Computational Models of Bipolar Disorder

Sarah Kayser

4th Year Project Report
Artificial Intelligence and Software Engineering
School of Informatics
University of Edinburgh

2022

Abstract

Bipolar disorder, a mental illness characterised by periods of extremely elevated and depressed mood, is one of the leading causes of disability in young people. However, the disorder's cause remains unknown. A rather young area of science, computational psychiatry, aims to introduce computational modelling techniques into the field of psychiatry, which could aid in the understanding of bipolar disorder. Several computational models that aim to represent the aforementioned mood fluctuations mathematically have been presented over the last two decades and are summarised in this report. The most recent approach, which is based on reinforcement learning, is reviewed as well as tested in greater depth, and subsequently extended.

Research Ethics Approval

This project was planned in accordance with the Informatics Research Ethics policy. It did not involve any aspects that required approval from the Informatics Research Ethics committee.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Sarah Kayser*)

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Peggy Seriès, for introducing me to the world of computational psychiatry, always giving me incredibly quick and valuable feedback on my drafts of this report, and guiding me through the entire process of working on this project!

Secondly, I would like to thank all the other members of the Seriès Lab for making me feel welcome in the research community.

Lastly, I would like to thank my friends (especially my best friend Clara) and my family for their continuous support throughout my degree!

Table of Contents

Chapter 1

Introduction

1.1 An overview of bipolar disorder

1.1.1 Introduction

Bipolar Disorder (formerly known as manic depression) is a recurrent and chronic mental health condition which affects between 1.5% and 4.8% of the world's population [Angst, 2008]. It is characterised by periods of mania or hypomania (extremely elevated mood) and periods of depression (extremely low mood) [Grande et al., 2016]. The typical age of onset is 18-25 years [Kawa et al., 2005].

Approximately 10–20% of people diagnosed with bipolar disorder take their own life, and around 30% of patients attempt suicide at least once [Müller-Oerlinghausen et al., 2002], [Novick et al., 2010]. Psychiatric as well as non-psychiatric comorbidities are common and include illnesses such as cardiovascular disease, anxiety, or substance misuse disorder [Grande et al., 2016].

Typical symptoms during a period of **depression** include feelings of hopelessness and worthlessness, a lack of energy and concentration, the loss of interest in everyday activities, sleep problems, a change in appetite, as well as suicidal thoughts [National Health Service, 2019a]. These symptoms are the sames ones that characterise Major Depressive Disorder (MDD).

Symptoms of **manic phases** include feeling happy and full of energy, having numerous new ideas and plans, being delusional, having hallucinations, not wanting to eat or sleep, and making potentially risky decisions, such as spending vast amounts of money. Patients often do not realise that they are experiencing a manic phase while it is happening, and may be shocked at their behaviour in hindsight. The symptoms for manic episodes of bipolar disorder significantly overlap with those of psychosis as well as schizophrenia [National Health Service, 2019b], [National Health Service, 2019c].

An episode of **hypomania** can be described as an elevation of mood, with the symptoms being similar to those typical for mania (excluding severe symptoms such as hallucinations or delusions). The main difference to a manic episode is that patients do not

experience social or occupational dysfunction and, by definition, do not have to be admitted to hospital [Phillips and Kupfer, 2013].

Mixed episodes are periods during which patients experience both manic and depressive symptoms [Grande et al., 2016].

Most patients with bipolar disorder go through phases of 'normal' functioning and mood changes, which are called **Euthymia** [Mason et al., 2017].

1.1.2 Types of bipolar disorder

According to the Diagnostic and Statistical Manual for Mental Disorders (DSM-V) [American Psychiatric Association, 2013], there are six subtypes of bipolar disorder: Bipolar I, Bipolar II, Cyclothymia, Substance-induced and medication-induced bipolar disorder, as well as Other specified and unspecified bipolar and related disorders. The different types of the disorder vary in cause and intensity of the depressive and manic (or hypomanic) episodes [American Psychiatric Association, 2002].

Bipolar I is characterised by at least one episode of full-blown mania or a mixed episode. Usually, patients have experienced one episode of major depression as well [Phillips and Kupfer, 2013].

The diagnosis of **Bipolar II** is given to patients who have experienced at least one episode of major depression and hypomania.

Patients with **cyclothymia** have experienced at least two years of hypomanic as well as depressive episodes. The periods of mood changes which they go through do not, however, fulfill the criteria for episodes of hypomania, mania, or major depression.

Substance-induced and medication-induced bipolar disorder is characterised by the same symptoms as the previously mentioned types of bipolar disorder. It is, however, caused by a substance (i.e., a drug, such as cocaine, or medication) or another medical condition and the symptoms are not present when the causal element is removed [Vieta et al., 2018].

In order to be diagnosed with **Other specified bipolar or unspecified bipolar and related disorders**, patients must at least exhibit manic symptoms, but not meet the criteria for one of the other subtypes of bipolar disorder. The difference between other specified and other unspecified bipolar disorder is that for other specified bipolar disorder, "the clinician chooses to communicate the specific reason that the presentation does not meet the criteria for any specific bipolar and related disorder" [American Psychiatric Association, 2013].

1.1.3 Causes of bipolar disorder

Bipolar disorder is considered to be genetically complex and extremely heritable, with heritability estimated to be up to 85% [Vieta et al., 2018]. Twin studies have shown that the risk of having bipolar disorder given that an identical twin has it, is around 63% [Goodwin and Jamison, 2007].

Biological causes, such as mitochondrial dysfunction [Kato, 2007], the ”disruption of the molecular and cellular network [which regulates] energy production and expenditure” [Mansur et al., 2020], or dopamine dysregulation [Berk et al., 2007] have been hypothesised to contribute to the development of the disorder.

Impairments in the perception of reward and its interaction with mood have also been mentioned as possible cause [Mason et al., 2017].

Furthermore, there are several environmental and medical factors that increase the risk of developing bipolar disorder. These include, but are not limited to, (childhood) trauma, caesarean section delivery, maternal smoking during pregnancy, or substance abuse [Vieta et al., 2018].

Overall, it can be said that the causes of bipolar disorder (as well as the causes of many other mental disorders) remain poorly understood and constitute an active area of research.

1.1.4 Treatment of bipolar disorder

In general, there are three main approaches to the treatment of bipolar disorder: the pharmacological approach, the psychological approach, and the lifestyle approach [Grande et al., 2016].

Different types of therapy, such as Cognitive Behavioural Therapy (CBT), Interpersonal and Social Rhythm Therapy (ISRT), or Family-Focussed Therapy (FFT), as well as lifestyle changes (e.g., regular sleeping patterns, exercise, and limitation of the consumption of substances such as caffeine and alcohol) have been shown to be effective in managing and preventing the occurrence of the symptoms of bipolar disorder [Grande et al., 2016].

Furthermore, there is a range of pharmaceutical drugs that can be used to treat the symptoms of bipolar disorder. The most commonly used drugs for acute treatment are antipsychotics, antidepressants, as well as mood stabilisers, with Lithium being a popular choice [Grande et al., 2016]. Antidepressants can, however, induce hypomanic or manic episodes and are thus an unsuitable treatment option for some patients [Pacchiarotti et al., 2013]. This is especially critical when bipolar disorder gets misdiagnosed as unipolar depression (which, due to the length and severity that depressive episodes can have, it often does) [Vieta et al., 2018].

1.2 Computational psychiatry and its relevance

Computational Psychiatry is a rather young, but nevertheless fast-growing, field of science. It uses models of brain function to specify the mechanisms of psychopathology, which are typically described in computational or mathematical ways [Friston et al., 2014]. Hence, the main aim of computational psychiatry is to apply computational modelling techniques as well as theoretical approaches to psychiatric questions [Seriès, 2020].

Since computational psychiatry makes use of generative models (i.e., probabilistic descriptions of how high-level causes generate low-level data), it is mechanistic. This means that these computational models describe the underlying processes that take place and can therefore make predictions about events taking place in the real world [Adams et al., 2015].

One can only understand mental illnesses by linking several interacting levels - cells, circuits, cognition, behaviour, as well as the patient's social and physical environment [Huys et al., 2016]. This is also linked to the core assumption of the Research Domain Criteria (RDoC) project, which is a framework that aims to explain mental illnesses through the use of different constructs (e.g., cells, neural circuits, physiology, and behaviour) rather than the symptomatology alone and was thus proposed as an alternative to the DSM by the US National Institute of Mental Health [Cuthbert, 2022]. Due to the complexity of the field of psychiatry, computational approaches can help researchers examine possible relationships between different factors, lead to new insights, and guide experiments. Computational psychiatry can therefore be a valuable resource that helps scientists in their research of the psychopathology of mental disorders, including bipolar disorder [Seriès, 2020].

1.3 An overview of reinforcement learning

Reinforcement, i.e., reward, has been studied by animal psychologists for more than 60 years [Russell and Norvig, 2010]. Reinforcement Learning is a branch of Artificial Intelligence - one of the three basic paradigms of Machine Learning, to be exact (the other two being supervised and unsupervised learning) [Dunjko et al., 2016]. It addresses how agents (which can be artificial or natural) learn to maximise reward and minimise punishment in environments that involve different states as well as transitions between states [Montague et al., 2012].

Reinforcement learning models can be used to describe how learning processes work for agents such as humans or other animals [Niv, 2009]. The task of reinforcement learning is to utilise observed rewards in order to learn an (almost) optimal policy for the environment that the agent is in. In other words, the agent tries to maximise the cumulative sum of the rewards of its decisions. Reinforcement learning algorithms use a process of trial and error to estimate state-action values and make decisions - i.e., pick optimal sequences of actions [Russell and Norvig, 2010].

There are several studies which investigate the neural processes associated with reinforcement learning in the context of bipolar disorder [Niv, 2009], [O'Doherty et al., 2015]. They hypothesise (and provide supporting data for the fact) that reward hypersensitivity [Whitton et al., 2015] and white matter hyperintensity [McDonald et al., 1999], amongst other factors, could be potential causes for bipolar disorder.

1.4 An overview of mood modelling

Mood describes a person's long-term feelings which do not change quickly and are not directed towards a specific object or person [Collenette et al., 2017]. Emotions are caused by an external stimulus or event and are thus short-term, directed feelings [Deligianni et al., 2019].

Mood modelling is predominantly being used to study the influence of people's moods on decision making tasks, as well as to predict people's mood changes over time [Collenette et al., 2017]. It has been used to study several mental illnesses, such as major depressive disorder [Sharp et al., 2020] or schizophrenia [Yasui-Furukori, 2012].

Chapter 2 of this report describes and compares several different computational models of bipolar disorder that aim to model a patient's mood over time. The aim of chapters 3 and 4 is to simulate a range of dynamics that Mason et al.'s (2017) model of mood fluctuations in bipolar disorder is able to account for and extend it.

1.5 Relevance of this honours project

Bipolar disorder is one of the leading causes of disability among young people [Grande et al., 2016]. Hence, it is crucial to improve doctors' as well as people's understanding of it in order to make accurate predictions about patients' mood changes, intervene when danger occurs, and ameliorate the lives of patients.

Computational Psychiatry is a promising field of research that can provide tools for the modelling and prediction of mood fluctuations in bipolar disorder. Therefore, the comparison and critical evaluation of several computational models of bipolar disorder, as well as the reimplementation and extension of [Mason et al., 2017]'s model are a step into the direction of improvement of understanding and therefore preventing the occurrence of extreme symptoms of bipolar disorder.

Chapter 2

Literature review – Models of mood fluctuations in bipolar disorder

2.1 Introduction

This chapter aims to give a comprehensive overview of the existing models designed to display and/or explain mood fluctuations in bipolar disorder. Bipolar disorder has a rather complicated pathology, with several factors (e.g., genetic factors, cellular processes, and cognition) having been proposed as factors contributing to the development of it. Therefore, it is not possible (yet) to model one specific biological system and research has largely been focussed on mood modelling, the directly observable and defining feature of bipolar disorder [Cochran et al., 2018].

Daugherty et al. were the first researchers who attempted to represent mood fluctuations in bipolar disorder with the help of mathematical equations. They published their first paper in 2003 and completed their in work in 2009. Nana (2009) also built upon Daugherty et al.'s (2003) work (without referencing it [Wirkus and Porter, 2009]) and therefore was the first researcher to propose a model of mood fluctuations in patients with bipolar disorder. Both models are described in the section below.

There are two fundamentally different approaches to modelling mood changes in bipolar disorder (as well as other biological rhythms). Firstly, one can use collected mood data from patients over a certain course of time and fit mathematical functions to the data. Secondly, one can try to find potential regulatory mechanisms which are able to produce oscillations between the two poles of bipolar disorder and later assess clinical accuracy. The models described below lay their focus on the latter approach, as the data available is rather limited in terms of amount of patients. Additionally, the current approach is to use self-assessment and report of mood, rather than more clinically accurate approaches [Goldbeter, 2013].

2.2 Nana (2009) & Daugherty et al. (2009): Limit cycle oscillators

Nana (2009) notes that even though bipolar disorder episodes can be rather erratic, there exist patters of recurrence to the mood fluctuations. These can be simplified with a function modelling them as periodic oscillations between manic and depressive episodes. He attempts to model the mood fluctuation of a single patient suffering from bipolar II disorder using a negatively damped harmonic oscillator. A damped harmonic oscillator portrays a vibrating system whose amplitude of vibrations decreases as time increases and can be used to model several biological and physical effects, one simple example being the vibration of a guitar string after it has been hit.

The author assumes small damping, causing the function's oscillations to remain approximately periodic and not decreasing continuously (as higher damping would).

The system is made up of two coupled equations which describe the amplitude and phase of solutions. The equations' fixed points resemble limit cycles for the system, indicating the maximum mood levels in either direction.

Furthermore, Nana (2009) attempts to model the aggregated effect of treatment (psychotherapy as well as different kinds of medication) by making use of a parametric forcing function (i.e., an external force which is only dependent on time, and not on any of the other parameters). He notes that the equations are lacking a time-delay function which should be used to model the delayed effectiveness of treatment.

His model is represented graphically in figure 2.1.

