**Supplemental Information**

**Randomization method and blinding**

The randomization code was drawn up by a researcher not involved in the study using an online randomization tool ([www.sealedenvelope.com](http://www.sealedenvelope.com)). Randomisation was stratified for gender and blocked with a block size of 4. In order to maintain blinding, the drug/placebo were placed into containers labelled with each participant ID at the start of the study by a researcher not involved in the study, according to the randomization schedule. The participants and study researchers were blinded to intervention until the end of the study. All assessments were administered by study researchers who were blind to group allocation.

**Task Descriptions**

*Facial Expression Recognition Task (FERT).* Task stimuli consisted of a series of facial expressions associated with six basic emotions: anger, disgust, fear, happy, sad, surprise and neutral. Each emotion also had a range of different intensity levels by using faces morphed between neutral (0%) and full intensity (100%) in 10% steps. For each emotion, there were 4 faces for each intensity level, totaling 280 emotional facial expressions randomly presented in the center of a display screen, each for 500 ms, divided across 4 blocks. Participants were required to indicate the emotional expression on the face by pressing a correspondingly labelled button. Outcome measures were % accuracy for recognizing facial expressions for each emotion and average reaction time for correct responses for each emotion.

*Emotional Categorisation Task (ECAT).* A series of 60 words, selected to be either positive or negative descriptors of personality, were randomly presented, each for 500ms. The personality descriptive words were matched for word length and ratings of frequency and meaningfulness. Participants were required to indicate whether they would like or dislike to be referred to as each personality descriptive word by pressing a correspondingly labelled button. Outcome measures were % accuracy for categorization and average reaction time for correct responses for positive versus negative self-referent personality descriptive words.

*Facial Dot-Probe Task (FDOT).* During each trial of the FDOT, a pair of faces was presented vertically prior to the appearance of a dot-probe. There were 96 trials in total, with 48 trials showing a positive facial expression (happy) and 48 trials showing a negative facial expression (fearful), each time paired with a neutral facial expression. Additionally, the pairs of faces could either be unmasked and presented for 500 ms or briefly presented for 14 ms and then replaced by a jumbled face mask. The pairs of faces were presented in a random order, divided across 4 blocks. The dot-probe appeared either in the same position as the emotional facial expression (congruent trials) or in the opposite position to the emotional facial expression (incongruent trials). The dot-probe comprised a pair of dots that could either be vertically or horizontally aligned and participants were instructed to indicate the alignment of the dot-probe by pressing a correspondingly labelled button as quickly as possible. Attentional vigilance scores were calculated for each participant by subtracting the mean reaction time from trials when probes appeared in the same position as the emotional face (congruent trials) from trials when probes appeared in the opposite position to the emotional face (incongruent trials). Positive values reflect attention towards the emotional face (vigilance) and negative values reflect attention away from the emotional face (avoidance).

*Emotional Recall Task (EREC).*Participants were given a surprise free recall task where they were asked to write down as many of the self-referent words that they could remember from the ECAT within a 4 minute time limit. The totals for positive versus negative self-referent words correctly recalled (hits) and falsely recalled (false alarms) were calculated.

*Emotional Recognition Task (EMEM).* 60 personality characteristic targets that appeared in the ECAT, and 60 novel personality characteristic distractors (30 positive, 30 negative, totalling 120 words) were presented in a random order at screen centre for 500ms. For each word, participants had to indicate 'Yes' if the word had appeared earlier in the ECAT for categorisation, or 'No' if they were novel, by pressing a correspondingly labelled button. Reaction time, accuracy, and false alarms were measured for positive and negative words.

*Rey Auditory Verbal Learning Task:* The RAVLT was used to measure declarative verbal memory. Participants were read 15 concrete nouns at a rate of one word per second (List A, e.g. Drum, Curtain). Participants were asked to immediately verbally recall as many items as they could, in any order. This was repeated a further four times, comprising five acquisition trials. Participants were then read a second set of 15 nouns (List B), and asked to recall words from this second list only. Immediately following List B, participants were asked to recall List A (short-delay), and once more after a delay of approximately 15 minutes, during which the probabilistic instrumental learning task was completed (long-delay). Number of words correct, repetitions (correct words recalled more than once in the same acquisition trial), and intrusions (incorrect words not present in the list) were measured. In the recognition memory component of the AVLT, participants were read a 50-item list comprising the 15 List A target words and 35 distractors. After each word, participants were asked to indicate verbally whether the word appeared in List A. Number of hits and false alarms were measured.

