqueira, A.S.S., Biella, M.M., Borges, M.K., Mauer, S., Apolinario, D., Alves, T.C. de T.F., Jacob-Filho, W., Oude Voshaar, R.C., Arahamian, I., 2021. Decision-making executive function profile and performance in older adults with major depression: a case-control study. *Aging & Mental Health* 0, 1–7. <https://doi.org/10.1080/13607863.2021.1950617>
- Small, D.M., Gregory, M.D., Mak, Y.E., Gitelman, D., Mesulam, M.M., Parrish, T., 2003. Dissociation of Neural Representation of Intensity and Affective Valuation in Human Gustation. *Neuron* 39, 701–711. [https://doi.org/10.1016/S0896-6273\(03\)00467-7](https://doi.org/10.1016/S0896-6273(03)00467-7)
- Smith, D.V., Delgado, M.R., 2015. *catj* 3, 361–366. <https://doi.org/10.1016/B978---0---12---397025---1.00255---4>
- Smith, S.M., 2002. Fast robust automated brain extraction. *Human Brain Mapping* 17, 143–155. <https://doi.org/10.1002/hbm.10062>
- Smith, D. V., Hayden, B.Y., Truong, T.K., Song, A.W., Platt, M.L., Huettel, S.A., 2010. Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *Journal of Neuroscience* 30, 2490–2495. <https://doi.org/10.1523/JNEUROSCI.3319-09.2010>
- Smoski, M.J., Felder, J., Bizzell, J., Green, S.R., Ernst, M., Lynch, T.R., Dichter, G.S., 2009. fMRI of alterations in reward selection, anticipation, and feedback in major depressive disorder. *Journal of Affective Disorders* 118, 69–78. <https://doi.org/10.1016/j.jad.2009.01.034>
- Smoski, M.J., Rittenberg, A., Dichter, G.S., 2011. Major depressive disorder is characterized by greater reward network activation to monetary than pleasant image rewards. *Psychiatry Research: Neuroimaging* 194, 263–270. <https://doi.org/10.1016/j.psychresns.2011.06.012>
- Sokoloff, L., Reivich, M., Kennedy, C., Rosiers, M.H.D., Patlak, C.S., Pettigrew, K.D., Sakurada, O., Shinohara, M., 1977. The [14c]deoxyglucose Method for the Measurement of Local Cerebral Glucose Utilization: Theory, Procedure, and Normal Values in the Conscious and

- Anesthetized Albino Rat1. *Journal of Neurochemistry* 28, 897–916.  
<https://doi.org/10.1111/j.1471-4159.1977.tb10649.x>
- Sperry, S.H., Walsh, M.A., Kwapil, T.R., 2020. Emotion dynamics concurrently and prospectively predict mood psychopathology. *Journal of Affective Disorders* 261, 67–75.  
<https://doi.org/10.1016/j.jad.2019.09.076>
- Spielberger, C., Gorsuch, R., Lushene, R., Vagg, P., Jacobs, G., 1983. *Manual for the State-Trait Anxiety Inventory (Form Y1 – Y2)*, Palo Alto, CA: Consulting Psychologists Press;
- Spijker, J., Bijl, R.V., de Graaf, R., Nolen, W.A., 2001. Determinants of poor 1-year outcome of DSM-III-R major depression in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta psychiatrica Scandinavica* 103, 122–30.
- Steele, Kumar, P., Ebmeier, K.P., 2007. Blunted response to feedback information in depressive illness. *Brain* 130, 2367–2374. <https://doi.org/10.1093/brain/awm150>
- Stoy, M., Schlagenhauf, F., Sterzer, P., Berman, F., Hägele, C., Suchotzki, K., Schmack, K., Wrase, J., Ricken, R., Knutson, B., Adli, M., Bauer, M., Heinz, A., Ströhle, A., 2012. Hyporeactivity of ventral striatum towards incentive stimuli in unmedicated depressed patients normalizes after treatment with escitalopram. *J Psychopharmacol* 26, 677–688.  
<https://doi.org/10.1177/0269881111416686>
- Subramaniapillai, M., Mansur, R.B., Zuckerman, H., Park, C., Lee, Y., Iacobucci, M., Cao, B., Ho, R., Lin, K., Phan, L., McIntyre, R.S., 2019. Association between cognitive function and performance on effort based decision making in patients with major depressive disorder treated with Vortioxetine. *Comprehensive Psychiatry* 94.  