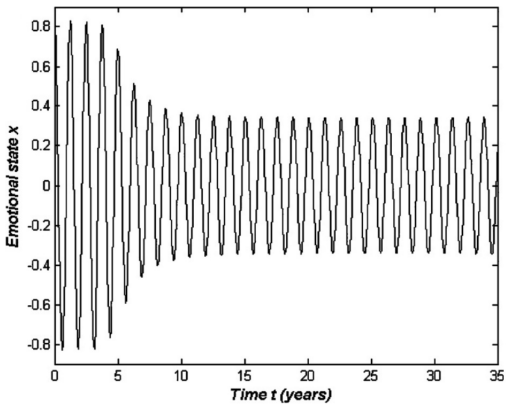

Figure 2.1: Limit cycle oscillator model of mood fluctuations in bipolar disorder (including treatment), as proposed by Nana (2009)

As one can see, the system modelled above describes the emotional state variations in one patient suffering from bipolar II disorder over the course of several years, with treatment causing the mood fluctuations to go from a larger to a smaller limit cycle and the patient's mood therefore not reaching the extreme values it did without treatment.

Daugherty et al. (2009) also make use of limit cycle oscillators to model bipolar II disorder as their work is largely based on their research published in 2003, which

Nana (2009) uses in his paper as well. In addition to modelling the dynamics of an individual bipolar II disorder patient, Daugherty et al. (2009) also examine two models that describe the interaction between a pair of patients suffering from bipolar II disorder, e.g., siblings.

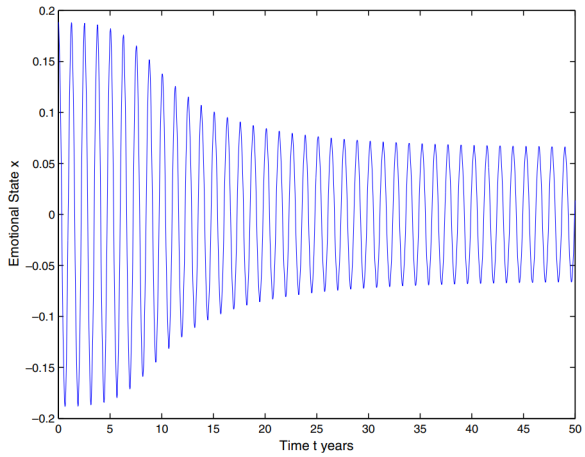

Figure 2.2: Limit cycle oscillator model of mood fluctuations in bipolar disorder (including treatment, with a delayed onset), as proposed by Daugherty et al. (2009)

The shape of the function in figure 2.2 is clearly similar to Nana's (2009). However, the parameters in this model were adjusted in order to include the delayed onset of the effect of treatment. The patient starts to receive treatment between the years 5 and 10 in both models, so it can be clearly seen that the smaller stable limit cycle is reached significantly later in Daugherty et al.'s (2009) model.

The authors note that bipolar disorder is often diagnosed at later stages in a person's life, however, the onset of it can be in early childhood. It can be said that this model is purely theoretical and does not assume any differences in a patient's mood swings that are related to their age, i.e., the mood fluctuations are just dependent on the time that has passed since an arbitrary point in time (i.e., the onset of the disease, which corresponds to year 0 in the model), which can be at any point in a person's life.

Daugherty et al.'s (2009) modelling of the interaction between two treated patients with bipolar II disorder yielded the following results: Assuming that the individuals generally have similar ranges of mood and mood cycling rates, they tend to remain in-phase, i.e., they start entering hypomanic as well as depressive episodes approximately at the same point in time. Additionally, when they are currently experiencing a euthymic period, they are both likely to stay in it. However, since there was little clinical data available at the time, the accuracy of the models attempting to model these interactions could not be analysed. Nevertheless, the results provide predictions about those interactions and can thus be used further in future research.

Both of the models described above solely focus on modelling the mood swings in one/two interacting bipolar disorder patients. They do not examine the underlying causes for the disorder and also assume regular oscillations of mood – which is not

necessarily realistic, as patients frequently experience euthymic periods as well as irregular fluctuations of mood.

Nevertheless, the models provide a basic framework for mood modelling in bipolar disorder, and therefore lead the way to the development of more sophisticated models.

2.3 Bonsall et al. (2011): An autoregressive model

Bonsall et al. (2011) note that apart from the manic and depressive episodes, some bipolar disorder patients also suffer from chronic week-to-week mood instability (as shown in figure 2.3). This characteristic of the disorder is usually not the main focus of research, but is important for the development of treatment approaches and methods as well.

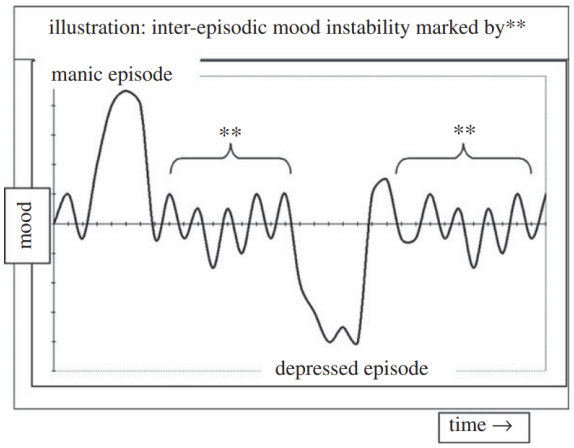

Figure 2.3: High-level diagram showing what typical mood fluctuations that a bipolar disorder patient experiences might look like – i.e., manic and depressive episodes as well as inter-episodic mood instability

Bonsall et al. (2011) propose a time-series approach in order to represent mood fluctuations in bipolar disorder patients. They make use of self-rated mood data from 23 patients over the course of 220 weeks and use that data to show that the mood variability is best described by non-linear time series processes. These processes are different depending on whether the patient is prone to inter-episodic mood fluctuations. The authors fit several autoregressive (AR) models (i.e., models that represent time-varying processes whose output variable depends linearly on the previous values as well as a stochastic term – in simpler terms, models that make use of past behaviour in order to predict future behaviour [Chan, 1993]) to the data.

In an AR(0) model, there is no dependence between terms. In an AR(1) model, the value of the current term depends on the previous term (and noise); in an AR(2) model it depends on the previous two terms and noise, and so forth. Bonsall et al. (2011) conclude that in the group consisting of patients with rather stable week-to-week mood, patients' mood changes are best predicted with AR(1), so only the previous week's mood score is needed to predict the current week's mood score. For patients who exhibit

more unstable mood, an AR(2) model is used for predictions, meaning that the mood scores of the previous two weeks are necessary for predicting the current week's mood score. This is the first study which makes use of a statistical time-series framework in order to interpret the longitudinal patters of clinical data from bipolar disorder patients.

2.4 Bonsall et al. (2015): A relaxation oscillator framework

Bonsall et al.'s (2015) aim is to find out to what extent the dynamics of self-reported mood in bipolar disorder patients can be described and understood by using a relaxation oscillator framework. Relaxation oscillators have been applied to a range of processes within biology, and can therefore be useful in terms of explaining bipolar disorder through the use of the 'biological hierarchy', i.e., levels such as molecular processes, the circadian rhythm, and mood fluctuations. The use of this hierarchy - which is based on the assumption that psychological processes can be explained by making use of the same methods that are used for explaining mechanical as well as physiological processes - is also known as the mechanistic approach to cognition, or "mechacognitive approach".

In order to validate this approach, Bonsall et al. (2015) test their model against mood data from several individuals with diagnosed bipolar disorder. They show that the mood fluctuations can be driven by the combination of effects from dynamics described by relaxation oscillators and noise (i.e., stochastic variability). The authors note that relaxation oscillators have one major advantage in comparison to limit cycle oscillators: They are nonlinear periodic dynamic systems which are characterised on different timescales. This results in the model being able to capture one important characteristic of bipolar disorder – the existence of larger changes in state (corresponding to (hypo)manic and depressive episodes), as well as time intervals with small changes in state (corresponding to inter-episode mood variability). However, they mainly lay their focus on depressive episodes, as this is the dominant state that bipolar disorder patients are in (see figure 2.4).

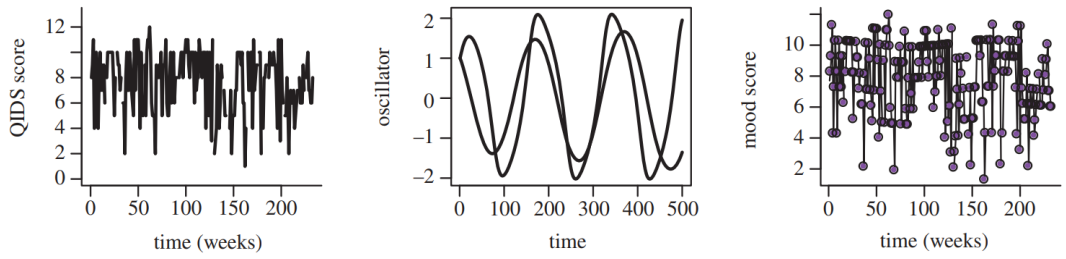

Figure 2.4: Depressive symptomatology from one patient over time (depicted by the QIDS score), predicted relaxation oscillator dynamics, and final predicted mood dynamics by making use of each score shown in diagram 1 in order to predict the next mood score

Bonsall et al. (2015) conclude that several coupled relaxation oscillators can describe

and thus aid in the understanding of the biological, psychological, as well as social rhythms of bipolar disorder. Therefore, this approach is another method of understanding mood fluctuations in bipolar disorder and can aid in the development of a more in-depth understanding of how the impacted processes on different biological levels are linked.

2.5 Goldbeter (2011) & Goldbeter (2013): Mutual inhibition of mania and depression

Goldbeter (2011 & 2013) bases his work on the assumption that mania and depression – the two poles in bipolar disorder – are mutually exclusive and inhibit each other. This is an effect that can be observed in various areas of biology, some as simple as two different animal populations competing for the same resources [Volterra, 1931].

His phenomenological model predicts bistability if mutual inhibition is strong enough, i.e., the system is in either the manic or the depressive state and can suddenly switch between the two stable states. The author proposes two mechanisms that – when used in conjunction with mutual inhibition – add oscillations to the system. These oscillations can have several different shapes (unlike in Nana's (2009) and Daugherty et al.'s (2009) models), such as periodic oscillations with either similar or different durations of the manic or depressive episodes, as well as less intense oscillations around a certain state before a periodic change of higher amplitude to the other state. Hence, the self-sustained oscillations in the system are able to model the rather sudden switches between manic and depressive episodes.

Goldbeter notes that both mania and depression are complex, and therefore multidimensional traits. However, for simplification purposes, they are represented by continuous variables in this model and the system can thus be represented by two ordinary differential equations. These equations are coupled through terms which represent mutual inhibition. Additionally, the coupling of the two equations is mediated by intermediate factors in Goldbeter's model.

As mentioned above, there are two proposed mechanisms which can cause the transition from bistable to oscillatory behaviour of the system. Mutual inhibition combined with cross-activation of mania and depression (represented by the aforementioned intermediate factors) is one way of introducing this behaviour. This means that mania (M) and depression (D) directly inhibit each other as well as indirectly activate each other after a certain time delay. In other words, M and D activate each other through the intermediate factors F_M and F_D and inhibit each other directly. Hence, the shape of the oscillations is determined by the relative time course of the intermediate factors that have the activating effect (see figure 2.5).

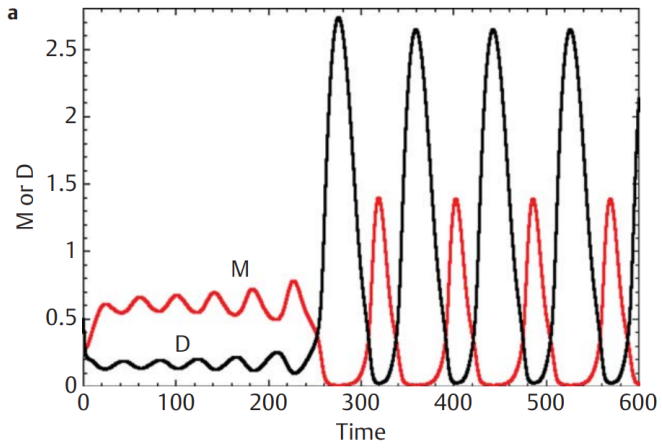

Figure 2.5: Fluctuations in Mania (M) and Depression (D) in Goldbeter's (2011) model: Mutual inhibition in combination with cross-activation

Secondly, mutual inhibition can be paired with auto-inhibition of mania and depression in order to produce oscillatory behaviour from former bistability. In this case, the intermediate factors F_M and F_D inhibit the growth of the variables M and D, respectively (see figure 2.6).

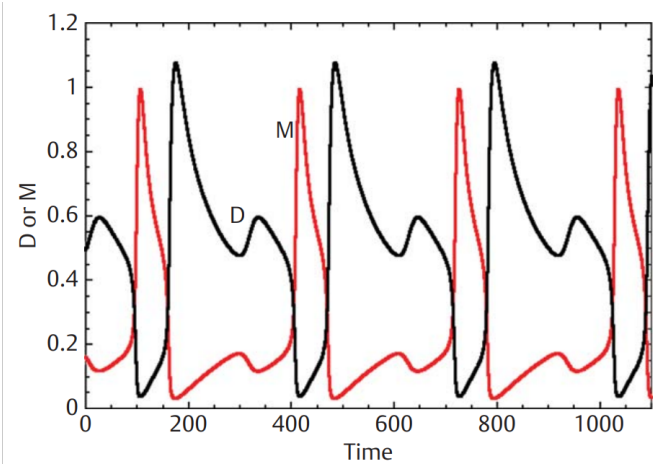

Figure 2.6: Fluctuations in Mania (M) and Depression (D) in Goldbeter's (2011) model: Mutual inhibition in combination with negative auto-regulation through intermediate parameters (F_M and F_D)

As one can see, this rather simple exploratory model is able to represent dynamical behaviour that represents clinical observations, e.g., sudden switches between mania and depression, or the fact that mania can be induced by antidepressants. This effect, however, is modelled by instant mood changes here, which is not realistic as the onset of the effects of antidepressants is generally delayed.

Since this is a model that is not based on specific clinical data, Goldbeter (2013) notes that the main current challenge is to find regulatory mechanisms on a neurocellular level

which have been or might be linked to mood disorders and therefore correspond to the proposed mathematical mechanisms.

2.6 Steinacher & Wright (2013): Modelling behavioural activation levels

Steinacher and Wright (2013) propose a minimal mathematical model of behavioural activation levels, which have been shown to correlate with (hypo-)manic as well as depressive episodes in bipolar disorder patients. They predict that when the non-linearity dimension of the model is increased, multistability as well as sudden (i.e., "switch-like") transitions between activation levels in a patient occur. In other words, the model is capable of reproducing core features of bipolar disorder, including the fluctuations between (hypo-)manic and depressive episodes, as well as other features such as slower recovery time after negative or positive events in bipolar disorder patients than in healthy individuals.

