Title: Bipolar symptoms and lithium treatment affect neural signatures of adaptation of risk-taking to past outcomes during reward-guided decision-making

Jacqueline Scholl, Priyanka Panchal, Natalie Nelissen, Lauren Z Atkinson, Nils Kolling, Kate EA Saunders, John Geddes, Matthew FS Rushworth, Anna C Nobre, Paul J Harrison, Catherine J Harmer

Abstract

Introduction:

Mood instability in bipolar disorder (BD) is poorly understood. Here we examined cognitive and neural mechanisms related to these fluctuations and how they are changed with the mood stabilizer lithium.

Methods:

We recruited volunteers with low (n=37) or high (n=40) risk of BD (using the Mood Disorder Questionnaire, MDQ). We also recruited patients with BD who were assigned (randomized, double-blind) to six weeks of lithium (n=19) or placebo (n=16) after a two-week baseline period. Participants completed mood ratings daily over 50 (healthy) or 42 (BD) days, as well as a risky decision-making task and one functional magnetic resonance imaging session. The task measured adaptation of risk taking to past outcomes (increased risk aversion after a previous win, 'outcome history').

Results:

While the low MDQ group was risk averse after a win, this was less evident in the high MDQ group and least so in the patients with BD. Neurally, 'outcome history' was linked to medial frontal pole activation at the time of the decision. Corresponding to the behavioural effect, this activation was reduced in the high MDQ vs. the low MDQ group. While lithium did not reverse the pattern of BD in the task, it changed reward processing in the dorsolateral prefrontal cortex.

Discussion:

Healthy participants' modulation of risk-taking in response to reward outcomes was reduced by risk of BD and BD. These results provide a model for how reward may prime escalation of risk-related behaviours in bipolar disorder and how mood stabilising treatments may work.

Introduction

Bipolar disorder (BD) is typically characterized by episodes of depression or mania, lasting weeks and months. Lithium is the most effective mood stabiliser for management of BD, reducing the frequency of both manic and depressive episodes (1). While, fluctuating mood episodes have traditionally be seen as lasting weeks or months, more recent work has shown

that, in fact, patients with BD show large day-to-day fluctuations in mood even when symptoms are in the non-clinical range (2). A specific aspect of this mood instability – volatility - is affected by lithium treatment (3). Understanding the processes underpinning these fluctuations may help us develop and assess more effective treatment approaches.

To understand mood fluctuations in BD, recent work has used a computational-psychiatry approach drawing from the field of reinforcement learning. This work has suggested that destabilizing positive feedback cycles between mood and perceptions of rewards may contribute to BD (4–6): In people with subclinical symptoms of BD, positive or negative surprises were found to affect the neural and behavioural responses to reward and punishments. In particular, symptoms were associated with an increase in reward value after a positive surprise. This kind of reward sensitivity has been linked to later changes in mood, suggesting a route by which escalation of reward responses may translate into clinical symptoms (7). In patients, however, work so far has focused on lab-based measures of mood and its impact on behaviour on the scale of minutes.

Beyond learning, research on decision-making has also revealed that behaviour shows temporal interdependencies. For example, people show 'biases' such as 'loss chasing' (8) (taking more risks to recover losses). While these behaviours have often been interpreted as biases, from more recent conceptual work from an ecological perspective (9–12), they could be seen as useful for achieving homeostasis between needs for different rewards. Using momentary ecological monitoring has revealed homeostatic, mood stabilizing, behaviour. In healthy controls, when mood fluctuates, people self-report using strategies to re-establish mood homeostasis such as engaging with aversive activities when they are in a good mood (13). It has been shown that this strategy is reduced in people with depression or low mood (14). However, it is yet unclear whether homeostatic behaviour is also reduced in BD.

Studies in healthy participants and patients show that there is a wide network of frontal and striatal areas in the brain that process rewards and are crucial for motivation (15,16) and potentially important for understanding mood disorders (17,18) and homeostasis (9,10,12). However, how BD and mood instability change signals in these networks related to homeostasis is still unclear.

Here, we have built on these findings to test whether a gradient of mood elevation, was linked to behavioural measures of reduced active homeostatic behaviour in a decision-making task, measured as changes in decision strategies (risk taking) from trial to trial in response to reward/loss outcomes. For this, we recruited 40 healthy volunteers with risk of BD, i.e. a history of mood elevation, assessed using the mood disorder questionnaire (MDQ (19)), 37 volunteers without mood elevation, and 35 patients with BD. To assess whether behaviour and naturally occurring daily-life mood fluctuations were related, participants completed up to 50 longitudinal testing sessions at home. To understand the neural mechanisms of homeostatic stabilizing behaviour, we measured brain activity with fMRI. To test the causal effect of a commonly prescribed mood-stabilizing drug, lithium, 19 patients were randomly assigned for six weeks to lithium and 16 to placebo in a double-blind design.

METHODS

Participants

Participants were recruited in two separate studies (see below) approved by the local ethics committee: Oxford University Ethic Committee (MSD-IDREC-C2-2014-023) and Soutch Central – Oxford A Research Ethcis Committee (15/SC/0109). Participants gave informed consent and were reimbursed for taking part in the study. The intervention study was registered: https://doi.org/10.1186/ISRCTN91624955.

Volunteers with history of high or low mood elevation: Participants were recruited through local advertisement and from pools of previous participants who had consented to be re-contacted. Seventy-seven were included in the study. Participants were recruited into two age- and gender-matched groups depending on their scores from the Mood Disorders Questionnaire, MDQ (19): 'low MDQ' group: <5 points (n=37); 'high MDQ' group: ≥ 7 points and additionally indicating that several of these symptoms happened during the same period of time (n=40). Detailed exclusion criteria appear in supplementary method 1A. A higher score on the MDQ indicates a history of mood elevation and risk of BD.

Patients with BD: Participants were recruited through the BD Research Clinic (Oxford). All participants met criteria for BD-I (n=7), BD-II (n=27) or BD not otherwise specified (BD-NOS, n=1), based on structured clinical interview. Full exclusion criteria are provided in the supplementary materials [1B]. Participants were assigned to placebo (n=16) or lithium (n=19), see below.

Study design

Volunteers. We measured participants' mood and behaviour in a cognitive task longitudinally five times a week over ten weeks. Brain activity during the same task was measured during an MRI scan. The data here were part of a larger study (supplementary method 1B).

Patients with BD. This study was a randomised, 6-week, double-blind, placebo-controlled trial (20). See supplementary method [1B] for full information. All participants underwent a two-week pre-randomization phase during which they completed the cognitive task and mood ratings daily at home. Due to logistic challenges, for some participants this phase lasted longer than two weeks. For the next phase, which lasted 6 weeks, participants were randomly assigned to receive either lithium or placebo in a double-blind design. The first 10 participants were fully randomly assigned to avoid predictability, while for subsequent participants, an algorithm was used to minimize differences in age (<25 or >25 years) and gender between the two groups. In the lithium group, participants were titrated to doses producing plasma levels of 0.6-1 mmol/L (see supplementary methods 1C for dosing details). Data from one MRI session are ported here (n= 23 participants were fMRI compatible).

Throughout, we performed two types of group comparisons. First, we compared across history of mood elevation severity/ risk of BD (i.e. group as ordered factor (21) in regressions, Low MDQ < High MDQ < patients with BD), subsequently referred to as 'mood elevation gradient'. Second, we tested for the effects of lithium treatment as drug (lithium/placebo) x time point (baseline/post) interactions.

