

Value-based decision-making between affective and non-affective memories

Erdem Pulcu

erdem.pulcu@psych.ox.ac.uk

Oxford University

Calum Guinea

Oxford University

Hannah Clemens

Oxford University

Catherine Harmer

University of Oxford <https://orcid.org/0000-0002-1609-8335>

Susannah Murphy

Oxford University

Article

Keywords:

Posted Date: February 23rd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1334268/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Yes there is potential Competing Interest. This work is funded by a joint grant from UK Medical Research Council and Janssen Pharmaceuticals awarded to CJH and SEM. CJH has received consultancy fees from P1vital Ltd., Janssen Pharmaceuticals, Sage Therapeutics, Pfizer, Zogenix and Lundbeck. SEM has received consultancy fees from P1vital, Zogenix, Sumitomo and Janssen Pharmaceuticals. EP has received consultancy fees from Janssen Pharmaceuticals. CJH and SEM hold grant income from UCB Pharma, Zogenix and Janssen Pharmaceuticals. CJH and SEM hold grant income from a collaborative research project with Pfizer. CG and HC do not declare any conflict of interest.

Version of Record: A version of this preprint was published at iScience on March 1st, 2024. See the published version at <https://doi.org/10.1016/j.isci.2024.109329>.

Abstract

Affective biases can influence how past events are recalled from memory. However, the mechanisms underlying how discrete affective events shape memory formation and subsequent recall are not well understood. Further understanding these processes is important given the central role of negative biases in affective memory recall in depression and antidepressant drug action. In order to capture cognitive processes associated with affective memory formation and recall, we studied value-based decision-making between affective memories in two within-subject experiments (n=45 and n=74). Our findings suggest that discrete affective events, created by large magnitude wins and losses on a Wheel of Fortune (WoF) outcomes, influence affective memory formation processes during reinforcement-learning (RL). 24 hours after learning, we show that healthy volunteers display stable preferences during value-based recall of affective versus non-affective memories in a binary decision-making task. Computational modelling of these preferences demonstrated a positive bias during value-based recall, induced by previously winning in the WoF. Across two complimentary analyses, we further showed that value-based decision-making between affective memories engages the pupil-linked central arousal systems, leading to pupil constriction prior to choice, and differential pupil dilation after the decision onset depending on the valence of the chosen options. Taken together, we demonstrate that mechanisms underlying human affective memory systems can be described by RL and probability weighting models. This approach could be used to understand the organisation of affective memories in humans, and as a translational assay to study the effects of novel antidepressants.

Introduction

Human social life is arguably the most complex in the animal kingdom, enriched by our ability to express and infer from others a wide spectrum of emotions. The breadth of this affective repertoire, along with our tendency to process positive and negative information asymmetrically^{1,2}, introduces biases that can shape not only our present experiences, but also how we recall events from the past. For example, the study of eyewitness memory has highlighted that centrally relevant details from emotional events (e.g. the characteristics of a criminal or the weapon used in a crime) are remembered more accurately than non-affective content³. Humans have asymmetries in affective information processing, known as “affective biases”. In healthy volunteers, affective bias is frequently observed in favour of positive events across multiple domains including perception, attention, reinforcement learning (RL), and memory^{2,4,5}. Conversely, in psychiatric conditions such as major depressive disorder (MDD), negative affective biases (i.e. preferential processing of negative relative to positive information)⁶⁻¹⁰ have been shown to play a role in the development and maintenance of symptoms¹¹⁻¹³. Whilst these trait-like biases in affective processing are well characterised, it is less well understood how discrete affective events (such as an unexpected positive or negative occurrence) induce biases that can influence learning and subsequent memory recall.

Previous work in humans has used Wheel of Fortune (WoF) paradigms to investigate the influence of discrete affective events (e.g. a large monetary win or loss) on subsequent value-guided choice. Using this approach, Eldar and Niv (2015) demonstrated in a sub-clinical cohort that individuals who scored highly on a measure of mood instability preferred slot machines encountered after a WoF when they had won the draw, whereas those who lost the WoF draw preferred slot machines that preceded it. More recently, in a study investigating the effects of serotonin on learning and memory processes, a similar effect was reported in healthy volunteers who demonstrated a preference for abstract fractals encountered after a WoF win, than for those encountered before a WoF loss (Michely et al 2020). Importantly, in both of these studies the expected values of the stimuli on either side of the WoF draw were comparable¹⁴ and the differences in choice preference were therefore driven by the influence of the affective context created by the unexpected positive or negative affective event.

These findings are consistent with evidence from animal assays, which have demonstrated that the affective context in which information with equal expected value is learnt can influence subsequent memory-guided value-based decisions in rodents. Interestingly, these affective biases can be pharmacologically altered, for example the rapid antidepressant agent ketamine has been shown to attenuate the influence of negatively biased memory on behaviour choice (Stuart, et al. ¹⁵). These effects contrast with those seen with conventional antidepressants which affect the encoding of new information in this rodent model, but leave former negatively biased memory intact^{16,17}. These findings highlight that affective biases in memory recall are malleable and may be an important neuropsychological mechanism through which rapid acting antidepressants exert their clinical effects.

The aim of the current studies is to extend this work and address some key outstanding questions about representations of reward value encoded under different affective influence. First, we investigated the stability and endurance of the influence of affective context during learning on subsequent value-based decision making, by testing whether experimentally induced changes in emotional state influence human value-based decision-making 24 and/or 48 hours later. Second, we applied a computational approach to further understand the mechanisms underlying affective biases. Recent RL studies have demonstrated that negative affective biases may develop even in healthy volunteers as a rational response to environmental contingencies¹⁸ and relate to poor filtering of informative negative experiences from uninformative ones¹⁹. In the current study, we analysed participant choice behaviour with a well-established computational model of value-guided choice, which posits that choice preference between binary options with different reward probabilities can be expressed in terms of weighted probabilities²⁰. Using this model, we tested an a priori hypothesis that a non-clinical population would overall display a positive bias, indicated by a preference for shapes encoded after winning on the WoF. Third, we used pupillometry to extend recent findings that have demonstrated that the information content of negative affective events engages the pupil-linked central arousal systems¹⁸. We investigated whether value-based recall of affective memories also engages central arousal systems. We used a model-based approach in our analysis of pupillary data and tested the prediction that subjective values which guide value-based decision making between affective memories will significantly influence pupil dilation.

Results

Participants and demographics

45 healthy volunteers took part in Study 1 and 74 healthy volunteers took part in Study 2. A summary of participant demographics and baseline self-report questionnaire data is presented in Table 1.

Table 1
Participant Demographics and questionnaire scores in the two studies.

| Measure | Study 1 (n = 45) | Study 2 (n = 74) |
|--|------------------|------------------|
| | Mean ± SD | Mean ± SD |
| Age | 29.33 ± 7.98 | 33.94 ± 7.95 |
| Gender, female | 31 (69%) | 62 (84%) |
| Years of education | 16.53 ± 3.00 | N/C |
| Trait-STAI | 32.51 ± 9.11 | 32.76 ± 10.94 |
| State-STAI | 29.31 ± 7.39 | 29.83 ± 11.1 |
| BDI | 3.56 ± 4.00 | 5.15 ± 5.46 |
| BAS Drive | 6.64 ± 2.31 | 8.2 ± 2.75 |
| BAS Fun | 7.16 ± 2.01 | 8.45 ± 2.71 |
| BAS Reward | 12.02 ± 1.71 | 13.05 ± 2.59 |
| BIS | 15.69 ± 3.41 | 16.15 ± 3.92 |
| MDQ | 3.02 ± 3.65 | 4.08 ± 3.68 |
| PANAS positive affect | 34.64 ± 6.97 | 31.54 ± 10.4 |
| PANAS negative affect | 17.67 ± 12.40 | 12.21 ± 4.51 |
| Trait STAI, Spielberger State-Trait Anxiety Inventory, trait form; State STAI, Spielberger State-Trait Anxiety Inventory, state form; BDI, Beck Depression Inventory; BAS, Behavioural Activation; BIS, Behavioural Inhibition; MDQ, Mood Disorder Questionnaire; PANAS, Positive and Negative Affect Schedule. | | |

N/C, not collected.