*N-Back:* The N-back provides an assessment of working memory capacity. Working memory load was manipulated by using three levels of complexity: 1-, 2-, 3-back, in a task design adapted from Mannie et al (1). The task comprised 160 trials, separated into 16 blocks. In this task, participants are required to indicate whether a letter presented on the screen (the “target” stimulus) matches a previously present letter (the “cue”) stimulus. Participants were required to compare the target letter to the cue letter that had been presented *n* trials ago (where *n* was one, two, or three in the 1-back, 2-back, and 3-back conditions respectively). Performance was compared to a control, 0-back, condition where participants were asked to press 'Same' if they saw the letter 'X', and 'Different' in response to any other letter. To minimize visual and phonological strategies, only phonologically closed letters were presented in upper and lower case. Thus, only the following letters were presented: b, B, d, D, p, P, t, T, v, V. Participants were instructed to ignore the case of letters and respond to indicate whether the target matched the cue using buttons on a keyboard labelled ‘same’ and ‘different’. Each block consisted of a sequence of 10 consonants varying in case. Letters were presented for 500msec with a fixed interstimulus interval of 1500ms. A fixation cross was presented for 5000ms at the end of each block. Prior to each task block, an instruction screen (0-, 1-, 2-, 3-back) was presented until the participant made a button press. A 4000ms fixation screen separated the instruction from the onset of the first letter. Participants first completed a 0- back, 1-back, 2-back, and then 3-back block. Condition order for the remaining 12 blocks was randomized. Mean accuracy and response latency for each condition were recorded.

*Contextual Cueing:* Implicit contextual learning was assessed using a visual search task similar to that described in (2). Participants were required to report the orientation of a target (‘T’) which is hidden amongst an array of distractor stimuli (‘L’s). On half of the trials, the target and distractor array is a repeat of one previously presented and on the other half of trials the array is novel. On each trial, a central fixation dot is presented for 500ms, followed by a stimulus array which includes the target stimulus ‘T’ and 11 distraction stimuli (‘L’s) that can appear in different orientations (rotation in all possible 90° angles) in 48 pre-specified locations. Participants were required to indicate the orientation of the target (which is rotated 90° to the left or right) using a button press while keeping central fixation. The stimulus array was presented for 700ms (or until participant response). There was a variable inter-trial interval (800ms, 1000ms, 1400ms). Reaction times greater than 2000ms were not recorded. Participants completed one practice block (48 trials) prior to the main experiment where the T appeared a reddish color to aid its location amongst distractors. Across the practice block, the T faded to white until it was indistinguishable in color from the distractors. Following the practice, 10 blocks of 48 trials (480 trials total) were presented. Presentation order of repeated and novel arrays was randomized within each block. At the end of each block, the average reaction time and accuracy rate for that block were presented on the screen as feedback for the participant. Median reaction time and mean accuracy were calculated for repeat and novel array trials separately. In order to measure contextual learning, difference scores (novel vs repeat) were computed for the first (five blocks) and second (five blocks) half of the task. Increased accuracy and decreased reaction times in repeated trials compared with novel trials are indicative of learning.

*Probabilistic Instrumental Learning Task:* A schematic of the task is illustrated in Figure S1. Participants completed three runs of the task and began each run with £1. Task stimuli were 6 pairs of symbols taken from the 'Agathoda' and 'Kryptonian' fonts. In each run, one pair of symbols was associated with winning £0.20, of which one was a high probability (0.7) win symbol, and one was a low probability (0.3) win symbol. The other pair of symbols was associated with losing £0.20, with one symbol associated with a 0.7 probability of losing, the other with a 0.3 probability of losing. High and low probability symbols in the pairs had reciprocal probabilities of 0.3 and 0.7 respectively of resulting in no change. Presentation of win and loss pairs was randomized across trials in each run. On each trial, one symbol appeared to the left, and one to the right of a central fixation cross for 4000ms. The position of each symbol to the left or right was random. Participants were asked to use the left and right arrow keys to pick the symbol they believed was most likely to win (or least likely to lose) them money, with the aim of maximizing their monetary pay off. They were informed that the total amount won would be added to their reimbursement for study participation. Once a choice had been made, a box appeared around the chosen symbol and the outcome (won £0.20, lost £0.20, no change) was displayed, as well as their running monetary total for that run. Participants first completed 10 practice trials to ensure they understood the task, featuring symbols that did not appear in the main experiment. Total amount won, total amount lost, end total, symbol choice, and choice consistency were recorded. In order to provide temporal information about reward learning in the two groups, learning curves were created, displaying the proportion of participants on each trial that chose the correct (high probability) symbol in the win condition, and the incorrect (high probability) symbol in the loss condition. Lastly a reinforcement learning model was fit to participant choice data (3). This model contained separate learning rate and decision temperature parameters for the win and loss outcomes, and was used to complement the non-model based analysis by further characterizing the effect of the drug on the task.

