



Published in final edited form as:

*Neuroimage*. 2018 September ; 178: 266–276. doi:10.1016/j.neuroimage.2018.05.039.

## Is the encoding of Reward Prediction Error reliable during development?

Hanna Keren<sup>a,\*</sup>, Gang Chen<sup>b</sup>, Brenda Benson<sup>c</sup>, Monique Ernst<sup>d</sup>, Ellen Leibenluft<sup>e</sup>, Nathan A. Fox<sup>f</sup>, Daniel S. Pine<sup>c</sup>, Argyris Stringaris<sup>a</sup>

<sup>a</sup>Mood Brain and Development Unit, Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA

<sup>b</sup>Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA

<sup>c</sup>Section on Development and Affective Neuroscience, Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA

<sup>d</sup>Section on the Neurobiology of Fear and Anxiety, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA

<sup>e</sup>Section on Mood Dysregulation and Neuroscience, Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA

<sup>f</sup>Department of Human Development and Quantitative Methodology, University of Maryland, College Park, 9000, Rockville Pike, Bethesda, MD, USA

### Abstract

Reward Prediction Errors (RPEs), defined as the difference between the expected and received outcomes, are integral to reinforcement learning models and play an important role in development and psychopathology. In humans, RPE encoding can be estimated using fMRI recordings, however, a basic measurement property of RPE signals, their test-retest reliability across different time scales, remains an open question. In this paper, we examine the 3-month and 3-year reliability of RPE encoding in youth (mean age at baseline =  $10.6 \pm 0.3$  years), a period of developmental transitions in reward processing. We show that RPE encoding is differentially distributed between the positive values being encoded predominantly in the striatum and negative RPEs primarily encoded in the insula. The encoding of negative RPE values is highly reliable in the right insula, across both the long and the short time intervals. Insula reliability for RPE encoding is the most robust finding, while other regions, such as the striatum, are less consistent. Striatal reliability appeared significant as well once covarying for factors, which were possibly confounding the signal to noise ratio. By contrast, task activation during feedback in the striatum

\*Corresponding author. Hanna.Keren@nih.gov (H. Keren).

Conflicts of interest  
None.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.neuroimage.2018.05.039>.

is highly reliable across both time intervals. These results demonstrate the valence-dependent differential encoding of RPE signals between the insula and striatum, and the consistency of RPE signals or lack thereof, during childhood and into adolescence. Characterizing the regions where the RPE signal in BOLD fMRI is a reliable marker is key for estimating reward-processing alterations in longitudinal designs, such as developmental or treatment studies.

## Keywords

Reliability; Prediction Error; Reward; Development; Adolescence; fMRI

---

## Introduction

Encoding of Reward Prediction Error (RPE), the difference between the expected and received reward value, can be estimated using fMRI in humans and its alterations are thought to be involved in developmental and psychopathological processes. Yet, a basic measurement property of the RPE, its test-retest reliability, remains to be established. In this paper, we examine RPE reliability in young people (mean age at baseline =  $10.6 \pm 0.3$  years), across 3 months and across 3 years.

The RPE is an important learning signal that helps organisms to maximize wins and minimize losses through value computations (Schultz 1998, 2006, 2013, 2016, 2017; Sutton and Barto, 1998; Rolls et al., 2008; Diederer et al., 2016; Schultz et al., 2017). An RPE arises whenever the outcome of an action is different from what was predicted. In situations where the outcome is better than predicted, the RPE is positive and is associated with an increased likelihood of the behavior that led to the reward to re-occur. If the reward falls below what was predicted, a negative RPE occurs along with a decrease of the likelihood of repeating the same behavior. The RPE has been extensively studied in animals and found to be encoded by mesolimbic dopaminergic neurons (Olds and Milner, 1954; Corbett and Wise, 1980; Schultz et al., 1993; Bayer and Glimcher, 2005; Pan et al., 2005; Cohen et al., 2012; Averbeck and Costa, 2017).

Functional magnetic resonance imaging (fMRI) has made it possible to localize the RPE encoding in the human brain. A recent meta-analysis of such studies indicates that the RPE is encoded in a distributed network, positive RPEs seem to be primarily represented in the striatum and negative RPEs are primarily encoded in the insula (Liu et al., 2007; Palminteri et al., 2012; Garrison et al., 2013). This has opened the way for examining the role of RPEs in sensitive stages of development, such as adolescence, and in psychopathology. Developmentally, increasing evidence suggests that reward sensitivity increases in adolescents, and, indeed, positive RPE signals in the striatum and negative RPE signals in the insula, seem to peak in adolescents compared to children or adults (Cohen et al., 2010; Somerville and Casey, 2010; Lamm et al., 2014; Smith et al., 2014; Braams et al., 2015). In psychopathology, alterations in the processing of RPEs have been proposed to be centrally involved in a range of psychiatric disorders (Murray et al., 2008; Moutoussis et al., 2015; Radua et al., 2015; Ubl et al., 2015; Schmidt et al., 2016; Rothkirch et al., 2017; White et al., 2017), including depression and schizophrenia.

