Timing matters: primacy effects of experience on mood revealed using a novel Mood-Machine-Interface

Hanna Keren[1 ł], Charles Y. Zheng[2], David C. Jangraw[3], Katharine Chang[1], Aria Vitale[1], Dylan Nielson[1], Robb B. Rutledge[4,5], Francisco Pereira[2], Argyris Stringaris[1]

Abstract

Mood is thought to integrate across our experiences, yet we do not know how the relative timing of past events shapes how we feel in the moment. Here, we investigate the relationship between the timing of previous experiences and mood by combining a novel closed-loop mood controller alongside computational modelling and neural data. We first present the development of a Mood Machine Interface which allows us to individualize rewards in real-time in order to generate substantial mood transitions, across healthy as well as depressed, adolescents and adults. We then show that early-experiences have a larger effect on mood than recent ones, and that the longer one is exposed to a given context, the harder it is for new events to change mood. We find that ACC neural activity underlies the influence of early experiences on mood. This provides a neuro-computational account of mood regulation by early events and suggests new directions for individualized mood interventions.

[1] Mood Brain and Development Unit, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. [2] Machine Learning Team, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. [3] Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. [4] Max Planck University College London Centre for Computational Psychiatry and Ageing Research, London, United Kingdom. [5] Wellcome Centre for Human Neuroimaging, University College London, London, United Kingdom.

[ł] Correspondence and requests for materials should be addressed to H.K. (Hanna.Keren@nih.gov; 9000 Rockville Pike, Bethesda, MD, USA).

Mood is thought to integrate over the history of our experiences[1]. Yet, it remains unclear how the timing of our experiences influences how we feel in the moment. Existing models of mood[2-5] assume that recent experiences are most important in shaping mood. However, it is commonly found that early events, such as poor performance at the beginning of a job interview, or a mishap at the start of the day, can have a lasting negative impact on our mood overshadowing more recent positive experiences[6-9]. Resolving this question has two important implications. First, knowledge of environmental events timing could help predict their impact on individuals' mood at any given time point. Second, knowledge of the relative influence of event timing may help optimize the design of artificial environments (clinical or otherwise) that induce mood changes.

To address this central issue of the relation between previous experiences and mood, we took the following steps:
First, we built a novel closed-loop paradigm, which we term a Mood Machine Interface (MMI), by adding a proportional-integral controller into a standard gambling task. This allowed us to shift individual mood parametrically to arbitrary values, and thus investigate the relation between mood changes and past experiences across strongly directional (positive or negative) environments.
Second, we developed a novel computational model which operationalized the theory that mood can be primarily influenced by early experiences. We show that the model which gave the highest weight to early experiences, by considering expectation as the long-term average (LTA) of previous outcomes, was superior to a broad range of alternative models. This was judged on both training error and streaming prediction error, and across two independent samples, one of which was a preregistered, confirmatory analysis. These results were also robust to participants' initial mood (depressed vs normothymic participants) and developmental status (adolescents vs adult participants).
Third, we showed with fMRI that the influence of previous events on mood is encoded by neural signals in the brain region suggested to be involved in mood regulation, the anterior cingulate cortex (ACC)[10-13]. In particular, the weights of the LTA expectation parameter, which defined the extent of the influence early events have on mood, were related to ACC activity preceding mood rating.

This work shows how applying computational and engineering tools can address some of the challenges involved in studying complex phenomena such as mood[14,15], and bring new insights to questions relevant to psychopathology[16,17]. Specifically, our findings provide a novel way to induce large mood shifts as well as a neuro-computational account of the relationship of mood to the history of rewarding and negative experiences.

1. Developing the MMI paradigm

The need for a new paradigm

To study the impact of previous events on mood, we needed to develop a paradigm in which each participant receives parametrically quantifiable stimuli, of sufficient magnitude to be influential on their mood. Crucially, we also required a method unlikely to change mood due to demand effects (i.e., when participants become aware of the experimental goal and respond to comply with it). We addressed these requirements by developing a closed-loop controller that adjusted reward values in real-time according to individual mood response. This strategy can compensate for a low mood response or adaptation to stimuli over time. Recent findings suggested that Reward Prediction Errors (RPE) (i.e., the difference between expected and received outcome), influence momentary fluctuations in mood. Therefore, we set this paradigm to parametrically manipulate the value of RPEs to shift mood. The Mood-Machine-Interface (MMI) we built to meet these requirements ensures that experimental experiences continue to influence mood, and therefore it can uncover which of the experiences are most influential.

The MMI paradigm

The MMI generated mood transitions by adjusting RPEs in a closed-loop circuit. Figure 1 provides a schematic description of the paradigm: Each trial consisted of a choice about whether to gamble between two monetary values or to receive a certain amount instead of gambling. The task RPE value was considered as the difference between the outcome and the mean of the two gamble values (“RPE Value”). It was possible to modify the RPE value by manipulating the values of the gamble possibilities and therefore also the received outcome. The aim of the task was to shift mood towards the Mood Target value. The participant provided mood ratings every 2-3 gambling trials by moving a cursor along a scale between unhappy and happy mood (b. “Mood Rating”). To shift mood towards the target value, following each mood rating the algorithm recalculated the Mood Error, the difference between the last mood rating and the mood target (c. “Mood Error”). The Mood Error value was translated to the RPE value of the consecutive trials, using a proportional-integral (PI) control algorithm (d. “PI Control”), a combination of proportional and integral accounts of the mood error (see Methods for details). Using this algorithm, the next RPE value (e. “Next RPE”) was estimated to compensate for the mood error RPE, such that it increased when mood reactivity was too low.

To investigate both positive and negative mood transitions, we split the experiment into three blocks, each with a different mood target: the goal of the first block was to increase mood towards the target of 100% (highest) mood value; the second block decreased mood by setting the target to 0% (lowest) mood level; and again in the third block, the goal was to increase the mood to a target of 100% (highest) mood. To maintain the unpredictability of outcomes, 30% of trials were small RPE values, incongruent to the block valence (small negative RPEs during first and third increasing-mood blocks and small positive RPEs during the second decreasing-mood block). Thus, during the first and third blocks 70% of trials delivered the better value if participants chose to gamble, while during the second block, 70% of trials delivered the lowest option of the gamble.

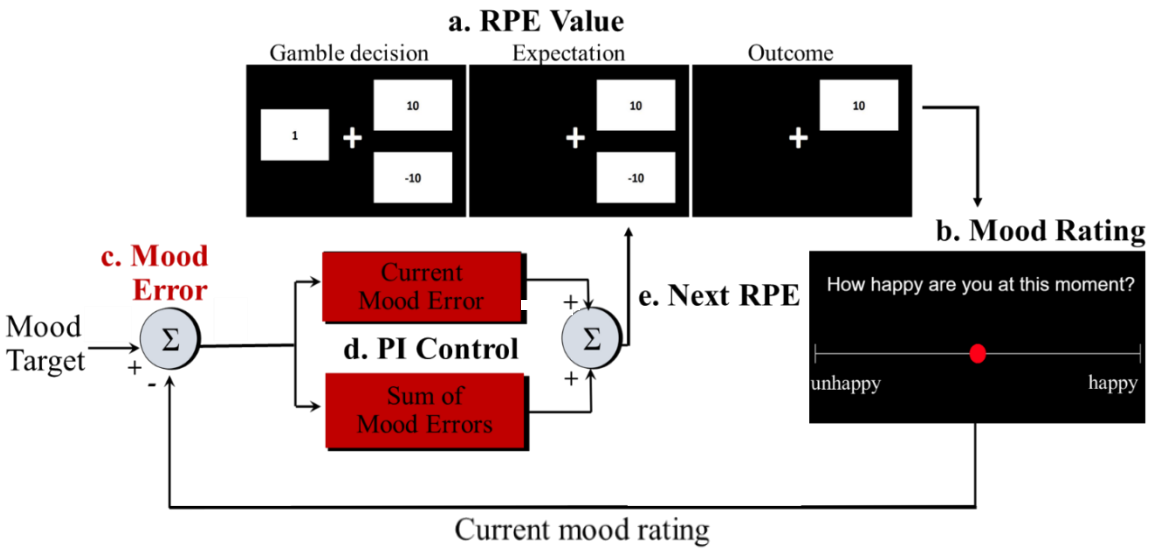

Figure 1: The Mood-Machine-Interface paradigm.
The experiment consisted of three blocks, each with 27 gambling trials. Values for the gamble decision **(a),** were presented for 3 seconds, followed by an anticipation phase of 4 seconds when the chosen option remained on the screen and then 1 second of outcome presentation. The control algorithm used the mood rating as an input **(b)**. The mood error computed from the current mood rating was used to determine the next RPE value using a proportional-integral (PI) algorithm **(c-e)**: a smaller change in mood towards the mood target resulted in an increase in RPE value on the following trials.

Validation of mood ratings

The first ratings (before participants completed any gambling trials) were strongly correlated with continuous measures of depressed mood (with MFQ measure in adolescent sample: CC=-0.62, p=2.62e-8, CI=[-0.75, -0.44]; with CESD in adult confirmatory sample: CC=-0.69, p=7.12e-13, CI=[-0.79, -0.56]) and in strong concordance with the gold standard psychiatric interview (KSADS) in distinguishing between patients with MDD and healthy volunteers (on a mood scale of 0 to 100, mean initial mood of healthy = 74, $sd_{healthy}$= 15, was significantly higher than the mean initial mood of MDD = 60, sd_{MDD}=19, t=-3.36, df=69, p=0.0012, Cohen's d effect size= 0.97). This is consistent with a previous study by Rutledge et al.[18] where the baseline mood, which was similarly rated, correlated with the continuous depression scores BDI and PHQ. This indicates that our mood assessment question is valid, though we have no way of indicating the general mood state, though we have no way of validating the repeated mood assessment during the task.

MMI induced shifts in mood

Our closed-loop MMI paradigm generated consistent mood transitions in both positive (upwards) and negative (downwards) directions compared to baseline mood (Figure 2). Most individuals showed a similar trend in mood values (n=72 adolescents; Figure 2a). The significant group effect can be appreciated by the individually normalized mood change, averaged across all participants (Figure 2b). We found a significant effect of time, both linear and squared, on mood in a linear mixed-effects model (t_{time}=-12.8 and t_{time^2}=11.6, p<0.001). The mean effect size (±SD) of mood modification within block was: 1.92 (±1.12) for increasing mood, -2.9 (±0.9) for decreasing mood and 3.02 (±0.92) for increasing mood again following the low mood phase. The RPE values

required to modify mood differed across individuals, reflecting individual differences in RPE responsiveness (Figure 2c). Moreover, it is apparent that these values kept enhancing over time (the slopes of RPE values were significantly larger than zero as tested with a t-test with $p<0.05$; the mean±SD RPE values across participants: block1=0.42±0.27; block 2=-0.53±0.36; block3=0.51±0.31). RPE values varied because they were individualized in real-time to participant's mood response dynamics (recalculated every 2-3 trials following each mood rating; see Figure S1 for a summary of all task parameters). This is important in face of previous works[4,19], which showed that positive mood increases the subjective value of rewards, and therefore reduces the experienced RPE. Our data provides another support for those findings, as the controller had to enhance RPE values over time to keep increasing/decreasing mood.

As a result, the closed-loop MMI could generate substantial mood changes across individuals, even though participants varied in their mood response and initial mood state (as evident in Figure 2a). We also included participants diagnosed with Major Depressive Disorder (MDD) in our lab-based discovery sample, in order to test the MMI paradigm on a wide range of initial moods. Even though depressed participants (n=43 of 72) started from significantly lower mood values ($p<0.001$), they showed similar effect sizes for both positive and negative mood changes ($p>0.05$ in a t-test; see Methods for participants' characteristics).

