nature human behaviour

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41562-019-0628-0

In the format provided by the authors and unedited.

Altered learning under uncertainty in unmedicated mood and anxiety disorders

Jessica Aylward[1], Vincent Valton[1], Woo-Young Ahn[2], Rebecca L. Bond[1], Peter Dayan[3], Jonathan P. Roiser[1] and Oliver J. Robinson[1,4]*

[1]Neuroscience and Mental Health group, Institute of Cognitive Neuroscience, University College London, London, UK. [2]Department of Psychology, Seoul National University, Seoul, Korea. [3]Gatsby Computational Neuroscience Unit, University College London, London, UK. [4]Research Department of Clinical, Educational and Health Psychology, University College London, London, UK. *e-mail: oliver.j.robinson@gmail.com

Supplement

Supplementary Methods 1

Prospect valence learning (PVL) models

These models are both reported in Ahn et al[1]. We reproduce the description here and as such there may be consistencies across the text. Both models are identical with the exception of the learning rules. The utility of each trial (t) of each net outcome x(t) is calculated according to:

$$(1) \qquad utility(t) = \begin{matrix} x(t)^{\alpha} & \text{if } x(t) \geq 0 \\ -\lambda|x(t)|^{\alpha} & \text{if} x(t) < 0 \end{matrix}$$

Where α governs the shape of the utility function while λ determines the sensitivity to losses compared to gains.

Based on the outcome of the chosen option, the expectancies of the decks were computed using a learning rule. In the *igt_pvl_decay* model, the expectancies of all decks are discounted on each trial and then the expectancy of the chosen deck is updated by the current outcome utility:

$$(2) \qquad Ej(t+1) = A \cdot Ej(t) + \delta j(t) \cdot utility(t)$$

Where A determines how much the past expectancy is discounted and δj(t) is a dummy variable which is 1 if deck j is chosen and 0 otherwise.

In the *igt_pvl_delta* model, the expectancy of only the selected deck is updated and the expectancies of the other decks remain unchanged:

$$(3) \qquad Ej(t+1) = Ej(t) + W \cdot \delta j(t) \cdot (utility(t) - Ej(t))$$

W determines how much weight is placed on past experiences of the chosen deck vs. the most recent selection from the deck. The probability of choosing each deck j is then determined by the softmax (where θ sensitivity governs the degree of exploitation vs. exploration):

$$(4) \qquad Pr[D(t+1) = j] = \frac{e^{\theta \cdot Ej(t+1)}}{\sum_{k-1}^{4} = 1 e^{\theta \cdot Ek(t+1)}}$$

Supplementary Results 1

Examining the individual parameters from the four prior model for the bandit4arm_lapse model, we found a main effect of diagnosis only on the lapse and punishment learning rate parameters (Table 4 in main text) reiterating the same pattern seen in the winning two prior model. Of note, a similar pattern was seen on punishment learning rates in the model without the lapse parameter under threat (punishment learning rate under threat HDI 0.05-0.3; **Supplementary Table 2**); but, interestingly, not in the safe condition (HDI -0.19-0.27), although this model was not favoured in the model-comparison.

We also assessed whether the best patients and controls were better fit using a single or separate learning rates. We showed that in both groups, a model including two learning rates (bandit4arm_lapse / bandit4arm_lapse_decay) was better than one with only a single learning rate (bandit4arm_singleA_lapse; **Supplementary Table 3**)

Supplementary Results 2

Post-hoc Analyses

Possible confounds

A post-hoc exploratory analysis in the symptomatic group alone revealed no significant effect of specific diagnosis (8 MDD, 8 GAD, 28 MDD/GAD) on the lapse (*lapse* $F_{(2,41)}$=0.54,p=0.59, η^2=0.026; *lapse_decay* $F_{(2,41)}$=0.60,p=0.55, η^2=0.029), punishment learning rate (*lapse* $F_{(2,41)}$=0.50,p=0.61, η^2=0.024; *lapse_decay* $F_{(2,41)}$=0.56,p=0.58, η^2=0.027), reward sensitivity (*lapse* $F_{(2,41)}$=0.048,p=0.95, η^2=0.002; *lapse_decay* $F_{(2,41)}$=0.32,p=0.73, η^2=0.015), punishment sensitivity (*lapse* $F_{(2,41)}$=1.88,p=0.17, η^2=0.084; *lapse_decay* $F_{(2,41)}$=2.6,p=0.08, η^2=0.11) or decay (*lapse_decay* $F_{(2,41)}$=0.14,p=0.87, η^2=0.007) parameters. Moreover, Bayesian ANOVA equivalents provided better model evidence for the null model *over* a model including diagnosis (*lapse:* lapse: BF_{10}=0.2; punishment learning rate: BF_{10}=0.3; reward sensitivity: BF_{10}=0.2; punishment sensitivity: BF_{10}=0.6; *lapse_decay*: lapse: BF_{10}=0.3; punishment learning rate: BF_{10}=0.3; reward sensitivity: BF_{10}=0.2; punishment sensitivity: BF_{10}=0.96; decay: BF_{10}=0.2).