Yet, despite the importance of measuring RPE in fMRI, a fundamental psychometric property remains unexamined, namely its test-retest reliability across different time scales. Test-retest reliability studies are critical for distinguishing true signal changes from other sources of measurement instability (Maitra et al., 2002; Bennett and Miller, 2010, Raemaekers et al., 2012; Herting et al., 2017). Evaluating change over time is critical for understanding developmental processes as well as psychopathology. If RPE fMRI signal is to be helpful in understanding the contribution of reward processing in these areas, then its reliability needs to be established. It is critical to understand that reliability does not represent constancy or lack of change in a measure. For example, brain activity of individuals can change over time, yet still be reliable if the rank order between those individuals in relation to the mean is maintained. This fact can also be intuited from the original formulation of the intra-class correlation (ICC) coefficient given by Fisher (1954):

$$ICC = \frac{1}{Ns^2} \sum_{n=1}^N (x_{n,1} - \bar{x})(x_{n,2} - \bar{x}) \quad (1)$$

where  $\bar{x}$  is the pooled mean,  $N$  is the number of subjects, and the variance is given by:

$$s^2 = \frac{1}{2N} \left[ \sum_{n=1}^N (x_{n,1} - \bar{x})^2 + \sum_{n=1}^N (x_{n,2} - \bar{x})^2 \right] \quad (2)$$

The difference of each individual value at each time point ( $x_{n,1}$ ,  $x_{n,2}$ ) is subtracted from the overall mean of the measurement occasion. It is also obvious from this formulation that reliability is inversely related to within-subject variance. When studying temporal changes, there are several sources of variance that can decrease the signal to noise ratio (SNR), such as decay in equipment calibration, or individual differences in motion parameters (Green and Swets, 1974; Horowitz and Hill, 1980; Cover and Thomas, 1991; Herting et al., 2017). Given that such noise can accumulate differentially over different time scales, it is important to estimate reliability across diverse intervals. So far, no study has addressed RPE reliability in young ages and even more so across different intervals. There have been two reports about reliability of other reward signals during adolescence (Braams et al., 2015; Vetter et al., 2017). These studies report low reliability values in mid-brain regions, where reward related signals would be typically expected. Both studies examine reliability over a single long test-retest interval of two years, which could be more influenced by cumulative errors.

In this work, we seek to establish the reliability of RPE signals across both a short (several months) and a long (several years) test-retest interval during development. We do so by using the ICC coefficient, which informs the within-subject variance relative to the total measurement variability (Bartko, 1966; Shrout and Fleiss, 1979; McGraw and Wong, 1996). For example, the popular version ICC(2,1) is defined as:

$$ICC = \frac{\sigma^2_{within\ subjects}}{\sigma^2_{within\ subjects} + \sigma^2_{between\ subjects} + \sigma^2_{error}} \quad (3)$$

As obvious from this formulation of reliability, the smaller the other sources of variability in the denominator (i.e., the between-subject variance and the measurement error), the higher (i.e., closer to 1) the within-subject reliability. We estimate the ICC using a two-way random-effects modeling approach, sometimes also referred to as a multilevel or hierarchical model, which is a powerful statistical method for estimating individual trajectories of change over time. Even though calculating the ICC measure using the ANOVA framework has been widely adopted, the application of LME methodology to ICC has several advantages in some aspects of computation where limitations are present under the ANOVA framework. Specifically, the variances for the random effects components and the residuals are directly estimated through optimizing the restricted maximum likelihood (REML) function, and thus the ICC value is computed with variance estimates instead of with their mean square counterparts under ANOVA. Therefore, in conjunction with the theoretical quantities, the estimated ICCs are nonnegative by definition. Missing data can be naturally handled in LME because parameters are estimated through the optimization of the (restricted) maximum likelihood function, where a balanced structure is not required. Moreover, incorporating confounding effects is available through adding more fixed-effects terms into the model. This LME approach for ICC has previously been implemented in the program 3dLME (Chen et al., 2013) for voxel-wise data analysis in neuroimaging. In this context, the fMRI BOLD signal change is modeled linearly via the random intercept (initial state) and slope (trajectory of change). Hence, the ICC(2,1) model is an LME case with two crossed random-effects terms. The randomization of both terms differentiates the between- and within-subject variances, enabling the estimation of within-subject reliability (Singer and Willett, 2003; Chen et al., 2013).

In this paper, we examine RPE signaling and its reliability using the “Piñata” task, a child-friendly version of the Monetary Incentive Delay (MID) task. The Piñata task has been previously shown to evoke robust reward-related fMRI BOLD activations in children and adolescents (Helfinstein et al., 2013; Lahat et al., 2016). The task elicits larger negative than positive RPE values, which occur due to “no win” outcomes in win trials. This is because in this paradigm task parameters are adjusted online to maintain a ratio of 66% of successful trials for all subjects, inducing an expectation of more positive outcomes than negative outcomes. Therefore, “no wins”, when they occur, tend to induce larger RPEs relative to wins (as the latter are more expected). Subjects conducted this task in fMRI at three time points. The baseline scan (mean age  $10.6 \pm 0.3$  years) is compared to a repeat scan following  $3 \pm 2.24$  months and another scan following  $33.6 \pm 9.36$  months. As a first step, we demonstrate that behavioral performance of subjects across all visits is reliable and confirm that negative RPEs predominate in this task across the three scans. For the calculation of RPE values, we follow previous studies which defined the expected value as the product of reward magnitude and the success probability (Staudinger et al., 2009; Chase et al., 2015; Ubl et al., 2015). We compare different modeling approaches for estimating the expected success probability, where each model assumes different influence of previous outcomes on the expected value. We address the question of how RPE encoding is distributed in the brain, at each one of the three scans. RPE values are used as a parametric modulator of brain activity during the reward feedback times. We test the hypothesis that negative RPEs are represented mostly in the insula while striatal regions activity is correlated to positive RPE

values. We then ask whether the identified RPE signals are reliable, over three time points during development, separated by a three month and a three year test-retest interval. These results are then compared to the reliability pattern of other task activations.