Most participants (90%, 65/72) were unaware of the manipulation, as we estimated using a follow-up questionnaire (in a scale between 0-3, the average rating for whether the task was unfair was 0.36±0.69SD with 7/72 subjects indicating 'agree' or 'strongly agree').

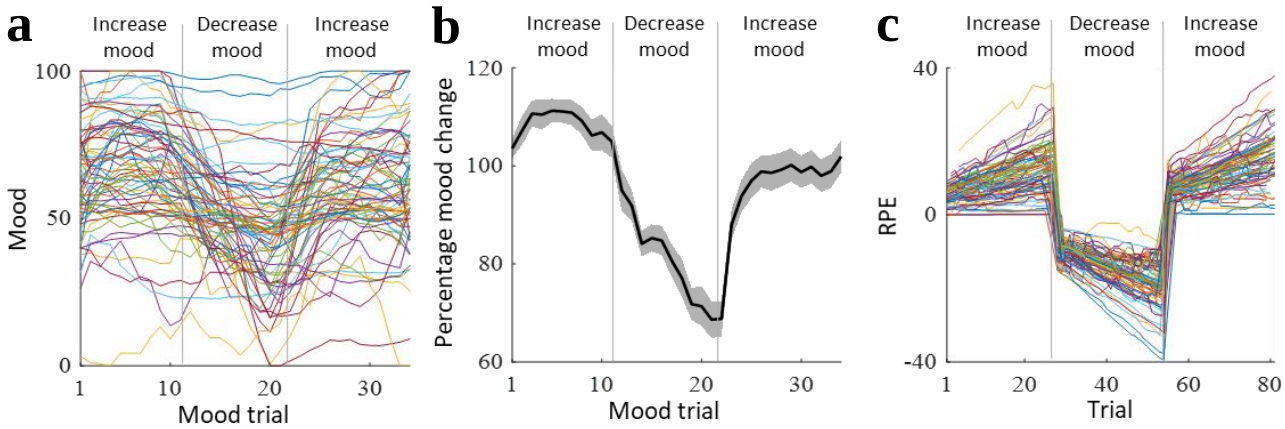

Figure 2: Mood manipulation to a pattern of high-low-high mood, by adjusting RPE values. **(a)** Mood ratings of each participant, shown by a separate curve (values are smoothed with a moving average with a 5 steps window). **(b)** The percentage of mood change from the first mood rating, averaged across all participants. Mood ratings were first individually transformed to a percentage from each participant's starting mood rating. Then all percentile mood curves were averaged (shaded region is standard error of the mean). See Figures S2-S3 for similar plots separated by healthy and depressed participants. **(c)** The RPE values generated during the task: individualized values were adjusted depending on each participant's mood response. Each curve is of a single participant and presents the 70% congruent RPE values determined by the control algorithm (not shown are the 30% incongruent RPE values that were included for maintaining surprise and so participants are unaware of the manipulation, see methods for further details).

The efficient generation of mood transitions also allowed us to uncover features of mood dynamics. In particular, we found that each participant's degree of mood change in the positive direction was highly correlated to the degree of change in the negative direction (correlation coefficient between positive and negative mood changes was 0.77, with $p<0.001$). This valence symmetry in mood change was also evident by participants returning at the end close to their initial mood value (the Intra-subject Correlation Coefficient measure, ICC, between first and last mood rating, was 0.64 with $p<0.001$).

MMI respective changes in behavior and neural activity

We find substantial differences in behavior between blocks (Figure S4): participants were more likely to gamble and were faster in making their choices under the positive context blocks (the first and third blocks), but were less likely to gamble and responded more slowly in the negative context block (the second block). These changes in behavior relate to previous accounts of the interaction between mood and gambling as well as other cognitive functions[20,21]. A linear mixed effects model showed a significant effect of both time and time squared on response times ($t_{time}=8.19$ and $t_{time^2}=-7.48$, $p<0.001$) and probability to gamble ($t_{time}=-6.76$ and $t_{time^2}=6.26$, $p<0.001$) and also a t-test between blocks was significant (for response times of block1 vs block2: p=2e-4, t=-5.6, and block3 vs block2: p=2.1e-5, t=-4.5; for gamble probability of block1 vs block2: p=1.4e-7, t=5.8, and block3 vs block 2: p=1.39e-5, t=4.6). Moreover, RPE encoding during the task (the correlation between brain activity and RPEs, evaluated using a parametric modulation with RPE values), changed depending on block and the reward-context. As shown in Figures S5a-c and S6, RPE encoding changed in regions which have been previously shown[22-24] to encode RPE values in a valence dependent manner (striatum and insula). Additionally, neural activity in the striatum during the period preceding mood rating was correlated with the subsequent mood rating value (Figure S5-d), in congruence with previous accounts of mood relations to striatal activity[2,25,26].

Replication of results

We replicated the MMI induction results of significant mood and behavioral changes, in an independent confirmatory sample (pre-registered analysis, see Methods for link). Specifically, the MMI was run in 80 adults using Amazon Mechanical Turk, and showed a significant effect of time and time squared on mood ($t_{time}=-11.4$ and $t_{time^2}=11.29$, $p<0.001$) and on decision behavior (RT: $t_{time}=7.4$ and $t_{time^2}=-8.5$, $p<0.001$, gambling probability: $t_{time}=-11.59$ and $t_{time^2}=11.77$, $p<0.001$). See Methods and Figure S7 for further information.

<u>*2. Modelling the impact of previous experiences on mood*</u>

The LTA mood model

Using the MMI, we then developed a computational model to investigate the influence of previous experiences on mood. To test this, we pit two models against each other: The *Standard Mood Model* as developed by Rutledge et al. in 2014[2], which has a recency weighting of the influence of previous events on mood (see equations 1-3); versus the new *Long-Term-Average (LTA) model*, where previous events have a primacy weighting and expectations are formed over the average of all previous outcomes (equations 4-6).

These models included two dynamic terms: the expectation term I and the RPE term (denoted *R*), which is the surprise relative to this expected value.

Specifically, in the standard model, the expectation at trial t is defined as

$$(8)\ E_t = \frac{H_t + L_t}{2}$$

Where H and L are the two gamble values and the RPE term, R, is defined as

$$(2)\ R_t = A_t - E_t$$

where A is the trial outcome value, and then the model for mood is:

$$(3)\ M_t = \epsilon_t + M_0 + \beta_C \sum_{k=0}^{t-1} \gamma^k C_{t-k}\, I(G_{t-k} = 0) + \beta_{\mathrm{E}} \sum_{k=0}^{t-1} \gamma^k E_{t-k}\, I(G_{t-k} = 1) + \beta_R \sum_{k=0}^{t-1} \gamma^k R_{t-k}\, I(G_{t-k} = 1)$$

where ϵ_t is a random noise variate with some unknown distribution (we may assume it to be Normal with mean 0 and standard deviation σ), M_0 is the participant's baseline mood, $\gamma \in (0,1)$ is an exponential discounting factor, C is the non-gamble certain amount, G is an indicator of gambling chosen trials with I being the trials index, β_C is the participant's sensitivity to certain rewards during non-gambling trials, β_E is the participant's sensitivity to expectation and β_R is the sensitivity to surprise during gambles.
For the LTA model, we define the expectation as:

$$(4)\ E_t = \frac{1}{t} \sum_{i=1}^{t} A_i$$

which is the average of all received outcomes A_i, over all trials from the first to current trial t.
We define RPE again as

$$(5)\ R_t = A_t - E_t$$

and then the long-term average model for the mood is:

$$(6)\ M_t = \epsilon_t + M_0 + \beta_E \sum_{k=0}^{t-1} \gamma^k E_{t-k} + \beta_R \sum_{k=0}^{t-1} \gamma^k R_{t-k}$$

where β_E and β_R are the participant's sensitivity to expectation and to surprise, respectively. Note that here we did not distinguish between gambling and non-gambling trials, which was another divergence from the standard model.

Our results showed that with the LTA model, the influence of earlier outcomes on mood is stronger than of recent outcomes (as demonstrated in figure 3a), in contrast to the strong weighting of more recent events in the standard model (Figure 3b). We also validated the primacy weighting of the LTA model with a t-test which showed that the weight of the first four events was significantly larger than of the last four events (with p=0.0036, t=8.34, CI=[0.08,0.18]). The stronger weight of earlier outcomes emerged from two separate aspects of the LTA model: First, that one's expectation for the next reward was based on the average of all previously received outcomes and, secondly, that mood was determined not only by the current expectation but also by past

expectations. Additionally, the LTA model showed that the extent of induced mood changes depended mainly on the expectation term, as the weights of the expectation term E were significantly higher than the weights of the R term (p=3.8e-4, t=3.7, CI=[0.0016, 0.0051]; see Figure S8 for the model parameters weights across all participants).

Model comparison

The LTA model outperformed the standard model when comparing the training error of the two models (Wilcoxon signed-rank test, p = 0.0204; training error for LTA: median MSE = 0.00262, IQR = 0.00496, training error for the standard model: median MSE = 0.00313, IQR = 0.00656). Moreover, the LTA model performed better in streaming predictions, where the ability of the model to predict within-participant subsequent mood ratings, not used in fitting the model, was evaluated. As depicted in Figure 3c, the LTA model was superior compared to the standard model specifically in accounting for transitions between moods; this was captured by comparing the error of streaming-prediction between the two alternative models (Wilcoxon signed-rank test, p< 0.0001; LTA streaming prediction error: median MSE = 0.0040, IQR = 0.00747, standard model streaming prediction error: median MSE = 0.01153, IQR = 0.0122). The streaming prediction by the LTA model is also exemplified in the rightmost panel of figure 3c, showing the alignment of predicted and rated mood data in a single participant.

Testing for other possible relations between previous experiences and mood

We then tested this finding by implementing an additional model, which allowed us to parameterize and fit the optimal distribution of weights over all previous events. This model was flexible in respect to which previous events are most influential on mood, as a result of adding the following parameters to the expectation term (E_t): a decay of the influence of previous outcomes w_i, the number of trials included in the history of outcomes, $\mathrm{t_{max}}$, and the decay constant $\tau \geq 0$, defined as follows:

$$(7)\ E_t = \frac{1}{t}\sum_{i=1}^{t_{max}} w_i A_i$$

$$(8)\ w_i = \frac{\tau^i}{\sum_{j=1}^{t_{max}} \tau^j}$$

The term for R and the mood model were defined similarly to the LTA model. This model fitted best with τ=0.8 and a decay of 0.01, which resulted again in early events having the strongest weights and weighting monotonically decreasing as events are more recent (see Methods for further details).

We also tested the LTA model against several additional models: the LTA model with separate gamble and non-gamble terms, LTA model without exponentially accumulated average, a model with both LTA and standard model terms, a model where expectation is the average of LTA expectation and standard model expectation, and different combinations of the standard model terms and the standard model with combined gamble and non-gamble values. The LTA model outperformed all these alternative models with a significantly lower MSE, as shown by a p<0.05 in the Wilcoxon signed-rank test (see Supplement for the description of all alternative models).