'Wheel of fortune' task

Trial structure. On each trial of the task, participants were given two options shown side-by-side. In the at-home version, these were wheels of fortune (Figure 1A). In the fMRI version, they were instead presented as bars. Each option had three attributes: probability of winning vs. losing (size of green vs. red area), magnitude of possible gain (number on green area, 10 to 200), and magnitude of possible loss (number on red area, also 10 to 200). After participants chose one option, the wheel of fortune started spinning and then randomly landed on either win or loss. Finally, participants were shown their updated total score. The experiment was designed so that most choices were difficult, i.e., the options were very similar in value (90% of choices were not more than 20 points apart; 76% not more than 5 points apart), Figure 1B.

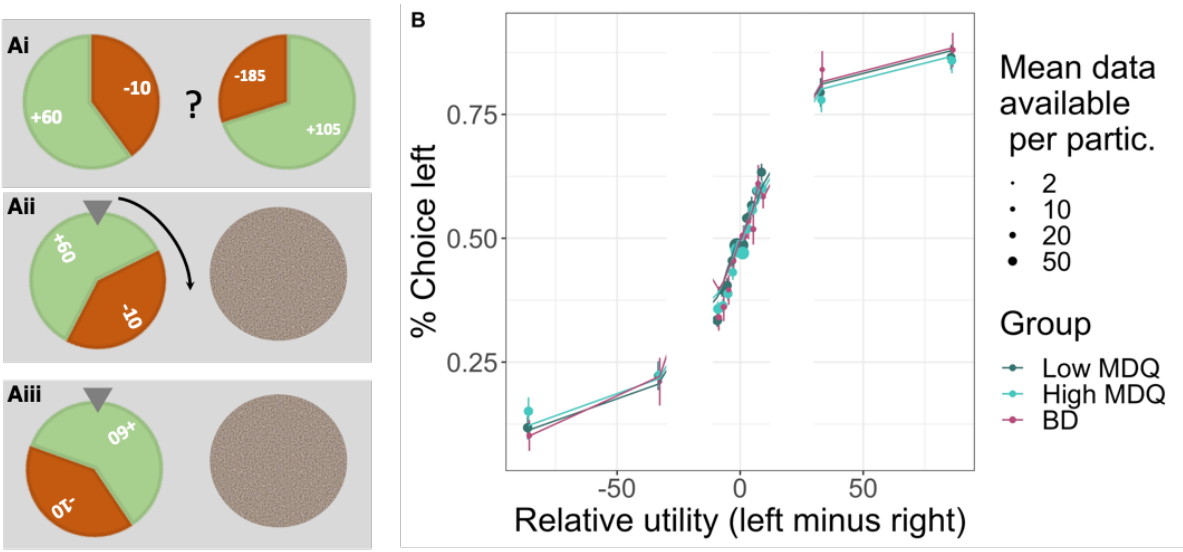

Figure 1. Task design and longitudinal behaviour. *A) On each trial, participants chose between two gambles ('wheel of fortune') that differed in their probability of winning or losing points and in the number of points that could be won or lost. Once participants had chosen an option, the alternative was hidden, and the chosen wheel started spinning until finally landing on the win or loss. B) Participants' choices (left vs. right option) were guided by the relative utilities (reward utility - i.e., probability * magnitude – minus loss utility): the higher the utility of the left option, the more it was chosen. The computational model (lines) captured behaviour (dots with error bars) well. Data were combined across all testing sessions (up to 50) per participant (20 trials per session). Error bars show the standard error of the mean, and the size of the dots indicates the number of data points available.*

Timings and number of trials. Each day, participants were asked to rate their positive and negative mood using the Positive and Negative Affect Schedule – Short Form, PANAS-SF (22). They were also asked to give an overall rating of their mood ('How are you feeling', referred to here as 'Happiness VAS') using a slider ranging from 'very unhappy' (red sad face drawing) to 'very happy' (green smiley face). They then played 20 trials of the task. After the task, they repeated the happiness VAS.

In the fMRI scanner, participants played 100 trials. All timings were jittered. From the onset of options until participants could make a choice: 1-2 sec; delay between participants' response and outcomes shown: 2.7 to 7.7 sec; duration of outcome shown: 1-3 sec; duration of total score shown: 1-9 sec; ITI: 1-9 sec.

Behaviour

Behavioural data were analysed in R (23) (version 4.0.2) and Matlab. R-packages: Stan (24) , BRMS (25,26), dplyr (27), ggpubr (28), sjPlot (29), compareGroups (30), emmeans (31), ggsci (32).

Model-free analyses of behaviour: We analysed participants' choices without and with computational models. First, to test that participants could perform the task, i.e., that their choices were sensitive to value, we binned their choices (% left vs. right option) according to the overall utility difference between the two options (i.e., left vs. right reward utility minus loss utility).

To test sensitivity to risk of losses, as has been previously reported to be affected in BD (33,34), we refined the binning of choices (as above) by further splitting the data according to win and loss utility (i.e. probability * magnitude).

We next analysed behaviour for adaptions of risk taking to past outcomes. One way this can be measured is by considering how participants change their behaviour – here risk-taking – based on win/loss outcomes on previous trials ('outcome history' effect), as we and others have previously done in a learning context (16,35). One might think that rational behaviour could mean that the outcome of one trial should not affect the choice in the next trial. While this might be rational in a simple computer task, it is plausible that participants approach the task with mechanisms adapted to decision making in the real world in which there are sequential dependencies between decisions or extended temporal contexts that might affect which outcomes will follow choices. For example, in an ecological framework where participants try to achieve homeostasis (9,12), they might be willing to accept higher risks of losing to make up for losses just encountered, or vice versa, be more avoidant of future losses after having received a reward outcome. To measure this in our experiment, we considered in the binning of choices (as above) whether choices were influenced by the previous trial's win/loss outcome.

To compare groups, instead of a standard ANOVA procedure which tests for any differences between groups, we tested for a systematic effect, i.e. a gradient of mood elevation (group as ordered factor (21), Low MDQ < High MDQ < BD patients) in linear regressions, also controlling for age and gender. Models used the BRMS toolbox interface for Stan (supplementary methods 2). For this and all subsequent analyses, we used Bayesian Credible Intervals (36) to establish significance by the 95% CI not including zero.

Computational models

Decision making. We used a computational model to capture participants' choices. The model first computed the overall value ('utility') of each option, then made a choice (left or right option) depending on which option had the better utility, but also allowing for some random choice behaviour (17,37).

First, the model compared the options' utilities as displayed at the time of choice on the current trial, i.e., probability (prob) x magnitude (mag). We allowed for individual differences in sensitivity to the loss vs. reward utility (λ). Participants' 'outcome history' was captured by a parameter (γ) that changed the weighting of the loss utility on the current trial depending on whether the previous trial's outcome was a win or a loss (i.e., γ>0 means increased sensitivity to losses after a win on the previous trial).

$$Utility_{left} = Prob * Mag_{rew} - \left(\lambda + \gamma * PrevOutc_{\frac{win}{loss}}\right) * (1 - Prob) * Mag_{loss}$$

To decide which option to choose, the model compared the utilities of the left and right options taking into account each participant's 'randomness' (inverse temperature (β), with higher values indicating choice consistency and lower choice randomness):

$$p\left(Choice_{left}\right) = \frac{1}{1 + e^{\beta * (Utility_{right} - Utility_{left})}}$$

Models were validated using simulations (Table S1 and supplementary methods 2A).