The model includes a variable which represents the level of behavioural activation, i.e., the tendency of a patient to interact with their environment. This parameter is regulated by two feedback mechanisms – a self-activation (positive) as well as an auto-regulatory (negative) feedback mechanism, representing the tendency to keep the behavioural activation parameter at a normal level (see figure 2.7). This way of formalising mechanisms is commonly used in approaches explaining regulatory networks on a cellular or molecular level.

Steinacher and Wright (2013) predict that the level of nonlinearity in the behavioural activation system corresponds to the stage of the disorder a patient is in. In other words, the higher the degree of nonlinearity, the higher the tendency for developing bipolar disorder, as well as the number of previous (hypo-)manic/depressive episodes is.

This model does not make use of memory of previous states.

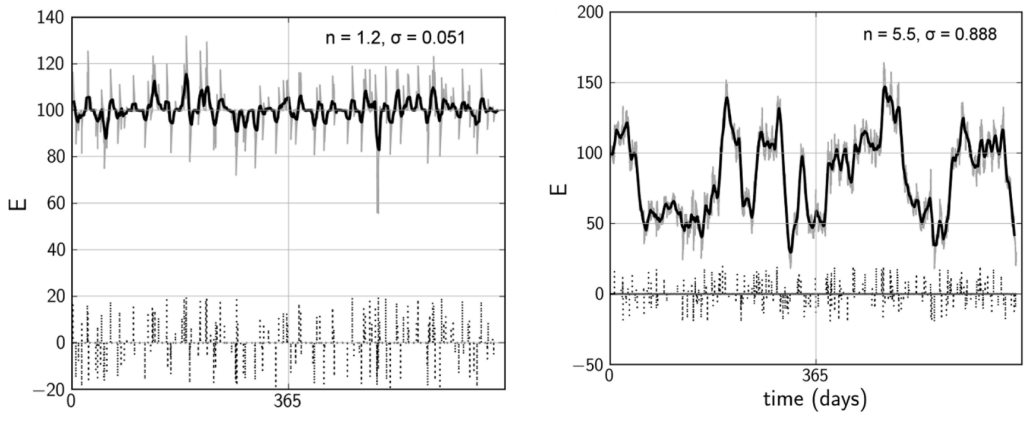

Figure 2.7: Diagrams showing fluctuation in behavioural activation (E) when performing stochastic simulations for different values for parameter n (which represents level of nonlinearity). Dotted lines display activating and inhibiting events over the course of 2 years.

As it can be inferred from the diagrams provided above, the proposed model goes beyond simple oscillatory dynamics and is able to provide a description of inter-episodic as well the general rather irregular (but still cyclic) mood fluctuations.

2.7 Cochran et al. (2018): Two-dimensional stochastic processes

Cochran et al. (2018) review several mathematical modelling frameworks used to model mood fluctuations in bipolar disorder, including the ones described above. Additionally, they introduce two new models, which model mood with two variables (in contrast to the models described above (with the exception of Goldbeter's (2011 & 2013)). They note that modelling mood with two variables instead of a single one provides the opportunity to capture known features of bipolar disorder, including the existence of mixed states (i.e., states where a patient experiences manic as well as depressive symptoms) as well as direct transitions between the depressive and the manic state, without a period of euthymia in between.

Their models use a manic variable as well as a depressive variable in order to model mood. In the first model, each variable satisfies a stochastic differential equation called "double-well" model, which is a physical model that has been used to describe the behaviour of particles that randomly move between two energy wells. If noise is absent, multistability occurs in this model.

The second model that Cochran et al. (2018) describe is two-dimensional as well; however, multistability cannot be achieved. In this model, mood is modelled by the aforementioned variables as well, but with a two-dimensional Ornstein-Uhlenbeck process. This process is a modified version of the random walk model, with the particle having a greater attraction towards a central location, the further away from the centre it is. An example of the fluctuations of depression and mania in a bipolar disorder patient can be seen in figure 2.8.

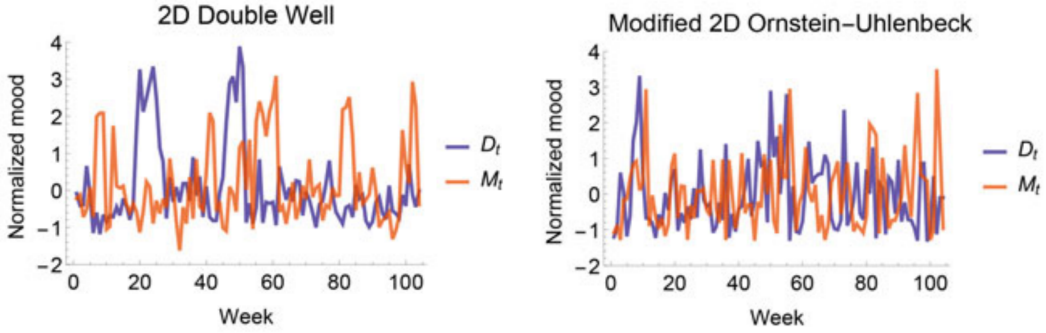

Figure 2.8: Modelled mood fluctuations (described by the depressive variable D and the manic variable M) in the two models proposed by Cochran et al. (2018)

2.8 Radulescu & Niv (2019): Mood fluctuations as a cause of reward prediction errors

Radulescu and Niv are two researchers belonging to a group of researchers who have been investigating how reinforcement learning models can be applied to bipolar disorder. Other significant papers suggesting this approach are [Eldar and Niv, 2015], [Eldar et al., 2016], [Mason et al., 2017], as well as [Chang and Chou, 2018] (who base their research on the aforementioned authors' findings). Chapter 3 of this report contains a detailed description of the approach. In summary - as stated in chapter 1.3 - reinforcement learning models are based on an agent maintaining a model of the external world, with rewards and punishments attached to different actions. The model gets updated after every action that is taken and the agent tries to maximise cumulative reward.

Radulescu and Niv (2019) state that mood can be defined as one of the features of the brain's valuation system, or rather, the running average of recently encountered reward prediction errors (RPEs). They find that bipolar disorder patients are – on average – more likely to take on effort when pursuing reward. This is consistent with the finding that bipolar disorder patients believe reward to be more abundant than it actually is in their environment and thus encounter larger reward prediction errors which can explain more intense mood fluctuations.

Hence, bipolar disorder patients might have a skewed internal state representation resulting in optimistic expectations about future reward, which then lead to greater mispredictions and surprises that cause mood to fluctuate.

2.9 Chang & Chou (2018): Mood fluctuations as a cause of mood sensitivity

Chang and Chou (2018) also focus on the interactions between expectations and mood and describe a variant of the model described in chapter 3 [Mason et al., 2017], [Eldar et al., 2016], [Eldar and Niv, 2015] in their paper.

In their model, mood behaves in an oscillatory way when the mood sensitivity parameter exceeds a particular threshold value. They also include the effects of different amplitudes of changes in mood following positive or negative events, also known as "asymmetric mood sensitivity".

Chang and Chou's (2018) model slightly differs from the one proposed by Eldar et al. (2016). Eldar et al. (2016) assume that the mood recovery rate (following an extremely positive or negative event) is the same as the mood learning rate. Chang and Chou (2018), however, use a different parameter for the two rates, allowing them to vary independently and therefore providing more flexibility. In general, it can be said that their model is slightly simpler mathematically, but is still able to capture the main features of bipolar disorder which are captured by the model proposed by Eldar et al. (2016) as well.

As stated above, Chang and Chou's (2018) model transitions to limit cycle behaviour

after surpassing a threshold value, which signals the transition to a bipolar state (see figure 2.9). Furthermore, their model predicts that asymmetric mood sensitivity can induce either depression or mania (which is an observation that has been clinically found as well) and hence be the cause for bipolar disorder.

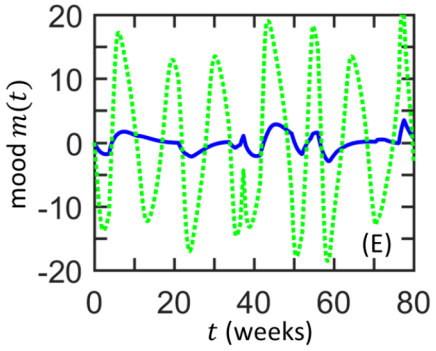

Figure 2.9: When mood sensitivity lies above a certain threshold value, the model is able to represent mood being much more oscillatory (green line) in bipolar disorder patients than in normal individuals (blue line)

As stated previously, antidepressants can induce manic episodes in bipolar disorder patients. When this effect is interpreted as an increase in mood sensitivity towards positive events, Chang and Chou's (2018) model predicts that the following manic episode happens earlier and is more intense than it would be without medication. Furthermore, assuming that the treatment is not changed, the frequency of manic episodes increases as well. This is represented graphically in figure 2.10.

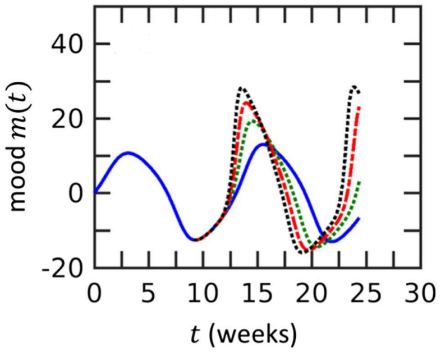

Figure 2.10: The effect of antidepressant dosage on the intensity and frequency of subsequent manic episodes (blue line: no medication, green line: light dosage, red line: medium dosage, black line: high dosage)

2.10 Balasubramani & Chakravarthy (2019): The role of the basal ganglia network

Balasubramani and Chakravarthy (2019) use a decision making framework in order to investigate how the basal ganglia network could potentially be involved in causing

mood oscillations in bipolar disorder. This network is a group of subcortical nuclei that are involved in variety of processes, including motor control, emotions, and – more specifically – reward prediction measures. This study therefore tries to explain the underlying pathological neural dynamics in bipolar disorder, and goes beyond a simple display of mood oscillations. Since impaired decision making may be linked to pathological mood states, the different mood states in bipolar disorder can be approached via a decision making framework.

Balasubramani and Chakravarthy (2019) state that the main factor which contributes to the development of bipolar disorder is the dysfunction of the serotonergic and dopaminergic system which is influential for the perception of reward and well as decision making and is controlled by the basal ganglia network.

They propose three models that make use of the notion of utility – a combination of the concepts of value and risk.

The task framework that the authors chose for their investigation is a two-armed bandit task. Bandit tasks are an approach which is extensively used within probability theory as well as machine learning. The name's origin are one-armed bandits, which are a type of slot machine. The gambler's aim is to maximise the cumulative reward earned by several lever pulls, each yielding a reward from a specific probability distribution which is unknown to the gambler. By doing so, they face a trade-off between 'exploiting' the machine with the highest expected reward and 'exploring' other machines in order to learn more about their associated rewards [Berry and Fristedt, 1985].

As stated above, Balasubramani and Chakravarthy (2019) make use of a two-armed bandit problem, with probabilistic positive and negative outcomes attached to each of the arms (i.e., states). In other words, one arm is linked to a positive (rewarding) outcome, and the other one to a negative (punitive) outcome, both outcomes having the probability 0.5. They then apply the task to three action selection models (or rather, one model but three different levels of abstraction) of the basal ganglia network, and thus try to identify the parameters which contribute to the occurrence of mood fluctuations in bipolar disorder. The first variant is a rather simplistic reinforcement learning model which uses a softmax model in order to make choices. The second variant is an extension of the first model, and is a neural network model of the basal ganglia network which can be related to neurobiological processes (dopamine and serotonin neuromodulation in the decision making process). Variant three consists of a reduced model system with two equations, which is capable of describing the main dynamics of the first two models and includes also specific cellular level models, making it more biologically realistic and understandable.

Under healthy conditions, the model chooses actions associated with positive outcomes and avoids actions associated with negative outcomes. Under bipolar conditions, however, it displays (slow) oscillations in its choices.

After mathematically analysing the dynamical system, Balasubramani and Chakravarthy (2019) find that decreased risk sensitivity levels and significantly increased reward sensitivity are the main factors that are responsible for the occurrence of bipolar oscillations between manic and depressive episodes.

2.11 Other research regarding mood modelling and prediction in bipolar disorder

There have been numerous other studies in the field of mood assessment, modelling, and forecasting in the context of bipolar disorder.

Hadeghi et al. (2013) propose that a forced Duffing oscillator (which is a non-linear second order differential equation) could be used to model mood fluctuations in bipolar disorder, due to its ability to capture typical patterns of mood fluctuations.

In order to produce more biologically realistic models, a connection between phenomenological models and biochemical mechanistic approaches which make use of specific cellular processes should be established. Frank (2013) suggests that one plausible biochemical approach for this connection is the interaction of the pathways of protein kinases A and C, which have been shown to affect an individual's mood.

Bayani et al. (2017) suggest that there exists a link between circadian activity rhythms and transitions in episodes in bipolar disorder. Their computational model could therefore be used to determine early warning signs that predict severe mood fluctuations in bipolar disorder patients.

Busk et al. (2020) make use of self-assessed mood data in order to predict future mood values. They use a hierarchical Bayesian model for the forecasting of mood values, as their data shows a correlation between mood values observed over a short period of time (4-7 days).

2.12 Conclusion & Discussion

This chapter's aim was to give the reader a comprehensive overview of the existing models of mood fluctuations in bipolar disorder. These models usually make one of two assumptions: bistability or biological rhythm [Cochran et al., 2018]. Models assuming bistability are based on the assumption that there are two stable states that represent the different phenotypes of mania and depression. The mood fluctuations are induced by random external shifts caused by life events [Steinacher and Wright, 2013], [Cochran et al., 2018]. Biological rhythm models are rooted in the assumption that there exist intrinsic oscillations in a patient's brain and the mood fluctuations occur without the existence of external perturbations [Daugherty et al., 2009], [Goldbeter, 2011], [Bonsall et al., 2015], [Eldar and Niv, 2015], [Mason et al., 2017], [Radulescu and Niv, 2019].

Overall, it can be said that the models mentioned above constitute interesting approaches to modelling the dynamics of mood in bipolar disorder. However, there are some significant shortcomings – some of which are mentioned in the papers themselves – that are summarised below.