## Methods

### Participants

Participants were drawn from a longitudinal cohort. Specifically,  $n = 23$  subjects contributed to the first scan and to at least one of the repeated scans. The initial scan (visit 1) was followed by a repeated scan, either  $3 \pm 2.24$  (visit 2,  $n = 18$ ) or  $33.6 \pm 9.36$  (visit 3,  $n = 16$ ) months later. All subjects participated in at least two visits, as follows: visit 1 and visit 2 ( $n = 9$ ); visit 1 and visit 3 ( $n = 7$ ); visit 2 and visit 3 ( $n = 1$ ); visit 1 and visit 2 and visit 3 ( $n = 7$ ). Exclusion of subjects from the analyses was based on excessive motion or technical deficiencies of the data ( $n = 2$  for visit 2 and for visit 3); see Table 1 for included subjects' information. The results were controlled for a possible impact of the number of scans during the third visit, as seven of the subjects participated in all 3 visits - hence we also considered separately only subjects who scanned twice. The reliability results for this additional analysis were consistent with the results we present (see Table-s6 in the Supplement). The reliability of the subjects which participated in three scans (for visit 1 to 3) was quite noisy, possibly due to including only 7 subjects, so we cannot conclude or address how this impacts on reliability with this data set. All participants provided informed assent, and participants' guardians provided informed consent. The study was approved by the Institutional Review Boards of the National Institute of Mental Health and the University of Maryland, College Park.

### Piñata fMRI task

Participants completed the fMRI Piñata task (Fig. 1), a child-friendly version of the Monetary Incentive Delay (MID) task (Helfinstein et al., 2013; Lahat et al., 2016). The task was administered using E-Prime (Psychology Software Tools, USA). Reward incentives in the piñata task differ from the MID task as there are no negative expected values (Knutson et al., 2000), but rewards range from no to large rewards. Participants had to 'whack' piñatas by pressing a button as fast as possible, to earn the presented stars. Each trial included three stages: the anticipation stage was comprised of the cue presentation phase, where participants saw the piñata partially revealed at the top of the screen with the number of stars visible inside (cue, 1500 ms), and a cue-free anticipatory period that varied between 1000 and 2000 ms. In the response stage, the piñata dropped to the center of the screen and the participant made a speeded button press (target). The target appeared for a variable period of time, followed by a delay period, such that the combined duration of both target and delay was in total 1500 ms; in the feedback stage (1500 ms), participants either saw the piñata cracked open with won stars falling (positive feedback), or the intact piñata swinging away (a loss feedback). The task consisted of one practice run of 22 trials, followed by six task runs of 22 trials each, for a total of 132 task trials. Trials were divided evenly between the four incentive levels (0 stars, 1 star, 2 stars and 4 stars) for a total of 33 trials at each incentive level. Participants received money based on the amount of stars they earned, up to \$15 with a minimum of \$3, plus an additional \$3 for every 47 stars they captured. A real-

time algorithm was used to maintain a fixed 66% success level, which adjusted the duration of target image presentation in each trial, to increase or decrease success level.

### **Behavioral data analysis**

Reliability of in-scanner behavior was estimated for reaction time (RT), using the average response time across all trials of each subject. The ICC(2,1) value was estimated over each of the test-retest intervals (visit 1 to 2; visit 1 to 3).

### **fMRI data acquisition**

Participants were scanned in a General Electric (Waukesha, WI, USA) Signa 3 T 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 two-button response box. Forty-seven sagittal slices (3.0-mm thickness) per volume were obtained using a T2\*-weighted echo-planar sequence (echo time, 25 ms; flip angle, 50°; 96 × 96 matrix; field of view, 240 mm; in-plane resolution, 2.5 mm × 2.5 mm; repetition time was 2300 ms). A total of 77 vol were collected in each run. 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: 124 1.2-mm axial slices; repetition time, 8100 ms; echo time, 32 ms; flip angle, 15°; 256 × 256 matrix; field of view, 240 mm; in-plane resolution, 0.86 mm × 0.86 mm; NEX, 1; bandwidth, 31.2 kHz.

### **fMRI data processing**

Analysis of fMRI data was performed using Analysis of Functional and Neural Images (AFNI) software version 2.56 b (Cox, 1996). Echo-planar images (EPI) were visually inspected to confirm image quality and minimal movement. 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 subject'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 an event-related design. Time series for each individual were analyzed using multiple regression (Neter et al., 1996). The entire trial was modeled using a gamma-variate basis function, including five cue events (0 star cues, 1 star cues, 2 star cues, 4 star cues and cues from premature response trials), the target event and the feedback event. 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). For RPE modulation of reward feedback event, each feedback time in that regressor was multiplied by the respective trial base RPE value. Our regressor of interest was the RPE modulation of the feedback event.

Region-of-interest (ROI) approach was used to analyze average activations or reliability, of pre-defined regions (coordinates were derived from the Talairach atlas). To analyze RPE encoding in the ROI level, the individual whole-brain RPE encoding maps were masked for

the right insula, left insula and striatum (a combination of bilateral caudate and putamen). A mean RPE encoding value was then calculated per region by averaging the values of all voxels in that region. To illustrate the reliability of the RPE encoding in the right insula, we extracted the mean RPE encoding value across the reliable voxels in the right insula of each subject, for each of the time points (as presented in Fig. 5C and D).