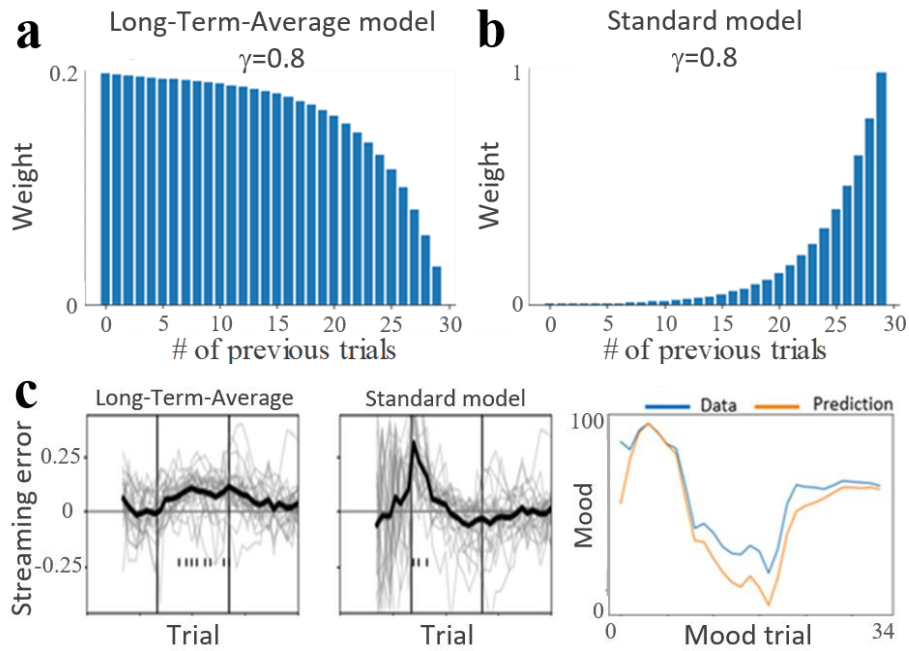

Figure 3: The Long-Term-Average model

Model weights of previous events: **(a)** In the LTA model earlier outcomes had a stronger influence on mood. **(b)** In contrast, in the standard model there was a stronger influence of most recent outcomes on the current mood. **(c)** Model comparison using streaming prediction performance, on mood ratings not used for fitting the model. The two left panels show the trial-wise error in predicting mood for all participants (average across participants is depicted by a black line), and the rightmost panel shows an example participant mood data aligned with predicted mood values using the LTA model.

Replication of results

We replicated these computational findings on the independent confirmatory online sample, where we showed again that the LTA model outperformed the standard model in training error and streaming prediction of the MMI data (tested using a Wilcoxon signed-rank test with $p<0.001$ for both tests; training error for LTA: median MSE = 0.00627, IQR = 0.00891; training error for the standard model: median MSE = 0.00724, IQR = 0.01164; streaming prediction error for LTA: median MSE = 0.01015, IQR = 0.01386; streaming prediction error for the standard model: median MSE = 0.01609, IQR = 0.02169).

3. Neural correlates of the effect of experience on mood

The neural level model

Finally, we sought evidence of a neural basis for the relationship between mood and early events, as defined by the expectation term of the LTA model. To this end, we searched for neural activity correlated with the subject level weights from the LTA model. Specifically, we ran a whole-brain, group-level ANOVA (3dMVM in AFNI[27]) with the weights of the LTA model as between-participant covariates of neural activity (each participant's neural activity was represented by a single whole-brain image of activation across all trials). We examined neural activity related to three different aspects of the task: activation during the pre-mood rating period (the question "How

happy are you at this moment?" is presented but the mood-rating option is not available yet), mood rating encoding (mood as a parametric regressor) during the pre-mood rating period, and stimulus based RPE encoding (RPE as a parametric regressor) during outcome period. Since these are three separate tests, we added a Bonferroni correction to the multiple comparison correction, which resulted in a p-value threshold of 0.005/3= 0.0017.

Neural activity related to the effect of past events on mood

We found a significant positive correlation between the weight of the expectation term (β_E) from the LTA model and neural activity during the pre-mood rating period, focused in the ACC (figure 4a). This implies that this region can regulate mood changes by mediating the influence of previous outcomes on mood. We found no relations of neural activity at this pre-mood rating time with other LTA model parameters. For task RPE encoding, we found a correlation with the mood intercept term (namely a positive correlation between mood intercept and RPE encoding in the right insula: 208 voxels, t=4.1, p=0.0017; and a few smaller clusters in the brain stem and paracentral lobule). We found no significant relations of neural encoding of mood values with the LTA model parameters. Moreover, we found no neural relations with the weights of the standard model; we show that the relation with the LTA expectation term is significantly stronger than with the standard model term, by contrasting between the neural relations with each of the models (Figure 4b).

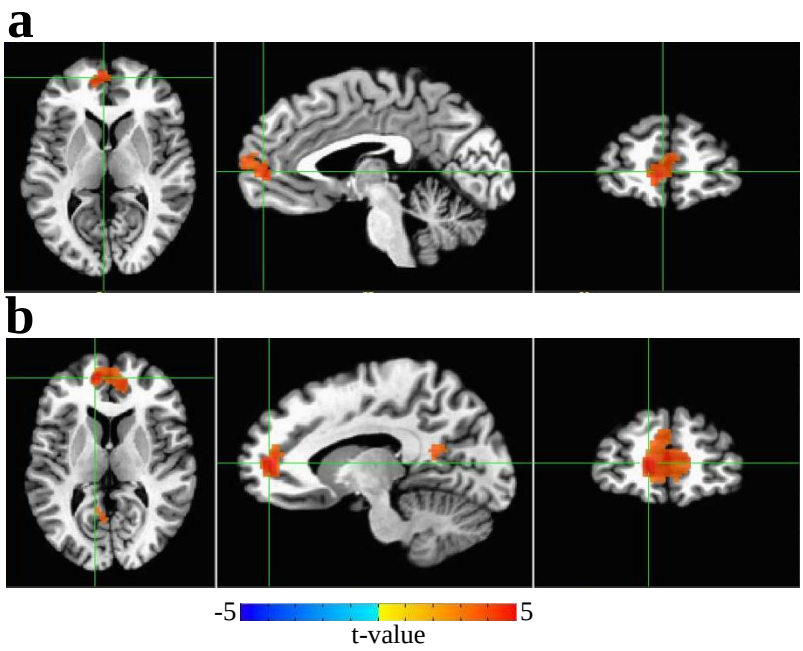

Figure 4: **(a)** Neural correlates of the individual LTA model expectation weights. The expectation term, i.e., the average of previous outcomes, β_E, estimated from mood ratings correlated at the whole brain level to individual neural activation preceding the mood rating in the ACC (corrected for multiple comparisons and a Bonferroni correction for the three 3dMVM models we ran: ACC cluster: 132 voxels, t=4.28, thresholded at p≤0.0017, [-3,52,6]). **(b)** A contrast between the relation of brain activity with each of the models' expectation term weights, $\beta_E(\text{LTA}) > \beta_E(\text{Standard Model})$. This contrast showed a significant difference between the models in the relation to brain signals (ACC cluster: 529 voxels, t=5, [-11,49,9]; corrected for multiple comparisons and a Bonferroni correction for the three 3dMVM models we ran with a threshold of p<0.0017).

Discussion

We provide a neuro-computational account of how mood is influenced by the history of previous experiences. We addressed the question – which past experiences shape mood the most? To answer this question, we developed the MMI, an individualized reward-based paradigm which can generate substantial mood transitions, in both positive and negative directions, across healthy, depressed, adolescent and adult participants. The MMI can uniquely adjust in real-time the reward and punishment intensities to ensure these are continuously influential on mood. We used this paradigm to show that mood changes are dominated by early experiences which have a large and persistent effect on mood also when environments change. Moreover, we show that this relationship is mediated by neural activity in the ACC.

Generating parametric mood transitions

We developed the MMI paradigm to shift mood parametrically to allow quantification and modeling of mood changes. While prior mood induction approaches are qualitative[28-30], the MMI generates a quantitative manipulation, which can adjust the reward and punishment intensities to influence mood. Such a closed-loop strategy is commonly used in engineering for controlling complex systems (whether it is room temperature or car velocity), and it can also be found endogenously, in controlling hormone levels, for example[31-33]. The importance of using this system-level approach is in overcoming the strong between-individual variability in mood ratings at baseline (as evident by a range of values between 3.1-100%, SD=18.9%, χ^2=90.5, $p<0.001$, and as we can also expect across different ages[34]), and the differential responsiveness of participants to RPEs (as depicted by Figure 2c). Moreover, to be able to determine which parameters influence mood changes, it was also important that the MMI manipulation was not prone to demand effects, where participants respond to satisfy the experimental goal. Some of the best-known mood manipulations face this limitation, for example by asking individuals to imagine situations that evoke certain feelings[30]. We can conclude that overall participants were not aware of the MMI strategy, as they 1) rated the task as fair (90% rated as fair, see methods) and 2) continued making nonoptimal gamble choices across the task. Since the MMI algorithm compensated for low mood levels or low mood responsiveness, we could shift the mood of both healthy and depressed individuals, and therefore find general characteristics of mood changes. We did not focus in this work on differences between healthy and depressed mood dynamics, although future extensions of the MMI could address such questions as well. Moreover, we present one possible implementation of the paradigm, but other approaches—including open loop (i.e., using predefined stimuli intensities)—might be useful too. Other types of reward stimuli, for example, could be parametrized to modify mood and characterize relations between mood and other reward modalities (e.g., social rewards). Additional methods for tracking mood could be implemented in the future, improving the validity of subjective mood ratings. Whilst the efficiency of the MMI paradigm is useful for studying mood shifts in a parametric way in the laboratory, it also opens the way for a new type of devices with a clinical utility. It is conceivable that this approach could be extended to shift mood in depression, but this requires a study of the effects of the MMI on mood outside the trial setting[35].

The importance of past events in forming mood changes

We developed a computational model to test the influence of previous events on mood changes. Previous models suggested that mood depends primarily on recent events[2-5]. In those models, mood either followed recent RPEs or recent outcomes (the finding of a stronger effect of outcomes was considered as possibly due to the task having no clear expectation phase). However, as the LTA model suggests, early events have a strong influence on mood, and moreover, the influence of more recent events decreases monotonically over time. Also, when fitting a group of models which allowed the influence of all previous events to vary (by parametrizing the discounting of previous outcomes within the expectation term), we found a similarly strong relation between early events and mood. Specifically, the LTA model gave higher weight to early events by considering expectation as a long-term average of all previous outcomes. It outperformed the other tested models using two different performance criteria (i.e., goodness of fit as well as streaming prediction of mood using only a subset of previous mood ratings). These results were also replicated with a preregistered analysis of data collected online from a different age group (adult participants). This encapsulates in quantitative terms long-standing intuitions about mood, namely that it integrates over the history of rewarding or punishing events[1]. Moreover, from a clinical perspective, the LTA model formulates how early negative events can have a long-lasting effect on mood[36-38]. Importantly, it shows that the longer one is in any given reward (or punishment) context, the less influence each additional event could have on mood. Therefore, this model holds implications for how mood disorders may be conceptualized.

Neural mediation of the influence of previous outcomes on mood

We then linked these findings to neural activity and showed that the relationship between mood and previous events has a neural underpinning. Specifically, the weight of the expectation term, which captures the influence of early experiences on mood, was positively correlated to activation of the ACC (during the presentation of the mood question, prior to mood rating phase). This region has been often indicated as involved in mood regulation[10-13,39]. Moreover, previous studies showed that ACC activation mediated decision making relative to previous outcomes[40-43]. Therefore, our results formulize a specific role for the ACC in emotion regulation – where it is mediating between past early experiences and mood changes. This study also demonstrates the strength of integrating model-driven parameters with fMRI data[44,45]: The model-based analysis enabled us to move beyond localizing task-related signals to identifying a neural relation to the complex experience-mood interaction (otherwise not detectable by typical behavioral measures). Nonetheless, fMRI imaging data faces several analytic limitations. For example, there is a limited capability to distinguish between neural encoding of temporally overlapping processes. Running the MMI in the future with imaging techniques of higher temporal resolution might enable a better distinction between spatial and temporal encoding of outcome versus mood, for example. Moreover, detection of relevant activations of small regions, such as the Amygdala, is limited in our fMRI recording and analysis. These results suggest a quantitative measure of the ACC mediation between past experiences and mood, which opens the door for neural-based characterization of individual mood states.