Model Fitting. To allow fitting of individual sessions (20 trials), a Bayesian approach was implemented that allowed specifying priors for each parameter (supplementary methods 2B).

To assess group differences, we then entered these session-wise parameters into hierarchical regressions (using BRMS). This enabled consideration of parameters that might change over the days of testing, as well as individual differences in the means and variability (standard deviation) across sessions of parameters. For example:

Mean: invTemp(β) ~ 1 + day + group+ Age + Gender + (1 + day | ID),

And error term: sigma ~ 1 + group + Age + Gender + (1|ID))

The effect of lithium (vs. placebo) was tested analogously:
Mean: invTemp(β) ~ 1 + day + group*pre/post+ Age + Gender + number_days_baseline (1 + day | ID)

These models were used for group comparisons of mean parameters (supplementary methods 2B+C). Variabilities were not compared as model validation (table S1) suggested poor recovery. Mood data (positive and negative PANAS, happiness VAS) were analysed using

similar regressions (supplementary methods 2C) to assess group differences in mood (mean or variability) or the relationship between task outcomes and changes in happiness VAS.

MRI acquisition

Data from all 77 healthy volunteers and 13 patients with BD were collected on a 3T Siemens Magneton Trio. Data from 9 patients with BD were collected at a different site using a Siemens Magneton PRISMA. Group comparisons include scanner as a control regressor. Scan protocols were carried out following (16), supplementary methods 3A.

FMRI analysis – whole-brain

General approach. Data were pre-processed using FSL ((38), supplementary methods 3B). Data were pre-whitened before analysis (39). Statistical analysis was performed at two levels. At the first level, we used an event-related GLM approach for each participant. On the second level, data were analysed using a mixed-effects model using FSL's FLAME 1 (40,41) with outlier de-weighting. The main-effect images are all cluster-corrected ($p<0.05$ two-tailed) with the inclusion threshold of $z < 2.3$.

Regression designs. Firstly, at the time of the decision, we looked for neural activity correlating with the utility (predicted reward, loss) of the choice. Second, at the time of the outcome of the gamble, we looked for neural activity related to the processing of the outcome (win/loss as continuous regressor). Decision and outcome-related activity could be dissociated due to jitter used in the experimental timing (see (16)). Third, as a key measure of interest, we looked at whether there was a history effect at the time of the choice (i.e., previous trial's gamble win/loss outcome (16,42), analogous to the behavioural findings). Full design information is included in the supplementary methods [3C] and Figure S2.

Group-level comparisons. We compared the low vs high MDQ groups (n=77) in a second-level analysis (statistical thresholds as described above). For the patients with BD, only 23 participants were available. Therefore, group comparisons were first performed in the regions of interest (ROIs) derived from comparisons of the healthy volunteers (see below). As exploratory analyses, groups were also compared at the whole-brain level. Scanner assignment was included as a control variable.

ROI analyses. Mean brain activations (z-stats) were extracted for each participant. These were used to illustrate group differences and also to perform independent statistical tests (e.g., ROIs of clusters defined based on group differences of high vs. low MDQ could be used to test group differences between lithium and placebo). For this, non-hierarchical Bayesian regressions were used, also controlling for age and gender. Brain activations were also correlated with behavioural measures. For this, effects of age, gender and group (and for the patients with BD: number of days in the baseline phase) were first removed using regressions from both neural and behavioural measures. As correlations of primary interest, activity from brain areas of significant differences between the groups was correlated with significant behavioural measures.

RESULTS

We recruited four groups of participants in two separate studies (Table 1).

	Low MDQ	High MDQ	BD lith	BD pla	Low vs high MDQ (p-value)	BD lith vs. pla (p-value)
	N=37	N=40	N=19	N=16		
Age	25.0 (6.61)	25.0 (7.06)	28.8 (9.81)	35.1 (13.8)	0.974	0.137
Gender:					0.998	1
F	24 (64.9%)	27 (67.5%)	11 (57.9%)	9 (56.2%)		
M	13 (35.1%)	13 (32.5%)	8 (42.1%)	7 (43.8%)		
MDQ	1.11 (1.31)	9.32 (1.67)	NA	NA	<0.001	NA
Handedness:					0.757	NA
Right	32 (86.5%)	32 (80.0%)	NA	NA		
Ambidext	0 (0.00%)	1 (2.50%)	NA	NA		
Left	5 (13.5%)	7 (17.5%)	NA	NA		
Diagnosis:					0.119	0.527
BDI	0 (0.00%)	0 (0.00%)	3 (15.8%)	4 (25.0%)		
BDII	0 (0.00%)	3 (7.50%)	16 (84.2%)	11 (68.8%)		.
BD NOS	0 (0.00%)	2 (5.00%)	0 (0.00%)	1 (6.25%)		
None	37 (100%)	35 (87.5%)	0 (0.00%)	0 (0.00%)		
Additional diagnosis:					0.236	NA
Depression	0 (0.00%)	4 (10.0%)	NA	NA		
Depression & Past alcohol dependence	0 (0.00%)	1 (2.50%)	NA	NA		
Depression & Past panic disorder	0 (0.00%)	1 (2.50%)	NA	NA		
Depression & PTSD	0 (0.00%)	2 (5.00%)	NA	NA		
Past alcohol dependence	0 (0.00%)	1 (2.50%)	NA	NA		
None	37 (100%)	31 (77.5%)	NA	NA		
# Behav. days	46.7 (3.63)	44.9 (6.61)	NA	NA	0.146	NA
# Behav. days (pre)	NA	NA	11.8 (6.72)	12.4 (5.04)	NA	0.771
# Behav. days (post)	NA	NA	24.5 (7.85)	28.5 (9.32)	NA	0.182
Has longitudinal data: Yes	37 (100%)	38 (95.0%)	19 (100%)	16 (100%)	0.494	1
Has FMRI data: Yes	37 (100%)	40 (100%)	13 (68.4%)	10 (62.5%)	1	0.992

Table 1. *Participant demographics. Statistical tests are two-tailed p-values and refer to comparisons between the two groups of healthy participants with low or high mood elevation ('Low vs. high MDQ') and between the two groups of patients with BD randomized to lithium or placebo ('Lith vs. pla'). Values are the mean and standard error of the mean. Abbreviations: '# Behav. Days' – number of days of behavioural data available (20 trials per day), '# Behav. days (pre)' – number of days in the baseline phase for the patients with BD, '# PANAS days' – number of days with mood scores (PANAS, positive affect negative affect scale, short form) available. 'MDQ' – Mood disorder questionnaire. 'Has longitudinal data: Yes' – percentage of participants from whom longitudinal data (i.e., sessions at home) were available. Diagnoses: 'BD-I' – bipolar I disorder; 'BD-II' – bipolar II disorder; 'BDNOS' – bipolar disorder not otherwise specified; 'PTSD' – post traumatic stress disorder. For the patients with BD, comorbid disorders were not measured. Note that in the low and high MDQ groups, diagnoses were only based on SCID, not on a full clinical examination.*

General performance

Participants completed longitudinal daily behavioural test sessions at home, consisting of 20 trials of a gambling task and mood self-reports. In the task (Figure 1A), participants needed to choose repeatedly between two gambles (wheels of fortune), considering the probabilities of winning or losing points and the number of points that could be won or lost. Participants

in all groups performed the task well (Figure 1B), selecting options with higher values more frequently.