To validate the reliability of specific ROIs, the whole-brain ICC maps were masked for the right insula, left insula and striatum (combination of bilateral caudate and putamen). ICC values of all voxels in each of these regions were extracted and the number of voxels crossing the ICC threshold was estimated per region. This extraction was also used to estimate the number of reliable voxels shared between the two time intervals.

We summarize the reliability of reward-related regions, relying on a previous metaanalysis where a Reward-Network, comprised of 11 anatomical regions, is defined (Bartra et al., 2013, Satterthwaite et al., 2015, Pan et al., 2017).

### Reward Prediction Error (RPE) computational models

For the RPE modulation analysis, a single-subject model was generated with RPE values as the parametric linear modulator of the BOLD signals in the respective feedback times. This implementation has been previously described for the MID task (Staudinger et al., 2009). The RPE was calculated as follows:

$$\text{Expected value (EV)} = \text{Magnitude} \times \text{Probability} \quad (4)$$

$$\text{Reward Prediction Error (RPE)} = \text{Outcome} - \text{EV} \quad (5)$$

The Magnitude was ascribed the number of stars (cue) and the outcome was the actual received amount. Probability was set to a fixed success probability of 66%, in accordance with the real-time tracking algorithm. There were 7 possible RPE values: Table 2

Moreover, to control for the impact of the chosen computational model for RPE calculation, we compared the results to several dynamic expectation models, where the success probability is modified individually according to previous outcomes per trial. These models account for a stronger influence of outcomes in most recent trials, which decays exponentially in time. Each model considered a different value for the decay rate, reflecting how many previous outcomes are still influential for the expected value.

This weighted probability was realized using the recursive formula:

$$\text{Expected value (EV)}_{n+1} = \text{Magnitude} \times P_n \quad (6)$$

$$P_n = \left(1 - e^{-\frac{t_n - t_{n-1}}{\tau}}\right) \cdot y(n) + e^{-\frac{t_n - t_{n-1}}{\tau}} \cdot P_{n-1} \quad (7)$$

Where  $y(n)$  is the current trial success, such that  $y(n) = 1$  if the subject hit successfully in trial  $n$  within the predefined response interval, and  $y(n) = 0$  otherwise. The exponential function decay rate defines the length of the time window in which previous trials are still considered influential. This feature is determined in the equation by the exponent power, which is the ratio between a single trial duration ( $t_n - t_{n-1}$ ) and a chosen probability estimation time window,  $\tau$ . By setting this rate to different values we created models spanning either 5, 10 or 20 previous trials.

### Test-retest reliability analyses

One set of analyses aimed at analyzing the behavioral reliability of the Piñata task, using the performance measure of mean Response Time. The ICC estimate was implemented following McGraw and Wong (1996) in Matlab (R2017a, Mathworks, MA, USA), across visit 1 to 2 and visit 1 to 3.

Reliability analysis of fMRI data, was conducted by implementing a two-way random-effects model, with the random variables being subject and visit, in the AFNI program 3dLME (Chen et al., 2013; Chen et al., 2017). These ICC estimates were parallel to ICC(2,1) as denoted by Shrout and Fleiss (1979). The decision of defining the explanatory variable in the model as random effects was determined in light of the interchangeability of the factor levels. In this case, we assume that there is no systematic difference across the factor levels, in contrast to case-control designs (e.g. patient-controls) which are typically handled as fixed effects because of the lack of exchangeability. Similarly, the visit is treated as random effects as there is no systematic difference in the conditions between the three visits.

We examined using this approach the test-retest reliability of the RPE signals during the reward feedback phase. Modulation was implemented for BOLD activity during the feedback event. The inputs to the fMRI voxel-wise ICC analyses were the modulation strength estimates from the individual analyses. The whole brain level ICC threshold was set to 0.45, corresponding to the mid-value of the range defined as a “fair” ICC (Cicchetti, 2001). Next, significant clusters were corrected for multiple comparisons in the whole brain level, using Monte Carlo simulations and a mixed autocorrelation function (3dClustSim in AFNI using the acceptable ACF model, which simulates noise volume assuming the ACF is given by a mixed-model), which produced a threshold of 102 voxels for corrected  $p < 0.05$  (uncorrected  $p < 0.005$ ). Analysis of ICC significance was conducted in AFNI at the whole brain level, using Fisher’s  $r$  to  $z$  transformation voxel-wise, as modified in McGraw and Wong (1996):

$$z_i = \frac{1}{2\sigma} \log \frac{1 + (k - 1)ICC_i}{1 - ICC_i} \quad (8)$$

$$\sigma^2 = \frac{k}{2(n - 2)(k - 1)} \quad (9)$$

Where  $i$  is the voxel index,  $k$  the number of repeated scans and  $n$  is the number of subjects. The whole brain maps of  $z$  scores were then thresholded for  $p < 0.05$ . Group-level one



sample *t*-test images of RPE encoding, were produced per visit and thresholded similarly (corrected to  $p < 0.05$  using a minimal cluster size of 102 voxels and  $p < 0.005$ ).

Impact of motion on reliability results, was tested using the same LME model with mean subject motion as a between subject covariate.

We also examined whether reliability of the RPE signal differs from the reliability of other task activations. For this purpose, we analyzed activity evoked during feedback receipt and also looked separately at activation due to loss feedback (as RPE values in this task are mainly negative). In this analysis, the feedback event was not modulated by RPE values and was considered versus baseline (the non-modeled phases of the task).