Firstly, the majority of models does not include a link between the parameters mentioned in the models and a possible biological counterpart. This clearly limits the explanatory power of the models; however, it provides pointers for future

research [Steinacher and Wright, 2013], [Balasubramani and Chakravarthy, 2019]. In other words, it might be possible to find direct biological links for the variables, which therefore constitutes a concrete goal for future research in the field.

Secondly, the different types of bipolar disorder as well as the different mood states that a patient can be in are not explored equally. For instance, Bonsall et al. (2011) only focus on bipolar II disorder, Bonsall et al. (2015) lay their focus on depressive episodes, and in general, research tends to focus on full-blown manic and depressive episodes. Inter-episode mood instability is being explored and modelled less often (e.g., it is not mentioned at all in early models [Nana, 2009], [Daugherty et al., 2009]).

Furthermore, as stated in the introductory section, there is a lack of precise quantitative mood data. This is due to a lack of a precise measurable biological factor which causes the mood fluctuations in bipolar disorder [Goldbeter, 2013], as well as the absence of a way of studying long-term mood fluctuations in a lab setting (i.e., with controlled environmental conditions).

In summary, it can be said that the field of mood modelling, especially in the context of bipolar disorder, is still in its infancy; however, some promising findings have been presented. Once the existing computational models have been refined and expanded, they could be used as a powerful tool in clinical settings [Cochran et al., 2018]. Accurate predictions of future mood episodes might be especially impactful, as this would make early intervention possible, increasing patients' (and their relatives') quality of life significantly.

Chapter 3

Mood fluctuations caused by RPEs - A reinforcement learning model

3.1 Introduction

The concepts of reinforcement learning and reward prediction errors (RPEs) have been used to study animal and human learning for several decades. The first, and perhaps the most influential model of reinforcement learning is the Rescorla-Wagner model, which was published in 1972 [Rescorla and Wagner, 1972]. The model attempts to describe (Pavlovian) conditioning (i.e., the association between a signal and a following stimulus) by formalising the amount of learning that occurs at each iteration of conditioning.

The main aspect of reinforcement learning models is that they work on the basis of iterative value updates; i.e., the difference between the predicted and observed reward is used to learn to make better reward predictions.

Since 1972, the field has developed further and more reinforcement learning models have been published. However, the first reinforcement learning model used to model bipolar disorder was not published until 2017. This section aims to introduce this model, as it lays the foundations for further research in this area.

3.2 The theoretical and experimental basis of Mason et al.'s (2017) model

As stated above, Mason, Eldar, and Rutledge's (2017) model of bipolar disorder is the first model which tries to explain the disorder by making use of reinforcement learning as well as learning impairments that are caused by high mood bias.

Their work is largely based on a paper by Eldar and Niv and well as a paper by Eldar et al. which were published in 2015 and 2016, respectively [Eldar and Niv, 2015], [Eldar et al., 2016]. These papers describe the experimental foundations as well as the mathematical details of Mason et al.'s (2017) model.

Eldar and Niv (2015) were interested in testing whether mood manipulations could modulate reward sensitivity. They conducted an experiment with 56 human participants in order to investigate the effect of unexpected outcomes on the mood of the participants and the valuation of following outcomes. In order to achieve this, they compared the valuation of small rewards ($0.25) given by a slot machine before and after a wheel of fortune draw that had larger rewards ($7) attached to it. They hypothesised that if an unexpected outcome (winning/losing $7) has an impact on the mood of the participants and influences the perception of subsequent events, winning $7 in the wheel of fortune draw should increase the participants' mood and make them value rewards received from the slot machine played afterwards higher than rewards received from the slot machine played before and vice versa.

In order to study the correlation between patients suffering from unstable mood/bipolar disorder and the perception of rewards, all participants completed the International Personality Pool – a version of the Hypomanic Personality Scale – which is a self-report measure that correlates with frequency of mood fluctuations and the risk of developing bipolar disorder [Meyer, 2002].

Eldar and Niv's (2015) results show that participants' mood and the valuation of rewards received from the slot machine played after the wheel of fortune draw were in fact valued higher than rewards received from the slot machine played before in participants who reported a tendency to mood instability (i.e., for participants with a high HPS score). Participants with low HPS scores were not significantly influenced by the result of the wheel of fortune draw. Furthermore, when asked to choose between the slot machine encountered before and after the wheel of fortune draw, participants with high HPS scores favoured the slot machines they encountered after the draw if they had won $7 or the one from before if they had lost $7. There was no significant trend in participants with low HPS scores.

If those biased choices actually stem from a biased perception of reward, there should be a bias in neural responses to rewards in the striatum as well. The striatum is a brain area where blood-oxygen-level dependent (BOLD) signals corresponding to RPE signals that influence learning as well future choices have been identified. Eldar and Niv (2015) found that there was a correlation between higher HPS scores and stronger BOLD responses to small rewards in participants who won the wheel of fortune draw, as well as weaker responses to small rewards for the participants who lost the draw.

Thus, it can be concluded that participants with a more unstable mood tend to be affected more by an unexpected outcome and their reward perception is more biased than participants' with a low HPS score.

Mood bias is not inherently detrimental to a person's life. Moderate mood bias helps people adjust to their changing environment and enables them to make realistic predictions about the world. Mason et. al (2017) claim that in individuals suffering from bipolar disorder, however, their mood influences reward perception strongly, which then leads to mispredictions about the world. This shows in recursive cycles which cause expectations, mood, as well as behaviour to escalate to extremes.

In general, it can be said that the higher the mood bias parameter, the higher a patient's

(hyper-)sensitivity to reward (when their mood is high) is. Conversely, the lower the patient's mood, the larger their hyposensitivity to reward is.

Eldar and Niv (2015) formalise the relation between emotional state and reward perception with a reinforcement learning model. The core mechanism of this model is that positive surprises (i.e., prediction errors) improve mood and negative surprises lower mood.

They define mood as the running average of recent outcomes, which is based on other research that aims to find a formal way of expressing mood mathematically, despite the lack of a definition that makes it easy to do so [Katsimerou et al., 2014], [Marinier et al., 2009]. This formalisation enables mood to change gradually (due to the aggregated effect of multiple values) or more drastically as a response to a single extreme value, corresponding (for example) to a traumatic event.

The model suggests that the bidirectional interaction between perceived outcomes and emotional state can generate mood instability and is therefore relevant for mood modelling in bipolar disorder.

This can be illustrated by the following diagram (figure 3.1) that is included in Mason et al.'s (2017) paper.

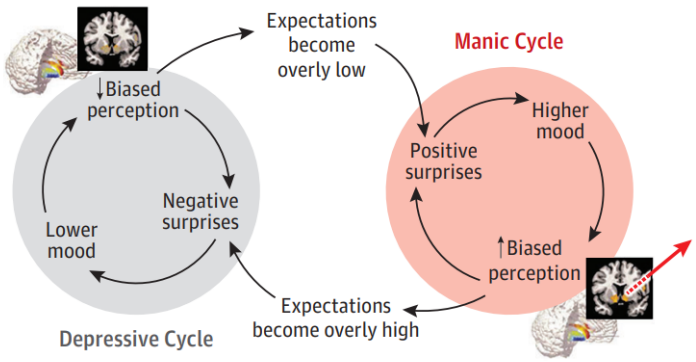

Figure 3.1: Diagram providing a high-level description of the feedback cycles causing mood instability to occur in bipolar disorder patients

Eldar et al. (2016) show that the coupling between mood and reward (expectations) can not only cause the aforementioned positive feedback dynamics, but also that its degree correlates with patients' self-reported hypomanic tendencies [Radulescu and Niv, 2019].

3.3 Formulas defining the reinforcement learning model

In general, updates of the expected value of a certain stimulus are made via the reward prediction error (denoted by δ). The reward prediction error is the difference between actual reward (r) and the expected reward (v), so $\delta = r - v$. Eldar and Niv (2015) modify this formula slightly in order for the model to be able to capture prediction errors based

on perceived reward instead of actual reward. The reward prediction error can thus be calculated as shown in equation 3.1.

$$\delta = r_{perceived} - v \tag{3.1}$$

The perceived reward ($r_{perceived}$) in this formula is calculated by multiplying actual reward (r) with a mood bias parameter (f^m) that is able to capture good, bad, and neutral mood (see equation 3.2).

$$r_{perceived} = r \cdot f^m \tag{3.2}$$

The parameter m denotes bad ($-1<m<0$) or good ($0<m<1$) mood. f is a positive constant that represents the direction (positive or negative mood bias) as well as the intensity of the mood bias. If f=1, mood does not influence the reward perception. If $0<f<1$, rewards are perceived smaller when the patient is in a good mood and larger when the patient is in a bad mood (which is also called negative feedback). Hence, if $f>1$, rewards are perceived larger when the patient is in a good mood and smaller when the patient is in a bad mood (which is also called positive feedback).

Updates of the expected reward (v) are made via iterative value updates using the standard reinforcement learning equation (see equation 3.3). Apart from the expected reward, the formula includes the reward prediction error (δ) as well as the learning rate (η_v), which denotes the step size of the value updates, i.e., how quickly the model updates the values.

$$v = v + \eta_v \cdot \delta \tag{3.3}$$

As mentioned previously, Eldar and Niv (2015) make the assumption that mood fluctuations are a direct result of recent reward prediction errors. They formalise this by using the parameter h, which denotes the recent prediction error history. The update of this value is tracked using the learning rate (i.e., step size) η_h (see equation 3.4).

$$h = h + \eta_h \cdot (\delta - h) \tag{3.4}$$

In order to constrain the values of m within -1 and 1, Eldar and Niv (2015) use the sigmoid function tanh (see equation 3.5).

$$m = tanh(h) \tag{3.5}$$

This reinforcement learning model is the most recent approach to modelling bipolar disorder as well as understanding the potential causes of it.
Reinforcement learning is a well-researched concept that has been used to successfully explain a range of different behaviours seen in animal experiments [Russell and Norvig, 2010]. Furthermore, numerous studies have established links between different mental illnesses and impairments in reinforcement learning. Examples for this are depression [Brown et al., 2021] and schizophrenia [Deserno et al., 2013], whose symptomatology - as stated in section 1.1.1 - overlap with the symptoms of bipolar disorder. Hence, it is worthwhile to explore this approach further by reimplementing Eldar and Niv's (2015) reinforcement learning model and investigating which specific dynamics that can be observed in the different subtypes of bipolar disorder

it can and cannot account for. Additionally, discovering how the individual model parameters influence the modelled mood can give pointers for future research regarding the biological processes influencing a patient's mood.

The implementation of the model which was done for this honours project uses Python 3 (and its standard libraries) and is based on the code written by Kacki (2021) .

3.4 Simulations

The first part of the simulations (sections 3.4.1-3.4.5) does not include noise and therefore displays the basic mood trends that a bipolar disorder patient might experience. In other words, the mood fluctuations that are described by the basic version of the model are oscillations (of different frequency and intensity) around the mean mood of 0 (i.e., neutral mood). In the following diagrams, the x-axis is labelled with the word "Iterations", with one iteration being one update of all the parameters. In real life, this can correspond to any unit of time, with one iteration corresponding to one day being a sensible choice.

Eldar and Niv (2015) use 300 iterations in their simulations. In the simulations done for this report (figure 3.3 onwards), the number of iterations was set to 1000 in order to capture the behaviour of the functions for a higher amount of iterations (i.e., for a longer period of time).

3.4.1 The role of the mood bias parameter f: Investigating Eldar and Niv's (2015) findings

The authors state that when using a mood bias parameter (f) of 1 and a reward value (r) of 10, the expected reward value converges to the actual reward value quickly and consequently, mood does not fluctuate. When setting f=1.2 and using the same actual reward value (r=10), however, the authors state that expected reward value as well as mood oscillate and hence do not converge, as it can be seen in figure 3.2. This is consistent with their claim that mood fluctuations in bipolar disorder arise due to heightened mood bias.

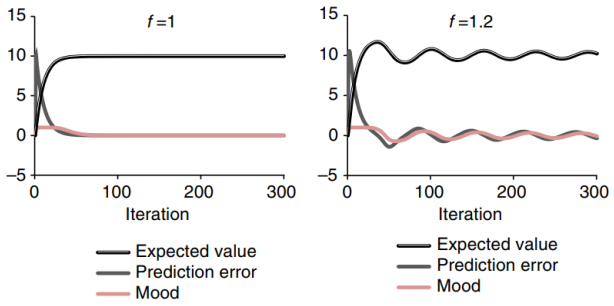

Figure 3.2: Simulations of the interaction between learning and mood, using the reinforcement learning model described above [Eldar and Niv, 2015]

Eldar and Niv (2015) do not mention which values they use for the learning rates η_v and η_h. The above result for f=1 can be reproduced for any learning rate, one example (for η_v=0.1 and η_h=0.1) is shown in figure 3.3.

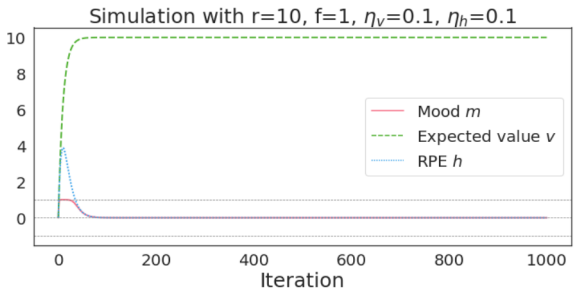

Figure 3.3: The interactions between learning and mood for f=1 when the number of iterations n=1000

For f=1.2, however, the behaviour of the mood function depends of the learning rates. When η_v=0.1 and η_h=0.1 (which looks similar to the second function in figure 3.2), the function clearly converges, which can be observed due to the scaling of the x-axis (see figure 3.4).

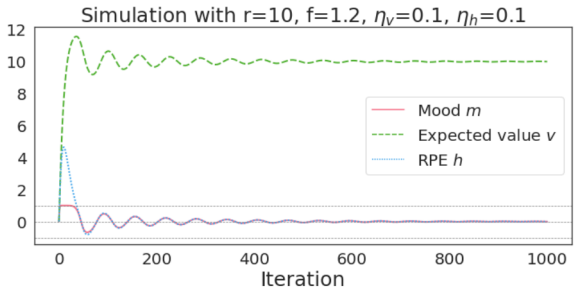

Figure 3.4: The interactions between learning and mood for f=1.2 when the number of iterations n=1000

However, when increasing η_h (in the first diagram in figure 3.5, η_h=0.2 was used) or when decreasing η_v (in the second diagram in figure 3.5, η_v=0.05 was used), the oscillatory behaviour of the function continues until infinity (and hence, it does not converge).