Sensitivity to potential losses

To test whether sensitivity to potential losses vs. wins when gambling was reduced with a mood elevation gradient (low MDQ < high MDQ < patients with BD), we examined how much participants' choices took the value differences between options into account. We found that indeed, the mood elevation gradient was related to reduced sensitivity to losses compared to wins (Figure 2 Ai, group*win/loss dimension* utility bin: mean=0.33, 95% CI = [0.06; 0.61]), driven by both an increased sensitivity to wins (group*utility bin: mean 0.24, 95% CI = [0.08; 3.99]) and a decreased sensitivity to losses (group* utility bin: mean = -0.15, 95% CI = [-0.30; -0.01]). To quantify the behavioural effect precisely while controlling for potential confounds, we used computational modelling. We built a stochastic decision-making model that described participants' choices as being based on the reward and loss utilities of the two options while allowing for individual differences in how people made decisions. The model captured participants' sensitivity to losses (vs. wins) as a parameter (λ). We found again that the higher the mood elevation, the lower the sensitivity to losses (Figure 2Aii, Table S2, mean =-0.27, 95%CI = [-0.49; -0.05], driven mainly by a decrease in the group of patients with BD compared to the low/high MDQ groups). Lithium vs. placebo did not affect this (Figure 2Aiii).

Outcome history effects

We next analysed how participants adapted their risk taking across trials based on win or loss outcomes in the previous trial ('outcome history effect'). Low MDQ participants had the strongest outcome history effects, i.e., reduced risk taking (avoidance of potential losses) after a previous win. This was reduced with the mood elevation gradient (Figure 2A mean = -0.4, 95% CI = [-0.74; -0.06).

In the computational model, outcome history effects were captured as a parameter (γ) that described to what extent participants were less sensitive to potential losses after a win on the previous trial. We found again that the mood elevation gradient reduced outcome history effects (Figure 2B, Table S2, mean=-0.05, 95% CI=[-0.11; -0.0001]).

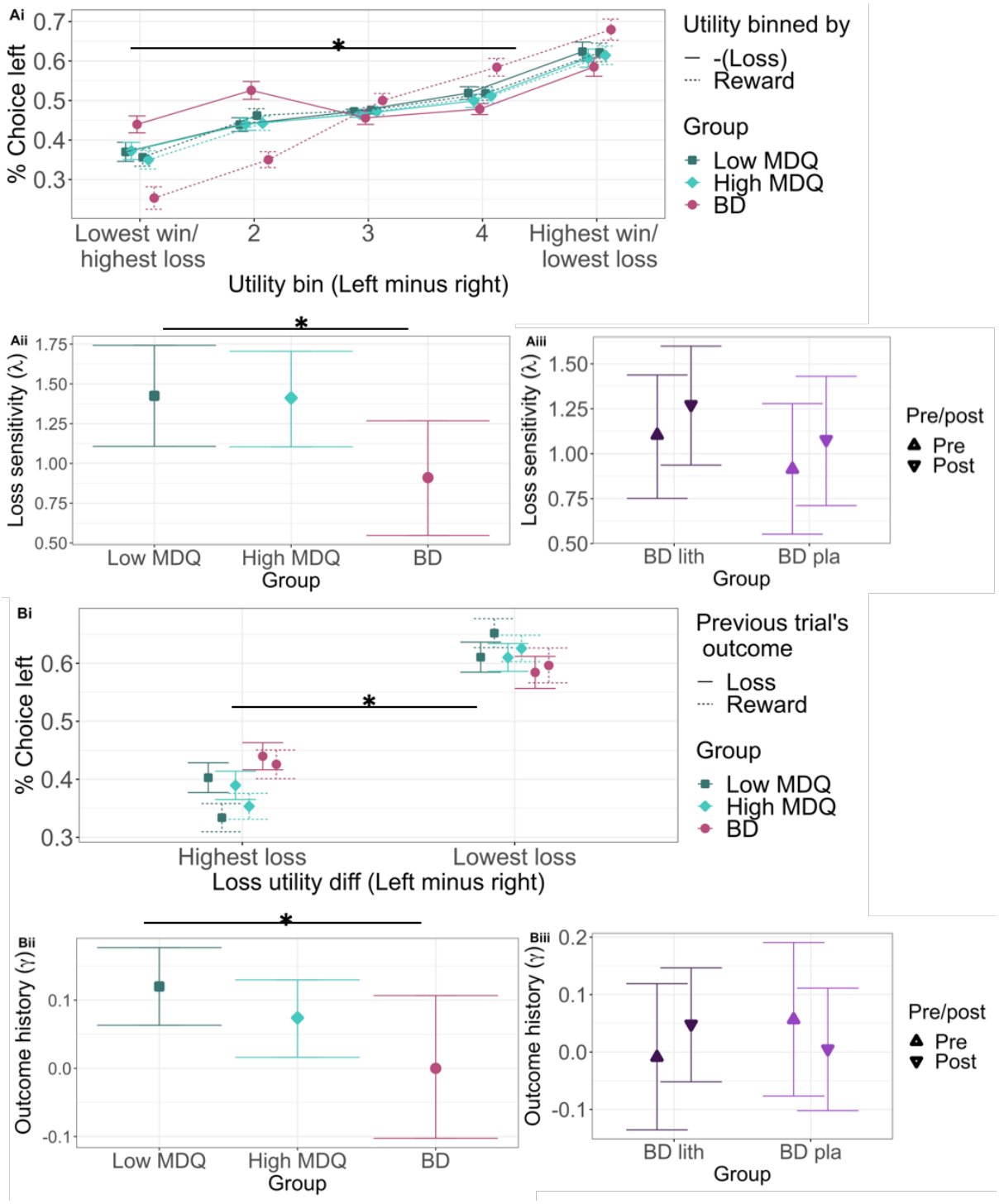

Figure 2. Group differences in longitudinal behaviour and mood. *A) Loss sensitivity. Ai) Illustration of sensitivity of choices to loss/reward utility – as utility increases for the left compared to the right option, participants are more likely to choose the left option. For low/ high MDQ participants, this increase in choice probability is similar for the reward or loss dimension. In contrast, patients with BD show decreased sensitivity to losses vs. rewards (the loss curve is shallower). Aii) The loss sensitivity model parameter reflects the group difference observed in the binned data (Ai). Aiii) Lithium (vs. placebo) does not affect loss sensitivity (group [lithium/placebo] * time [pre/post] interaction).B) Outcome history. Bi) Illustration of participants' increased risk-taking after a loss (compared to a win) on the*

previous trial ('outcome history'), indicated by the distance between the % choice values split by previous trial's outcome. This difference is smaller with mood elevation gradient. Bii) The outcome-history model parameter (γ) differed between the groups capturing the effects observed in the binned behaviour (Bi). Low MDQ participants showed the most and the patients with BD showed the least outcome history effects. Biii) Lithium (vs. placebo) does not affect outcome history effects. A full list of comparisons of parameters for the groups is shown in Tables S2 (longitudinal data) and Table S4 (fMRI session data). Relationships between parameters measured longitudinally over weeks or in the lab during the fMRI session are shown in Table S5. ii) and iii) show conditional effects from regression models, roughly equivalent to means, controlling for regressors of no interest.