## Results

### Behavioral reliability

Reliability of task behavioral measures is examined to determine whether participants' responding patterns are stable between scans. As shown in Fig. 2A, mean RT values are highly reliable between both the short (visit 1 to 2; ICC = 0.91,  $p = 0.001$ ) and long (visit 1 to 3; ICC = 0.85,  $p = 0.007$ ) scanning intervals.

### Task RPE values

We first test whether indeed negative RPE values are more dominant in this paradigm. We demonstrate this in Fig. 3B, by the cumulative sum of RPEs during the experiment (the sum of all preceding RPE values, per trial), of each of the subjects. It is apparent that most of the curves are turning negative, showing that higher negative RPE amplitudes are intrinsic to the task.

### RPE encoding

RPE values are used as a parametric modulator for feedback phase at the individual level, followed by a group *t*-test at the whole brain level. The resulting voxel-wise effect images, reflect the strength of the linear correlation between BOLD activation (during reward feedback) and RPE value on that trial. Positive values indicate stronger activation for *higher* RPE values, while negative values reflect increased activity for *lower* RPEs. As shown in Fig. 4A (and supplementary Table-s1), during visit 1 RPE signal is represented positively in the striatum (depicted by an orange arrow) and negatively in the insula (depicted by the blue arrow). During visit 2 there is a significant negative association between RPE values and insula activation, but the striatal signal falls below correction threshold (Fig. 4B). During visit 3, both RPE activations in the insula and in the striatum are detected, but when using an uncorrected threshold (Fig. 4C).

To compare the group level effect across visits, we contrasted between the group whole-brain images (visit 1 versus visit 2, visit 1 versus visit 3 and visit 2 versus visit 3). Following FWE correction of the contrasted images, we find that RPE encoding is significantly higher during the first visit compared to the follow-up visits, in frontal and temporal regions. Of the reward-network regions, the caudate and thalamus show significantly better encoding, during the first visit relative to the third (see Figure s1 and supplement table-s1, for detailed

results of contrasting RPE encoding between the three visits). When comparing RPE encoding between the three visits in the individual ROI level (Fig. 4D), we do not find the differences we identified by contrasting the whole-brain group images: in this analysis both insular and striatal mean activations do not change significantly over the three time points. This was tested with a series of t-tests across each pair of visits (for the striatum, visits 1 and 2:  $p = 0.14$ ,  $t = 1.4$ ,  $z = 0.99$ ; visits 1 and 3:  $p = 0.47$ ,  $t = 0.79$ ,  $z = 0.07$ ; visits 2 and 3:  $p = 0.72$ ,  $t = -0.37$ ,  $z = 0.52$ ). The differences which are observed in the whole brain group level might be also caused by differences in the spatial spread of active voxels, limiting the detection of the same clusters following minimal cluster corrections. As explained in more detail by Aron et al., (2006), this can be expected also due to the nonlinearity of the thresholding of whole-brain images, which can exaggerate very small differences in the signal or noise to substantial differences between thresholded images.

### RPE encoding reliability

Reliability of RPE encoding in the whole brain level is shown in Fig. 5 (lower panels show the voxels which cross the z score threshold) and summarized in supp. Table 2. A significant cluster of a reliable RPE signal is found in the right insula between both visit 1 to 2 (Fig. 5A;  $ICC_{\text{peak}} = 0.731$ ) as well as 1 to 3 (Fig. 5B;  $ICC_{\text{peak}} = 0.735$ ). Only RPE encoding in the right insula was significantly reliable across both time intervals across all tested models. By contrast, no other of the eleven reward network regions (left striatum, right striatum, ventromedial prefrontal cortex, left insula, posterior cingulate, ventral tegmental area, anterior cingulate, pre-supplementary motor area, left thalamus, right thalamus) was significant using the standard model. In some but not across all additional models, other brain regions, listed in the supplement, were found to be significantly reliable. For example, striatal results are crossing statistical threshold when excluding premature response trials. We summarize the reliability of reward-related regions, by listing whether each of the Reward-Network regions is reliable or not, in Supplement Table-s9 (according to the regions which appeared as significantly reliable in the whole-brain ICC analysis). Fig. 5C –D demonstrate how reliability in the right insula reflects a stable rank order at the individual RPE encoding level.

To validate the anatomical region of reliability we used an ROI extraction of ICC values. We find that the number of voxels which are exceeding the ICC threshold in the insula is 122 and 334, across visit 1 to 2 and 1 to 3, respectively. These values indicate that the region is reliable across both time intervals. It should be noted, that these values are extracted based on anatomical location and do not consider how spatially contiguous voxels are, hence we can expect differences from the cluster sizes detected with a whole-brain minimal cluster correction.

It should be also stressed, that reliability of activation is not necessarily influenced by differences in activation across visits: a signal can significantly change over time but still be reliable, if the change is consistent within the group and subjects maintain the same rank order.

### Controlling for confounding factors

We compared the described reliability results to the reliability of RPE signals when RPE values are calculated using a dynamic model of the expected success probability. This model accounts for the influence of recent outcomes on our expectation to win, with a temporal decay in their impact. We find that insula reliability is high and robust for both time intervals across all of the tested models. Striatal results are crossing statistical threshold but inconsistently, being significantly reliable between visit 1 to 2 for the model which considers the last 20 trials outcomes, and between visit 1 to 3 when considering the last 5 trials (see Table-s3 and Figure s2 for a detailed description of the results).