Our mood would likely become low after a job interview, if we did not get the answer right to an important question. As the LTA model suggests, getting that question wrong at the beginning of the interview will have a more persistent negative impact on our mood, which is harder to override. Similarly, this computational notion relates to why we consider it important to start off with the right foot, why we hope to make a good first impression, how we can decide to stop watching a movie after seeing just the beginning, and in this very context- why we tend to invest so much effort in choosing the right first sentence for a manuscript. Moreover, there is empirical evidence that it is feasible to use just a first instance of an interaction to draw inferences about people's character[46,47]. In the clinical context, our model suggests that the potential to elevate individual's mood might be constrained by early experiences. Moreover, it shows how over time new events can have a smaller influence relative to past experiences. Thus, to increase mood we would need to adjust the intensity of rewards based on individual history. As we show, a closed-loop approach can be useful in generating such individualized events that are potent enough to change mood. The paradigm, model, and results we present here are important for understanding mood changes and suggest new directions for mood interventions.

Methods

1. Participants

Lab-based discovery sample: 80 adolescents (70.5% females, mean age = 15.4 (± 1.4 SD) years, 43 participants diagnosed with Major Depressive Disorder, mean depression score MFQ=5.8 (± 6 SD)) participated in the study for monetary compensation. In this sample, the study included completing the MMI task in an fMRI scanner and answering debriefing questions after completion. Participants were compensated for scanning and they also received a separate bonus between $5 to $35, in proportion to the points they earned during the task. Participants were screened for eligibility and inclusion criteria were the capability to be scanned in the MRI scanner and not satisfying diagnosis criteria for disorders other than depression according to DSM5. Overall, five participants were excluded from analyses due to incomplete data files, and three additional participants were excluded due to repeatedly rating a single fixed mood value for an entire block of the task, reaching a final sample of n=72. Every participant received the same scripted instructions and provided informed consent to a protocol approved by the NIH Institutional Review Board.

Online confirmatory sample: 80 participants recruited from Amazon Mechanical Turk (MTurk) system completed the MMI task (41.2% females, mean age = 37.7 ± 11.2). We analyzed the data in a pre-registered framework provided by OSF, to confirm our lab-based results. The MTurk Worker ID was used to distribute a compensation for completing the task of $8 and a separate task bonus between $1 to $6, according to the points gained during the task. The study population was ordinary, non-selected adults of 18 year of age or older. Participants were not screened for eligibility, all individuals living in the US and who wanting to participate were able to do so. Participants were restricted to doing the task just once. Three participants were excluded from analyses due to an error in the task script where mood ratings were inconsistently spread along the 3 blocks.

Age	
Average	15.49
SD	1.48
Sex	
Male	17
Female	55
Race	
Asian	8
Black/African American	9
Indian/Alaska Native	1
Multiple Race	10
Hawaiian/Pacific Islander	0
Unknown	2
White	42
MFQ Score	
Average	5.81
SD	5.98
Diagnosis	
Healthy Volunteer	29
MDD	43

Table 1: Lab-based discovery sample demographics (n=72)

Age	
Average	37.76
SD	11.23
Sex	
Male	46
Female	34
CESD score	
Average	12.99
SD	12

Table 2: Online confirmatory sample demographics (n=80)

2.The Mood-Machine-Interface (MMI) paradigm

Task design:
This task was designed to manipulate mood to target values by modifying Reward Prediction Error values in real-time based on participant mood ratings. The task consisted of 3 blocks, each comprised of 27 gambling trials and 11-12 mood ratings. Blocks were separated by a short break, typically less than 1 min long. In the lab-based task (done in the fMRI scanner) each block lasted about 8 min (24 min for the whole task), while the online task on MTurk had shorter inter-trial intervals and therefore took 15 min to complete (5 min for each block; all other characteristics remained unchanged). Before starting the task, participants were instructed how to rate their mood and to make gamble choices as quickly as possible. They were not told whether potential gamble outcomes were or were not equally likely. Moreover, participants were told that the payment they will receive at the end would be proportional to the number of points they gain during the task.

RPE values – the input to the task from the controller:
During the task participants received different RPEs using a simple gambling task. In each trial participants made a choice whether to gamble or not. An RPE value was generated per trial by first presenting the participant with 3 potential monetary outcomes: one is certain and two are potential outcomes from a gamble. Outcomes for chosen gambles are revealed after a brief delay period. The RPE was constructed by the difference between the received reward and the trial's expectation (the mean of the two gamble values). Each trial consisted of three phases: (1) Choice: 3 seconds during which the participants presses left to get the certain value or right to gamble between two values (using a four-button response device); (2) Anticipation: only the chosen certain value or the two gamble options remain on the screen, for 4 seconds. (3) Outcome: A feedback of the outcome value is presented for 1 second, followed by an inter-trial-interval of 2-8 seconds. Participants completed 81 trials, divided into three blocks of 27 trials each.

Mood ratings – the output measure from the task:
Participants also rated their mood after every 2-3 gambling trials. The mood rating consisted of two separate phases: (1) Presentation of the mood question: "How happy are you at this moment?" for a random duration between 2.5-4 seconds. (2) Rating mood by moving a curser along a scale labeled "unhappy" on the left end and "happy" on the right end. Each rating started from the initial location of the center of the scale, and participants had a time window of 4 seconds to rate their mood. Participants were instructed to move the cursor by holding down continuously the left or right button. The final cursor position was taken as their mood rating. Each rating was followed by a 2-8 seconds jittered interval.

Real-time modification of RPE values:
We based the mood shifting algorithm on the finding that momentary mood reflects cumulative RPEs. If this relationship is monotonic, accumulating positive RPE values will increase mood and accumulating negative RPE values will lower mood. We therefore developed an algorithm which recalculates in real-time the RPE value for the next trial that would be predicted to achieve a desired mood change. This strategy of a closed-loop control is formally derived from control of non-linear systems in engineering. It is used to bring a system to a new state and maintain it (whether the state is room temperature, velocity in car cruise control, angle in flight control etc.). A similar circuit is found also endogenously (most hormonal systems are negative feedback

loop systems) and in modern medical therapeutics. The development of the exact algorithm parameters and gains of control included a trial and error relative to mood changes observed in a group of pilot participants. During this process we searched for the optimal step of RPE manipulation, that is minimal on the one hand but still shows a sufficient mood response on the other.

Specifically, in each iteration of mood rating, the current mood at time t (M) was compared to the block mood target value (M_T), being 100% in the first and third blocks and 0% during the second mood-decreasing block. To bring the mood value as close as possible to the target value M_T, the algorithm aimed at minimizing the error between the rated mood and the target mood value (M_E):

$$(1)\quad M_E(t) = \begin{cases} \dfrac{M_T - M(t)}{M_{max} - M_{min}}, & M_T = 100\%\ of\ mood\ scale \\ \dfrac{M(t) - M_T}{M_{max} - M_{min}}, & M_T = 0\%\ of\ mood\ scale \end{cases}$$

where the resulting M_E is a value ranging between [0-1], M_T is the mood target value and $M_{max} - M_{min}$ is the range of possible mood values which is between 0-100%.
The M_E value was then mapped to a change in the task RPE value, using a PI controller. This control algorithm enabled the calculation of an efficient updating required for the next trials, by using both a proportional and an integral error term. Importantly, the integral error term (the sum of previous M_E values), enhances the RPE modification when mood remains in the same distance from the Target value (the proportional term in such a case would provide repeatedly the same RPE value, which can limit the efficiently of the manipulation). This term was reset at the beginning of each block to avoid a carryover of mood errors across the positive and negative mood changes.
According to M_E the next 2-3 trials RPE value (t+1) was recalculated, such that the larger this error, the stronger was the modification of the RPE value, as follows:

$$(2)\ RPE(t+1) = \begin{cases} RPE_{baseline} * M_E(\mathrm{t}) + \sum_{1}^{t} M_E\,, & IsCongruent\ \leq 70\% \\ RPE_{baseline} * M_E(\mathrm{t}) + \dfrac{\sum_1^t M_E}{3}\,, & IsCongruent\ > 70\% \end{cases}$$

where $RPE_{baseline}$ is a fixed value that was pre-calibrated while developing the task to the value of 14 points (so RPE values change in a minimal yet sufficient step size), and ISCongruent is a randomly selection such that 70% of trials are congruent with the control algorithm and 30% are incongruent - providing an RPE value with the opposite sign to the block context (negative during the first and third mood-increasing blocks and positive during the second mood-decreasing block). These incongruent RPE values were set to be smaller in amplitude, by decreasing the integral cumulative error value by a factor of 3 in equation (2), and an additional reduction in the outcome value as shown in equation (5)). As a result, for example during the first mood increase block, on average the size of these incongruent RPE values was -1.5±0.8SD. Moreover, to maintain unpredictability also the location of the higher gamble value H(t) and the lower value L(t), was randomly assigned to appear at either the upper or lower gamble value squares. These two gamble

values were recalculated per each of the 2-3 trials until the next mood rating, using the same new RPE(t+1), as follows:

$$L(t+1) = H(t+1) - RPE(t+1) \tag{3}$$

where H(t+1) was a value randomly reassigned from a list ranging between [-1.5,14] with a step size of 0.2.
Then the certain value (CR) which appeared on the left side, was derived according to the two gamble values (while ensuring that the certain value cannot provide a reward larger than 2 points):

(4)

$$CR(t+1) = \begin{cases} \dfrac{L(t+1) + H(t+1)}{2}, & CR(t+1) \leq 2 \\ \dfrac{L(t+1)}{2}, & CR(t+1) > 2 \end{cases}$$

And last, which outcome value (A) is going to be received when a gamble is chosen in the next trial, was assigned relative to the block number and whether this is a congruent trial or not:

(5)

$$A\,(t+1) = \begin{cases} H(t+1), & block\ =\ 1{,}3;\ ISCongruent\ \leq 70\% \\ L(t+1), & block\ =\ 2;\ ISCongruent\ \leq 70\% \\ \dfrac{L(t+1)}{4}, & block\ =\ 1{,}3;\ ISCongruent\ > 70\% \\ \dfrac{H(t+1)}{4}, & block\ =\ 2;\ ISCongruent\ > 70\% \end{cases}$$

Hence a feedback loop was created between the input to the system and the output measure (see Figure1 for illustration of the whole circuit). This cycle continued throughout the task, with each new mood rating used to update the reward values for the next series of 2-3 trials. This setup generated personalized “reward environments” in each block, as the task stimuli were calculated online and were not pre-determined like in conventional paradigms.

Importantly, most participants were unaware of the manipulation. After doing the task participants rated their agreement with the statement “The task was unfair” (possible answers were 0-1-2-3, where 0=strongly disagree and 3=strongly agree). The average rating value across participants was 0.36 (SD=0.69), while 6/72 participants rated the value 2 (agree), and a single participant from the sample rated the value of 3 (strongly agree). Other questions in the debriefing form validated that participants had no technical difficulties in rating their mood nor significant issues with the scanning experience.