In the patient group, symptoms did not correlate with IQ (BDI: r(41)=0.15, p=0.1 [95%CI -0.05, 0.5]; trait anxiety: r(41)=0.27, p=0.36 [95%CI -0.2, 0.5]) or age (BDI: r(41)=-0.062, p=0.7 [95%CI -0.4, 0.3]; trait anxiety: r(41)=-0.15, p=0.3 [95%CI -0.4, 0.2]). Moreover, IQ did not correlate with the lapse (*lapse:* r(41)=-0.09 [95%CI -0.39, 0.22], $logBF_{10}$=-1.5, p=0.59, *lapse_decay:* r(41)=-0.03 [95%CI -0.34, 0.28], $logBF_{10}$=-1.6, p=0.85), punishment learning rate (*lapse*: r(41)=-0.3 [95%CI -0.56, 0.006], $logBF_{10}$=0.15, p=0.055, *lapse_decay*: r(41)=-0.23 [95%CI -0.50, 0.09], $logBF_{10}$=-0.64, p=0.15) or decay (*lapse_decay:* r(41)=-0.03 [95%CI -0.34, 0.28], $logBF_{10}$=-1.1, p=0.30) parameters.

Additional model independent analyses

We analysed reaction time in a 2(Threat, Safe) x 2(Post Stay, Post Switch) x 2(asymptomatic, symptomatic) repeated measures ANOVA. There was a main effect of diagnosis (F(1,130)=12, p<0.001, η^2=0.08), because patients were slower overall, and a main effect of preceding trial (F(1,130)=19, p<0.001, η^2=0.12), as all participants were slower following a switch trial, but no other main effects or interactions (all ps>0.4).

We also analysed total amount won and lost across conditions and groups in a 2(Threat, Safe) x 2(Wins, Losses) x 2(asymptomatic, symptomatic) repeated measures ANOVA. There was no main effect of diagnosis (F(1,130)=0.8, p=0.36, η^2=0.006), interaction between diagnosis and valence (F(1,130)=0.014, p=0.91, η^2<0.001), or any other interaction (all ps>0.5).

Supplementary Table 1: *All Bayes Factors from the Bayesian repeated measures ANOVA in the 'model agnostic task analysis':*

Model Comparison

Models	P(M)	P(M|data)	Log(BF_M)	Log(BF_{10})	error %
Null model (incl. subject)	0.053	1.561e-40	-88.768	0.000	
Condition	0.053	1.700e-41	-90.985	-2.217	1.135
Outcome	0.053	0.764	4.063	91.388	1.763
Condition + Outcome	0.053	0.097	0.660	89.326	7.225
Condition + Outcome + Condition ✻ Outcome	0.053	0.014	-1.398	87.356	2.306
Diagnosis	0.053	1.796e-41	-90.930	-2.162	1.097
Condition + Diagnosis	0.053	2.041e-42	-93.105	-4.337	2.549
Outcome + Diagnosis	0.053	0.092	0.605	89.276	2.658
Condition + Outcome + Diagnosis	0.053	0.010	-1.675	87.083	1.654
Condition + Outcome + Condition ✻ Outcome + Diagnosis	0.053	0.002	-3.511	85.255	3.574
Condition + Diagnosis + Condition ✻ Diagnosis	0.053	3.293e-43	-94.929	-6.161	2.451
Condition + Outcome + Diagnosis + Condition ✻ Diagnosis	0.053	0.002	-3.308	85.458	5.230
Condition + Outcome + Condition ✻ Outcome + Diagnosis + Condition ✻ Diagnosis	0.053	3.123e-4	-5.181	83.587	5.823
Outcome + Diagnosis + Outcome ✻ Diagnosis	0.053	0.016	-1.260	87.492	2.117
Condition + Outcome + Diagnosis + Outcome ✻ Diagnosis	0.053	0.003	-2.904	85.861	40.280
Condition + Outcome + Condition ✻ Outcome + Diagnosis + Outcome ✻ Diagnosis	0.053	2.788e-4	-5.294	83.473	4.514
Condition + Outcome + Diagnosis + Condition ✻ Diagnosis + Outcome ✻ Diagnosis	0.053	3.241e-4	-5.144	83.624	3.106
Condition + Outcome + Condition ✻ Outcome + Diagnosis + Condition ✻ Diagnosis + Outcome ✻ Diagnosis	0.053	5.157e-5	-6.982	81.786	5.252
Condition + Outcome + Condition ✻ Outcome + Diagnosis + Condition ✻ Diagnosis + Outcome ✻ Diagnosis + Condition ✻ Outcome ✻ Diagnosis	0.053	1.392e-5	-8.291	80.476	2.912

Supplementary Table 2: Group and condition effects on the full bandit4arm_lapse model
Values represent 95% highest density intervals (HDI) lower bound and upper bound). If the HDI does not encompass zero, we consider there to be a meaningful difference between the groups/conditions. We find a main effect of group on the punishment learning rate and lapse parameters (in bold).

	Symptomatic – Control				**Threat - safe**			
	Threat		**Safe**		**Anxious**		**Healthy**	
Reward Sensitivity	-5.71	1.63	-1.31	10.69	-12.00	0.76	-2.33	3.57
Punishment Sensitivity	-4.83	6.72	-4.48	21.52	-21.51	7.40	-1.80	3.32
Reward Learning Rate	-0.13	0.25	-0.14	0.26	-0.27	0.28	-0.07	0.11
Punishment Learning Rate	**0.11**	**0.45**	**0.08**	**0.55**	-0.30	0.25	-0.08	0.10
Lapse	**0.01**	**0.23**	**0.12**	**0.34**	-0.25	0.02	-0.07	0.08

Supplementary Table 3: Comparing learning rate models across groups*. We demonstrate that both groups are better fit (lower leave one out information criterion (LOOIC) in bold) by a model including two learning rates and a decay parameter. In other words, group differences are not driven by the different groups using different models.*

	LOOIC Anxious	**LOOC Healthy**
bandit4arm_lapse	44446	83771
bandit4arm_singleA_lapse	44679	84442
bandit4arm_lapse_decay	**44097**	**81997**

Supplementary References

1 Ahn, W.-Y. *et al.* Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. *Frontiers in psychology* **5**, 849 (2014).