Mood

Finally, an advantage of the behavioural data being collected at home was that we could relate daily moods to task-based behaviour. As reported previously (3,43) and similar to other studies (2,44,45) groups differed in their instability (standard deviation) of mood. The low MDQ group showed the lowest and the patients with BD the highest (positive PANAS: mean= 0.22, 95%CI = [0.11; 0.33]; negative PANAS: mean=0.64, 95%CI = [0.45; 0.83], Table S2A, Figure S1A). Lithium did not affect instability when using our measure of standard deviation here (table S2B), though note that using a measure of Bayesian volatility, lithium has been found to increase volatility of positive mood (3). Looking at the relationship between task outcomes and mood (happiness VAS), across all groups, happiness was increased by reward and decreased by loss (mean = 0.42, 95%CI = [0.31; 0.52]), but this did not differ mood elevation (mean =-0.06, 95% CI = [-0.15, 0.03]).

Neural results

Neural data were available for 77 volunteers and 23 patients. Across volunteers, brain activations to reward and loss utility during decisions (Figure 3A) and at the receipt of outcomes (Figure 3C) activated brain evaluation networks, including ventromedial prefrontal cortex (vmPFC), ventral striatum, dorsal anterior cingulate cortex (dACC), insula (Table S6). Next, we tested whether, related to the outcome history effect, there was brain activity when participants made a choice related to what had happened previously. Indeed, we found that activity in a network including the ventral striatum, vmPFC and medial frontal pole (FPm) related to the outcome of the previous trial (Figure 3B, Table S6).

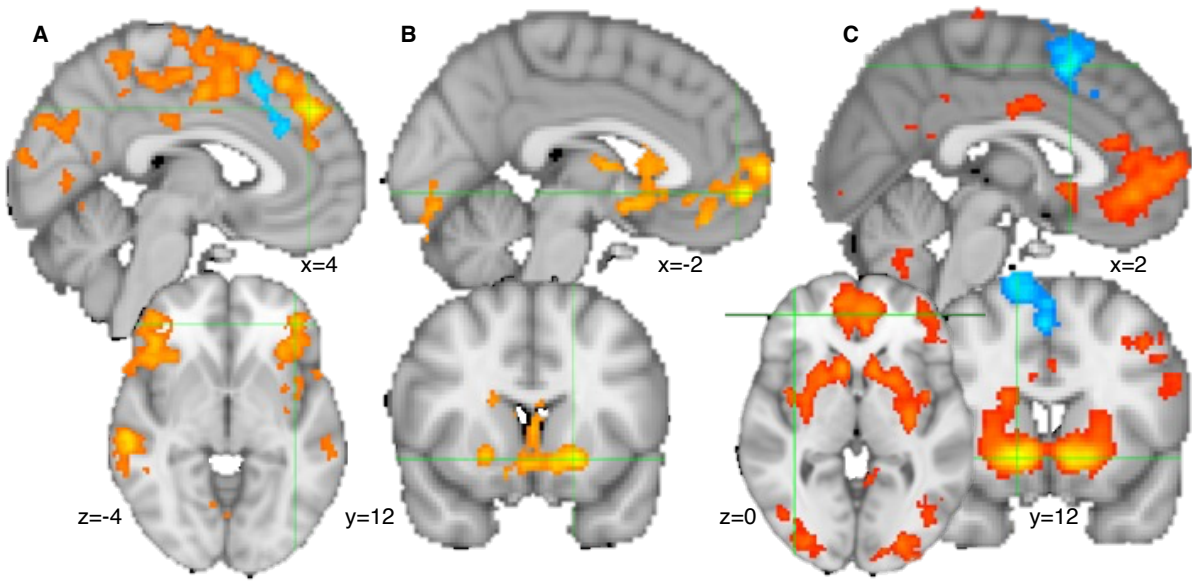

Figure 3. Neural activity during gambling. *A) At the time of the decision a wide network of areas activated with relative (chosen minus unchosen) reward utility (orange), while loss relative utility activated the anterior cingulate cortex (blue). B) At the time of the decision, the last trial's outcome (points won or lost) activated areas including vmPFC and ventral striatum (orange). C) At the time of the outcome (win or loss received), the outcome (points won or lost) activated areas including vmPFC, FPm, and ventral striatum (red/orange) and deactivated the pre-supplementary area. All results are cluster-corrected at p<0.05, two-tailed, with inclusion cut-off z>2.3. See Table S6 for the full list of results. Data were combined across both volunteer groups (low and high MDQ).*

Next, we compared the low and high MDQ groups. Activity for the last trial's outcome was higher for the low MDQ vs high MDQ group in FPm (Figure 4A-B, Table 2A, p=0.038, whole-brain cluster corrected). In other words, while all participants showed activity in vmPFC/FPm, in low MDQ participants the cluster extended further into FPm. Moreover, the stronger the activity for the last trial's outcome in this area, the stronger the behavioural outcome history effect (Figure 4C, r=0.24, p=0.017, partial correlation after correction for control variables and group; without correction: r=0.28, p=0.005). Lithium vs. placebo participants' activity did not differ in this area (mean=0.64, 95% CI = [-0.23; 1.44]).

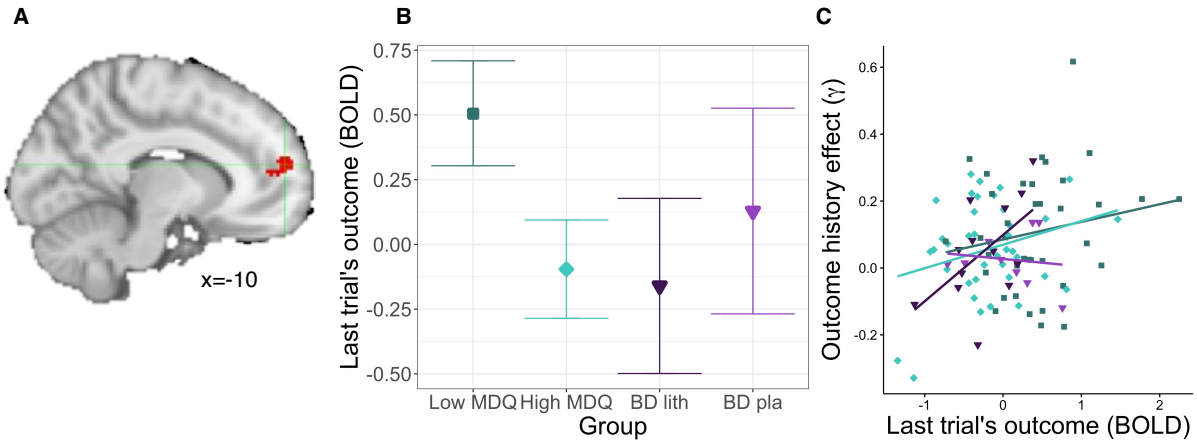

Figure 4. Group differences in signals related to outcome history effects. *A) Activation with last trial's outcome at the time of the current trial's decision differed between the low and high MDQ groups in the medial frontal pole (FPm; x=-10, y=56, z=16; p=0.038, n=77, cluster-corrected, Table 2A). In the low MDQ group, the activation with the last trial's outcome that is found across both groups (Figure 3B) extends further dorsally. B) This group difference was driven by the low MDQ group showing stronger activation than the high MDQ group in FPm (Figure shows conditional effects from regression model, roughly equivalent to means, controlling for regressors of no interest). There was no significant difference between activations comparing lithium and placebo groups. C) This FPm activity correlated with the longitudinally measured outcome history parameter. Colours match those of groups in B.*