3. The LTA model

Model comparison:
We started by comparing two alternative models, the standard mood model and the LTA model. The criteria for comparison between the models were:
(1) the model fit, by testing whether the LTA model has statistically significantly lower training error on the 3-block data. It should be also considered that the LTA model has one fewer parameter, which indicates it provides a better description of mood fluctuations in our task.
(2) Streaming prediction (MSE). We tested whether the LTA model has statistically lower streaming prediction error on the 3-block data. The streaming prediction error of a model is defined as the average error when predicting the t-th mood rating using a model fit on the first t-1 mood ratings.
Both comparisons were tested using the Wilcoxon signed-rank test, with the one-sided null hypothesis MSE_LTA >= MSE_Standard model. We chose a one-sided null because the conservative null would be that the new approach is equal or worse than the existing approach. The Wilcoxon signed-rank test tests the null hypothesis that two related paired samples come from the same distribution. In particular, it tests whether the distribution of the differences x - y is symmetric about zero.

Model fitting:
All models were fit using Tensorflow. We use the following notations to define the models:
Let s = 1, … , n_s index the subjects from 1 to n_s, the number of subjects;
t=1,2,…, the trial of the game (1 is the first round);
$C_s(t)$, the certain, non-gamble value, which was possible to choose on this trial;
$H_s(t)$, the maximal gambling amount for the *t*th trial;
$L_s(t)$, the minimal gambling amount for the *t*th trial;
$G_s(t)$, whether or not the subject took the gambling option, 1=Gamble, 0=Certain choice;
$A_s(t)$, the actual value the subject received at the end of the trial;
$M_s(t)$, subject's mood rating.
Mood ratings are rescaled between 0 and 100. This value is missing in trials where the participant was not prompted to rate mood.
The general form of the models (both LTA and Standard model) can be described by the following parametric model:

$$(6)\quad \widehat{M_s}(t) = \mu_s + \sum_{v=1}^{p} \beta_{v,s} \sum_{l=0}^{t} \gamma_s^k X_{v,s}(t-l)$$

where *s* indexes the subject, *t* is the trial, *v* is one of *p* time-varying variables, μ_s is the subject-specific baseline mood, and $\beta_{v,s}$ are subject-specific coefficients for each time-varying variable (note that we constrain $\beta_1,\ldots,\beta_3 \geq 0$).
For instance, in the standard model, the model has p=3 time-varying variables:
X_1 is the certain amount (C) in rounds where the subject did not gamble,

$$(7)\quad X_{1,s}(t) = C_s(t)\, I(G_s(t) = 0)$$

X_2 is the expected gamble (E_t) in rounds where the subject did gamble,

$$(8)\ \ X_{2,s}(t) = \frac{H_s(t) + L_s(t)}{2} I(G_s(t) = 1)$$

X_3 is the reward prediction error (RPE) in rounds where the subject did gamble (R_t),

$$(9)\ \ X_{3,s}(t) = I(G_s(t) = 1)\left(A_s(t) - X_{2,s}(t)\right)$$

Meanwhile, in the Long-Term Average (LTA) model, there are p=2 time-varying variables:
X_1 is the average of the previous actual amounts (E_t in the main text),

$$(10)\ \ X_{1,s}(t) = \frac{1}{t}\sum_{i=1}^{t} A_s(i)$$

X_2 is the reward prediction error (RPE) with respect to the average actual amount (R_t),

$$(11)\ X_{3,s}(t) = A_s(t) - X_{1,s}(t)$$

And we again constrain $\beta_1, \beta_2 \geq 0$.
In order to facilitate optimization, we further re-parameterized γ_s by defining

$$(12)\ \gamma_s = \frac{1}{1 + \exp(-\xi_s)}$$

so that ξ_s is an unbounded real number.

We found that the use of group-level regularization was necessary in order to stabilize the estimated coefficients. This took the form of imposing a variance penalty on ξ and a variance penalty on each coefficient β_v. The empirical variance is defined as

$$(13)\ Var(\xi) = \frac{1}{n_s}\sum_{s=1}^{n_s}\left(\xi_s - \bar{\xi}\right)^2$$

where $\bar{\xi}$ is the group mean:

$$(14)\ \bar{\xi} = \frac{1}{n_s}\sum_{s=1}^{n_s}\xi_s$$

Likewise, we define $Var(X_v)$ for v = 1,…, p.
The objective function is therefore

$$(15)\ \text{minimize} \sum_{s=1}^{n_s}\sum_{t\in T}\left(\widehat{M_s}(t) - M_s(t)\right)^2 + \lambda_\xi Var(\xi) + \lambda_\beta \sum_{v=1}^{p} Var(\beta_v)$$

where T is the set of trials where $M_s(t)$ was defined. Optionally, one can also discard the first few trials in T to minimize window effects (we require $t \geq 11$).

Model development process:
For both the LTA and standard model, we chose the group regularization constants by creating simulated datasets with realistic parameters and selecting the regularization parameters from a grid that had the best performance. The grid consisted of powers of 10 from 0.001 to 10000. For both models, the regularization parameters with the best performance in recovering the simulation ground truth were $\lambda_\xi = 10$ and $\lambda_\beta = 100$. The LTA model represented the best fit also among the family of models with a variable form of the distribution of the weights of previous events on mood (see Figure S9 for examples of possible distributions that were realized via this generic model).
The model development used only an initial sample of 40 randomly selected participants (from the discovery sample of 72 adolescents). Then results were then confirmed on 60 random participants of this adolescent sample. Then, we held another separate replication analysis on a sample of 80 adults doing the task online (the entire analytic plan for this replication was pre-registered).

3. Neural activity analysis

fRMI data acquisition:
Participants in the adolescent discovery sample, performed the task while scanning in a General Electric (Waukesha, WI, USA) Signa 3-Tesla magnet. Task stimuli were displayed via back-projection from a head-coil mounted mirror to a screen at the foot of the scanner bed. Foam padding was used to constrain head movement. Behavioral data were recorded using a hand-held FORP response box. Forty-seven oblique axial slices (3.0-mm thickness) per volume were obtained using a T2-weighted echo-planar sequence (echo time, 30 ms; flip angle, 75°; 64 × 64 matrix; field of view, 240 mm; in-plane resolution, 2.5 mm × 2.5 mm; repetition time was 2000 ms). To improve the localization of activations, a high-resolution structural image was also collected from each participant during the same scanning session using a T1-weighted standardized magnetization prepared spoiled gradient recalled echo sequence with the following parameters: 176 1-mm axial slices; repetition time, 8100 ms; echo time, 32 ms; flip angle, 7°; 256 × 256 matrix; field of view, 256 mm; in-plane resolution, 0.86 mm × 0.86 mm; NEX, 1; bandwidth, 25 kHz.

Data preprocessing:
Analysis of fMRI data was performed using Analysis of Functional and Neural Images (AFNI) software version 2.56 b[48]. Standard pre-processing of EPI data included slice-time correction, motion correction, spatial smoothing with a 6-mm full width half-maximum Gaussian smoothing kernel, normalization into Talairach space and a 3D non-linear registration. Each participant's data were transformed to a percent signal change using the voxel-wise time series mean blood oxygen level dependent (BOLD) activity. Images were analyzed using a mixed event-related and block design. Time series were analyzed using multiple regression[49], where the entire trial was modeled using a gamma-variate basis function. The model included the following task phases: Choice time: an up to 3 seconds interval, from the presentation of the 3 monetary values to the button press, left for the certain amount or right to gamble. This phase was covered by three regressors in the model, a separate one for each block. Anticipation time: the interval from making the choice to gamble to receiving the gamble outcome. Outcome time: a 1 second interval where the received outcome is shown; split to 3 regressors, a separate one for each block. The Mood Question time: a variable

interval between 2.5-4 seconds, when the mood question is presented although the option to rate mood is still disabled. Rating time: a 4 seconds interval when participants rate their mood.
The model also included six nuisance variables modeling the effects of residual translational (motion in the x, y and z planes), rotational motion (roll, pitch and yaw) and a regressor for baseline plus slow drift effect, modeled with polynomials (baseline being defined as the non-modeled phases of the task).
To this model we also added two parametric modulators: trial-wise RPE values (RPE was set to zero when the certain value is chosen instead of gambling; modulation is of the respective outcome times) and mood ratings (modulating respectively the times preceding the mood rating, i.e., the mood question phase).
Echo-planar images (EPI) were visually inspected to confirm image quality and minimal movement.
Statistical significance at the group level, was determined using 3dClustSim (the latest acceptable version in AFNI with an ACF model) which generated a corrected to $p<0.05$ voxel-wise significance threshold of $P < 0.005$ and a minimal cluster size of 100 voxels. Region-of-interest (ROI) approach was used to determine the individual average RPE encoding in the striatum (coordinates were derived from the Talairach atlas). First, the unthresholded individual whole-brain RPE encoding maps were masked for the striatum region (bilateral putamen) and then the mean RPE encoding value across all voxel values included in that region was extracted.

4. Statistical testing of MMI effects

We applied a linear mixed effects model to estimate the significance of mood, behavior and neural activity changes over time. This model enabled the estimation of the across-participants significance of mood change while controlling for the within-participant variability in mood change slopes and intercepts, defined as a random effect. Specifically, the independent variable was the response variable of interest (mood, behavioral measure, ROI neural activation) and the dependent variable time (trial index) and time squared, with the two different time variables considered as random effects. For example, the model for estimating mood change was formalized as:

$$\text{(6)}\quad \text{Mood} \sim \text{time} + \text{time2} + (\text{time} + \text{time2} \mid \text{participant})$$

This model was applied across the three blocks of the experiment and P-values were considered significant at $p<0.05$.

1 Nettle, D. & Bateson, M. The evolutionary origins of mood and its disorders. *Curr Biol* **22**, R712-721, doi:10.1016/j.cub.2012.06.020 (2012).

2 Rutledge, R. B., Skandali, N., Dayan, P. & Dolan, R. J. A computational and neural model of momentary subjective well-being. *Proc Natl Acad Sci U S A* **111**, 12252-12257, doi:10.1073/pnas.1407535111 (2014).

3 Katsimerou, C., Redi, J. A. & Heynderickx, I. A Computational Model for Mood Recognition. *Lect Notes Comput Sc* **8538**, 122-133 (2014).

4 Eldar, E., Rutledge, R. B., Dolan, R. J. & Niv, Y. Mood as Representation of Momentum. *Trends in Cognitive Sciences* **20**, 15-24, doi:10.1016/j.tics.2015.07.010 (2016).

5 Vinckier, F., Rigoux, L., Oudiette, D. & Pessiglione, M. Neuro-computational account of how mood fluctuations arise and affect decision making. *Nat Commun* **9**, 1708, doi:10.1038/s41467-018-03774-z (2018).

6 Fox, N. A., Henderson, H. A., Marshall, P. J., Nichols, K. E. & Ghera, M. M. Behavioral inhibition: Linking biology and behavior within a developmental framework. *Annu Rev Psychol* **56**, 235-262, doi:10.1146/annurev.psych.55.090902.141532 (2005).

7 Rothbard, N. P. & Wilk, S. L. Waking up on the Right or Wrong Side of the Bed: Start-of-Workday Mood, Work Events, Employee Affect, and Performance. *Acad Manage J* **54**, 959-980, doi:10.5465/amj.2007.0056 (2011).

8 Dvir, Y., Ford, J. D., Hill, M. & Frazier, J. A. Childhood Maltreatment, Emotional Dysregulation, and Psychiatric Comorbidities. *Harvard Rev Psychiat* **22**, 149-161, doi:10.1097/Hrp.0000000000000014 (2014).