To control for behavioral effects on reliability estimates, in particular due to losses following prepotent responses, we excluded those trials from the RPE modulation. We find that under this restriction the insula remains reliable as before (sTable-s4). Moreover, once these are removed the reliability in the striatum becomes significant at the pre-determined level (visit 1 to 2: right insula and striatum  $ICC_{\text{peak}} = 0.7$ ; visit 1 to 3: right insula and striatum:  $ICC_{\text{peak}} = 0.77$ ; left putamen:  $ICC_{\text{peak}} = 0.78$ , see Table-s4 for detailed results).

We also considered the status of puberty along the three time points using the puberty Tanner scale (Tanner, 1986), which shows values of (average  $\pm$  SD):  $1.42 \pm 0.48$  during the first two visits, and  $2.46 \pm 0.28$  during the third. When using these puberty scores as an explanatory variable in the ICC LME model (over visit 1 to 3), we find that in addition to the consistent reliability of RPE signals in the insula (here on left and right sides), reliability in the striatum becomes significant at the pre-determined level (right putamen; see detailed results in Table-s5).

To account for the potential effects of individual motion differences, we repeated all ICC analyses with the motion parameter as a covariate. Mean subject motion per visit was used as a between subject covariate. This modification did not influence the pattern or significance of results.

### Feedback phase activations reliability

We then analyze reliability of general activation during feedback times - not modulated by RPE values. As shown in Fig. 6A, the reliability pattern differs and the striatum is in fact the most reliable region (reliability is across visits 1, 2 and 3:  $ICC_{\text{peak}} = 0.89$  when using a more stringent threshold of  $ICC > 0.6$  for a better anatomical separation), while insula is less reliable (a reliable cluster appears only under a lower threshold).

Interestingly, when analyzing the activations during feedback of loss events only, activation is now negative in the striatum and positive in the insula (see Figure s3). This observation means that at the times of presentation of loss feedback, striatum becomes less activated, while the insula shows increased activation (positive correlation to loss feedback). It is important to stress that there is no straightforward relation between the insula being activated synchronously with loss feedback *times*, and the negative correlation of the amplitude of this insular activity with the values of loss RPE.

Analyzing the reliability of feedback phase signals, considering loss events only, shows as well the highest reliability in striatal regions (Fig. 6B, visit 1 to 2  $ICC_{\text{peak}} = 0.73$ ; visit 1 to 3 focus in left putamen of  $ICC_{\text{peak}} = 0.79$ , cluster extends also to left insula and there is no right insula cluster), as also depicted clearly by contrasting these maps with the respective RPE encoding reliability maps (Fig. 6C). Overall, the reliability of activation during feedback for this task seems to be highest in striatum as per theory, while the reliability of RPE signals is specifically high in the insula.

## Discussion

The current study examined the test–retest reliability of RPE fMRI signals, during the transition from childhood to adolescence. Two test-retest time intervals were used, one spanning several months and the other several years. Results show the distributed encoding of RPEs, being maximal in the insula for negative RPE values whilst focused in the striatum for positive RPEs. These insular negative RPE signals are highly reliable across both time intervals, suggesting its potential utility as a marker for tracking aberrations in loss-processing and punishment-based learning, during development or disease.

We first validated the behavioral reliability of the piñata task, which is a modified version of the MID task adapted for use in pediatric populations. We then addressed RPE encoding in the brain, by analyzing the linear modulation between RPE values and brain activity during the feedback phase. In keeping with previous findings (Liu et al., 2007; Palminteri et al., 2012; Garrison et al., 2013), RPE values were negatively correlated to insular activity, and a positive correlation was found between RPE and striatal activity. In this valence-dependent dissociation, striatal activity increases as RPE values are more positive, while insular activity increases as RPE values become more negative. This encoding pattern does not seem to change over the three time points.

We then showed that the signal for negative RPE values in the right insula is reliable. This does *not* mean that there are no changes in encoding of negative RPEs over the three time points (and hence across that developmental period), but that changes occur similarly across the subjects. Hence, a subject who was on the high end of RPE signal strength during childhood, would maintain this rank relative to the group following several months, as well as following several years in adolescence. Such consistency requires low measurement noise and no individual changes which differ significantly from the trend of the group. When reliability is low on the other hand, the measure is not consistent across subjects and the group average estimates could be erroneously enhanced or blunted. Therefore, the reliability of RPE encoding as observed in the insula, indicates that changes found in the group average level, are consistent with changes in the individual level. Hence, estimating reliability provides a necessary validation of the significance of group level activation changes. In addition, further regions comprising frontal and parietal areas, show sufficient reliability for either one or both intervals. The anterior cingulate cortex for example, is a region comprising the fronto-striatal dopaminergic pathways, the connectivity of which influences reward-processing and regulation (Gotlib et al., 2010).