A Low vs high MDQ groups	x	y	z	max z-score	p-value (2-tailed)	# voxels
Last trial's win/loss magnitude (signed) at choice						
Low > high MDQ						
Medial frontal pole (FPm), area 9m [2]	-10	56	16	3.52	0.0378	398
B Bipolar lithium vs. placebo groups - exploratory						
Win/loss magnitude (signed) at outcome						
Placebo > Lithium						
Dorsolateral prefrontal cortex (Area 46 [5]) and lateral frontal pole [2], Inferior frontal sulcus (IFS), (right)	38	48	0	3.49	0.00898	503

Table 2. Group comparisons. *Statistics for group comparisons in Figures 4-5. A) Comparisons of the low vs high mood elevation volunteers (Figure 4). B) Comparisons of the patients with BD assigned to placebo or lithium (Figure 5). All cluster-based thresholded, inclusion threshold: z=2.3, significance p<0.05 two-tailed. The maximum z-value of the cluster, the p-value and number of voxels are given for each cluster. Anatomical labels are based on: [1] (46)) [2] (47), [3] (48), [4] (49), [5] (50).*

As exploratory analyses (due to low sample sizes in the clinical FMRI groups), we next compared lithium vs. placebo treatment at the whole-brain level. We found that patients receiving placebo had stronger activity related to the outcome of gambles in an area spanning dorsolateral prefrontal cortex (dlPFC, area 46) and lateral frontal pole (Figure 5A-B, Table 2B, p=0.009).

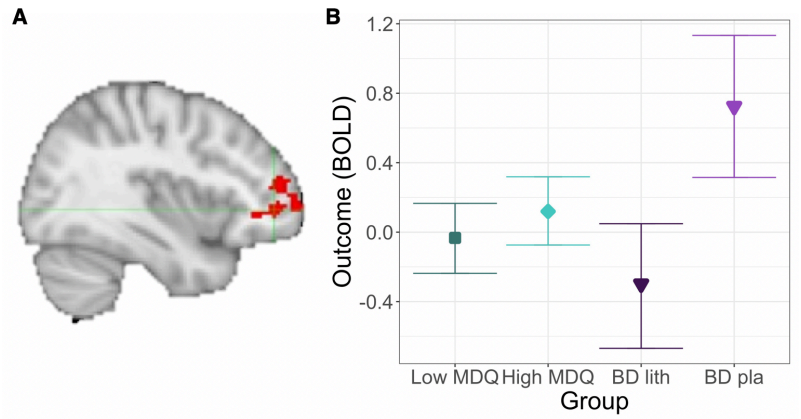

Figure 5. Exploratory whole-brain group differences in the patients with BD for gamble outcome signal. *A) Outcome related activity differed between the placebo and the lithium participants in an area including dorsolateral prefrontal cortex and lateral frontal pole (whole-brain cluster-corrected, Table 2B). This effect is illustrated in B).*

DISCUSSION

We designed a study to test the computational and neural correlates of adaptations of risk-taking to gains and losses in bipolar disorder (BD) and treatment with lithium. We included participants along a gradient of history of mood elevation ranging from volunteers with low risk of BD or mood instability (low MDQ group), to volunteers high risk of BD, to patients with BD. In the patients, we tested the effect of lithium treatment in a placebo-controlled double-blind design. We measured how much participants adapted their risk-taking following reward outcomes in a risky decision-making task ('outcome history effects'). We measured behaviour both longitudinally over up to 50 days and during a brain imaging (FMRI) session. We found that the low MDQ group showed 'outcome history effects'. Specifically, after a win on a trial, they were more risk averse. This was reduced across the mood elevation gradient (lowest in patients with BD). Neurally, outcome history was related to the representation of past information in a large swath including ventromedial prefrontal cortex (vmPFC) and medial frontal pole (FPm). In low MDQ volunteers, this brain signal extended further dorsally into FPm.

First, we replicated previous findings (33,34) that mood elevation decreased sensitivity to potential losses (vs wins).Then, we went further looking at adaptation of risk taking to past outcomes. We found that healthy volunteers without mood elevation showed sequential dependencies between their choices and previous trials' outcomes. This was not strictly rational in our task since outcomes for gambles across trials were independent. Therefore, these sequential effects could be understood as a bias, e.g., 'chasing' as has been described

in the gambling literature (8). However, an alternative view is that 'biases' observed in the lab are actually functionally appropriate in more naturalistic environments (17,51–53). For example, in natural environments, which are experienced continually rather than in discrete trials and in which different types of rewards (e.g. food, water) need to be accumulated or a homeostatic setpoint needs to be reached, it would make sense to adapt behaviour according to previous outcomes (9,10,12,54,55). Here we find that in the absence of mood elevation, volunteers show a kind of 'active homeostatic behaviour', i.e., they modulate their risk-taking depending on the previous trial's outcome. This tendency was lower in the high MDQ group and lowest in patients with BD. Reduced homeostatic behaviour could lead to unstable moods as in the healthy population mood has been found to be regulated through behaviour (13). Relatedly, in patients with BD, purposefully regulating behaviour during the prodromal periods has been shown to reduce the risk of relapse (56).

We focused on whole-brain analyses for the low/high MDQ volunteer sample due to the larger sample size. Decision-making and the processing of outcomes produced a typical pattern of activation (15,57–59) in areas including dorsal anterior cingulate cortex, striatum and vmPFC. However, there were no group differences in any of these signals, matching our behavioural results. We next looked for brain activity related to the modulation of risk taking with 'outcome history'. We found that at the time when people made decisions, there was activity representing the last trial's outcome in an area spanning vmPFC to FPm. This is similar to previous findings in a learning context of between-trial activities (16,35,42). This signal extended more dorsally into FPm in low MDQ volunteers. Furthermore, the stronger this signal, the stronger the modulation of risk taking by outcome history. In this region, lithium did not affect brain activity. In an exploratory analysis, we also compared the brain activity of patients with BD with lithium or placebo. Patients given placebo showed larger outcome-related activity in dorsolateral prefrontal cortex, while under lithium this activity was similar to the non-clinical groups. Previously, a dampening of reward responses with lithium has been reported in the ventral striatum in healthy volunteers (60).

It has been proposed that changes to reward processing, particularly reward hypersensitivity, are central to BD (61,62). While self-reports strongly support this (63), behavioural findings are more mixed: though some have found differences in reward-based decision-making or risk-taking (64,65), the interpretation has been hampered by concerns about the tasks (64), non-computational analysis strategies focusing on 'correct' answers (65), or similar effects being seen in other disorders. Neurally, while there have been suggestions that BD is accompanied by increased reward-related brain signals (66), opposite findings have also been observed (67). One possible explanation could be that these neural effects are sensitive to patients' mood states or medication (67). Recently, computational work has suggested how differences in the interplay between reward processing and mood could lead to feedback loops and fluctuations in mood, including spiralling out into manic and depressive episodes (4,6,7). Our findings here complement these results by revealing the process underlying changes in active homeostatic behaviour. In the future, it would be interesting to see how these different processes interact or whether there are distinct clusters of patients as has been proposed for other facets of decision-making (68).