9 Vidal-Ribas, P. *et al.* Bidirectional Associations Between Stress and Reward Processing in Children and Adolescents: A Longitudinal Neuroimaging Study. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **4**, 893-901, doi:10.1016/j.bpsc.2019.05.012 (2019).

10 Bush, G., Luu, P. & Posner, M. I. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci* **4**, 215-222 (2000).

11 Stevens, F. L., Hurley, R. A. & Taber, K. H. Anterior cingulate cortex: unique role in cognition and emotion. *J Neuropsychiatry Clin Neurosci* **23**, 121-125, doi:10.1176/appi.neuropsych.23.2.121 (2011).

12 Etkin, A., Egner, T. & Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn Sci* **15**, 85-93, doi:10.1016/j.tics.2010.11.004 (2011).

13 Rudebeck, P. H. *et al.* A role for primate subgenual cingulate cortex in sustaining autonomic arousal. *Proc Natl Acad Sci U S A* **111**, 5391-5396, doi:10.1073/pnas.1317695111 (2014).

14 Gu, X. S. Computational Models of Emotion. *Routledge Hbk Philos*, 436-451 (2019).

15 Pessoa, L. Embracing integration and complexity: placing emotion within a science of brain and behaviour. *Cogn Emot* **33**, 55-60, doi:10.1080/02699931.2018.1520079 (2019).

16 Paulus, M. P., Huys, Q. J. & Maia, T. V. A Roadmap for the Development of Applied Computational Psychiatry. *Biol Psychiatry Cogn Neurosci Neuroimaging* **1**, 386-392, doi:10.1016/j.bpsc.2016.05.001 (2016).

17 Huys, Q. J. M., Maia, T. V. & Paulus, M. P. Computational Psychiatry: From Mechanistic Insights to the Development of New Treatments. *Biol Psychiatry Cogn Neurosci Neuroimaging* **1**, 382-385, doi:10.1016/j.bpsc.2016.08.001 (2016).

18 Rutledge, R. B. *et al.* Association of Neural and Emotional Impacts of Reward Prediction Errors With Major Depression. *JAMA Psychiatry* **74**, 790-797, doi:10.1001/jamapsychiatry.2017.1713 (2017).
19 Eldar, E. & Niv, Y. Interaction between emotional state and learning underlies mood instability. *Nat Commun* **6**, 6149, doi:10.1038/ncomms7149 (2015).
20 Otto, A. R., Fleming, S. M. & Glimcher, P. W. Unexpected but Incidental Positive Outcomes Predict Real-World Gambling. *Psychol Sci* **27**, 299-311, doi:10.1177/0956797615618366 (2016).
21 Okon-Singer, H., Hendler, T., Pessoa, L. & Shackman, A. J. The neurobiology of emotion-cognition interactions: fundamental questions and strategies for future research. *Front Hum Neurosci* **9**, doi:10.3389/fnhum.2015.00058 (2015).
22 Palminteri, S. *et al.* Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. *Neuron* **76**, 998-1009, doi:10.1016/j.neuron.2012.10.017 (2012).
23 Garrison, J., Erdeniz, B. & Done, J. Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev* **37**, 1297-1310, doi:10.1016/j.neubiorev.2013.03.023 (2013).
24 Keren, H. *et al.* Is the encoding of Reward Prediction Error reliable during development? *Neuroimage* **178**, 266-276, doi:10.1016/j.neuroimage.2018.05.039 (2018).
25 Heller, A. S. *et al.* Sustained striatal activity predicts eudaimonic well-being and cortisol output. *Psychol Sci* **24**, 2191-2200, doi:10.1177/0956797613490744 (2013).
26 Admon, R. & Pizzagalli, D. A. Corticostriatal pathways contribute to the natural time course of positive mood. *Nat Commun* **6**, 10065, doi:10.1038/ncomms10065 (2015).
27 Chen, G., Adleman, N. E., Saad, Z. S., Leibenluft, E. & Cox, R. W. Applications of multivariate modeling to neuroimaging group analysis: a comprehensive alternative to univariate general linear model. *Neuroimage* **99**, 571-588, doi:10.1016/j.neuroimage.2014.06.027 (2014).
28 Westermann, R., Spies, K., Stahl, G. & Hesse, F. W. Relative effectiveness and validity of mood induction procedures: A meta-analysis. *Eur J Soc Psychol* **26**, 557-580, doi:Doi 10.1002/(Sici)1099-0992(199607)26:4<557::Aid-Ejsp769>3.0.Co;2-4 (1996).
29 Zhang, X., Yu, H. W. & Barrett, L. F. How does this make you feel? A comparison of four affect induction procedures. *Front Psychol* **5**, 689, doi:10.3389/fpsyg.2014.00689 (2014).
30 Mikita, N., Mehta, M. A., Zelaya, F. O. & Stringaris, A. Using arterial spin labeling to examine mood states in youth. *Brain Behav* **5**, doi:10.1002/brb3.339 (2015).
31 Levine, S. W. *Control Handbook: Control System Advanced Methods, 2nd Edition*. (CRC Press, 2011).
32 Bally, L. *et al.* Closed-Loop Insulin Delivery for Glycemic Control in Noncritical Care. *New Engl J Med* **379**, 547-556, doi:10.1056/NEJMoa1805233 (2018).
33 Uluseker, C. *et al.* A closed-loop multi-level model of glucose homeostasis. *Plos One* **13**, doi:ARTN e0190627 (2018).
34 Heller, A. S. & Casey, B. J. The neurodynamics of emotion: delineating typical and atypical emotional processes during adolescence. *Dev Sci* **19**, 3-18, doi:10.1111/desc.12373 (2016).

35 Heller, A. S. *et al.* The Neurodynamics of Affect in the Laboratory Predicts Persistence of Real-World Emotional Responses. *J Neurosci* **35**, 10503-10509, doi:10.1523/JNEUROSCI.0569-15.2015 (2015).
36 Douglas, K. R. *et al.* Adverse childhood events as risk factors for substance dependence: Partial mediation by mood and anxiety disorders. *Addict Behav* **35**, 7-13, doi:10.1016/j.addbeh.2009.07.004 (2010).
37 Raby, K. L., Roisman, G. I., Fraley, R. C. & Simpson, J. A. The Enduring Predictive Significance of Early Maternal Sensitivity: Social and Academic Competence Through Age 32Years. *Child Dev* **86**, 695-708, doi:10.1111/cdev.12325 (2015).
38 Lewis-Morrarty, E. *et al.* Infant Attachment Security and Early Childhood Behavioral Inhibition Interact to Predict Adolescent Social Anxiety Symptoms. *Child Dev* **86**, 598-613, doi:10.1111/cdev.12336 (2015).
39 Etkin, A., Buchel, C. & Gross, J. J. The neural bases of emotion regulation. *Nat Rev Neurosci* **16**, 693-700, doi:10.1038/nrn4044 (2015).
40 Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. Learning the value of information in an uncertain world. *Nat Neurosci* **10**, 1214-1221, doi:10.1038/nn1954 (2007).
41 Scholl, J. *et al.* The Good, the Bad, and the Irrelevant: Neural Mechanisms of Learning Real and Hypothetical Rewards and Effort. **35**, 11233-11251, doi:10.1523/JNEUROSCI.0396-15.2015 %J The Journal of Neuroscience (2015).
42 Wittmann, M. K. *et al.* Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nat Commun* **7**, 12327, doi:10.1038/ncomms12327 (2016).
43 Scholl, J. *et al.* Excitation and inhibition in anterior cingulate predict use of past experiences. *Elife* **6**, doi:10.7554/eLife.20365 (2017).
44 O'Doherty, J. P., Hampton, A. & Kim, H. Model-based fMRI and its application to reward learning and decision making. *Ann N Y Acad Sci* **1104**, 35-53, doi:10.1196/annals.1390.022 (2007).
45 Cohen, J. D. *et al.* Computational approaches to fMRI analysis. *Nat Neurosci* **20**, 304-313, doi:10.1038/nn.4499 (2017).
46 Ambady, N. & Rosenthal, R. Thin Slices of Expressive Behavior as Predictors of Interpersonal Consequences - a Metaanalysis. *Psychol Bull* **111**, 256-274, doi:Doi 10.1037/0033-2909.111.2.256 (1992).
47 Houser M. L. , H. S. M., Furler L. A. . Predicting Relational Outcomes: An Investigation of Thin Slice Judgments in Speed Dating. *Human Communication* **102**, 69 – 81 (2007).
48 Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* **29**, 162-173, doi:10.1006/cbmr.1996.0014 (1996).
49 Neter, J., Wasserman, W. & Kutner, M. H. *Applied linear statistical models : regression, analysis of variance, and experimental designs*. 3rd edn, (Irwin, 1990).

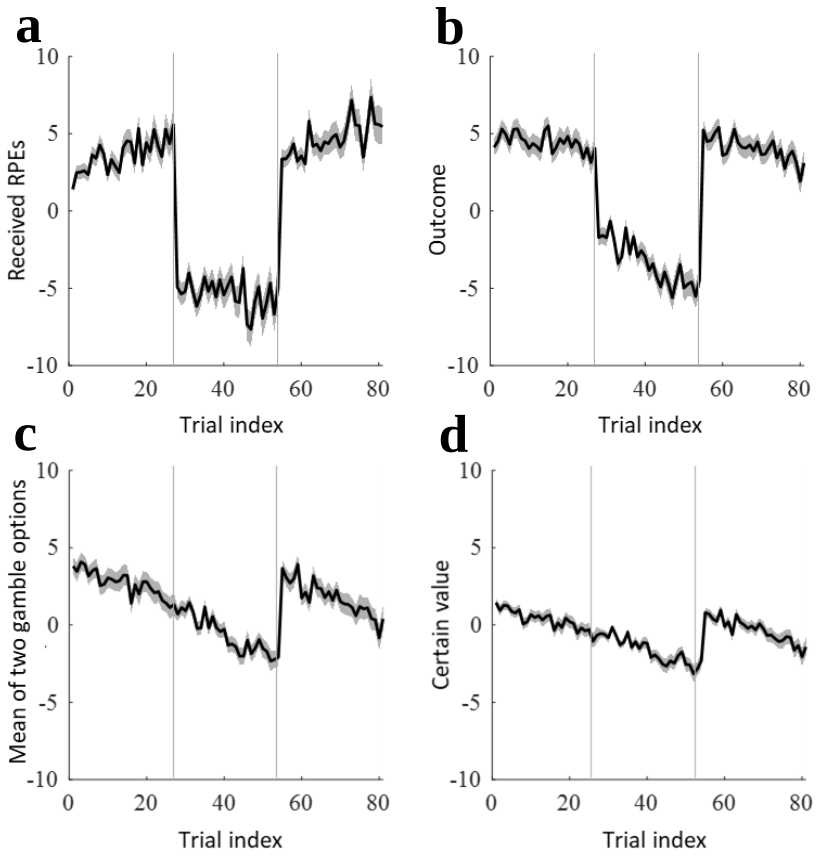

Figure S1: A summary of MMI task parameters across blocks. (a) RPE values which participants received along the task: in this presentation only trials where a gamble was chosen have an RPE value, trials where the non-gamble value was chosen have an RPE of zero. Moreover, both congruent (70% of trials) RPEs and incongruent values are presented. This presentation therefore only shows RPE values which were experienced in practice by participants, which differs from the control-intended RPE values (although the trend across and within blocks is overall similar). (b) Outcome values received during the task, of both gamble and non-gamble trials. (c) The mean between the two gamble options (the expectation EV term in the standard model by Rutledge et al., 2014). (d) The certain non-gamble values along the task. All curves are the average across all participants with SEM as the shaded region (values are derived from the lab-based exploratory sample of 72 adolescents). Vertical lines depict the three task blocks.