Experimental manipulation check for mood induction

To test whether the WoF manipulation influenced participants' mood, we compared their happiness ratings immediately before (pre-WoF) and immediately after (post-WoF) the draw. Overall, participants' ratings indicated that they felt significantly happier immediately after winning in the WoF and felt

significantly less happy immediately after losing ($F(2,64) = 9.388, p < .001$ for Study 1, and $F(1, 67) = 5.431, p = .023$ for Study 2, see Figure 2, further statistical analyses are available in Supplementary Results). Therefore, the WoF was effective at modulating mood in the expected direction.

Discrete affective events influence components of human reinforcement learning.

In Study 1 and 2 there was no significant effect of the WoF outcome (win or loss) on participants' tendency to choose the higher probability shapes during reinforcement learning (RL). (Study 1: $F(2,78) = 2.045, p = .136$, Study 2: $F(1,65) = 2.439, p = .123$, **Supplementary Figure 1**). However, analysis of learning behaviour with different RL models that can look beyond model-free summary statistics (**Supplementary Figure 2**), suggested that WoF outcome valence influenced components of participant choice behaviour. More specifically, our findings suggested that losing in the WoF reduced learning rates from negative prediction errors (PEs), whereas choice stochasticity and learning rates for +PEs were globally accelerated in post-WoF blocks (**Supplementary Figure 3**). Overall, these results were in line with the results emerging from their model-free analogues such as choice switch probability following a win or a no-win outcome during RL (further details in Supplementary Results and **Supplementary Figures 4**).

Value-based decision-making between affective memories reveals stable preferences

In Study 1 we evaluated participant choice preferences in a binary decision-making task in which the fractals from the learning stage were presented in random pairs (see Methods for details). We administered this preference test on two subsequent days (i.e. 24 and 48 hours post-learning) to be able to establish the stability of value-based decision-making between memories encoded in different affective contexts (2x4 rmANOVA: 2 preference days, 4 shape valence). There was no main effect of test day on participant choice behaviour (i.e. preference day 1 versus day 2, $F(1,39) = .303, p = .585$), indicating that value-based decision-making between affective memories remained stable (test-retest reliability coefficient 0.883). We observed a significant main effect of shape valence ($F(3,117) = 5.912, p = .001$). Shapes which were learnt following a WoF win or loss were selected more frequently than shapes learnt during the baseline (pre-WoF) block and those learnt following a neutral (blank) WoF (**Supplementary Figure 5A**). Specifically, shapes learnt after a WoF loss were not preferred significantly over shapes learnt after a WoF win (day 1: $t(86) = 1.1902, p = .24$; day 2: $t(86) = 1.867, p = .07$), but were preferred significantly over shapes learnt following a WoF blank outcome (day 1: $t(86) = 2.857, p = .005$; day 2: $t(86) = 2.267, p = .026$) and over shapes learnt during the pre-WoF baseline (day 1: $t(86) = 4.617, p < .001$; day 2: $t(86) = 4.313, p < .001$). Similarly, shapes learnt after a WoF win were preferentially chosen over WoF blank shapes (day 1: $t(86) = 1.689, p = .09$; day 2: $t(86) = 0.465, p = .643$) and over baseline shapes (day 1: $t(86) = 3.435, p < .001$; day 2: $t(86) = 2.418, p < .018$). The comparison between blank vs. baseline shapes was not significant (day 1: $t(86) = 1.644, p = .10$; day 2: $t(86) = 1.784, p = .078$). There was no significant main effect of WoF outcome order on participant choice behaviour ($F(5,39) = .364, p = .870$). Pairwise comparisons between equal value shape pairs are summarised in **Supplementary Table 1**.

We further investigated preferences between equal value shapes in Study 2. We observed that discrete affective events of comparable magnitude experienced during reinforcement learning in an experimental setting do not carry enough weight to make human learners negatively or positively biased across the board. After controlling for WoF order (e.g. whether participants experienced win or a loss outcome on Day 1) and shape identity order (i.e. whether shape pairs appeared after a win or a loss outcome during the learning stage) and individual differences in how well participants learned the reward probability of the environment during the learning phase, there was no significant main effect of valence [$F(1, 64) = .307, p = .582$], or reward probability [$F(5,320) = 1.542, p = .176$] or WoF order/shape identity on participant choice behaviour (all $p > .834$, **Supplementary Figure 6A**). Within individual comparisons, we observed that participants were significantly positively biased for win shapes associated with 60% reward probability ($t(73)=2.191, p=.03$), but none of the other comparisons at different reward probability levels were significant (all $|t(73)| < 0.7$, all $p > .48$). Although our experimental design did not allow us to decompose this effect any further, it is important to highlight that these shapes were farthest away in distance to the affective events (i.e. win and loss outcomes in the WoF draw) experienced during the learning/encoding stage, and were also associated with highest level of expected uncertainty among all the better shapes. Further analysis of participant choice behaviour raised the possibility that expected uncertainty of the reward environment may drive non-linear preferences between affective memories. This analysis is available in the next section.

Expected uncertainty of the reward environment may drive non-linear preferences between affective memories

In this section we report results focussing on the direction of the influence of affective events created by the WoF draw. Our findings rule out the possibility that discrete affective events systematically contaminate memories which are encoded prior to them, as there were no consistent preferences between the baseline shapes learned before a win or a loss outcome on the WoF draw ($F(1, 64) = .038, p = .846$, **Supplementary Figure 6B**).

Analysing participant choice behaviour in favour of affective memories (i.e. post-WoF shapes) against baseline shapes both of the same and the opposing days in a $2 \times 2 \times 6$ (valence \times same versus other day \times reward probability levels) rmANOVA indicated a significant main effect of reward probability ($F(5, 320) = 12.311, p < .001$) with no main effects of WoF and shape identity order, or any main effect of individual differences in learning (all $p > .274$). The main effect of reward probability was more evident for the high probability shapes, leading to non-linear preferences depending on the level of expected uncertainty in the learning environment. For example, irrespective of the WoF outcome valence, shapes learned immediately after the WoF draw (90% reward probability shapes) were preferred less relative to the equivalent baseline shapes, whereas those learned in the subsequent block (i.e. 75% reward probability shapes) were preferred consistently more than the equivalent baseline shapes (**Supplementary Figure 6C-F**). This observed nonlinearity emerging from the comparison of equal value shapes also contributes to the case

for using a non-linear probability weighting function to model participant choice behaviour that can capture preference biases across all comparisons.