<https://doi.org/10.1016/j.comppsy.2019.07.006>
- Takamura, M., Okamoto, Y., Okada, G., Toki, S., Yamamoto, T., Ichikawa, N., Mori, A., Minagawa, H., Takaishi, Y., Fujii, Y., Kaichi, Y., Akiyama, Y., Awai, K., Yamawaki, S., 2017. Patients with major depressive disorder exhibit reduced reward size coding in the striatum. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 79, 317–323.  
<https://doi.org/10.1016/j.pnpbp.2017.07.006>
- Thoma, P., Norra, C., Juckel, G., Suchan, B., Bellebaum, C., 2015. Performance monitoring and empathy during active and observational learning in patients with major depression. *Biological Psychology* 109, 222–231. <https://doi.org/10.1016/J.BIOPSYCHO.2015.06.002>
- Thompson, R.J., Mata, J., Jaeggi, S.M., Buschkuhl, M., Jonides, J., Gotlib, I.H., 2012. The Everyday Emotional Experience of Adults with Major Depressive Disorder: Examining Emotional Instability, Inertia, and Reactivity. *J Abnorm Psychol* 121, 819–829.  
<https://doi.org/10.1037/a0027978>
- Treadway, M.T., Bossaller, N.A., Shelton, R.C., Zald, D.H., 2012. Effort-based decision-making in major depressive disorder: A translational model of motivational anhedonia. *Journal of Abnormal Psychology* 121, 553–558. <https://doi.org/10.1037/a0028813>
- Treadway, M.T., Buckholz, J.W., Schwartzman, A.N., Lambert, W.E., Zald, D.H., 2009. Worth the ‘EEfRT’? The Effort Expenditure for Rewards Task as an Objective Measure of Motivation and Anhedonia. *PLoS ONE* 4, e6598. <https://doi.org/10.1371/journal.pone.0006598>
- Treadway, M.T., Waskom, M.L., Dillon, D.G., Holmes, A.J., Park, M.T.M., Chakravarty, M.M., Dutra, S.J., Polli, F.E., Iosifescu, D.V., Fava, M., Gabrieli, J.D.E., Pizzagalli, D.A., 2015. Illness Progression, Recent Stress and Morphometry of Hippocampal Subfields and Medial Prefrontal Cortex in Major Depression. *Biol Psychiatry* 77, 285–294.  
<https://doi.org/10.1016/j.biopsych.2014.06.018>
- Treadway, M.T., Zald, D.H., 2011. Reconsidering anhedonia in depression: lessons from translational neuroscience. *Neuroscience and biobehavioral reviews* 35, 537–55.  
<https://doi.org/10.1016/j.neubiorev.2010.06.006>
- Tremblay, L., Schultz, W., 1999. Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708. <https://doi.org/10.1038/19525>

- Tundo, A., de Filippis, R., Crescenzo, F.D., 2019. Pramipexole in the treatment of unipolar and bipolar depression. A systematic review and meta-analysis. *Acta Psychiatrica Scandinavica* 140, 116–125. <https://doi.org/10.1111/ACPS.13055>
- Tziortzi, A.C., Haber, S.N., Searle, G.E., Tsoumpas, C., Long, C.J., Shotbolt, P., Douaud, G., Jbabdi, S., Behrens, T.E.J., Rabiner, E.A., Jenkinson, M., Gunn, R.N., 2014. Connectivity-Based Functional Analysis of Dopamine Release in the Striatum Using Diffusion-Weighted MRI and Positron Emission Tomography. *Cerebral Cortex* 24, 1165–1177. <https://doi.org/10.1093/cercor/bhs397>