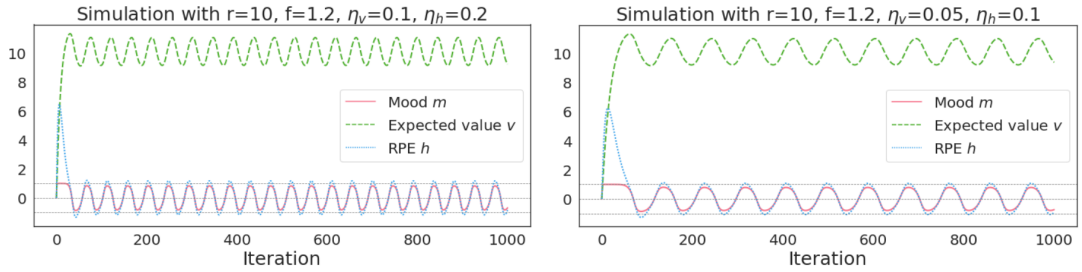

Figure 3.5: The interactions between learning and mood for f=1.2 when varying the learning rates η_v and η_h, for n=1000

When the mood bias parameter is sufficiently large, however, the oscillations' intensity increases and the function does not converge for any sensible value of the learning rates. One example of this is f=2; the corresponding graph is shown in figure 3.6.

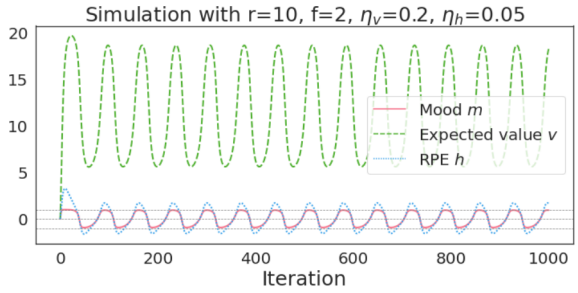

Figure 3.6: The interactions between learning and mood for f=2 when the number of iterations n=1000

As stated above, if f=1 (which means that mood does not influence learning), the mood function converges quickly. For slightly higher values of f (for r=10, e.g., f=1.2), the learning rates η_v and η_h determine whether the function oscillates. If f is sufficiently large, the function continuously oscillates regardless of the other parameter values. Mason et al. (2017) claim that different values of the mood bias parameter are able to produce different frequencies and intensities of oscillations and can therefore model different subtypes of bipolar disorder (see figure 3.7).

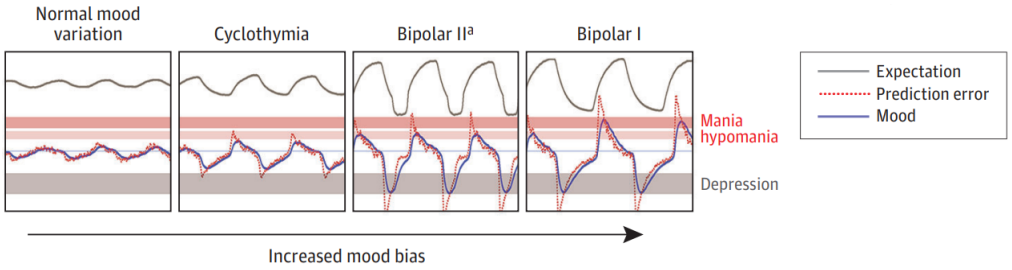

Figure 3.7: According to Mason et al. (2017), increasing the mood bias parameter produces functions modelling three subtypes of bipolar disorder

For visualisation purposes, the reward value r=3 was chosen for the simulations below.

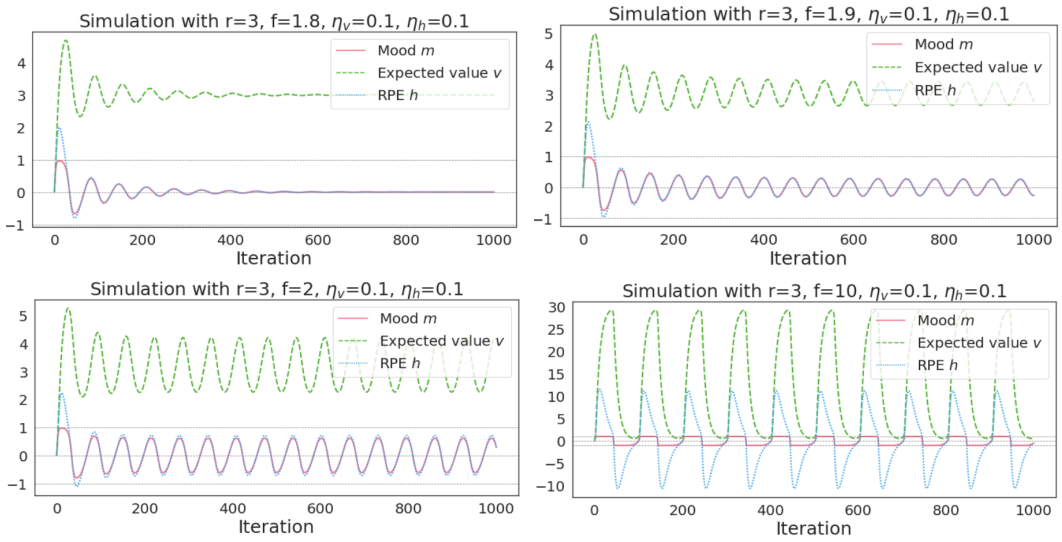

Figure 3.8: The interactions between learning and mood for increasing values of f when the number of iterations n=1000

As it can be seen in figure 3.8, increasing the mood bias parameter while keeping the other parameters constant predominately increases the intensity (i.e., the magnitude) of the mood fluctuations. One can observe that the number of manic and depressive episodes decreases slightly as the mood bias parameter increases; however, in order for this to have a significant effect, the mood bias parameter would have to increase drastically, which would result in the reward prediction errors rising to a level that results in the mood staying "at" (i.e., infinitely close to) its boundary values (-1 and 1) for several iterations at a time. This behaviour is not represented in figure 3.7 and thus, it can be said that the mood bias parameter cannot explain the mood variations observed in the different subtypes of bipolar disorder (as described by Mason et al. (2017)) by itself.

Furthermore, Mason et al.'s (2017) graphs (figure 3.7) show that the mood function reaches different maximum values (corresponding to hypomanic episodes in bipolar II disorder and manic episodes in bipolar I disorder) but the same levels of depression. This asymmetric behaviour of the function modelling bipolar II disorder can only be introduced by learning rate asymmetries (which are described in section 3.4.3), and not by solely varying the mood bias parameter. Mason et al. (2017) mention this fact in their paper, however, it is not mentioned in figure 3.7, making it potentially misleading.

It can be concluded that the mood bias parameter plays a crucial role in determining the shape of the function describing learning and mood, but the other parameters in the system can introduce oscillatory and other behaviour (e.g., asymmetric mood fluctuations) as well. However, when the mood bias parameter is sufficiently large, learning the correct reward value is impossible for the agent (regardless of the learning rates) and the mood function oscillates without converging. The sections below investigate the roles of the learning rates η_h, η_v, as well as the reward r further.

3.4.2 The role of the learning rate η_h

The learning rate η_h describes how quickly the reward prediction error history is updated following the most recent reward prediction error. Since mood is directly based on recent reward prediction errors in this model, the learning rate η_h is able to influence mood directly. As it can be seen in figure 3.9, varying η_h can thus either support the convergence of the function or introduce infinitely oscillatory behaviour. Hence, the model predicts that not only mood bias can lead to mood fluctuations, but the extent to which a patient (i.e., their overall reward prediction error and hence their mood) is influenced following an observed misprediction of reward has this ability as well.

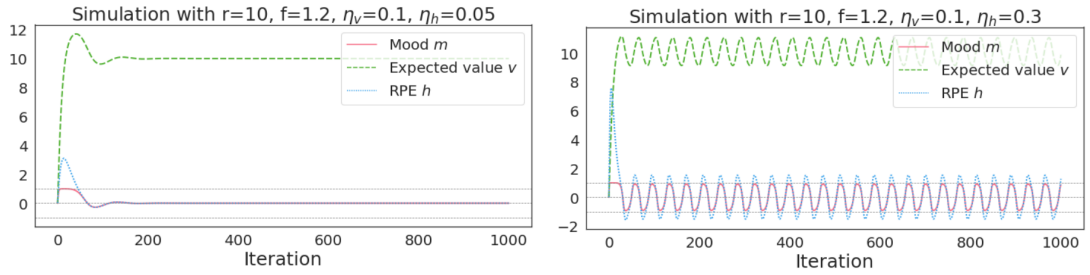

Figure 3.9: The interactions between learning and mood for different values of η_h when the number of iterations n=1000

3.4.3 The role of the learning rate η_v

As stated above, Mason et al. (2017) mention that asymmetries of the learning rate η_v are able to model the differences in the dynamics that are characteristic for bipolar I (manic, and usually some depressive episodes) and bipolar II (hypomanic (but not manic) and depressive episodes) disorder.

When an agent observes that they have mispredicted a reward (i.e., they have made a reward prediction error), they adjust their prediction for the next reward with the help of the learning rate η_v, which is a scalar that influences how drastically they change their reward prediction. When the learning rate for positive reward prediction errors (η_{v^+}) is higher than the one for negative reward prediction errors (η_{v^-}), the agent quickly adjusts their reward prediction upwards when positive surprises are encountered but only slowly adjusts it downwards when negative surprises are encountered. Thus, when the agent experiences a positive surprise, they increase their reward prediction significantly and the chance that a negative surprise in the next iteration is encountered is higher. Since the learning rate for negative surprises is lower, the agent continues to encounter negative surprises, which results in depressed mood.

The updates of the expected reward can thus be split up according the learning rates η_{v^+} and η_{v^-} as follows (equation 3.6).

$$v = \begin{cases} v + \eta_{v^+} \cdot \delta, \text{ if } \delta \geq 0 \\ v + \eta_{v^-} \cdot \delta, \text{ if } \delta < 0 \end{cases} \tag{3.6}$$

As it can be clearly seen in the simulations below (figure 3.10), mood is predominately

in the depressive range and only reaches hypomanic levels when the learning rate for positive reward prediction errors is higher than for negative reward prediction errors. Conversely, when the learning rate is lower for positive reward prediction errors, mood is predominantly in the manic range and does not reach extremely low values. Hence, the model predicts that when a person is more influenceable by a positive than a negative surprise, they tend to predominantly experience negative surprises and therefore depressed mood. As mentioned above, choosing values for η_{v+} and η_{v-} such that $\eta_{v+} > \eta_{v-}$ can aid in modelling the hypomanic episodes observed in bipolar II disorder.

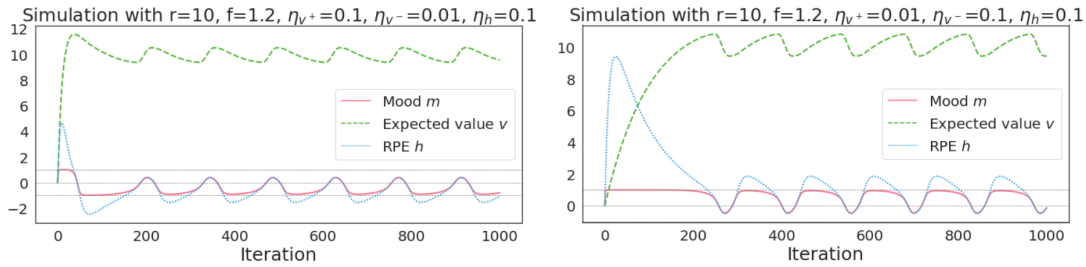

Figure 3.10: The interactions between learning and mood, with asymmetric learning rates η_{v+} and η_{v-} when number of iterations n=1000

3.4.4 The role of the reward r

The higher the reward, the harder it is for an agent with moderate mood bias to learn the correct reward value (see figure 3.11).

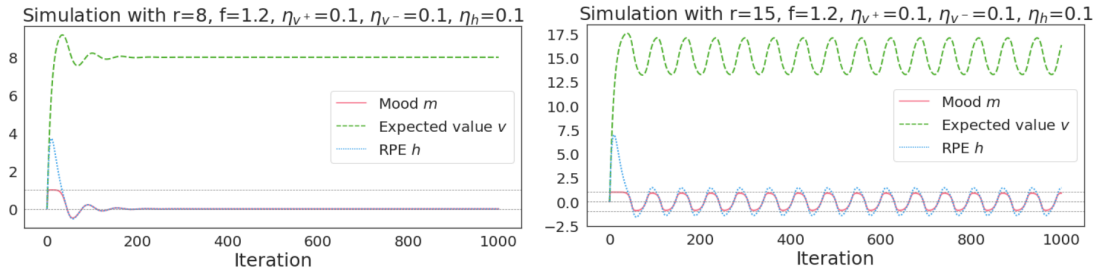

Figure 3.11: The interactions between learning and mood, with different amounts of actual reward when number of iterations n=1000

Due to the mathematical relationship of reward and mood bias in this model (see equation 3.2), increasing the reward results in lower values of the mood bias to be needed in order for oscillatory behaviour of the mood function to occur. This would mean that if the reward in an environment is high, one is more likely to experience mood fluctuations, i.e., if the reward is sufficiently high, even people with only slightly heightened mood bias parameters would experience bipolar mood fluctuations - a finding that has not been investigated to date.

3.4.5 Defining the functions piecewise accounts for additional dynamics

Johnson (2005) notes that a positive life event can cause a manic episode to occur in bipolar disorder patients and due to the linked positive feedback cycles in this framework, oscillatory behaviour of mood as well. By defining the parameters piecewise (i.e., increasing the reward at a certain point), Mason et al.'s (2017) model is able to capture this dynamic (see figure 3.12).

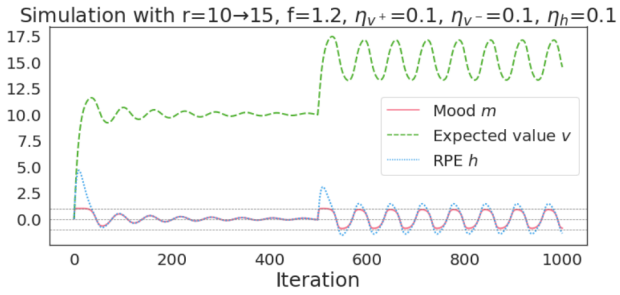

Figure 3.12: The interactions between learning and mood when the reward value is increased from r=10 to r=15 at n=500

It has to be noted that negative life events are able to trigger depressive episodes and thus oscillatory behaviour of mood as well. The model can account for the initial depressive episode, however, is unable to account for the subsequent mood oscillations (see figure 3.13).