Human affective memories are represented non-linearly

Due to a high number of equal value comparisons in the preference tests (reported in **Supplementary Table 1** and in **Supplementary Figure 6**), and considering the inherent stochasticity in participant choice behaviour, it is difficult to establish a bird's eye view on the organisation of human affective memories by solely relying on these comparisons. Furthermore, the interpretation of individual model-free contrasts may be problematic as only few direct comparisons out from a large number of combinations survive stringent Bonferroni corrections (Supplementary Table 1) and we were not able to single out individual comparisons as an a priori outcome measure due to lack of directly comparable studies in the literature on which such predictions may be based. To be able to look beyond individual comparisons and construct a model of human value-based recall of affective memories which we probed with 400+ trials involving many random shape pairs (e.g. win 90% vs other day baseline 10%), we further analysed participant choice behaviour in the preference tests with computational modelling (see Supplementary Methods and Materials for details). Here, we propose a simple value-based decision-making model with which participants choose between shape pairs based on the difference between their reward probabilities such that shapes with higher reward probabilities should be preferred over those with lower reward probabilities. In order to account for subjectivity in how each participant perceived reward probabilities associated with learned shapes we used a non-linear (i.e. an exponential logarithmic) 2-parameter probability weighting function. This model is described in Modelling Supplement (see Section 2, Equations 6-8).

It is important to highlight that in the preference tests, in a large majority of the trials, the expected value difference between the options were 0 (e.g. 60% win vs 60% loss shapes, **Supplementary Figure 7**), which would normally warrant random (i.e. 50-50) choices between these options, and consequently a benchmark log likelihood value of $-.69$ (i.e. $\log(.5)$) for any decision model. First, we tested how well our stochastic choice model for the preference test which relies on the probability weighting function, performs against this benchmark. Across both Study 1 and Study 2 this stochastic choice model performed significantly better than a random choice model (all $t > 8.2$, all $p < .001$), meaning that the model can capture the subjective valuations underlying binary decision-making between affective memories. This comparison also indicates that participants were attentive during the preference test and did not make decisions randomly, rather their decisions followed the proposed expected value function that relies on weighted probabilities. Our results demonstrate that where all trials and all possible comparisons between affective memories at different reward probability levels are concerned, discrete positive events (i.e. winning on a WoF draw relative to losing) influence subsequent value-based recall of memories associated with the better option during RL (i.e. shapes associated with higher reward probabilities which were sampled more frequently during the encoding stage, based on difference in area under the curve Study 1 $t(44) = 1.44$ and 2.40 , $p = .16$ and $.021$ (day 1 vs day 2 respectively); Study 2: $t(68) = 2.027$, $p = .047$, Figure 3). This affective influence occurs in a manner that further augments the subjective reward

probabilities of these options during value-based recall (Figure 3). Although there were some differences in the execution of preference tests between the 2 studies, we observed this positive bias in value-based memory recall and a differential effect of mood induction on preference consistently across 2 studies and 3 assessment time points, although with varying degrees of significance. Furthermore, this model-based AuC measure of positive bias correlated highly significantly with its model-free equivalent (i.e. the difference in the number of times high probability win shapes were chosen over loss shapes): $r(44)=.822$ and $.844$ in Study 1 with 14 abstract shapes (day 1 and day 2, respectively) and $r(73)=.652$ in Study 2 with 24 abstract shapes (all $p<.001$).

Finally, we explored the degree to which self-reported symptoms of depression in these nonclinical cohorts (23% had a 0 score on BDI) relate to our index of positive memory bias (i.e. AuC measure described above). We concatenated the data from both cohorts due to qualitatively overlapping results and used a multiple linear regression model with using BDI scores as a normalised regressor, while controlling for the effects of age, gender and WoF order. Intriguingly, the only significant regressor from this model is the intercept ($t(112)=2.458$, $p=.016$), whereas self-reported depression scores did not seem to influence this metric ($t(112)=-.246$, $p=.81$). A significant intercept in this model suggests that the AuC measure is greater than zero, and confirms positive memory bias.

Value-based decision-making between affective memories engages the pupil-linked central arousal systems

During the first preference test of Study 1, pupillometry data were collected across the entire decision process. We used a multiple linear regression model to quantify physiological response immediately before, during, and immediately after making choices between shapes learned following different WoF outcomes. Prior to choice, and even after controlling for the expected value difference between presented options as a proxy for choice difficulty, the expected value of chosen options estimated by the computational model reported above was significantly negatively correlated with pupil dilation ($t(38) = -2.48$, $p = .018$, Figure 4A). This means that choosing shapes associated with lower expected value leads to pupil dilation. After a choice had been made, affective memories had different physiological properties, and a 3×4 rmANOVA (3 levels of valence (i.e. win, loss, null WoF outcomes) \times 4 1-second timebins in the outcome delivery period) indicated a significant timebin by WoF outcome-valence interaction ($F(9,333) = 2.206$, $p=0.044$, Figure 4B, note that similarly significant results can be obtained if baseline shapes are also included in the model which would then be a 4×4 rmANOVA model). This interaction appeared to be primarily driven by a relatively higher pupil dilation in response to chosen shapes which were learned after winning in the WoF in early time bins (a main effect of valence $F(1,37)=3.782$, $p=.059$ for the comparison between wins and neutral between 0-2000ms), although none of the individual time bins showed a significant valence effect (all $p > 0.05$) and a crossover between the trajectories of loss versus neutral and baseline shapes which happens around 2500ms in the outcome delivery period.

Discussion

In the current studies, we investigated the mechanisms underlying value-based decision-making between affective memories formed under an RL protocol (Figure 1). Our findings from the preference tests suggest that human value-based decision-making between affective memories reveal stable preferences (**Supplementary Figure 5**). When only the pairwise comparisons of equal value options are concerned, discrete affective events of opposing valences do not carry enough weight to contaminate experimentally induced reward memories consistently across the reward probability spectrum (**Supplementary Table 1** and **Supplementary Figure 6**). However, when we also consider the global organisation of these affective memories (i.e. all cases where these memories were probed by randomly drawn options with not only equal but also different probabilities), our findings suggest that healthy volunteers retain positive biases for memories associated with better/higher probability options encoded through RL. We demonstrate that value-based decision-making between affective memories relies on nonlinear weighting of reward probabilities during recall (Figure 3). A model-based pupillometry analysis suggested that the expected value regressor from this decision model is encoded by the central arousal systems governing pupil size (Figure 4A). A complementary model-free pupil analysis suggested that pupil dilation is sensitive to the valence of affective memories in the early stages of outcome delivery during a binary preference test (Figure 4B). Taken together, these results illustrate that human memory-guided value-based decision-making is influenced by earlier experiences of discrete affective events and engages the pupil-linked central arousal systems prior to and after the decision onset.

In the current work, we investigated the degree to which nonclinical participants display a positive bias in value-based affective memory recall. Our reference points in designing this experiment were the preliminary work which utilised WoF paradigms in subclinical human populations and a rodent assay that assessed the impact of rapid versus traditional antidepressants on a single negative memory relative to a single control condition¹⁷. One of our key findings from Study 1 suggested that affective memory recall assessed 24 and 48 hours post-learning revealed stable preferences (**Supplementary Figure 5**), although some of the direct comparisons were significant only at the latter assessment point and below the statistical significance threshold in the former assessment point. This raises the possibility that as experimentally induced affective memories transition to be stored in the long-term memory, the differences between memories associated with positive and negative outcomes are further augmented. However, we did not anticipate this trajectory at the stage of experimental design and consequently we did not administer another preference test 72 hours later which could confirm this hypothesis.