The effect of changes in these two parameters can be seen on examination of the learning curves (Figure 3a and 3b): an increase in the learning rate will result in the initial upswing of the curve becoming steeper, whereas an increased in the inverse decision temperature will result in the learning curve plateauing at a higher value: that is, the participant more consistently selects the better option, as would be expected if the participant valued the outcome more.

**Computational Analysis of Probabilistic Learning Task**

The Q learning model described by Pessiglone and colleagues (3) was fitted to participant choice data. In this model the value of selecting the chosen shape was updated on every trial using the equation (NB the value of the unchosen shape was kept constant):

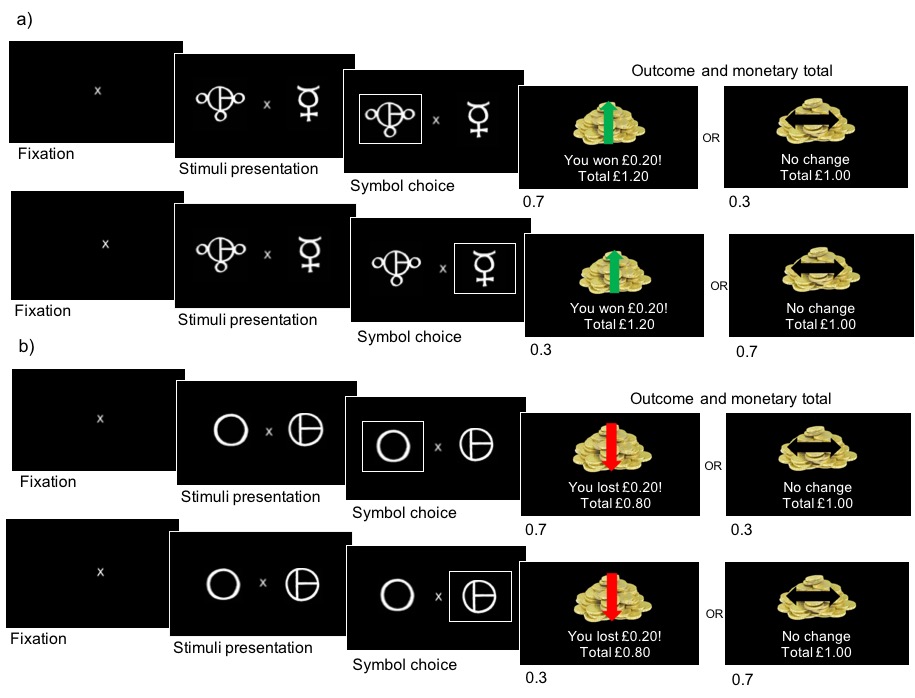
The above equation describes the learning process for shape “a”, in which is the value of shape “a” at time t, is the outcome observed for shape “a” at time t and is the learning rate. The Q values for each shape were initialized at 0 at the start of each block. Separate learning rates were estimated for win and loss trials. Finally the Q values of the two shapes in a pair were combined to produce a choice probability using the softmax function:

Where is the probability that the model will select shape “a” at time t and is the inverse decision temperature. Of note the inverse decision temperature in this equation functions identically to an “outcome sensitivity” parameter (i.e. a higher inverse decision temperature is the same as treating the outcome as if it had a higher value). Separate inverse decision temperature parameters were used for the win and loss outcomes. The parameters of this model were estimated, separately for the win and loss trials, by calculating the joint posterior probability of the two parameters, given participant choice, using a 100x100 grid and then deriving the expected value of each marginalized parameter (4). All parameters were log transformed before between subject analysis (5).