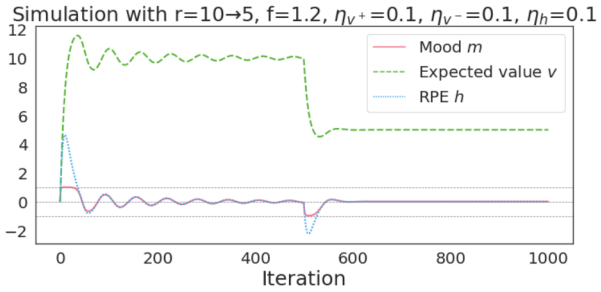

Figure 3.13: The interactions between learning and mood when the reward value is decreased from r=10 to r=5 at n=500

Assuming that Mason et al.'s (2017) claim is true and bipolar disorder is caused by an abnormally heightened mood bias parameter, substance and medication induced bipolar disorder could be understood as a sudden change in the mood bias parameter in the presence or absence of the causal element. As established in section 3.4.1, setting the mood bias parameter f to a sufficiently high value results in oscillatory behaviour of the mood function. Hence, by using the mood bias parameters f=1, f=1.5, and f=1 (where f=1.5 corresponds to the presence of the causal element and f=1 corresponds to its absence), the model is able to display the dynamics observed in patients suffering from this subtype of bipolar disorder (see figure 3.14).

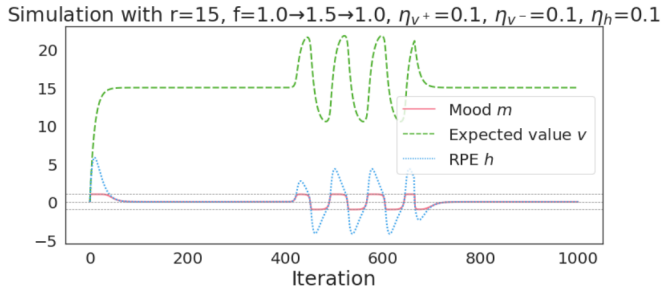

Figure 3.14: The interactions between learning and mood when the mood bias parameter f is changed from f=1 to f=1.5 to f=1 when n=1000

3.4.6 Adding noise to the system: 'Realistically' modelling different subtypes of bipolar disorder

In order to make the modelled mood fluctuations more realistic, noise can be introduced into the system. In this case, this was done by sampling the reward on each iteration from a Gaussian distribution with the mean value that was given as input to the function and a variance of 1. Figure 15 shows that the model is able to produce dynamics that are characteristic for the different subtypes of bipolar disorder. Generally, the boundaries for the different mood states were roughly defined as follows: $m > 0.5$ corresponds to manic episodes, $0.25 \leq m \leq 0.5$ to hypomanic episodes, -0.25 $\leq m \leq$ 0.25 to normal mood, and $m < 0.25$ to depressive episodes.

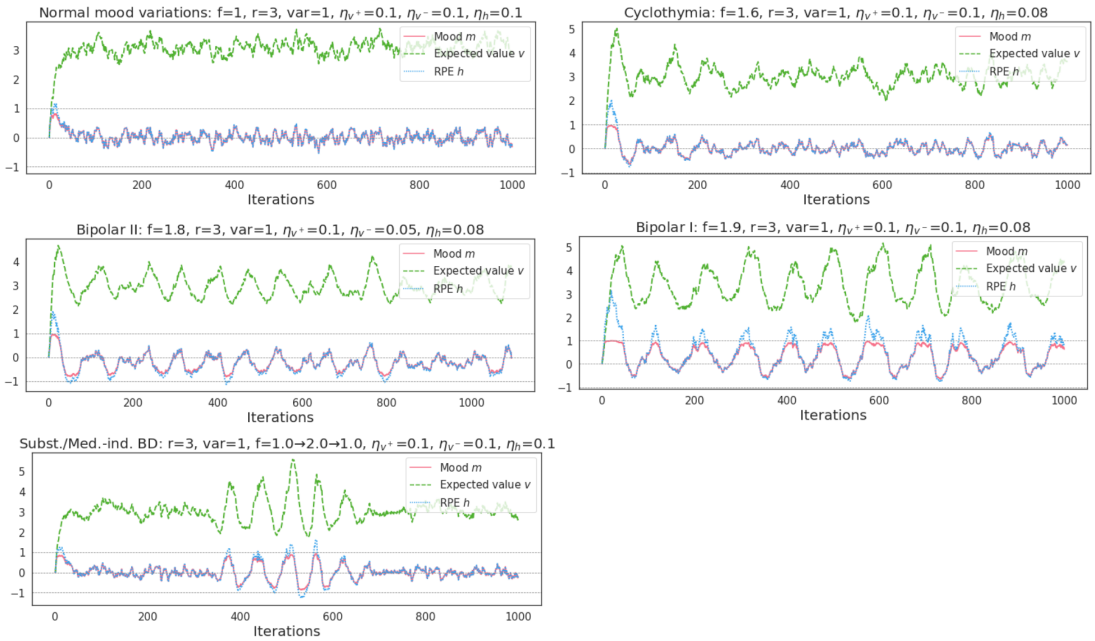

Figure 3.15: Mood fluctuations that are characteristic for normal mood fluctuations (top left), cyclothymia (top right), bipolar II disorder (middle left), bipolar I disorder (middle right), and substance & medication induced bipolar disorder (bottom left)

As it can be seen above, normal mood fluctuations were produced with a mood bias

parameter f=1, cyclothymia with f=1.6, bipolar II disorder with f=1.8, and bipolar I disorder with f=1.9. These values were chosen in order for the simulations to be consistent with Mason et al.'s (2017) claim that increasing the mood bias parameter in the order mentioned above causes the different mood dynamics. However, the mood bias parameter alone is unable to account for the different dynamics, which is why the other model parameters were also varied if deemed necessary for the visualisation of the different mood dynamics.

The higher the number of parameters that have to be varied in order for the model to produce realistic mood variations over time, the harder finding links between specific biological factors and the model parameters is. In other words, if the mood bias parameter was the only factor that could account for all the different dynamics, the probability of future research being able to find one measurable biological factor corresponding to the mood bias parameter in this theoretical model is rather high. Since the mood bias parameter is not able to account for all the dynamics and other parameters have to be varied as well, the process of finding biological links is significantly more complicated.

3.5 Evaluation & Conclusion

Mason et al.'s (2017) model of mood fluctuations in bipolar disorder is based on the approach that bipolar disorder patients' mood bias is too strong, which leads to positive feedback cycles that cause significant reward prediction errors and therefore manic and depressive episodes.

The model is mathematically rather simple, yet versatile. By modifying the different parameters and introducing noise to the system, one can generate a range of realistic mood dynamics that could be observed in patients suffering from one of the different subtypes of bipolar disorder. Furthermore, the model can also generate "normal", i.e., healthy, mood fluctuations (which are experienced by bipolar disorder patients during euthymic periods as well). Since the model parameters can be defined piecewise (i.e., the model parameters can be changed for a certain range of x-values, in this case iterations), all dynamics that can be generated individually can theoretically be used to define one mood function that has all of those (or a combination of several different) characteristics.

However, Mason et al.'s (2017) model has several shortcomings.

Firstly, mood is modelled one-dimensionally, meaning that it cannot be in the (hypo)manic or depressive range at the same time. Hence, Mason et al.'s (2017) model (unlike Goldbeter's (2011 & 2013) model) is not able to account for mixed episodes, during which a patient experiences both manic and depressive symptoms.

Secondly, the mood fluctuations generated by the model are periodic and any behaviour shown in the graphs above that seems to not be periodic is solely caused by the added noise (unless the function is defined piecewise). Patients can experience extended periods of "normal" mood fluctuations (euthymia), as well as an unequal number of (hypo-)manic and depressive episodes over a certain period of time, which the model is

unable to account for if not defined piecewise.

Furthermore, as stated in section 2.12, Mason et al.'s (2017) model is a generative model. This means that it does not describe observed biological data but rather generates dynamics that could potentially be observed during experiments. However, no measurable biological factor that is responsible for the mood fluctuations in bipolar disorder is currently known [Goldbeter, 2013]. Hence, it cannot be assessed whether the parameter settings that cause different mood dynamics to occur in the model correspond to specific biological factors and are thus realistic. Several parameters in the model are able to produce mood oscillations when varied and which one/which combination of them is the cause of the mood fluctuations in bipolar disorder remains unknown. Mason et al. (2017) note that investigating which of the parameters provides the best explanation of the oscillatory behaviour of mood in bipolar disorder patients is crucial as the different parameters correspond to different biological mechanisms and therefore have to be targeted in different ways when treating bipolar disorder.
The fact that there is no formal definition of mood that can easily be formalised mathematically adds to this problem as well [Katsimerou et al., 2014]. However, as stated previously, these core issues constitute specific pointers for future research and can potentially aid in the understanding of mood as a concept as well as the causes of mood fluctuations in bipolar disorder patients.

Lastly, Mason et al.'s (2017) model only accounts for the basic characteristic dynamics of several different subtypes of bipolar disorder. It does not include the effects of potential treatment (therapy, medication, as well as lifestyle changes – as described in section 1.1.4), or the fact that the condition tends to worsen over the course of time if it is not treated [IsHak et al., 2012].

In conclusion, the reinforcement learning model of mood fluctuations in bipolar disorder that was introduced by Eldar and Niv (2015), Eldar et al. (2016), and Mason et al. (2017) is flexible enough to account for a range of different mood dynamics experienced by bipolar disorder patients. However, it is impossible to model mixed episodes and the different dynamics that can be modelled can only be accounted for when varying several model parameters (and not just the mood bias parameter, as Mason et al. (2017) claim). This complicates the search for specific links to biological factors that are impaired in bipolar disorder patients. The effects of treatment as well as the fact that the condition tends to worsen over time if it is not treated, which are two major aspects of bipolar disorder patients' lives, are also not accounted for by Mason et al.'s (2017) model. Chapter 4 of this report therefore aims to expand the range of dynamics the model can account for by incorporating some of the aforementioned aspects.

Chapter 4

Extension of Mason et al.'s (2017) model

4.1 Modelling the effects of treatment

As stated in chapter 1.1.4, there are three main approaches to the treatment of bipolar disorder. All of them aim to reduce the intensity and ultimately prevent the occurrence of manic and depressive episodes. Since the effects of different types of medication are defined best (i.e., the most similar for patients overall), the following sections aim to model the effects of different types of medication (and a combination thereof).

In order to account for the changed mood dynamics as a result of medication, additional parameters, each influencing one of parameters in the set of equations, can be introduced. The updates at each iteration are made after in- or decreasing the parameters by the value that these parameters are set to (c_f for the mood bias parameter f (equation 4.1), $c_{\eta_{v^+}}$ for the learning rate η_{v^+} (equation 4.2), $c_{\eta_{v^-}}$ for the learning rate η_{v^-} (equation 4.3), and c_{η_h} for the learning rate η_h (equation 4.4)). Hence, before equations 3.1, 3.2, and 3.4-3.6 are used, the following individual parameter updates (equations 4.1-4.6) are made. In other words, at each iteration, the mood bias parameter f, as well as the learning rates η_{v^+}, η_{v^-}, and η_h can be in- or decreased individually, given that the new value is not below a specified threshold value (f=1 for the mood bias parameter, as this represents no influence of mood on learning; and 0.05 for the learning rates, since it is not realistic to assume that the learning rates can be completely 0 (one always updates their beliefs and expectations about the environment they are in following events that happen, even if just by a small amount)).

$$f = \begin{cases} f - c_f, \text{ if } f - c_f \geq 1 \\ f \text{ otherwise} \end{cases} \tag{4.1}$$

$$\eta_{v^+} = \begin{cases} \eta_{v^+} - c_{\eta_{v^+}}, \text{ if } \eta_{v^+} - c_{\eta_{v^+}} \geq 0.05 \\ \eta_{v^+} \text{ otherwise} \end{cases} \tag{4.2}$$

$$\eta_{v^-} = \begin{cases} \eta_{v^-} - c_{\eta_{v^-}}, \text{ if } \eta_{v^-} - c_{\eta_{v^-}} \geq 0.05 \\ \eta_{v^-} \text{ otherwise} \end{cases} \tag{4.3}$$

$$\eta_h = \begin{cases} \eta_h - c_{\eta_h}, \text{ if } \eta_h - c_{\eta_h} \geq 0.05 \\ \eta_h \text{ otherwise} \end{cases} \tag{4.4}$$

In order to make the simulations comparable and potentially more easily linkable to biological factors, solely the additional parameters introduced in equations 4.1-4.4 were varied. The other parameters were set to the following values: f=2, η_{v^+}=0.1, η_{v^-}=0.1, and η_h=0.1. The reward value for each iteration was sampled from a Gaussian distribution with a mean of 3 and a variance of 1.

4.1.1 The effects of mood stabilisers

Mood stabilisers are drugs that have the ability to reduce mania. They can, but do not necessarily, cause depressive symptoms to decrease [Simonetti et al., 2020]. This change in the mood dynamics can be achieved by setting the change in the learning rate from positive surprises to 0, while choosing a non-zero value (in this case 0.00007) for the change in the learning rate from negative surprises. Selecting these parameter values means that the learning rate from negative surprises is gradually reduced, resulting in the agent getting worse at learning from negative surprises (until the threshold value of η_{v^-}=0.05 is reached). In other words, with every iteration, the amount by which they update their reward prediction in case of a negative surprise decreases. If they encounter a negative surprise, they are thus increasingly likely to encounter subsequent negative surprises, resulting in depressed mood. Since the mood bias parameter f as well as the learning rate η_h are gradually reduced as well, the oscillatory behaviour of the system is diminished and thus, extreme manic episodes are prevented over time. Hence, mood is stabilised but depressive episodes do not disappear completely. The modelled dynamics are represented graphically in figure 4.1.