In our experimental protocol, we probed a much larger pool of affective memories relative to earlier human work in order to achieve a decent coverage of the probability spectrum to aid computational modelling of affective memory recall. For example, in Study 2 there were 24 abstract stimuli which could be uniquely paired with 23 other stimuli during the preference test, resulting in a total grid space of 552 combinations. When we consider this complexity and the global organisation of human affective memories, an overarching and reasonably conservative interpretation of our results is that nonclinical volunteers are overall positively biased in their value-based recall (Figure 3 all panels). We observed this positive memory bias more strongly for shapes with higher reward probabilities. It is highly likely that this

is due to our experimental design in which paired fractals during RL blocks did not have independent probabilities which would encourage sampling the lower probability shape, but these were set to p versus $1-p$. In this scenario, after participants identified the higher probability shape they would be more likely to continue sampling the higher probability shapes, meaning that these would also have stronger associations relative to lower probability shapes during memory recall. However, it is rather unlikely that our positive bias results can be explained by recency (shapes which are learned after the WoF and nearer in time to the preference test) or practice effects, considering that in our protocol we randomised (i) the order of WoF on different days (both Study 1 and 2), (ii) whether shape sets will be associated with post-win or post-loss blocks (to be more stringent in Study 2), (iii) the side of stimuli associated with post-win and post-loss WoF (in Study 1, this was no longer needed in Study 2).

We think our approach revealed only the tip of the iceberg when it comes to fully understanding the organisation of experimentally induced affective memories in humans. There is substantial evidence, although limited to *deterministic* stimulus-outcome associations formed under conditioning, demonstrating that human learners store abstract knowledge in a grid-like code, in a manner that is similar to how the firing of entorhinal grid cells reflect spatial navigation in laboratory animals^{21,22}. In our case, value-based recall demands navigating through an abstract reward probability space which is formed under a *stochastic* RL protocol and influenced by the valence of preceding affective events (i.e. WoF outcomes); it is therefore likely to have more uncertainty and nonlinearity in the way this information is stored. The nearest comparable study in the literature investigated how serotonin influence value-based recall between affective memories. Our model-free results are globally in line with the placebo group ($n=34$) from Michely et al., (2020). However, their limited coverage of the probability spectrum (ie. only focusing on 70 vs 30% reward contingencies), smaller sample size in the placebo group relative to our work and relying solely on model-free assessment of preference bias meant that they were not able to propose a computational model for value-based decision-making between affective memories, that can capture the nonlinearity and positive bias that we were able to demonstrate here.

The second cognitive process which we think is relevant for understanding value-based recall is memory replay²³. Previous research suggests that humans can simulate the timeline of events (e.g. remembering the loss on the WoF while recalling the reward probability associated with the better shape in the block immediately after) during memory recall and this can be detected through analysing the neural signature associated with different events happening in a sequence²⁴. For example, a recent study demonstrated that events which generate large magnitude prediction errors create boundaries in memory formation²⁵. In the context of our experimental protocol, the WoF draws were the affective events which arguably generated the largest magnitude of PEs and this might explain why we observed a nonlinearity in preferences for the some of the baseline shapes (**Supplementary Figure 6C-F**). Here, it is also worthwhile to note that our model-based analysis of individual RL blocks indicated that participants did not encode shape values through associations with their potential to generate large magnitude RPEs (i.e. Model 4), therefore it is more likely that in our experimental protocol event boundaries in memory emerged with respect to the WoF draw rather than learning individual reward associations within each block. Overall,

these questions about grid-like organisation of human memory²⁶ and memory recall through replay are timely topics within cognitive neuroscience and require further research, ideally using high-field MRI (further discussion available in Supplementary Methods and Materials).

Our results demonstrate that value-based decision-making between affective memories engages the pupil-linked central arousal systems, with a negative correlation indicating that the pupil dilates more to chosen shapes with a lower expected value (Figure 4A). This is in line with recent computational work which showed that expected values of chosen options are negatively correlated with pupil dilation²⁷. In the context of our study, this pupillary signal is independent of the uncertainty in the probabilistic association between shape identities and outcomes, because this kind of *irreducible* uncertainty peaks at 50% (i.e. maximum entropy). After the decision onset, the physiological response to affective memories is explained by a valence x time-bin interaction. Population averages of pupil traces for each outcome valence demonstrated that this significant interaction was driven by differential pupil dilation to negative versus neutral (i.e. blank WoF outcome) memories and between the early and late phase of the outcome delivery period (Figure 4B). Considering that pupil dilation is under the influence of a number of neurotransmitters such as norepinephrine, acetylcholine and serotonin²⁸, our current work may be useful for understanding the effects of psychotropic compounds on affective memories. Although there is preliminary evidence to suggest that selective serotonin reuptake inhibitors induce a specific positive bias during value-based recall²⁹, physiological correlates of this positive bias remain unknown. We think that future studies using imaging methods with high temporal resolution such as magnetoencephalography could be valuable in understanding neurotransmitter modulation of human memory systems.

Finally, it is important to highlight some limitations of the current study with a view to informing future experimental design. One limitation which we share with preliminary work in this area is that we cannot dissociate the effects of mood from arousal which may be associated with winning or losing in the WoF. This is a common limitation of existing literature and future studies may record physiological markers such as pulse or skin response during memory encoding to dissociate the effects of mood from physiological arousal. Secondly, in some of our model-free analysis of the preference tests we were not able to fully dissociate the interaction between WoF effects and the influence of different probability levels. This could be achieved by randomising the order of the learning blocks which we omitted for the current study as it would also mean recruiting a larger sample size which was not feasible. Finally, future studies may consider designing preference tests in which all shapes were paired with every other shape only once (e.g. $24 \times 23 = 552$ combinations in Study 2) instead of relying more heavily on equal value comparisons (Supplementary Figure 7). Although this approach may be more demanding for clinical research, it might better for computational modelling and moving away from relying on individual model-free comparisons to develop mechanistic models of affective memory retrieval.

In this work, we expanded on previous human and rodent studies, and endeavoured to provide a detailed account of mechanisms underlying human RL and decision-making between experimentally induced affective versus non-affective memories and their corresponding physiology. Previously this approach was used in order to understand the influence of traditional antidepressants on human memory retrieval,

and in rodents as a test bench to compare the effects of traditional versus rapid-onset antidepressants. We propose that combining the design of our Study 2 that involved the full coverage of the probability spectrum during learning and decision-making with pupillometry from Study 1 would have good translational utility in bridging the gap between rodent and human models of how compounds with different antidepressant properties influence human memory systems.

Materials And Methods

Participants

Forty-five (Study 1) and seventy-four (Study 2) English-speaking healthy participants were recruited from the general public using online and print advertisements around Oxfordshire, UK. All of the participants had normal or corrected to normal vision and did not report a present or past psychiatric diagnosis, nor any serious medical condition that could impact their study participation. Participants were excluded if they were currently using psychotropic medication. Participants received monetary reimbursement for their time (£50) plus additional payment depending on their task performance across the learning and decision-making components of the experiment (£33.26-£38.40, mean±SD £37.25±0.90). The study was approved by the University of Oxford Central Ethics Committee (CUREC; ethics approval reference: R66705/RE001). All participants completed an informed consent form conforming to the Declaration of Helsinki.