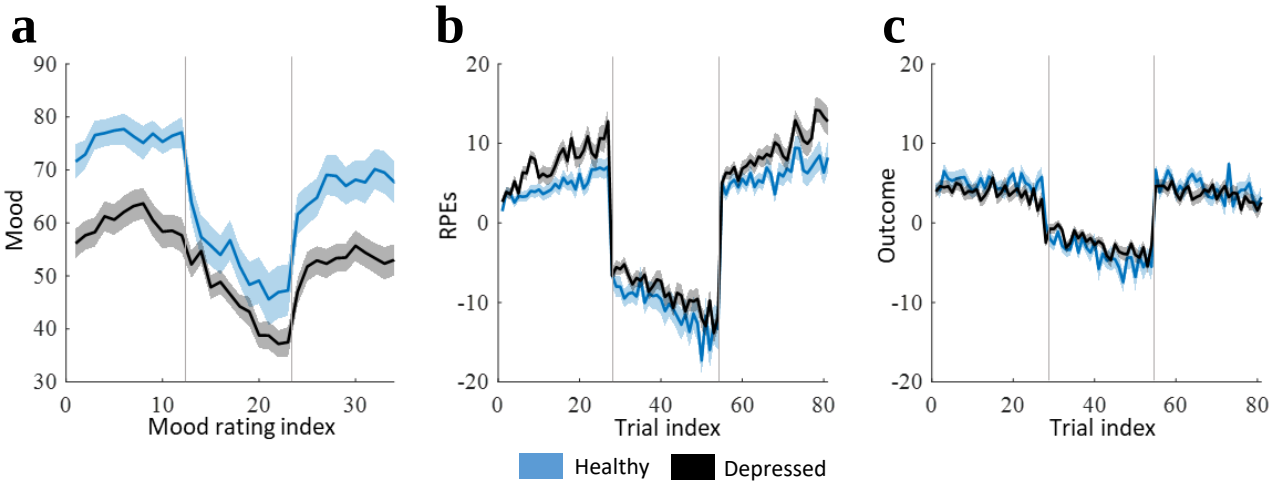

Figure S2: A summary of MMI mood ratings and task parameters across blocks, split between healthy and depressed participants. Depressed are participants that meet criteria for history or ongoing Major Depressive Disorder (according to DSM5, n=43 of the 72 lab-based adolescent participants). (a) Mood ratings, averaged across healthy (blue) and depressed (black). (b) RPE values as intended by the controller: in this presentation all trials RPE values, both congruent (70% of trials) RPEs and incongruent values, are presented. (c) Outcome values received during the task, resulting from both gamble and non-gamble trials. All curves are the average across all participants with SEM as the shaded region. All curves are the average across all participants with SEM as the shaded region. Vertical lines depict the three task blocks.

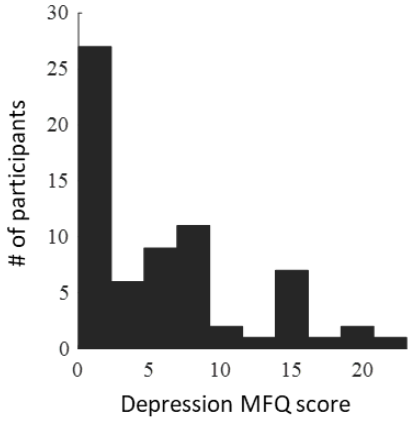

Figure S3: Distribution of the Depression score (MFQ), across all 72 lab-based adolescent participants.

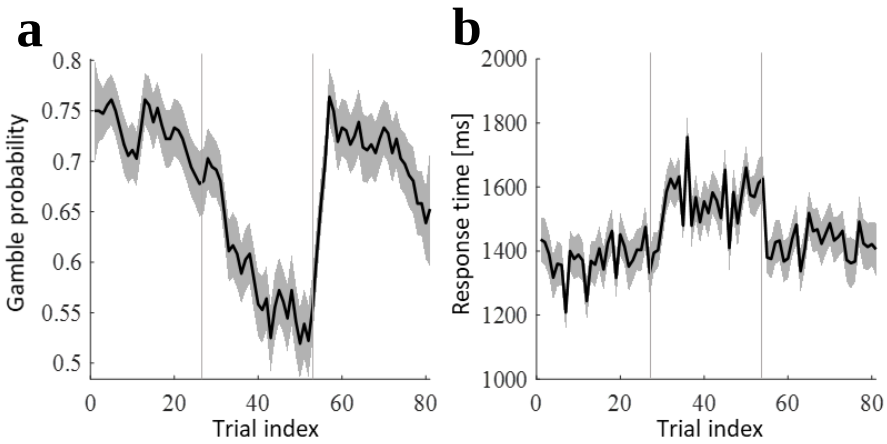

Figure S4: The MMI paradigm shows also substantial changes in behavior across blocks. (a) Gambling probability, averaged across all participants. Probability is calculated by a moving average of the binary vector of the choice to gamble or not (denoted by a value of 1 or 0, respectively). (b) Response times, that is the time from presentation of the three monetary values until participant is making the choice of whether to gamble or not (by pressing left or right). All curves are the average across all participants with SEM as the shaded region. Vertical lines depict the three task blocks.

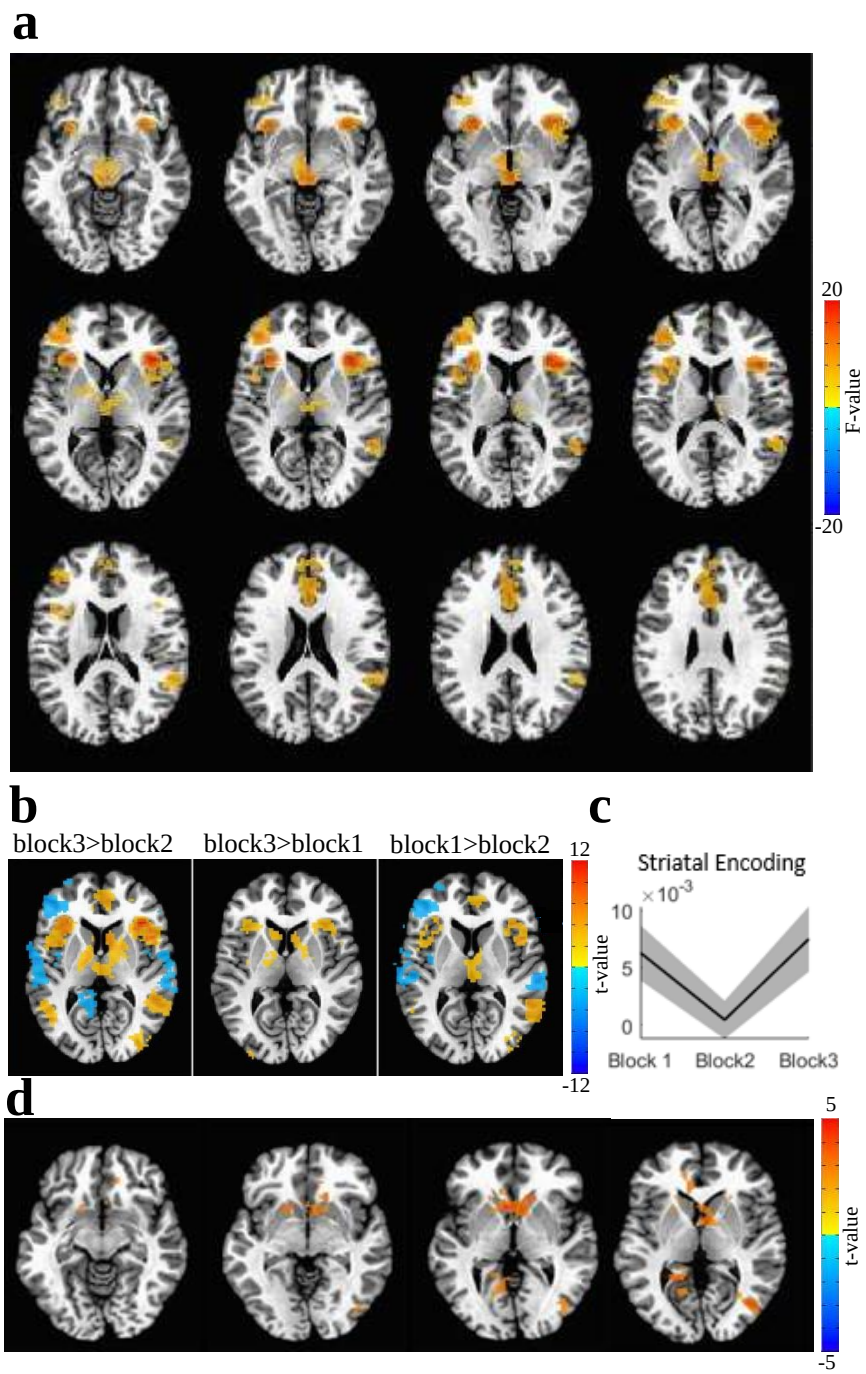

Figure S5: Respective changes in RPE and mood encoding during the MMI.
A significant block effect on RPE encoding. First, RPE encoding was derived using the trial RPE values as the parametric linear modulator of the BOLD signals of each block, during the outcome time (RPE=0 when the certain value is chosen rather than the gamble). Then a group level ANOVA model (3dMVM in AFNI) was run with the block index as the explanatory variable (ACC: F=11.16, 629 voxels, Insula: 437 voxels, F=19.88; cluster extends also over striatum: 358 voxels, F=15.11) (b) Comparison between RPE encoding across blocks at the whole brain group level. Leftmost image presents the contrast between RPE encoding during the third block of second mood increase versus the mood decrease block; the middle image contrasts between the two mood increase blocks; and the rightmost image between the first mood increase block and the mood decrease block. (c) Comparison between RPE encoding across the blocks at

the individual ROI level (n=67). Average activation in the anatomical mask of the striatum (left and right putamen, coordinates derived from the Talairach atlas). Shaded regions present SEM. (d) Mood encoding at the whole brain level: Mood encoding values are derived using the mood ratings as the parametric linear modulator of the BOLD signals during the pre-rating interval (at this interval, which lasts between 2.5-4 seconds, participants are presented with the mood question, but cannot rate their mood yet). Cluster peaks in the NAcc and covers the ACC and Caudate (337 voxels, t= 4.96).

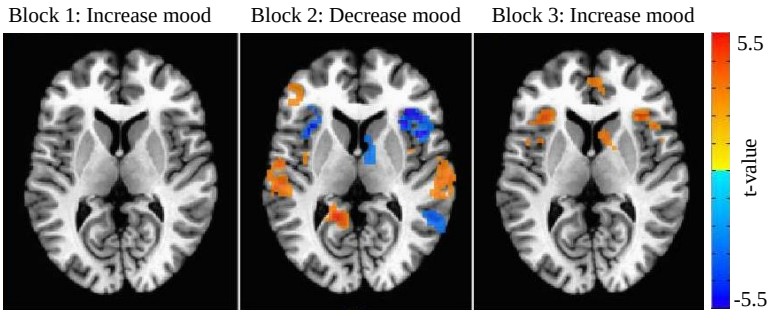

Figure S6: RPE encoding maps per block. RPE encoding is derived using the trial RPE values as the parametric linear modulator of the BOLD signals of each block, during the outcome time (RPE=0 when the certain value is chosen rather than the gamble). These images are corrected for multiple comparisons (see Methods).