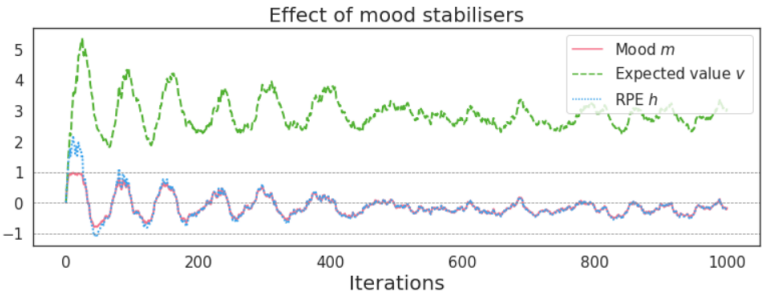

Figure 4.1: The effect of mood stabilisers - modelled with the following parameters: c_f=0.00005, $c_{\eta_{v^+}}$=0, $c_{\eta_{v^-}}$=0.00007, c_{η_h}=0.0001

Whether mood stabilisers affect a bipolar disorder patient's learning rate in reality has not been determined yet. However, links between learning and the effect of mood stabilisers have been proposed. For instance, protein kinase C, an enzyme that has been linked to the pathophysiology of bipolar disorder and that is inhibited by mood stabilisers [Saxena et al., 2017], has been shown to be activated in classical conditioning experiments where animals were exposed to a harmful stimulus [Noguès, 1997].

4.1.2 The effects of antidepressants

Antidepressants generally reduce the depressive symptoms a patient experiences. However, as stated above, they cause (hypo-)manic episodes in a significant number of patients, which is modelled by the function in figure 4.2. The parameters c_f and c_{η_h} were set to the same values as above (c_f=0.00005 and c_{η_h}=0.0001) for this simulation. Since mood stabilisers and antidepressants can be interpreted as roughly having 'opposite' effects, or rather, since they work on opposite ends of the mood scale (i.e., mood stabilisers decrease mania but do not necessarily decrease depressive symptoms, and antidepressants decrease depression but can induce mania), the values for $c_{\eta_{v^+}}$ and $c_{\eta_{v^-}}$ were reversed. Hence, $c_{\eta_{v^+}}$ was set to 0.00007 and $c_{\eta_{v^-}}$ was set to 0.

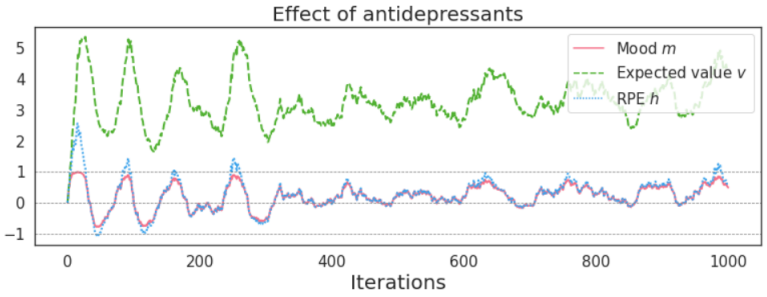

Figure 4.2: The effect of antidepressants - modelled with the following parameters: c_f=0.00005, $c_{\eta_{v^+}}$=0.00007, $c_{\eta_{v^-}}$=0, c_{η_h}=0.0001

4.1.3 The effects of mood stabilisers and antidepressants combined

When both mood stabilisers and antidepressants are used to treat bipolar disorder, a gradual reduction in severity of mood fluctuations can be observed over time [Goldberg and Ghaemi, 2005] (see figure 4.3). In order to model this, the parameters c_f and c_{η_h} were set to 0.00005 and 0.0001, respectively. The learning rate change parameters $c_{\eta_{v^+}}$ and $c_{\eta_{v^-}}$ were set to $c_{\eta_{v^+}}$=$c_{\eta_{v^-}}$=0.00001. The onset of the effect of psychotropic drugs is usually delayed by 2-4 weeks [Thompson, 2002], which is included in the model by reducing the values of f, η_{v^+}, η_{v^+}, and η_h (i.e., applying the formulas defined equations 4.1-4.4) from iteration 21 (corresponding to three weeks, if the assumption that one iteration represents one day is made) onwards. For all previous iterations, the given values for f, η_{v^+}, η_{v^+}, and η_h were used.

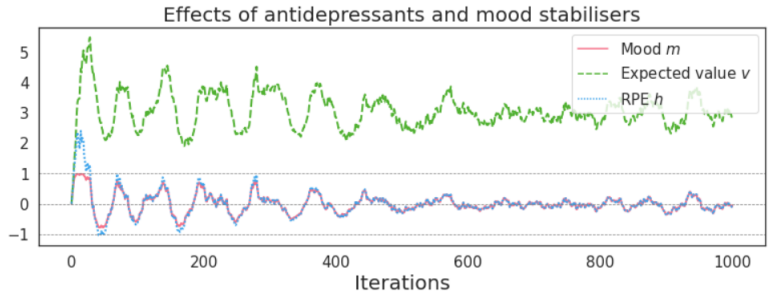

Figure 4.3: The effect of mood stabilisers and antidepressants - modelled with the following parameters: c_f*=0.00005,* $c_{\eta_{v^+}}$*=0.00001,* $c_{\eta_{v^-}}$*=0.00001,* c_{η_h}*=0.0001*

4.2 Modelling untreated bipolar disorder

If bipolar disorder is not treated, it tends to worsen over time [IsHak et al., 2012]. In order to model this effect, the parameters c_f, $c_{\eta_{v^+}}$, $c_{\eta_{v^-}}$, and c_{η_h} were set to negative values (c_f=-0.00005, $c_{\eta_{v^+}}$=-0.000007, $c_{\eta_{v^-}}$=-0.000007, and c_{η_h}=-0.0001), meaning that a gradual increase the parameters f, η_{v^+}, η_{v^-}, and η_h can be achieved. In order for the mood fluctuations to get visually worse over time, the mood bias parameter was initially set to f=1.2.

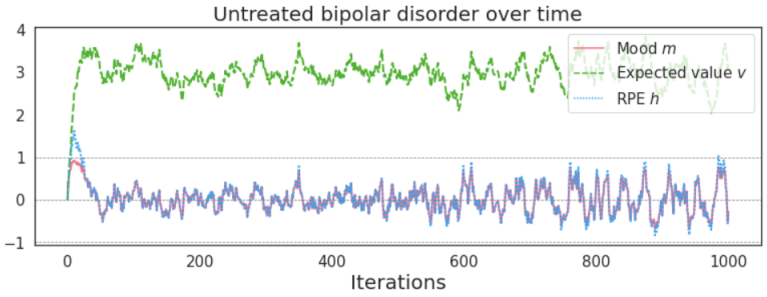

Figure 4.4: Mood fluctuations in untreated bipolar over time - modelled with the following parameters: c_f*=-0.00005,* $c_{\eta_{v^+}}$*=-0.000007,* $c_{\eta_{v^-}}$*=-0.000007,* c_{η_h}*=-0.0001*

4.3 Conclusion

In summary, the range of mood fluctuations in bipolar disorder that Mason et al.'s (2017) model can account for can be extended by the introduction of additional variables that modify the mood bias parameter f, as well as the learning rates η_{v^+}, η_{v^-}, and η_h. A reduction in f, η_h, and η_{v^-} models the effect of mood stabilisers over time, while a reduction in f, η_h, and η_{v^+} models the effect of antidepressants over time. This suggests that mood stabilisers reduce the learning rate from negative surprises over time, resulting in depressed mood (as described in section 3.4.3), and antidepressants reduce the learning rate from positive surprises, resulting in elevated mood. Assuming that mood bias is (at least partly) responsible for the mood fluctuations in bipolar disorder, the reduction in the mood bias parameter f leads to less oscillatory behaviour over time.

The decrease in η_h, which displays how much recent reward prediction errors influence the reward prediction error history (and thus, mood) has the same effect.

The modelling that was done in this chapter is based on Kacki's (2021) general approach that introduces a gradual reduction in the model's parameter values. The formulas chosen here (equation 4.1-4.4), however, result in a linear reduction of the model parameters (opposed to exponential decay), which - importantly - have lower thresholds. This results in the simulations done here remaining realistic for higher values of iterations than n=1000 as well. In other words, the learning rates cannot be reduced to zero (or infinitely close to zero).

Furthermore, Kacki's (2021) simulations predict approximately eight depressive and manic episodes in one year for an unmedicated patient, which is theoretically possible, however, most patients experience four or less episodes per year [Bauer et al., 2008]. Additionally, the author's simulations predict that the full effect of combined treatment with a mood stabiliser and an antidepressant is reached after approximately 3500 iterations, which - when making the assumption that one iteration corresponds to one day - corresponds to roughly 9-10 years. However, research shows that a patient usually becomes stable within several months and only years in rare cases [Lippard et al., 2020]. Hence, the parameters for the simulations presented in this report were adjusted in order to account for these facts.

Additionally, Kacki (2021) includes a scalar which is able to manipulate the mood value in equation 3.2. Since mood cannot be influenced directly, as it is dependent on the recent reward prediction error history, this was not included here. The author does not include a way to manipulate the mood bias parameter f, which - assuming that it is (partly) responsible for the mood fluctuations in bipolar disorder and the fact that the disorder is treatable - should be adjusted when modelling the effects of treatment (or the course of the disorder when it is not treated) and thus, the parameter c_f was included here.

In general, research tends to solely focus on the subtypes bipolar I and bipolar II disorder (when untreated) and the course of the disorder for an untreated patient is only modelled in Nana (2009) and Daugherty et al. (2009), with mood fluctuations gradually getting more extreme, until a large stable limit cycle is reached.

Chang and Chou (2018) propose another variant of Mason et al.'s (2017) model, which includes two separate parameters for mood learning and mood recovery (which is a parameter that drives mood towards the neutral level of zero). As shown in figure 2.10, they model the effect of antidepressants on a patient's mood for a short time period (30 weeks), and also do not mention the option of treatment with mood stabilisers (and therefore, combined treatment either).

The extensions of Mason et al.'s (2017) model that were done for this honours project thus aim to go into greater depth and be more realistic than the aforementioned ones.

Chapter 5

Conclusion

The aim of this honours project was to review existing frameworks which attempt to model mood fluctuations in bipolar disorder, as well as to provide an explanation and simulations of a recent framework ([Mason et al., 2017]) in greater depth.

Chapter 1 contains an overview of bipolar disorder (subtypes, causes, treatment options), as well as an introduction to the areas of computational psychiatry, reinforcement learning, and mood modelling.

The aim of chapter 2 was to give the reader a comprehensive overview of existing mood modelling frameworks for bipolar disorder. These are based on several different approaches, such as mood being modelled one-dimensionally by an autoregressive model [Bonsall et al., 2011], two-dimensionally by the mutual inhibition of two variables corresponding to mania and depression [Goldbeter, 2011], [Goldbeter, 2013], or mood fluctuations arising due to reward prediction errors (e.g., [Mason et al., 2017]).

The latter framework (which is the most recent approach to modelling the mood fluctuations in bipolar disorder) views the recent reward prediction errors as the cause of the mood fluctuations, i.e., the patient is unable to make realistic predictions about future rewards based on observed rewards and thus gets stuck in cycles that cause their mood to rise or decrease to extremes, also known as manic and depressive episodes. The resulting reinforcement learning model that was developed by Eldar and Niv (2015), Eldar et al. (2016), as well as Mason et al. (2017) was reimplemented and used for various simulations of certain characteristic mood dynamics that can be observed in bipolar disorder patients in chapter 3.

Using the concept of reinforcement learning in order to explain the mood fluctuations in bipolar disorder is a promising approach, as it has been successfully used to explain numerous processes observed in animal experiments and several links between impairments in reinforcement learning and mental illnesses have been identified. At its core, Mason et al.'s (2017) model is an oscillator, which several other models described in chapter 2 are as well. However, the framework is a promising approach to explaining mood fluctuations in bipolar disorder since it uses the well-researched and thus established concept of reinforcement learning and attempts to explain, and not only display, the mood fluctuations in bipolar disorder. Moreover, it provides a precise way

of modelling as well as explaining the different dynamics observed in bipolar I and bipolar II disorder (namely, manic vs. hypomanic episodes caused by learning rate asymmetries).

Chapter 4 contains an extension to Mason et al.'s (2017) model. It incorporates the effect of antidepressant treatment, mood stabiliser treatment, as well as combined treatment by introducing variables that have the ability to in- or decrease the mood bias parameter and the different learning rates. These variables can be used to model the course of the disorder if it is untreated as well, which is also shown in chapter 4. Thus, the extensions to the model (as described by equations 4.1-4.4) are able to expand the range of mood dynamics that Mason et al.'s (2017) reinforcement learning model of mood fluctuations in bipolar disorder is able to account for.

In summary, the models described in this report aim to represent different mood dynamics that can be observed in bipolar disorder patients, such as (hypo-)manic and depressive episodes, medication-induced (hypo-)manic episodes, as well as inter-episodic mood instability in some cases. While a wide range of dynamics can be accounted for, there is a lack of a strong link between the models and measurable biological factors, meaning that the models' validities cannot be determined. Some biological factors, such as the abnormalities in the serotonergic and dopaminergic system [Balasubramani and Chakravarthy, 2019] or the interaction of the pathways of protein kinases A and C [Frank, 2013], have been proposed as the mechanisms that could be responsible for the mood fluctuations in bipolar disorder; however, their validity still has to be determined with the help of large-scale studies. These further investigations would allow for the individual model parameters to be linked to specific biological factors and therefore unveil which model has the best potential to describe and predict mood fluctuations in bipolar disorder.

Computational psychiatry is a powerful tool that can be aid in the explanation and prediction of mood fluctuations in bipolar disorder (as well as other mental illnesses). Over the course of the last two decades, numerous computational models that provide promising mechanistic approaches to explaining mood fluctuations in bipolar disorder have been proposed. These models thus provide a number of specific pointers for future research in the field of bipolar disorder.

This is exceptionally important, as a better understanding of mental illnesses makes early intervention and treatment possible, which can help in the prevention of suicide and improve patients' (and their relatives') quality of life massively.

Bibliography

[Adams et al., 2015] Adams, R. A., Huys, Q. J. M., and Roiser, J. P. (2015). Computational psychiatry: towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery & Psychiatry*, 87:53–63.

[American Psychiatric Association, 2002] American Psychiatric Association (2002). *Practice guideline for the treatment of patients with bipolar disorder (revision)*. American Psychiatric Publishing.

[American Psychiatric Association, 2013] American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. American Psychiatric Publishing.