General Experimental Procedures

In Study 1, testing sessions took place over 5 consecutive days at the University of Oxford, Department of Psychiatry at Warneford Hospital. In the first visit, the participants were taken through a screening interview to assess their eligibility. Then, the participants responded to a set of demographic questions and completed a battery of psychological questionnaires. After the screening interview, the eligible participants continued with the first day of learning and completed 3 blocks of a simple RL task in order to learn the associations between shapes and rewards. In line with the aims of the study, participants' affective state was manipulated using a WoF paradigm adapted from Eldar and Niv (2015).

To probe value-based recall of affective memories, after the training days, we asked participants to make decisions in a two-option forced-choice (TOFC) preference task in which various combinations of the abstract shapes they had learned about were paired with each other (i.e. on the last 2 days of the lab-based study, and the last day of the online study). Although no explicit feedback was given to participants in the preference test, they continued to accumulate money based on the reward probability of the chosen shape (i.e. 90% chance of winning 2p if the participant selects the shape associated with 90% reward probability in the learning phase). In the preference tests accumulated money was calculated but not displayed on-screen to ensure that participants cannot use this information as a proxy for their decision performance which would otherwise contaminate preference choice behaviour with further learning.

In Study 2, testing sessions took place over 3 consecutive days and were delivered using an online platform (due to the global COVID-19 pandemic). We manipulated the reward probabilities in each RL block pre-and-post WoF in a balanced manner in order to investigate how discrete affective events influence human RL. Further details of experimental procedures and statistical analysis approach is written below. All tasks were presented on a laptop running MATLAB (MathWorks Inc) with Psychtoolbox (v3.1).

Training phase

In the lab-based study (Study 1) on each of the learning days, participants started with a baseline block (30 trials), followed by two more blocks (each 90 trials), with rest periods between the blocks. In the online study (Study 2), all blocks contained 40 trials each. The shapes in the RL tasks were selected from the Agathodaimon and Dingbat Cobogo fonts. During the RL blocks, participants were asked to learn the reward probabilities associated with each shape through trial and error. Participants were explicitly informed that on any given trial if one shape is rewarded the other one is not rewarded (such that the probabilities are p versus $1-p$). A green frame appeared around the rewarded option to provide feedback on participant choice. For every correct choice made, participants collected 2 pence and they were asked to accumulate as much reward as possible. Participants started with £15 on day 1 and a running total at the bottom of the screen was updated at the start of each subsequent trial.

In total, the participants were asked to learn reward probabilities associated with 14 shapes in the lab-based study and 24 shapes in the online study. Participant preferences between these shapes were later probed in the preference test which took place on the 4th and 5th days of the lab-based study and the 3rd /final day of the online study.

Wheel of Fortune

In order to influence the participants' affective state, a single WoF draw was used in the break between the first and second blocks on each of the learning days. Participants were told that they could either win money, lose money, or receive nothing (blank). Unknown to the participants, the draw was not random but fixed, such that in the lab-based study each participant won (+£15), lost (-£11), or received a blank outcome (£0) across the 3 days (see Figure 1 and legends about a detailed description of the experiment). In the online study, we excluded the blank condition and focused only on win (+ £14) and loss (-£7) outcomes, and adjusted their outcome magnitudes considering the commonly observed salience differences between wins and losses³⁰. To control for an effect of order, these outcomes were counterbalanced across participants. In Eldar and Niv³¹ (2015), individual trials were rewarded with 25 cents and participants could either experience the loss or win of \$7 after a single WoF draw (1/28 ratio). In contrast, we decided to increase the WoF outcome/reward ratio to ensure that these affective events (i.e. WoF outcomes) would feel more salient to participants. Therefore, in our experimental design, each correct prediction during RL led to a smaller reward (2p). As illustrated in Figure 1A, win and loss outcomes were sandwiched between large magnitudes of opposing outcomes, creating a near-miss

effect aimed at strengthening the affective impact of the WoF. We used these large magnitude WoF outcomes to experimentally induce negative or positive memory biases.

Happiness ratings

During the learning days, participants were asked to report their current happiness. In the lab-based study (Study 1), we used a Likert scale from 1-9 (with higher numbers indicating greater happiness), whereas in the online study (Study 2) we used a visual analogue scale. Participants were asked to provide a happiness rating at different time points during the experiment, for example before starting the first RL block, more critically immediately before and after the WoF, and at the end of their training on any given day.

Preference tests

In the last phase of the experiment, we conducted preference tests. In order to assess the test-retest reliability of participant preferences we conducted the preference test twice, once on the 4th day and again on the 5th day in the lab-based study. Participants were presented with random pairs of all shapes they encountered during the training days and on each trial, were asked to choose the shape with the higher reward probability. The aim was to examine whether a WoF induced change of affective state would bias their valuations of the learned shapes. Written and spoken instructions were given and subjects were presented with a print-out of all the shapes (14 in Study 1 (lab) and 24 in Study 2 (online)) before the task began, to provide them with an opportunity to recall the learned shapes. Within each study, shapes were randomly paired with each other to cover all possible combinations leading to 400/456 trials (lab-based versus online) presented across three blocks, and the order of these trials was randomised across participants. We were particularly interested in the pairs of shapes that had objectively identical reward probabilities but appeared after different WoF outcomes, thus should be encoded under different affective influence (Figure 1). Therefore, the majority of trials presented during the preference tests prioritised randomly pairing equal probability shapes such that the value difference between the options would be near 0 (**Supplementary Figure 7**).

In the preference tests (both Study 1 and 2), participants did not receive any explicit feedback about whether their choices were correct or not, so they had to rely on what they had previously learned. In order to prevent further learning, participants' running total did not appear on the screen during a block of trials, but they did continue to accumulate money based on whether their choice was correct and the reward probability of the chosen shape. Their running total was only displayed between blocks to provide an indication of their performance in the previous block. Administration of preference tests took roughly 30 minutes for the online study and 50 minutes for the in-lab study with pupillometry.

Questionnaire measures

In addition to the computerised task, participants were asked to respond to a series of self-report questionnaires on the first day of their testing, prior to completion of the first learning block. These

questionnaires included: (1) Beck Depression Inventory (BDI-II)³², a standard measure of depression; (2) Spielberger State-Trait Anxiety Inventory (STAI)³³, an anxiety measure comprising trait (anxiety proneness) and state (current state of anxiety) subscales; (3) Positive and Negative Affect Schedule (PANAS), assessing the feeling and expression of positive and negative emotions³⁴; (4) Behavioural Activation/Behavioural Inhibition (BIS/BAS), reflecting aversive motivation and appetitive motivation³⁵; (5) and the Mood Disorder Questionnaire (MDQ)³⁶, a screening tool for Bipolar Disorder.

Statistical analyses

We used repeated-measures analysis of variance (rmANOVA) models to investigate the effects of our experimental manipulations. All significant main effects and interaction terms were followed up by post hoc tests, corrected for multiple comparisons using Bonferroni correction. All analyses were conducted on MATLAB and SPSS v25. In the lab-based study, 6 participants were excluded from the analyses of pupillometry data due to signal dropout affecting more than half of the trials, and in total 2 participants were excluded from the behavioural analysis (one person reported intentionally selecting the lower probability shapes during the training phase, the other person dropped out from the study before preference test 2). In total, 5 participants were excluded from the online study, 4 individuals who did not give more than 4 unique mood ratings across 16 assessment points, and one individual who consistently selected the lower probability shapes more frequently than the higher probability shapes. Across all statistical models, reward probability, valence and behaviour in pre-post-WoF blocks were entered as within subject factors, whereas WoF order (e.g. whether participants experienced win or a loss outcome on the 1st day) and shape set which randomised the associations between shape identities and reward probabilities were entered as between subject factors.