Striatal encoding of RPE positive values seemed to be reliable under a less strict threshold, or when excluding premature response trials. Premature response trials may have added to the noise of the RPE signal, as there is no actual RPE experienced by the subject during such trials. The higher reliability of striatal signals once covarying for puberty status, could reflect a developmental effect on reward processing which should be further explored. The overall inconsistency of the reliability findings for the RPE signal in the striatum, may be also a reflection of the higher negative RPE amplitudes intrinsic in the task. If both RPE values and the associated brain signals have high amplitudes, a better signal to noise ratio should be expected and therefore a higher consistency (Welvaert and Rosseel, 2013). However, it is possible that the low reliability of striatal RPE encoding is not only due to the task design, as a previous study in adults showed a similar finding, even though using a task with a balanced relation between positive and negative RPE values (Chase et al., 2015). Another possibility is higher reliability due to a larger signal to noise ratio per the stronger activation (Caceres et al., 2009), hence a more dominant RPE encoding in the insula. In congruence, absolute striatal RPE encoding is lower compared to the insula in the second visit ( $p = 0.03$ ), possibly confounding striatal reliability (indeed found to be the lowest over visit 1 to 2). In contrast to the RPE signal, we show that during reward and loss feedback the striatum is the most reliable region, in keeping with prior theory and data. The higher reliability of the striatum is supported again by a larger absolute amplitude of the signal versus the insula, possibly increasing the signal to noise ratio and the consistency of activation (a peak beta of  $-0.23$  versus  $0.16$  in the insula). The two measures, feedback activation and RPE modulated effect, provide different information of brain activation. While striatum is more dominant in encoding occurrence and value of feedback, the insula, shows higher and more consistent association of activity to RPE values.

Reliability of RPE encoding in the insula, is high both over the short and the long test-retest intervals. This suggests that the RPE measure in the insula is not particularly prone to cumulative errors at longer time scales (up to the 3 years on average tested here). However, should be noted that the reliable voxels in the insula do not overlap between the two time intervals (there are 15 mutual voxels only). When lowering the ICC threshold (to  $ICC > 0.3$ ), a larger number of voxels overlap between both time intervals (103 voxels), but these are only 25% of the reliable voxels in the insula under this threshold. Possible reasons for this discrepancy, could be minor shifts in the insular voxels which predominantly encode RPEs, either between subjects (the two time intervals included different subjects), or over time. Future studies could address this question better, if using a design of multiple time intervals.

Establishing reliable signals over such different time intervals, is highly relevant to psychopathology as well as developmental studies. Treatment outcomes of drug or psychological therapies are typically measured over several weeks to a few months. RPE encoding alterations have been implicated in a range of psychiatric disorders (Murray et al., 2008; Moutoussis et al., 2015; Radua et al., 2015; Ubl et al., 2015; Schmidt et al., 2016; Rothkirch et al., 2017; White et al., 2017). For example, in depression, positive RPEs are found to be blunted, whilst negative RPEs in the insula appear increased (Chandrasekhar Pammi et al., 2015; Engelmann et al., 2017). Increased encoding of negative RPEs indicates higher sensitivity to unexpected losses, which in depression may characterize avoidance behavior (Luking et al., 2015; Hevey et al., 2017), as well as the impairment in reward-based

learning (Henriques and Davidson, 2000; Pizzagalli et al., 2005; Tavares et al., 2008; Whitmer et al., 2012; Vrieze et al., 2013). Reliable negative RPE encoding in the brain, therefore, could be used as a valid treatment parameter, to track intervention induced changes, both at a short and a long time scale.

Moreover, the identification of a reliable negative RPE signal is relevant to tracking developmental changes in reward processing. Previous studies have shown that positive RPE signals in the striatum are higher in adolescents, compared to children and young adults (Cohen et al., 2010; Galvan, 2010; Eppinger et al., 2013; Braams et al., 2015; Thomason and Marusak, 2017). This is thought to reflect the higher motivation for rewards and risk taking behaviors in adolescents. Other studies find negative RPE signals in the insula to increase during adolescence (Van Leijenhorst et al., 2010; Smith et al., 2014; Hauser et al., 2015), which has been taken to indicate of a higher sensitivity to losses. Such developmental changes can occur with different time scales across individuals, which would potentially compromise reliability. However, as we had the opportunity to analyze the reliability also over a more stable, short time interval, inter-subject differences over the longer time interval in this case, are probably less pronounced than the consistency of RPE encoding within the group.

It is important to note that while insular RPE signals appear to be consistent over time, a number of other reward-related regions don't show reliable activations under the primary model used in the paper. A possible explanation is that these regions are not predominantly involved in RPE encoding, and therefore have a low RPE signal to noise ratio which decreases the reliability of this signal.

This study has several strengths, such as the scanning at three different time points during development, modeling RPE signals and having a young age sample. However, the study should also be seen in the light of several limitations. First, the task used, whilst well suited for the study of negative RPEs, it doesn't have a good enough access to positive RPE signals reliability. Second, the relatively small sample size in this study means that our power to detect signals may be diminished. Third, the variability of time intervals across subjects, in both the short and the long intervals is large (although the two categories of time intervals are still separated by at least 1 year from each other).

Overall, this study addresses uniquely RPE encoding and its reliability during development, showing highest reliability for encoding negative RPEs in the insula. These results are relevant for studies aiming to estimate alterations in sensitivity to losses and punishment-based learning, across different time scales, whether during development or disease.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Funding

This work was supported by the National Institutes of Health, Intramural Research Program, grant number ZIAMH002957-01.