**Outliers**

For the ETB tasks (FERT, ECAT, FDOT, EMEM), trials were excluded if reaction times were greater than 6000ms or less than 200ms. In addition, trials on the FDOT were excluded if the reaction time was greater or less than 3 standard deviations from the individual participant’s mean reaction time. Across all tasks, if a participant’s performance met the criteria as an extreme outlier (data lying at more than three times the participants’ interquartile range above their third or below their first quartile), the analysis was repeated with and without their data included. However, given this did not change any of the results, data reported includes all participants.

Figure S1 – Probabilistic Instrumental Learning Task Schematic



*Participants chose from symbols presented either side of fixation. (a) Win condition: symbols were associated with a high probability (0.7) and low probability (0.3) of winning £0.20. (b) Loss condition: symbols were associated with a high probability (0.7) and low probability (0.3) of losing £0.20. High and low probability symbols had the reciprocal probabilities (0.3 and 0.7 respectively) of no change. Participants were presented with the monetary outcome following symbol choice.*

**Table S1: Mean scores on emotional tasks in prucalopride- and placebo-treated groups**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Task** | **Measure** | **Condition** | **Prucalopride**  **(N=19)** | **Placebo**  **(N=21)** |
| **Facial Expression Recognition Task (FERT)** | Accuracy (%) | Anger | 57.11 (7.23) | 54.64 (13.59) |
| Disgust | 65.79 (9.43) | 63.93 (12.46) |
| Fear | 55.53 (16.38) | 53.81 (15.54) |
| Happy | 82.11 (6.19) | 79.76 (5.91) |
| Sad | 65.26 (6.66) | 63.45 (7.92) |
| Surprise | 70.00 (4.86) | 69.05 (6.82) |
| Neutral | 83.68 (10.12) | 81.90 (10.78) |
| Reaction time (ms) | Anger | 1699.43 (399.79) | 1571.74 (182.99) |
| Disgust | 1697.90 (337.05) | 1561.58 (272.49) |
| Fear | 1887.38 (342.66) | 1807.79 (292.61) |
| Happy | 1315.43 (172.96) | 1227.74 (165.74) |
| Sad | 1407.62 (223.24) | 1369.00 (219.37) |
| Surprise | 1451.62 (242.26) | 1399.00 (246.66) |
| Neutral | 1268.48 (220.22) | 1326.74 (343.57) |
| **Emotional Categorization Task (ECAT)** | Accuracy (%) | Positive | 95.38 (5.62) | 93.18 (7.03) |
| Negative | 95.01 (4.16) | 92.38 (8.95) |
| Reaction time (ms) | Positive | 884.50 (217.85) | 901.76 (177.95) |
| Negative | 930.11 (192.51) | 973.86 (223.85) |
| **Facial Dot-Probe Task (FDOT)** | Masked  Attentional vigilance (ms) | Positive | -11.84 (33.98) | 2.43 (39.18) |
| Negative | -3.79 (36.08) | 2.52 (30.70) |
| Unmasked  Attentional vigilance (ms) | Positive | 10.58 (20.95) | -12.86 (41.17) |
| Negative | 9.89 (31.14) | -5.33 (46.24) |
| **Emotional Recall Task (EREC)** | Total items recalled | Positive | 6.47 (2.55) | 4.95 (2.44) |
| Negative | 6.00 (2.13) | 5.05 (2.40) |
| False alarms | Positive | 1.58 (1.43) | 3.38 (2.58) |
| Negative | 0.68 (1.0) | 1.48 (1.5) |
| **Emotional Recognition Task (EMEM)** | Accuracy (%) | Positive | 85.43 (10.38) | 83.81 (11.85) |
| Negative | 78.76 (13.71) | 72.38 (12.40) |
| Reaction time (ms) | Positive | 1209.89 (307.00) | 1088.48 (258.29) |
| Negative | 1240.95 (246.14) | 1201.67(335.55) |

*Values represent the mean with standard deviation in parentheses.*

1. Mannie ZN, Harmer CJ, Cowen PJ, Norbury R (2010): A functional magnetic resonance imaging study of verbal working memory in young people at increased familial risk of depression. *Biol Psychiatry*. 67:471-477.

2. Chun MM, Jiang Y (1998): Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cogn Psychol*. 36:28-71.

3. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006): Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*. 442:1042-1045.

4. Pulcu E, Browning M (2017): Affective bias as a rational response to the statistics of rewards and punishments. *Elife*. 6.

5. Browning M, Behrens TE, Jocham G, O'Reilly JX, Bishop SJ (2015): Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat Neurosci*. 18:590-596.