The current study also revealed an effect of lithium treatment on the neural response during rewarded outcome, making these measures more similar to healthy controls. Analysis of

subjective ratings from the same patient cohort revealed that lithium increased the volatility of positive affective experiences compared to placebo (3). This effect was proposed to reduce the persistence of positive affect in patients with bipolar disorder. A previous study also revealed effects of lithium on reward-related prediction errors in a healthy volunteer model (60). While the efficacy of lithium in the treatment of both manic and depressive episodes is well established, its mechanisms of action remain a subject of debate. Together our results highlight the importance of considering rewarded decision-making and learning perspectives to understand mood instability and the effects of lithium.

Contributions:
JS: Conceptualization, software, formal analysis, writing, original draft, reviewing and editing
PP: Investigation, writing, reviewing and editing
NN: Conceptualization, investigation, reviewing
NK: Conceptualization, reviewing and editing
LZA: Investigation, writing, reviewing and editing.
MFSR: Conceptualization, design, reviewing and editing.
CAN: Funding acquisition, conceptualization, supervision, design, reviewing and editing.
PJH: Funding acquisition, Conceptualisation, recruitment, reviewing and editing
JG: Funding acquisition, conceptualisation, design, supervision, reviewing and editing
KS: Investigation, project administration, reviewing and editing
CJH: Funding acquisition, conceptualization, design, supervision, reviewing and editing

Financial disclosure
The study was funded by a Wellcome Trust Strategic Award (CONBRIO: Collaborative Oxford Network for Bipolar Research to Improve Outcomes, reference No. 102,616/Z). JRG, CJH, PJH and KEAS are supported by the Oxford Health NIHR Biomedical Research Centre. MFSR is funded by the Wellcome Trust (221794/Z/20/Z). The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z). JS has been funded by the Institut National de la Santé et de la Recherche Médicale, the Biotechnology and Biological Sciences Research Council (BB/V004999/1, Discovery Fellowship) and Medical Research Council (MR/N014448/1, Skills Development Fellowship). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

JS, PP, NN, LZA, NK, JG, MFSR report no biomedical financial interests or potential conflicts of interest. CJH has received consultancy payments from P1vital, Lundbeck, Compass Pathways, IESO, Zogenix (now UCB). PJH reports receiving an honorarium for editorial work for Biological Psychiatry and Biological Psychiatry Global Open Science. ACN is non-executive director at the Oxford Health Foundation Trust. KEAS has received consultancy payment from Yale University.

References

1. Kessing LV, Bauer M, Nolen WA, Severus E, Goodwin GM, Geddes J (2018): Effectiveness of maintenance therapy of lithium vs other mood stabilizers in monotherapy and in combinations: a systematic review of evidence from observational studies. *Bipolar Disord* 20: 419–431.

2. Bonsall MB, Wallace-Hadrill SMA, Geddes JR, Goodwin GM, Holmes EA (2012): Nonlinear time-series approaches in characterizing mood stability and mood instability in bipolar disorder. *Proc R Soc B Biol Sci* 279: 916–924.

3. Pulcu E, Saunders KEA, Harmer CJ, Harrison PJ, Goodwin GM, Geddes JR, Browning M (2022): Using a generative model of affect to characterize affective variability and its response to treatment in bipolar disorder. *Proc Natl Acad Sci* 119: e2202983119.

4. Eldar E, Niv Y (2015): Interaction between emotional state and learning underlies mood instability. *Nat Commun* 6: 6149.

5. Eldar E, Rutledge RB, Dolan RJ, Niv Y (2016): Mood as representation of momentum. *Trends Cogn Sci* 20: 15–24.

6. Mason L, Eldar E, Rutledge RB (2017): Mood instability and reward dysregulation—a neurocomputational model of bipolar disorder. *JAMA Psychiatry* 74: 1275–1276.

7. Eldar E, Roth C, Dayan P, Dolan RJ (2018): Decodability of reward learning signals predicts mood fluctuations. *Curr Biol* 28: 1433–1439.

8. Zhang K, Clark L (2020): Loss-chasing in gambling behaviour: neurocognitive and behavioural economic perspectives. *Curr Opin Behav Sci* 31: 1–7.

9. Kolling N, Wittmann M, Rushworth MFS (2014): Multiple Neural Mechanisms of Decision Making and Their Competition under Changing Risk Pressure. *Neuron* 81: 1190–1202.

10. Korn CW, Bach DR (2015): Maintaining homeostasis by decision-making. *PLoS Comput Biol* 11: e1004301.

11. Hammack T, Cooper J, Flach JM, Houpt J (2017): Toward an Ecological Theory of Rationality: Debunking the Hot Hand "Illusion." *Ecol Psychol* 29: 35–53.

12. Juechems K, Balaguer J, Castañón SH, Ruz M, O'Reilly JX, Summerfield C (2019): A network for computing value equilibrium in the human medial prefrontal cortex. *Neuron* 101: 977–987.

13. Taquet M, Quoidbach J, Montjoye Y-A de, Desseilles M, Gross JJ (2016): Hedonism and the choice of everyday activities. *Proc Natl Acad Sci* 113: 9769–9773.

14. Taquet M, Quoidbach J, Gross JJ, Saunders KEA, Goodwin GM (2020): Mood Homeostasis, Low Mood, and History of Depression in 2 Large Population Samples. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2020.0588

15. Sescousse G, Caldú X, Segura B, Dreher J-C (2013): Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neurosci Biobehav Rev* 37: 681–696.

16. Scholl J, Kolling N, Nelissen N, Wittmann MK, Harmer CJ, Rushworth MF (2015): The good, the bad, and the irrelevant: neural mechanisms of learning real and hypothetical rewards and effort. *J Neurosci* 35: 11233–11251.

17. Scholl J, Klein-Flügge M (2018): Understanding psychiatric disorder by capturing ecologically relevant features of learning and decision-making. *Behav Brain Res* 355: 56–75.

18. Ng TH, Alloy LB, Smith DV (2019): Meta-analysis of reward processing in major depressive disorder reveals distinct abnormalities within the reward circuit [no. 1]. *Transl Psychiatry* 9: 1–10.

19. Hirschfeld RM, Williams JB, Spitzer RL, Calabrese JR, Flynn L, Keck Jr PE, *et al.* (2000): Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. *Am J Psychiatry* 157: 1873–1875.

20. Saunders KEA, Cipriani A, Rendell J, Attenburrow M-J, Nelissen N, Bilderbeck AC, *et al.* (2016): Oxford Lithium Trial (OxLith) of the early affective, cognitive, neural and biochemical effects of lithium carbonate in bipolar disorder: study protocol for a randomised controlled trial. *Trials* 17: 116.

21. Bürkner P-C, Charpentier E (2018): Monotonic effects: A principled approach for including ordinal predictors in regression models.

22. Watson D, Clark LA, Tellegen A (1988): Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 54: 1063.

23. R Core Team (2020): *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

24. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, *et al.* (2017): Stan : A Probabilistic Programming Language. *J Stat Softw* 76. https://doi.org/10.18637/jss.v076.i01

25. Bürkner P-C (2017): Advanced Bayesian multilevel modeling with the R package brms. *ArXiv Prepr ArXiv170511123*.

26. Bürkner P-C (2017): brms: An R package for Bayesian multilevel models using Stan. *J Stat Softw* 80: 1–28.

27. Wickham H, François R, Henry L, Müller K (2021): *Dplyr: A Grammar of Data Manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

28. Kassambara A (n.d.): ggpubr:‘ggplot2’Based Publication Ready Plots (2018). *R Package Version 02*.

29. Lüdecke D, Bartel A, Schwemmer C (2019): Package ‘sjPlot.’

30. Subirana I, Sanz H, Vila J (2014): Building Bivariate Tables: The compareGroups Package for R. *J Stat Softw* 57: 1–16.