[Angst, 2008] Angst, J. (2008). Bipolar disorder–methodological problems and future perspectives. 10(2):129–139.

[Balasubramani and Chakravarthy, 2019] Balasubramani, P. P. and Chakravarthy, V. S. (2019). Bipolar oscillations between positive and negative mood states in a computational model of basal ganglia. *Cognitive Neurodynamics*, 14(2):181–202.

[Bauer et al., 2008] Bauer, M., Beaulieu, S., Dunner, D. L., Lafer, B., and Kupka, R. (2008). Rapid cycling bipolar disorder - diagnostic concepts. *Bipolar Disorders*, 10(1p2):153–162.

[Bayani et al., 2017] Bayani, A., Hadaeghi, F., Jafari, S., and Murray, G. (2017). Critical slowing down as an early warning of transitions in episodes of bipolar disorder: A simulation study based on a computational model of circadian activity rhythms. *Chronobiology International*, 34(2):235–245.

[Berk et al., 2007] Berk, M., Dodd, S., Kauer-Sant'Anna, M., Malhi, G. S., Bourin, M., Kapczinski, F., and Norman, T. (2007). Dopamine dysregulation syndrome: implications for a dopamine hypothesis of bipolar disorder. 116(s434):41–49.

[Berry and Fristedt, 1985] Berry, D. A. and Fristedt, B. (1985). Introduction. In *Bandit problems*, pages 1–8. Springer Netherlands.

[Bonsall et al., 2015] Bonsall, M. B., Geddes, J. R., Goodwin, G. M., and Holmes, E. A. (2015). Bipolar disorder dynamics: affective instabilities, relaxation oscillations and noise. *Journal of The Royal Society Interface*, 12(112):20150670.

[Bonsall et al., 2011] Bonsall, M. B., Wallace-Hadrill, S. M. A., Geddes, J. R., Goodwin, G. M., and Holmes, E. A. (2011). Nonlinear time-series approaches in charac-

terizing mood stability and mood instability in bipolar disorder. *Proceedings of the Royal Society B: Biological Sciences*, 279(1730):916–924.

[Brown et al., 2021] Brown, V. M., Zhu, L., Solway, A., Wang, J. M., McCurry, K. L., King-Casas, B., and Chiu, P. H. (2021). Reinforcement learning disruptions in individuals with depression and sensitivity to symptom change following cognitive behavioral therapy. *JAMA Psychiatry*, 78(10):1113.

[Busk et al., 2020] Busk, J., Faurholt-Jepsen, M., Frost, M., Bardram, J. E., Kessing, L. V., and Winther, O. (2020). Forecasting mood in bipolar disorder from smartphone self-assessments: Hierarchical bayesian approach. *JMIR mHealth and uHealth*, 8(4):e15028.

[Chan, 1993] Chan, K. S. (1993). Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *The Annals of Statistics*, 21(1):520–533.

[Chang and Chou, 2018] Chang, S.-S. and Chou, T. (2018). A dynamical bifurcation model of bipolar disorder based on learned expectation and asymmetry in mood sensitivity. *Computational Psychiatry*, 2(0):205.

[Cochran et al., 2018] Cochran, A. L., Schultz, A., McInnis, M. G., and Forger, D. B. (2018). Testing frameworks for personalizing bipolar disorder. *Translational Psychiatry*, 8(1).

[Collenette et al., 2017] Collenette, J., Atkinson, K., Bloembergen, D., and Tuyls, K. (2017). Mood modelling within reinforcement learning. pages 106–113.

[Cuthbert, 2022] Cuthbert, B. N. (2022). Research domain criteria (RDoC): Progress and potential. *Current Directions in Psychological Science*, 31(2):107–114.

[Daugherty et al., 2009] Daugherty, D., Roque-Urrea, T., Urrea-Roque, J., Troyer, J., Wirkus, S., and Porter, M. A. (2009). Mathematical models of bipolar disorder. *Communications in Nonlinear Science and Numerical Simulation*, 14(7):2897–2908.

[Deligianni et al., 2019] Deligianni, F., Guo, Y., and Yang, G.-Z. (2019). From emotions to mood disorders: A survey on gait analysis methodology. *IEEE Journal of Biomedical and Health Informatics*, 23(6):2302–2316.

[Deserno et al., 2013] Deserno, L., Boehme, R., Heinz, A., and Schlagenhauf, F. (2013). Reinforcement learning and dopamine in schizophrenia: Dimensions of symptoms or specific features of a disease group? *Frontiers in Psychiatry*, 4.

[Dunjko et al., 2016] Dunjko, V., Taylor, J. M., and Briegel, H. J. (2016). Quantum-enhanced machine learning. *Physical Review Letters*, 117(13):130501.

[Eldar and Niv, 2015] Eldar, E. and Niv, Y. (2015). Interaction between emotional state and learning underlies mood instability. *Nature Communications*, 6(1).

[Eldar et al., 2016] Eldar, E., Rutledge, R. B., Dolan, R. J., and Niv, Y. (2016). Mood as representation of momentum. *Trends in Cognitive Sciences*, 20(1):15–24.

[Frank, 2013] Frank, T. (2013). A limit cycle oscillator model for cycling mood variations of bipolar disorder patients derived from cellular biochemical reaction equations. *Communications in Nonlinear Science and Numerical Simulation*, 18(8):2107–2119.

[Friston et al., 2014] Friston, K. J., Stephan, K. E., Montague, R., and Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *The Lancet Psychiatry*, 1(2):148–158.

[Goldberg and Ghaemi, 2005] Goldberg, J. F. and Ghaemi, S. N. (2005). Benefits and limitations of antidepressants and traditional mood stabilizers for treatment of bipolar depression. *Bipolar Disorders*, 7(s5):3–12.

[Goldbeter, 2011] Goldbeter, A. (2011). A model for the dynamics of bipolar disorders. *Progress in Biophysics and Molecular Biology*, 105(1-2):119–127.

[Goldbeter, 2013] Goldbeter, A. (2013). Origin of cyclicity in bipolar disorders: A computational approach. *Pharmacopsychiatry*, 46(Suppl. 01):S44–S52.

[Goodwin and Jamison, 2007] Goodwin, F. K. and Jamison, K. R. (2007). *Manic-depressive illness: bipolar disorders and recurrent depression*, volume 2. Oxford University Press.

[Grande et al., 2016] Grande, I., Berk, M., Birmaher, B., and Vieta, E. (2016). Bipolar disorder. 387(10027):1561–1572.

[Hadaeghi, 2013] Hadaeghi, F. (2013). Does "crisis-induced intermittency" explain bipolar disorder dynamics? *Frontiers in Computational Neuroscience*, 7.

[Huys et al., 2016] Huys, Q. J. M., Maia, T. V., and Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19(3):404–413.

[IsHak et al., 2012] IsHak, W. W., Brown, K., Aye, S. S., Kahloon, M., Mobaraki, S., and Hanna, R. (2012). Health-related quality of life in bipolar disorder. *Bipolar Disorders*, 14(1):6–18.

[Johnson, 2005] Johnson, S. L. (2005). Life events in bipolar disorder: Towards more specific models. *Clinical Psychology Review*, 25(8):1008–1027.

[Kacki, 2021] Kacki, M. (2021). Neurocomputational models of bipolar disorder and mood instability. Master's dissertation, The University of Edinburgh.

[Kato, 2007] Kato, T. (2007). Mitochondrial dysfunction as the molecular basis of bipolar disorder. 21(1):1–11.

[Katsimerou et al., 2014] Katsimerou, C., Redi, J. A., and Heynderickx, I. (2014). A computational model for mood recognition.

[Kawa et al., 2005] Kawa, I., Carter, J. D., Joyce, P. R., Doughty, C. J., Frampton, C. M., Wells, J. E., Walsh, A. E., and Olds, R. J. (2005). Gender differences in bipolar disorder: age of onset, course, comorbidity, and symptom presentation. 7(2):119–125.

[Lippard et al., 2020] Lippard, E. T. C., Weber, W., Welge, J., Adler, C. M., Fleck, D. E., Almeida, J., DelBello, M. P., and Strakowski, S. M. (2020). Variation in rostral anterior cingulate functional connectivity with amygdala and caudate during first manic episode distinguish bipolar young adults who do not remit following treatment. *Bipolar Disorders*, 23(5):500–508.

[Mansur et al., 2020] Mansur, R. B., Lee, Y., McIntyre, R. S., and Brietzke, E. (2020). What is bipolar disorder? a disease model of dysregulated energy expenditure. 113:529–545.

[Marinier et al., 2009] Marinier, R. P., Laird, J. E., and Lewis, R. L. (2009). A computational unification of cognitive behavior and emotion. *Cognitive Systems Research*, 10(1):48–69.

[Mason et al., 2017] Mason, L., Eldar, E., and Rutledge, R. B. (2017). Mood instability and reward dysregulation—a neurocomputational model of bipolar disorder. 74(12):1275.

[McDonald et al., 1999] McDonald, W. M., Tupler, L. A., Marsteller, F. A., Figiel, G. S., DiSouza, S., Nemeroff, C. B., and Krishnan, K. R. (1999). Hyperintense lesions on magnetic resonance images in bipolar disorder. 45(8):965–971.

[Meyer, 2002] Meyer, T. D. (2002). The hypomanic personality scale, the big five, and their relationship to depression and mania. *Personality and Individual Differences*, 32(4):649–660.

[Montague et al., 2012] Montague, P. R., Dolan, R. J., Friston, K. J., and Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, 16(1):72–80.

[Müller-Oerlinghausen et al., 2002] Müller-Oerlinghausen, B., Berghöfer, A., and Bauer, M. (2002). Bipolar disorder. 359(9302):241–247.

[Nana, 2009] Nana, L. (2009). Bifurcation analysis of parametrically excited bipolar disorder model. *Communications in Nonlinear Science and Numerical Simulation*, 14(2):351–360.

[National Health Service, 2019a] National Health Service (2019a). Symptoms - bipolar disorder. Available at https://www.nhs.uk/mental-health/conditions/bipolar-disorder/symptoms/, Accessed: 2021-10-19.

[National Health Service, 2019b] National Health Service (2019b). Symptoms - psychosis. Available at https://www.nhs.uk/mental-health/conditions/psychosis/symptoms/, Accessed: 2021-10-19.

[National Health Service, 2019c] National Health Service (2019c). Symptoms - schizophrenia. Available at https://www.nhs.uk/mental-health/conditions/schizophrenia/symptoms/, Accessed: 2021-10-19.

[Niv, 2009] Niv, Y. (2009). Reinforcement learning in the brain. 53(3):139–154.

[Noguès, 1997] Noguès, X. (1997). Protein kinase c, learning and memory: A circular determinism between physiology and behaviour. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 21(3):507–529.

[Novick et al., 2010] Novick, D. M., Swartz, H. A., and Frank, E. (2010). Suicide attempts in bipolar I and bipolar II disorder: a review and meta-analysis of the evidence. 12(1):1–9.

[O'Doherty et al., 2015] O'Doherty, J. P., Lee, S. W., and McNamee, D. (2015). The structure of reinforcement-learning mechanisms in the human brain. 1:94–100.

[Pacchiarotti et al., 2013] Pacchiarotti, I., Bond, D. J., Baldessarini, R. J., Nolen, W. A., Grunze, H., Licht, R. W., Post, R. M., Berk, M., Goodwin, G. M., Sachs, G. S., et al. (2013). The international society for bipolar disorders (ISBD) task force report on antidepressant use in bipolar disorders. 170(11):1249–1262.

[Phillips and Kupfer, 2013] Phillips, M. L. and Kupfer, D. J. (2013). Bipolar disorder diagnosis: challenges and future directions. 381(9878):1663–1671.

[Radulescu and Niv, 2019] Radulescu, A. and Niv, Y. (2019). State representation in mental illness. *Current Opinion in Neurobiology*, 55:160–166.

[Rescorla and Wagner, 1972] Rescorla, R. and Wagner, A. (1972). A theory of pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement. *Classical Conditioning: Current Research and Theory*.

[Russell and Norvig, 2010] Russell, S. J. and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall, 3rd edition.

[Saxena et al., 2017] Saxena, A., Scaini, G., Bavaresco, D. V., Leite, C., Valvassoria, S. S., Carvalho, A. F., and Quevedo, J. (2017). Role of protein kinase c in bipolar disorder: A review of the current literature. *Molecular Neuropsychiatry*, 3(2):108–124.

[Seriès, 2020] Seriès, P. (2020). *Computational psychiatry: A primer*. MIT Press.

[Sharp et al., 2020] Sharp, P. B., Miller, G. A., Dolan, R. J., and Eldar, E. (2020). Towards formal models of psychopathological traits that explain symptom trajectories. *BMC Medicine*, 18(1).

[Simonetti et al., 2020] Simonetti, A., Koukopoulos, A. E., Kotzalidis, G. D., Janiri, D., Chiara, L. D., Janiri, L., and Sani, G. (2020). Stabilization beyond mood: Stabilizing patients with bipolar disorder in the various phases of life. *Frontiers in Psychiatry*, 11.

[Steinacher and Wright, 2013] Steinacher, A. and Wright, K. A. (2013). Relating the bipolar spectrum to dysregulation of behavioural activation: A perspective from dynamical modelling. *PLoS ONE*, 8(5):e63345.

[Thompson, 2002] Thompson, C. (2002). Onset of action of antidepressants: results of different analyses. *Human Psychopharmacology: Clinical and Experimental*, 17(S1):S27–S32.

[Vieta et al., 2018] Vieta, E., Berk, M., Schulze, T. G., Carvalho, A. F., Suppes, T., Calabrese, J. R., Gao, K., Miskowiak, K. W., and Grande, I. (2018). Bipolar disorders. 4(1).

[Volterra, 1931] Volterra, V. (1931). Variations and fluctuations of the number of individuals in animal species living together. *Chapman, R.N., Ed., Animal Ecology*, pages 409–448.

[Whitton et al., 2015] Whitton, A. E., Treadway, M. T., and Pizzagalli, D. A. (2015). Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. 28(1):7–12.

[Wirkus and Porter, 2009] Wirkus, S. and Porter, M. A. (2009). Comment on "bifurcation analysis of parametrically excited bipolar disorder model". *Communications in Nonlinear Science and Numerical Simulation*, 14(6):2844.

[Yasui-Furukori, 2012] Yasui-Furukori, N. (2012). Update on the development of lurasidone as a treatment for patients with acute schizophrenia. *Drug Design, Development and Therapy*, page 107.