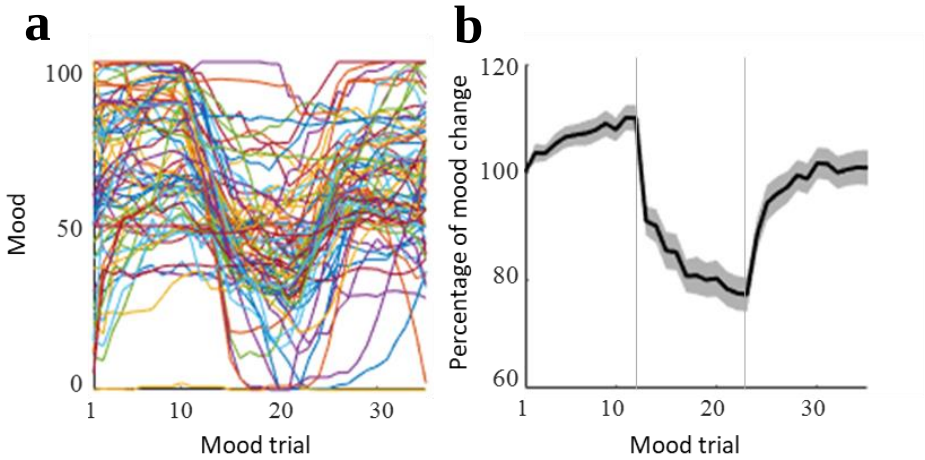

Figure S7: Replication of the MMI paradigm on the confirmatory adult sample, using Mechanical Turk online framework. (a) Mood ratings, each curve is of one participant (values are smoothed with a moving average of the last 5 steps). (b) The percentage of mood change (from the first mood rating), averaged across all participants. First, mood ratings are transformed to a percentage from individual starting mood rating. Then all percentile mood curves are averaged (shaded region is standard error of the mean). Vertical lines depict the three task blocks.

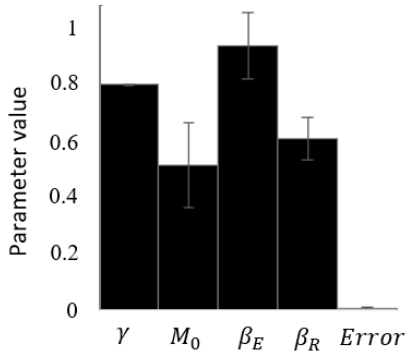

Figure S8: The LTA model parameters across subjects (error bars are standard deviation).

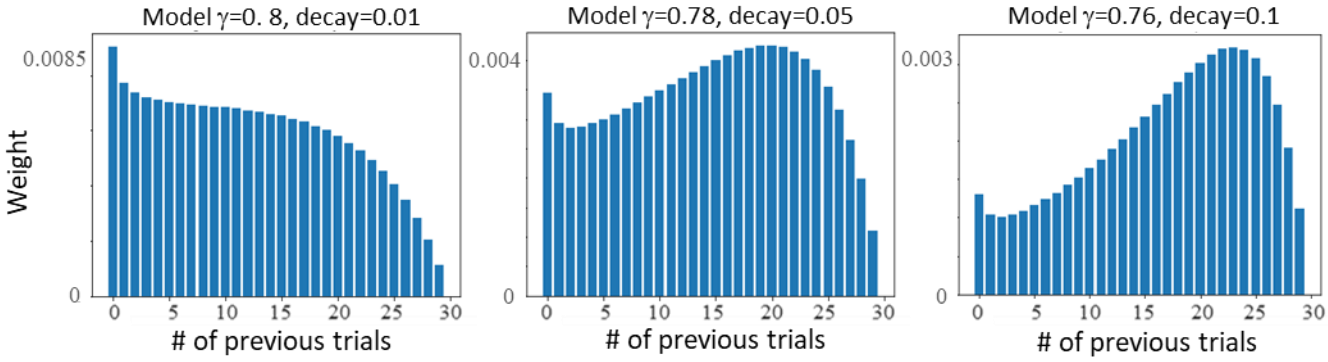

Figure S9: Model weights of previous events: Examples for various distributions of the influence of previous events on mood; these are achieved by varying the expectation term via the decay and the maximal number of previous trials parameters. A value of LTA gamma = 0.8 and a decay of 0.01, was the best fit, which results in a primacy weighting of the influence of previous events on mood.

Contrast	Region	Voxels	x_{peak}	y_{peak}	z_{peak}	Beta peak	Statistic (df =71)
	Group ANOVA model for the block effect on RPE encoding						F value
	Right Anterior Cingulate	629	-3.8	-21.2	23.8		11.16
	Right Insula	437	-33.8	-21.2	3.8		19.87
	Right Thalamus	358	-3.8	26.2	-3.8		15.17
	Left Middle Frontal Gyrus	312	41.2	-36.2	13.8		12.88
	Left Inferior Frontal Gyrus	308	26.2	-13.8	-8.8		17.61
	Right Precentral Gyrus	157	-33.8	-18.8	36.2		10.60
	Right Superior Temporal Gyrus	135	-53.8	48.8	11.2		11.20
	Right Posterior Cingulate	83	-21.2	56.2	6.2		10.23
	Right Cingulate Gyrus	80	-21.2	-1.2	28.8		8.48
	Right Caudate	59	-16.2	23.8	21.2		12.35
	Right Superior Temporal Gyrus	53	-48.8	23.8	-1.2		13.41
	Group one sample t-test of RPE modulation activation values						t value
block 3> block 2	Right Superior Temporal Gyrus (and right and left Insula, Caudate, Putamen)	4664	-51.2	-11.2	1.2	0.07	5.08
	Right Precentral Gyrus	1016	-23.8	18.8	66.2	-0.08	-3.12
	Left Middle Frontal Gyrus	504	46.2	-43.8	11.2	-0.13	-4.48
	Left Postcentral Gyrus	490	56.2	16.2	13.8	-0.04	-4.68
	Left Superior Frontal Gyrus	451	8.8	-36.2	51.2	-0.05	-2.92
	Right Superior Temporal Gyrus	407	-58.8	48.8	16.2	0.03	5.13
	Right Cingulate Gyrus	385	-1.2	21.2	38.8	0.04	4.41
	Left Precuneus	349	31.2	73.8	41.2	-0.06	-6.64
	Right Precentral Gyrus	291	-56.2	3.8	6.2	-0.05	-4.55
	Right Middle Occipital Gyrus	285	-36.2	83.8	13.8	0.05	3.85
	Left Posterior Cingulate	230	3.8	38.8	11.2	-0.07	-3.85
	Right Precuneus	164	-13.8	68.8	43.8	0.02	3.47
	Right Middle Temporal Gyrus	154	-51.2	16.2	-6.2	0.03	5.47
	Right Middle Frontal Gyrus	140	-31.2	-53.8	21.2	0.05	4.37
	Left Middle Frontal Gyrus	121	36.2	3.8	43.8	0.02	6.75
	Left Cerebelllar Tonsil	115	11.2	51.2	-36.2	0.03	3.86
	Left Middle Occipital Gyrus	115	41.2	66.2	8.8	0.02	3.05
	Left Inferior Temporal Gyrus	94	53.8	51.2	-11.2	-0.07	-4.76
	Left Superior Frontal Gyrus	92	31.2	-46.2	26.2	0.03	2.97
	Left Middle Frontal Gyrus	81	51.2	-8.8	36.2	-0.04	-3.38
	Right Subcallosal Gyrus	75	-21.2	-3.8	-13.8	-0.05	-3.90
	Left Parahippocampal Gyrus	70	13.8	16.2	-16.2	-0.04	-3.05
	Left Inferior Parietal Lobule	65	58.8	43.8	23.8	0.04	3.56
	Right Inferior Occipital Gyrus	56	-43.8	73.8	-6.2	0.02	3.26
	Right Insula	52	-26.2	31.2	21.2	-0.02	-3.64
	Left Inferior Occipital Gyrus	50	31.2	86.2	-6.2	0.05	3.37
	Right Superior Parietal Lobule	50	-16.2	56.2	63.8	-0.02	-3.21

block 3> block 1	Right Caudate	1190	-6.2	-8.8	13.8	0.26	2.98
	Right Cingulate Gyrus	282	-1.2	33.8	31.2	0.14	2.98
	Right Precentral Gyrus	192	-41.2	-3.8	36.2	0.13	2.95
	Right Angular Gyrus	136	-31.2	53.8	38.8	0.19	3.01
	Left Precentral Gyrus	133	36.2	-1.2	36.2	0.13	3.34
	Right Superior Frontal Gyrus	102	-36.2	-41.2	31.2	0.15	2.99
	Right Middle Temporal Gyrus	97	-53.8	18.8	-8.8	0.08	3.51
	Left Cerebellar Tonsil	61	3.8	58.8	-31.2	0.21	2.91
	Left Inferior Occipital Gyrus	60	28.8	91.2	-8.8	0.20	3.55
block 1> block 2	Left Middle Frontal Gyrus	2670	46.2	-28.8	28.8	-0.60	-3.77
	Right Medial Frontal Gyrus	622	-1.2	1.2	56.2	0.22	4.31
	Right Inferior Frontal Gyrus	478	-43.8	-18.8	-8.8	0.12	3.67
	Right Precentral Gyrus	422	-46.2	13.8	31.2	-0.11	-3.34
	Right Middle Occipital Gyrus	411	-31.2	83.8	18.8	0.29	3.14
	Right Parahippocampal Gyrus	401	-16.2	23.8	-13.8	-0.20	-3.34
	Right Thalamus (and right Caudate)	372	-1.2	28.8	6.2	0.47	3.01
	Left Insula (and left Putamen)	310	41.2	-6.2	3.8	0.14	4.77
	Left Superior Parietal Lobule	258	26.2	73.8	43.8	-0.21	-4.86
	Left Parahippocampal Gyrus	159	28.8	26.2	-13.8	-0.17	-3.58
	Right Superior Temporal Gyrus	156	-58.8	3.8	-1.2	-0.14	-5.86
	Right Cingulate Gyrus	141	-3.8	51.2	26.2	-0.11	-3.04
	Left Superior Temporal Gyrus	130	58.8	46.2	21.2	0.25	3.71
	Right Precuneus	127	-43.8	68.8	38.8	-0.32	-4.10
	Right Precentral Gyrus	111	-51.2	3.8	41.2	0.09	3.84
	Right Cingulate Gyrus	94	-23.8	26.2	28.8	-0.04	-3.04
	Right Middle Frontal Gyrus	79	-26.2	-13.8	43.8	-0.23	-3.19
	Right Cingulate Gyrus	77	-1.2	21.2	38.8	0.09	3.95
	Right Inferior Frontal Gyrus	62	-38.8	-36.2	-1.2	-0.14	-3.08
	Left Paracentral Lobule	57	3.8	38.8	51.2	-0.10	-3.05
Group one sample ttest of mood modulation activation values							
	Left Postcentral Gyrus	515	-51.2	-13.8	53.8	0.03	2.95
	Right Anterior Cingulate (covers right and left Caudate)	352	16	31	-11	0.01	3.12
	Left Parahippocampal Gyrus	159	-9	-49	4	0.01	3.06
	Left Middle Occipital Gyrus	147	44	-74	6	0.008	3.39
	Right Postcentral Gyrus	124	36	-26	46	-0.01	-4.36

Table-s1. Whole brain level results. Values depicted black are corrected for FWE to $p<0.05$ at the whole brain level, with $p<0.005$ and minimal cluster size of 102 voxels. Values in grey are showing clusters >50 voxels and below the 102 voxels threshold. Coordinates are according to the Talairach-Toutnoux Atlas.