Pupillometry

On the first day of preference tests (day 4) in the lab-based study, we collected pupillometry data to assess physiological response to affective memories during value-based decision-making. Participants' heads were placed at 70cm distance from the computer screen, stabilised by a chin rest. The eye-tracking system (Eyelink 1000 Plus; SR Research, Ottawa, Canada) was linked to the presentation computer through an ethernet connection. The sampling rate for pupillometry was set to 500 Hz recording from both eyes. The preference test was presented on a VGA monitor. A fixation cross marking the middle of the screen separated the presented pair of shapes.

Preprocessing of the pupillary data involved removing eye blinks that were identified using the built-in filter of the Eyelink system. A linear interpolation was implemented for all missing data points (including blinks). The resulting pupil trace was processed through a low pass Butterworth filter (cut-off of 3.75 Hz) and then z-transformed across the preference test session^{37,38}. In order to assess phasic response to task-related variables, we performed baseline correction by subtracting the mean pupil size during the 2-second baseline period prior to our epochs of interest (i.e. decision and outcome) from each time point in the post-stimuli presentation period. Individual trials were excluded from the pupillometry analysis if more

than 50% of the data from the outcome period had been interpolated^{18,37}. The preprocessing resulted in a single set of pupil time-series per participant containing pupil dilation data for each of the included trials.

Declarations

Funding and Disclosure

This work is funded by a joint grant from UK Medical Research Council and Janssen Pharmaceuticals awarded to CJH and SEM. CJH has received consultancy fees from P1vital Ltd., Janssen Pharmaceuticals, Sage Therapeutics, Pfizer, Zogenix and Lundbeck. SEM has received consultancy fees from P1vital, Zogenix, Sumitomo and Janssen Pharmaceuticals. EP has received consultancy fees from Janssen Pharmaceuticals. CJH and SEM hold grant income from UCB Pharma, Zogenix and Janssen Pharmaceuticals. CJH and SEM hold grant income from a collaborative research project with Pfizer. CG and HC do not declare any conflict of interest.

Author Contributions

EP, CG, CJH and SEM designed the study. EP, CG and HC collected the data. EP analysed the data. All authors contributed to writing of the manuscript. Funders did not have any input in study design, analysis approach or decision to disseminate the results.

References

1. Eil, D. & Rao, J. M. The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics* **3**, 114–138 (2011).
2. Rozin, P. & Royzman, E. B. Negativity bias, negativity dominance, and contagion. *Personality and social psychology review* **5**, 296–320 (2001).
3. Christianson, S.-Å. Emotional stress and eyewitness memory: a critical review. *Psychological bulletin* **112**, 284 (1992).
4. Sharot, T. The optimism bias. *Current biology* **21**, R941-R945 (2011).
5. Palminteri, S., Lefebvre, G., Kilford, E. J. & Blakemore, S.-J. Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. *PLoS computational biology* **13**, e1005684 (2017).
6. Elliott, R., Zahn, R., Deakin, J. W. & Anderson, I. M. Affective cognition and its disruption in mood disorders. *Neuropsychopharmacology* **36**, 153–182 (2011).
7. Gotlib, I. H. & Joormann, J. Cognition and depression: current status and future directions. *Annual review of clinical psychology* **6**, 285–312 (2010).
8. Leppänen, J. M. Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. *Current opinion in psychiatry* **19**, 34–39 (2006).

9. Mathews, A. & MacLeod, C. Cognitive vulnerability to emotional disorders. *Annu. Rev. Clin. Psychol.* **1**, 167–195 (2005).
10. Ressler, K. J. & Mayberg, H. S. Targeting abnormal neural circuits in mood and anxiety disorders: from the laboratory to the clinic. *Nature neuroscience* **10**, 1116–1124 (2007).
11. Harmer, C. J. Serotonin and emotional processing: does it help explain antidepressant drug action? *Neuropharmacology* **55**, 1023–1028 (2008).
12. Roiser, J. P., Elliott, R. & Sahakian, B. J. Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology* **37**, 117–136 (2012).
13. Lewis, G. *et al.* Variation in the recall of socially rewarding information and depressive symptom severity: a prospective cohort study. *Acta Psychiatrica Scandinavica* **135**, 489–498 (2017).
14. Eldar, E. & Niv, Y. Interaction between emotional state and learning underlies mood instability. *Nature communications* **6**, 6149 (2015).
15. Stuart, S. A., Butler, P., Munafo, M. R., Nutt, D. J. & Robinson, E. S. Distinct Neuropsychological Mechanisms May Explain Delayed- Versus Rapid-Onset Antidepressant Efficacy. *Neuropsychopharmacology* **40**, 2165–2174, doi:10.1038/npp.2015.59 (2015).
16. Stuart, S. A., Butler, P., Munafo, M. R., Nutt, D. J. & Robinson, E. S. A translational rodent assay of affective biases in depression and antidepressant therapy. *Neuropsychopharmacology* **38**, 1625–1635 (2013).
17. Stuart, S. A., Butler, P., Munafo, M. R., Nutt, D. J. & Robinson, E. S. Distinct neuropsychological mechanisms may explain delayed-versus rapid-onset antidepressant efficacy. *Neuropsychopharmacology* **40**, 2165–2174 (2015).
18. Pulcu, E. & Browning, M. Affective bias as a rational response to the statistics of rewards and punishments. *Elife* **6** (2017).
19. Pulcu, E. & Browning, M. The misestimation of uncertainty in affective disorders. *Trends in cognitive sciences* (2019).
20. Prelec, D. The probability weighting function. *Econometrica*, 497–527 (1998).
21. Constantinescu, A. O., O'Reilly, J. X. & Behrens, T. E. Organizing conceptual knowledge in humans with a gridlike code. *Science* **352**, 1464–1468 (2016).
22. Garvert, M. M., Dolan, R. J. & Behrens, T. E. A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *Elife* **6**, e17086 (2017).
23. Deuker, L. *et al.* Memory consolidation by replay of stimulus-specific neural activity. *Journal of Neuroscience* **33**, 19373–19383 (2013).
24. Liu, Y., Dolan, R. J., Kurth-Nelson, Z. & Behrens, T. E. Human replay spontaneously reorganizes experience. *Cell* **178**, 640–652. e614 (2019).
25. Rouhani, N., Norman, K. A., Niv, Y. & Bornstein, A. M. Reward prediction errors create event boundaries in memory. *Cognition* **203**, 104269 (2020).