31. Lenth RV (2021): *Emmeans: Estimated Marginal Means, Aka Least-Squares Means*. Retrieved from https://CRAN.R-project.org/package=emmeans

32. Xiao N (2018): *Ggsci: Scientific Journal and Sci-Fi Themed Color Palettes for “Ggplot2.”* Retrieved from https://CRAN.R-project.org/package=ggsci

33. Reinharth J, Braga R, Serper M (2017): Characterization of Risk-Taking in Adults With Bipolar Spectrum Disorders. *J Nerv Ment Dis* 205: 580.

34. Lasagna CA, Pleskac TJ, Burton CZ, McInnis MG, Taylor SF, Tso IF (2022): Mathematical Modeling of Risk-Taking in Bipolar Disorder: Evidence of Reduced Behavioral Consistency, With Altered Loss Aversion Specific to Those With History of Substance Use Disorder [no. 1]. *Comput Psychiatry* 6: 96–116.

35. Fischer AG, Bourgeois-Gironde S, Ullsperger M (2017): Short-term reward experience biases inference despite dissociable neural correlates. *Nat Commun* 8: 1690.

36. Morey RD, Hoekstra R, Rouder JN, Wagenmakers E-J (2016): Continued misinterpretation of confidence intervals: response to Miller and Ulrich. *Psychon Bull Rev* 23: 131–140.

37. Nassar MR, Frank MJ (2016): Taming the beast: extracting generalizable knowledge from computational models of cognition. *Curr Opin Behav Sci* 11: 49–54.

38. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM (2012): FSL. *NeuroImage* 62: 782–790.

39. Woolrich MW, Ripley BD, Brady M, Smith SM (2001): Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage* 14: 1370–1386.

40. Beckmann CF, Jenkinson M, Smith SM (2003): General multilevel linear modeling for group analysis in FMRI. *Neuroimage* 20: 1052–1063.

41. Woolrich MW, Behrens TE, Beckmann CF, Jenkinson M, Smith SM (2004): Multilevel linear modelling for FMRI group analysis using Bayesian inference. *Neuroimage* 21: 1732–1747.

42. Chau BKH, Sallet J, Papageorgiou GK, Noonan MP, Bell AH, Walton ME, Rushworth MFS (2015): Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit Assignment and Learning in Macaques. *Neuron* 87: 1106–1118.

43. Panchal, P, Nelissen, N, McGowen, N, Atkinson, LZ, Saunders, KEA, Harrison, PJ, *et al.* (in submission): Identifying mood instability and circadian rest-activity patterns using digital remote monitoring and actigraphy in participants at risk for bipolar disorder.

44. Holmes EA, Bonsall MB, Hales SA, Mitchell H, Renner F, Blackwell SE, *et al.* (2016): Applications of time-series analysis to mood fluctuations in bipolar disorder to promote treatment innovation: a case series [no. 1]. *Transl Psychiatry* 6: e720–e720.

45. Tsanas A, Saunders KEA, Bilderbeck AC, Palmius N, Osipov M, Clifford GD, *et al.* (2016): Daily longitudinal self-monitoring of mood variability in bipolar disorder and borderline personality disorder. *J Affect Disord* 205: 225–233.

46. Mars RB, Jbabdi S, Sallet J, O'Reilly JX, Croxson PL, Olivier E, *et al.* (2011): Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *J Neurosci* 31: 4087–4100.

47. Neubert F-X, Mars RB, Sallet J, Rushworth MFS (2015): Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proc Natl Acad Sci* 112: E2695–E2704.

48. Sallet J, Mars RB, Noonan MP, Neubert F-X, Jbabdi S, O'Reilly JX, *et al.* (2013): The Organization of Dorsal Frontal Cortex in Humans and Macaques. *J Neurosci* 33: 12255–12274.

49. Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N (2009): A probabilistic MR atlas of the human cerebellum. *NeuroImage* 46: 39–46.

50. Neubert F-X, Mars RB, Thomas AG, Sallet J, Rushworth MF (2014): Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron* 81: 700–713.

51. Hayden BY (2018): Economic choice: the foraging perspective. *Curr Opin Behav Sci* 24: 1–6.

52. Lieder F, Griffiths TL (2020): Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behav Brain Sci* 43.

53. Oaksford M, Chater N (2020): New Paradigms in the Psychology of Reasoning. *Annu Rev Psychol* 71: 305–330.

54. O'Reilly RC, Hazy TE, Mollick J, Mackie P, Herd S (2014): Goal-driven cognition in the brain: a computational framework. *ArXiv Prepr ArXiv14047591*.

55. Juechems K, Summerfield C (2019): Where Does Value Come From? *Trends Cogn Sci* 23: 836–850.

56. Lam D, Wong G, Sham P (2001): Prodromes, coping strategies and course of illness in bipolar affective disorder–a naturalistic study. *Psychol Med* 31: 1397–1402.

57. Liu X, Hairston J, Schrier M, Fan J (2011): Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev* 35: 1219–1236.

58. Rushworth MF, Kolling N, Sallet J, Mars RB (2012): Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr Opin Neurobiol* 22: 946–955.

59. Jauhar S, Fortea L, Solanes A, Albajes-Eizagirre A, McKenna PJ, Radua J (2021): Brain activations associated with anticipation and delivery of monetary reward: A systematic review and meta-analysis of fMRI studies. *PLOS ONE* 16: e0255292.

60. Volman I, Pringle A, Verhagen L, Browning M, Cowen PJ, Harmer CJ (2021): Lithium modulates striatal reward anticipation and prediction error coding in healthy volunteers. *Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol* 46: 386–393.

61. Alloy LB, Nusslock R, Boland EM (2015): The development and course of bipolar spectrum disorders: An integrated reward and circadian rhythm dysregulation model. *Annu Rev Clin Psychol* 11: 213–250.

62. Whitton AE, Treadway MT, Pizzagalli DA (2015): Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Curr Opin Psychiatry* 28: 7–12.

63. Katz BA, Naftalovich H, Matanky K, Yovel I (2021): The dual-system theory of bipolar spectrum disorders: A meta-analysis. *Clin Psychol Rev* 83: 101945.

64. Edge MD, Johnson SL, Ng T, Carver CS (2013): Iowa gambling task performance in euthymic bipolar I disorder: A meta-analysis and empirical study. *J Affect Disord* 150: 115–122.

65. Ramírez-Martín A, Ramos-Martín J, Mayoral-Cleries F, Moreno-Küstner B, Guzman-Parra J (2020): Impulsivity, decision-making and risk-taking behaviour in bipolar disorder: a systematic review and meta-analysis. *Psychol Med* 50: 2141–2153.

66. Nusslock R, Young CB, Damme KSF (2014): Elevated reward-related neural activation as a unique biological marker of bipolar disorder: Assessment and treatment implications. *Behav Res Ther* 62: 74–87.

67. Bart CP, Titone MK, Ng TH, Nusslock R, Alloy LB (2021): Neural reward circuit dysfunction as a risk factor for bipolar spectrum disorders and substance use disorders: A review and integration. *Clin Psychol Rev* 87: 102035.

68. Jiménez E, Solé B, Arias B, Mitjans M, Varo C, Reinares M, *et al.* (2018): Characterizing decision-making and reward processing in bipolar disorder: A cluster analysis. *Eur Neuropsychopharmacol* 28: 863–874.