- Uher, R., Perlis, R.H., Henigsberg, N., Zobel, A., Rietschel, M., Mors, O., Hauser, J., Dernovsek, M.Z., Souery, D., Bajs, M., Maier, W., Aitchison, K.J., Farmer, A., McGuffin, P., 2012. Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms. *Psychological Medicine* 42, 967–980. <https://doi.org/10.1017/S0033291711001905>
- Vallesi, A., Canalaz, F., Balestrieri, M., Brambilla, P., 2015. Modulating speed-accuracy strategies in major depression. *Journal of Psychiatric Research* 60, 103–108. <https://doi.org/10.1016/j.jpsychires.2014.09.017>
- van de Leemput, I.A., Wichers, M., Cramer, A.O.J., Borsboom, D., Tuerlinckx, F., Kuppens, P., van Nes, E.H., Viechtbauer, W., Giltay, E.J., Aggen, S.H., Derom, C., Jacobs, N., Kendler, K.S., van der Maas, H.L.J., Neale, M.C., Peeters, F., Thiery, E., Zachar, P., Scheffer, M., 2014. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences* 111, 87–92. <https://doi.org/10.1073/pnas.1312114110>
- van Eimeren, T., Ballanger, B., Pellecchia, G., Miyasaki, J.M., Lang, A.E., Strafella, A.P., 2009. Dopamine Agonists Diminish Value Sensitivity of the Orbitofrontal Cortex: A Trigger for Pathological Gambling in Parkinson's Disease? *Neuropsychopharmacology* 34, 2758. <https://doi.org/10.1038/sj.npp.npp2009124>
- Vinckier, F., Gourion, D., Mouchabac, S., 2017. Anhedonia predicts poor psychosocial functioning: Results from a large cohort of patients treated for major depressive disorder by general practitioners. *European Psychiatry* 44, 1–8. <https://doi.org/10.1016/j.eurpsy.2017.02.485>
- Vrieze, E., Pizzagalli, D.A., Demyttenaere, K., Hompes, T., Sienaert, P., de Boer, P., Schmidt, M., Claes, S., 2013. Reduced Reward Learning Predicts Outcome in Major Depressive Disorder. *Biological Psychiatry* 73, 639–645. <https://doi.org/10.1016/j.biopsych.2012.10.014>
- Walsh, A.E.L., Browning, M., Drevets, W.C., Furey, M., Harmer, C.J., 2018. Dissociable temporal effects of bupropion on behavioural measures of emotional and reward processing in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences* 373. <https://doi.org/10.1098/RSTB.2017.0030>
- Wang, Y.-Y., Wang, Y., Huang, J., Sun, X.-H., Wang, X.-Z., Zhang, S.-X., Zhu, G.-H., Lui, S.S.Y., Cheung, E.F.C., Sun, H.-W., Chan, R.C.K., 2022. Shared and distinct reward neural mechanisms among patients with schizophrenia, major depressive disorder, and bipolar disorder: an effort-based functional imaging study. *Eur Arch Psychiatry Clin Neurosci*. <https://doi.org/10.1007/s00406-021-01376-3>
- Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 54, 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Whitton, A.E., Reinen, J.M., Slifstein, M., Ang, Y.-S., McGrath, P.J., Iosifescu, D.V., Abi-Dargham, A., Pizzagalli, D.A., Schneier, F.R., 2020. Baseline reward processing and ventrostriatal dopamine function are associated with pramipexole response in depression. *Brain* 143, 701–710. <https://doi.org/10.1093/BRAIN/AWAA002>
- Wickens, J., Kötter, R., 1995. Cellular models of reinforcement, in: *Models of Information Processing in the Basal Ganglia, Computational Neuroscience*. The MIT Press, Cambridge, MA, US, pp. 187–214.