26. Peer, M., Brunec, I. K., Newcombe, N. S. & Epstein, R. A. Structuring Knowledge with Cognitive Maps and Cognitive Graphs. *Trends in Cognitive Sciences* (2020).
27. Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S. & Wyart, V. Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nature neuroscience* **22**, 2066–2077 (2019).
28. Faber, N. J. Neuromodulation of pupil diameter and temporal perception. *The Journal of Neuroscience* **37**, 2806 (2017).
29. Michely, J., Eldar, E., Martin, I. M. & Dolan, R. J. A mechanistic account of serotonin's impact on mood. *Nature communications* **11**, 1–11 (2020).
30. Ruggeri, K. *et al.* Replicating patterns of prospect theory for decision under risk. *Nature Human Behaviour*, 1–12 (2020).
31. Eldar, E. & Niv, Y. Interaction between emotional state and learning underlies mood instability. *Nat Commun* **6**, 6149, doi:10.1038/ncomms7149 (2015).
32. Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. & Erbaugh, J. An inventory for measuring depression. *Archives of general psychiatry* **4**, 561–571 (1961).
33. Spielberger, C. D., Gorsuch, R., Lushene, R., Vagg, P. & Jacobs, G. *Manual for the state-trait anxiety scale*. Consulting Psychologists (1983).
34. Watson, D., Clark, L. A. & Tellegen, A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* **54**, 1063 (1988).
35. Carver, C. S. & White, T. L. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of personality and social psychology* **67**, 319 (1994).
36. Hirschfeld, R. M. *et al.* Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. *American Journal of Psychiatry* **157**, 1873–1875 (2000).
37. Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X. & Bishop, S. J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature neuroscience* (2015).
38. Nassar, M. R. *et al.* Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature neuroscience* **15**, 1040–1046 (2012).

Figures

dependent on each other (probabilities add up to 1). Using trial and error, participants could infer the reward probability associated with each shape. This information could then be used to maximise their monetary reward. Participants started with £15 and their running total, displayed below the fixation cross for the duration of every trial, updated by 2p for each correct choice made. Incorrect choices did not have any monetary effect on participants' running total. Note that the visual appearance of the preference test was nearly identical to the learning task timeline shown. The only differences in appearance were: (i) shapes were randomly paired with each other, (ii) on each trial the unchosen shape would disappear from the screen as soon as a decision is made, and (iii) participants did not see their running total during the preference test to avoid having secondary learning effects.

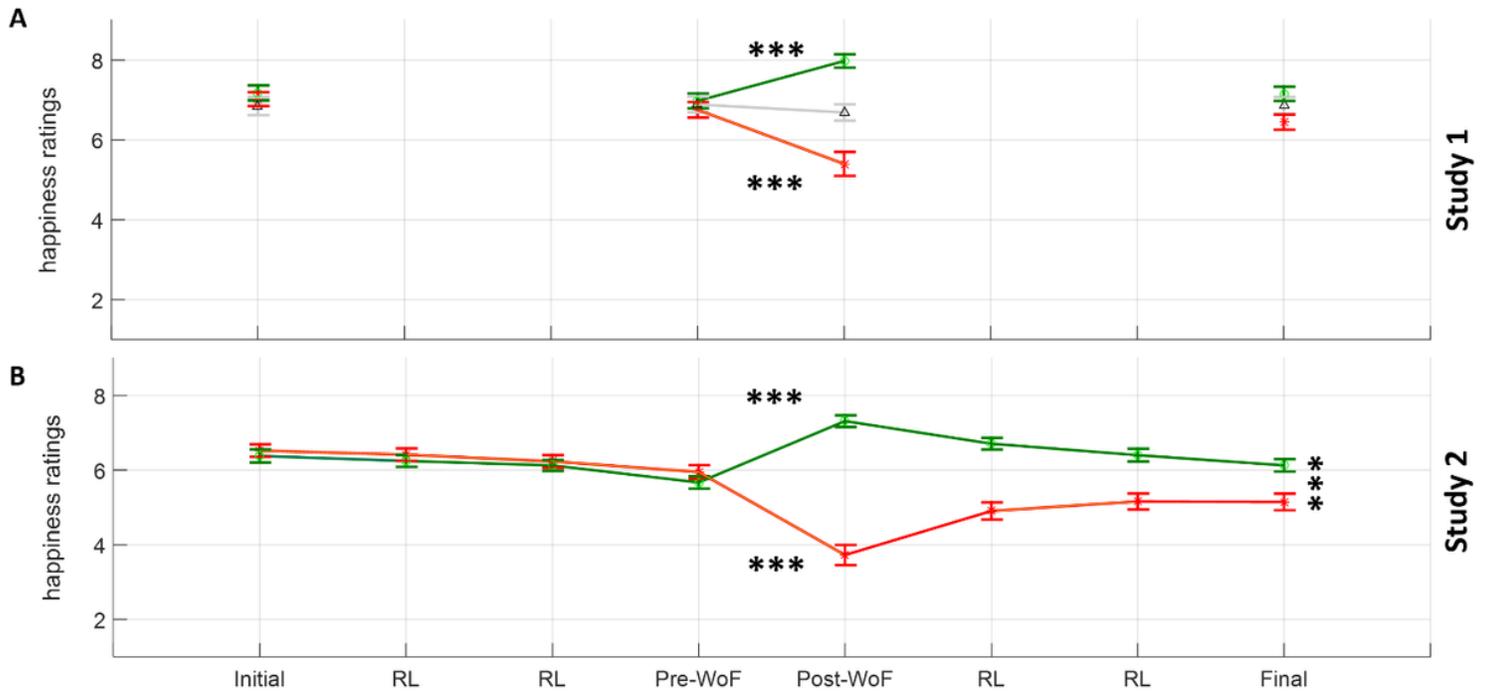


Figure 2

Happiness ratings. (A) In the lab-based study, participants' current happiness significantly improved post-WoF on a win day (green line), while significantly worsening on a loss day (red line). On the day with a blank WoF draw (grey line), there was no significant difference between pre- and post-WoF happiness ratings. Moreover, initial and final happiness ratings were comparable on all but the loss day. On the day with the loss outcome, although current happiness recovered significantly from post-WoF, it was still significantly lower than participants' initial rating. **(B)** In Study 2, mood was sampled more frequently than Study 1 (also see Figure 1A, blue diamonds for direct comparison between the studies). The findings related to mood change were replicated in the [online] Study 2. The green line indicates mood ratings on the training day with a win on the WoF, whereas the red line indicate the mood ratings on the loss day. Across all panels, error bars reflect ± 1 SEM. *** $p < .001$. RL blk: reinforcement learning block. In Study 2, the happiness ratings were made at the end of each block, apart from the initial rating which was made before participants started their daily training session.

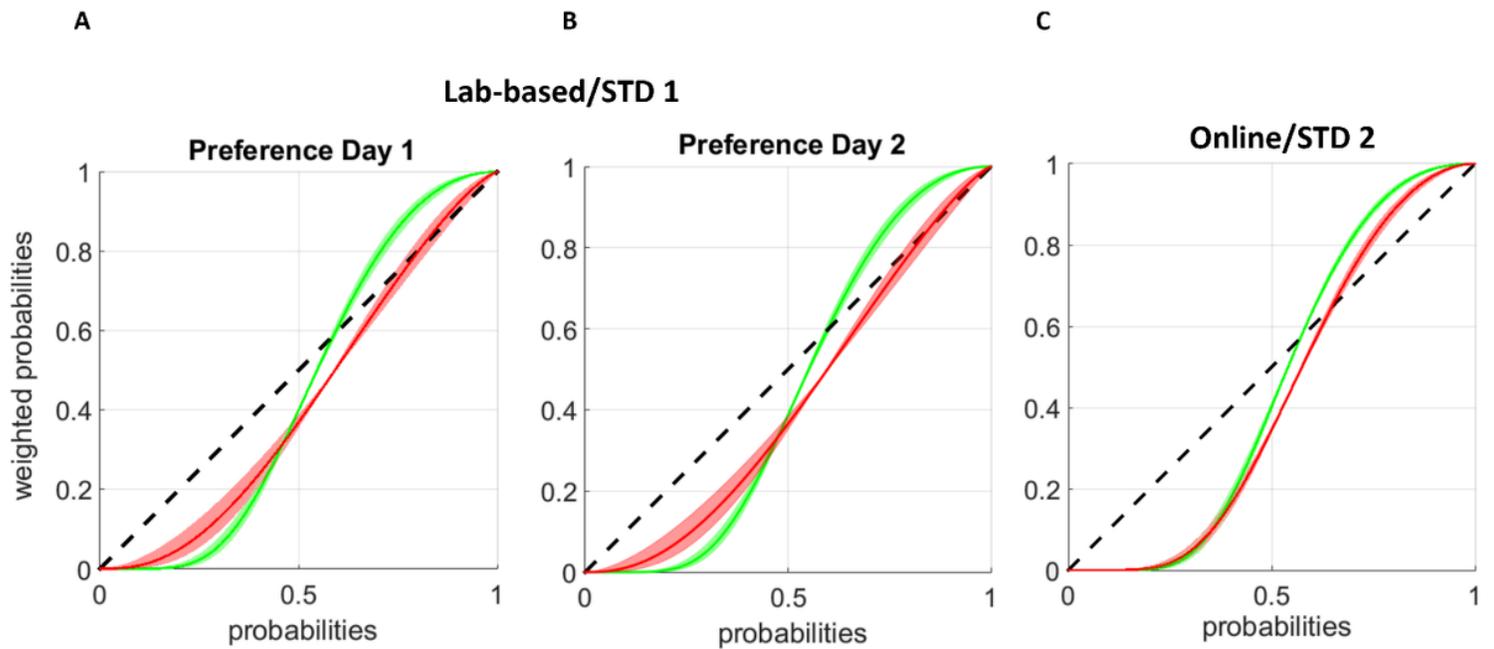


Figure 3

Probability weighting function demonstrating affective biases in memory-guided value-based decision-making. (A-B) In line with the model-free results reported in **Supplementary Table 1** and **Supplementary Figure 5**, model-based results show consistent effects between preference test days 1 and 2. The probability weighting function indicates that positive biases will have a stronger effect in contaminating neutral information associated with higher reward probability shapes (i.e. Baseline 70%), such that the reward probability associated with the baseline shape will be augmented when they are presented on the side associated with win shapes (green curve and SEM shading), whereas high probability baseline shapes will be under-weighted when they are presented on the side associated with loss shapes (red curve and SEM shading). Sidedness in stimuli presentation was counterbalanced across participants and between preference test days 1 and 2. This behaviour was reversed in the lower reward probability spectrum (i.e. for shapes associated with reward probability 30% and below). This crossover in the trajectory of the weighting function is also evident in the model-free comparisons reported in **Supplementary Table 1**. **(C)** Perceived probabilities during value-based recall between affective memories in the online study (Study 2) in which all stimuli were presented randomly on each side of the screen. The difference between win and loss trajectories become more evident for higher probability shapes which were sampled more frequently during the encoding stage. Shading around the population mean denotes ± 1 SEM.

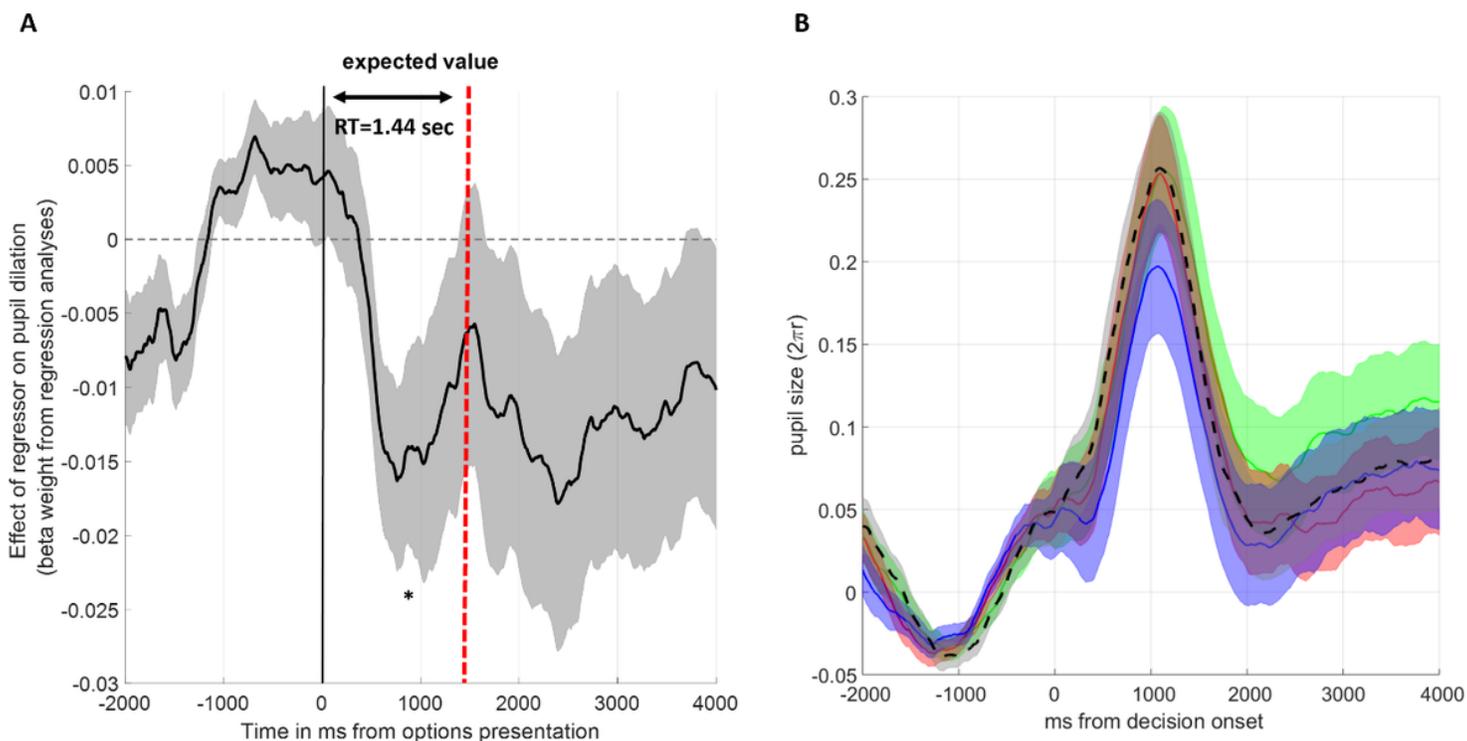


Figure 4

Pupillometry results. (A) A multiple linear regression analysis of the pupillary response suggests that pupil dilation negatively correlates with the expected values of chosen shapes based on the model shown in **Figure 4**. The vertical red dashed line marks the average response time (RT) during value-based decision-making. **(B)** During the outcome delivery period (i.e. once a decision has been made) win memories lead to a larger pupil dilation relative to neutral memories (the difference between green lines versus the blue line). The significant valence x time bin interaction seems to arise from differential pupillary time courses between negative (red line) versus neutral (blue line) and baseline (grey line) memories which crossover towards the end of the outcome delivery period. Note that as soon as the decision was made (0 ms), the unchosen option would disappear from the screen and chosen option was displayed for further 4 seconds. No feedback was given on the accuracy of choice to avoid additional learning during the preference test. In both panels, shading around the population mean denotes ± 1 SEM. * $p < .05$.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AffectiveMemoriesSupplementaryResults.docx](#)