- Willner, P., Lappas, S., Cheeta, S., Muscat, R., 1994. Reversal of stress-induced anhedonia by the dopamine receptor agonist, pramipexole. *Psychopharmacology* 115, 454–62. <https://doi.org/10.1007/bf02245568>
- Xie, C., Jia, T., Rolls, E.T., Robbins, T.W., Sahakian, B.J., Zhang, J., Liu, Z., Cheng, W., Luo, Q., Zac Lo, C.-Y., Wang, H., Banaschewski, T., Barker, G.J., Bokde, A.L.W., Büchel, C., Quinlan, E.B., Desrivières, S., Flor, H., Grigis, A., Garavan, H., Gowland, P., Heinz, A., Hohmann, S., Ittermann, B., Martinot, J.-L., Paillère Martinot, M.-L., Nees, F., Orfanos, D.P., Paus, T., Poustka, L., Fröhner, J.H., Smolka, M.N., Walter, H., Whelan, R., Schumann, G., Feng, J., Artiges, E., Aydin, S., Banaschewski, T., Barbot, A., Barker, G., Becker, A., Bezivin-Frere, P., Biondo, F., Bokde, A., Büchel, C., Chu, C., Conrod, P., Daelow, L., Dalley, J., Desrivieres, S., Dooley, E., Filippi, I., Fillmer, A., Flor, H., Fröhner, J., Frouin, V., Garavan, H., Gowland, P., Grimmer, Y., Heinz, A., Hohmann, S., Ihlenfeld, A., Ing, A., Isensee, C., Ittermann, B., Jia, T., Lemaitre, H., Lethbridge, E., Martinot, J.-L., Millenet, S., Miller, S., Miranda, R., Nees, F., Paillere, M.-L., Papadopoulos, D., Paus, T., Pausova, Z., Pentilla, J., Poline, J.-B., Poustka, L., Burke, E., Rapp, M., Robbins, T., Robert, G., Rogers, J., Ruggieri, B., Schumann, G., Smolka, M., Stringaris, A., van Noort, B., Walter, H., Whelan, R., Simon, R., Williams, S., Zhang, Y., 2021. Reward Versus Nonreward Sensitivity of the Medial Versus Lateral Orbitofrontal Cortex Relates to the Severity of Depressive Symptoms. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 6, 259–269. <https://doi.org/10.1016/j.bpsc.2020.08.017>
- Xie, W., Li, H., Luo, X., Fu, R., Ying, X., Wang, N., Yin, Q., Zou, Y., Cui, Y., Wang, X., Shi, C., 2014. Anhedonia and Pain Avoidance in the Suicidal Mind: Behavioral Evidence for Motivational Manifestations of Suicidal Ideation in Patients With Major Depressive Disorder. *Journal of Clinical Psychology* 70, 681–692. <https://doi.org/10.1002/jclp.22055>
- Yamamoto, K., Vernier, P., 2011. The Evolution of Dopamine Systems in Chordates. *Frontiers in Neuroanatomy* 5.
- Yang, X., Huang, J., Harrison, P., Roser, M.E., Tian, K., Wang, D., Liu, G., 2021. Motivational differences in unipolar and bipolar depression, manic bipolar, acute and stable phase schizophrenia. *J Affect Disord* 283, 254–261. <https://doi.org/10.1016/j.jad.2021.01.075>
- Yang, X., Huang, J., Zhu, C., Wang, Y., Cheung, E.F.C., Chan, R.C.K., Xie, G., 2014. Motivational deficits in effort-based decision making in individuals with subsyndromal depression, first-episode and remitted depression patients. *Psychiatry Research* 220, 874–882. <https://doi.org/10.1016/j.psychres.2014.08.056>
- Ye, Z., Hammer, A., Camara, E., Münte, T.F., 2011. Pramipexole modulates the neural network of reward anticipation. *Human Brain Mapping* 32, 800. <https://doi.org/10.1002/HBM.21067>
- Zarate, C.A., Payne, J.L., Singh, J., Quiroz, J.A., Luckenbaugh, D.A., Denicoff, K.D., Charney, D.S., Manji, H.K., 2004. Pramipexole for bipolar II depression: a placebo-controlled proof of concept study. *Biol Psychiatry* 56, 54–60. <https://doi.org/10.1016/j.biopsych.2004.03.013>
- Zhang, W.-N., Chang, S.-H., Guo, L.-Y., Zhang, K.-L., Wang, J., 2013. The neural correlates of reward-related processing in major depressive disorder: A meta-analysis of functional magnetic resonance imaging studies. *Journal of Affective Disorders* 151, 531–539. <https://doi.org/10.1016/j.jad.2013.06.039>
- Zou, Y.-M., Ni, K., Wang, Y.-Y., Yu, E.-Q., Lui, S.S.Y., Zhou, F.-C., Yang, H.-X., Cohen, A.S., Strauss, G.P., Cheung, E.F.C., Chan, R.C.K., 2020. Effort-cost computation in a transdiagnostic psychiatric sample: Differences among patients with schizophrenia, bipolar disorder, and major depressive disorder. *Psych J* 9, 210–222. <https://doi.org/10.1002/pchj.316>