## References

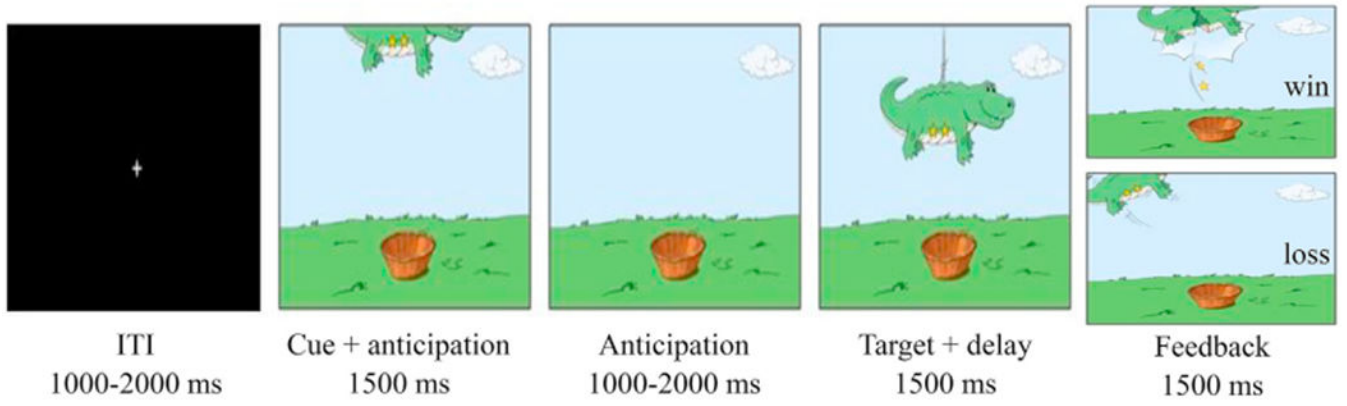
- Aron AR, Gluck MA, Poldrack RA, 2006 Long-term test-retest reliability of functional MRI in a classification learning task. *Neuroimage* 29 (3), 1000–1006. [PubMed: 16139527]
- Averbeck BB, Costa VD, 2017 Motivational neural circuits underlying reinforcement learning. *Nat. Neurosci* 20 (4), 505–512. [PubMed: 28352111]
- Bartko JJ, 1966 The intraclass correlation coefficient as a measure of reliability. *Psychol. Rep* 19 (1), 3–11. [PubMed: 5942109]
- Bartra O, McGuire JT, Kable JW, 2013 The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76, 412–427. [PubMed: 23507394]
- Bayer HM, Glimcher PW, 2005 Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47 (1), 129–141. [PubMed: 15996553]
- Bennett CM, Miller MB, 2010 How reliable are the results from functional magnetic resonance imaging? *Ann. N. Y. Acad. Sci* 1191, 133–155. [PubMed: 20392279]
- Braams BR, van Duijvenvoorde ACK, Peper JS, Crone EA, 2015 Longitudinal changes in adolescent risk-taking: a comprehensive study of neural responses to rewards, pubertal development, and risk-taking behavior. *J. Neurosci* 35 (18), 7226–7238. [PubMed: 25948271]
- Caceres A, Hall DL, Zelaya FO, Williams SCR, Mehta MA, 2009 Measuring fMRI reliability with the intra-class correlation coefficient. *Neuroimage* 45 (3), 758–768. [PubMed: 19166942]
- Chandrasekhar Pammi VS, Pillai Geethabhavan Rajesh P, Kesavadas C, Rappai Mary P, Seema S, Radhakrishnan A, Sitaram R, 2015 Neural loss aversion differences between depression patients and healthy individuals: a functional MRI investigation. *NeuroRadiol. J* 28 (2), 97–105. [PubMed: 25923684]
- Chase HW, Fournier JC, Greenberg T, Almeida JR, Stiffler R, Zevallos CR, Aslam H, Cooper C, Deckersbach T, Weyandt S, Adams P, Toups M, Carmody T, Oquendo MA, Peltier S, Fava M, McGrath PJ, Weissman M, Parsey R, McInnis MG, Kurian B, Trivedi MH, Phillips ML, 2015 Accounting for dynamic fluctuations across time when examining fMRI test-retest reliability: analysis of a reward paradigm in the EMBARC study. *PLoS One* 10 (5) e0126326. [PubMed: 25961712]
- Chen G, Saad ZS, Britton JC, Pine DS, Cox RW, 2013 Linear mixed-effects modeling approach to fMRI group analysis. *Neuroimage* 73, 176–190. [PubMed: 23376789]
- Chen G, Taylor PA, Haller SP, Kircanski K, Stoddard J, Pine DS, Leibenluft E, Brotman MA, Cox RW, 2017 Intra-class Correlation: Improved Modeling Approaches and Applications for Neuroimaging bioRxiv).
- Cicchetti DV, 2001 The precision of reliability and validity estimates re-visited: distinguishing between clinical and statistical significance of sample size requirements. *J. Clin. Exp. Neuropsychol* 23 (5), 695–700. [PubMed: 11778646]
- Cohen JR, Asarnow RF, Sabb FW, Bilder RM, Bookheimer SY, Knowlton BJ, Poldrack RA, 2010 A unique adolescent response to reward prediction errors. *Nat. Neurosci* 13 (6), 669–671. [PubMed: 20473290]
- Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N, 2012 Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482 (7383), 85–88. [PubMed: 22258508]
- Corbett D, Wise RA, 1980 Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: a moveable electrode mapping study. *Brain Res.* 185 (1), 1–15. [PubMed: 7353169]
- Cover TM, Thomas JA, 1991 *Elements of Information Theory*. Wiley, New York.
- Cox RW, 1996 AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res* 29 (3), 162–173. [PubMed: 8812068]

- Diederer KM, Spencer T, Vestergaard MD, Fletcher PC, Schultz W, 2016 Adaptive prediction error coding in the human midbrain and striatum facilitates behavioral adaptation and learning efficiency. *Neuron* 90 (5), 1127–1138. [PubMed: 27181060]
- Engelmann JB, Berns GS, Dunlop BW, 2017 Hyper-responsivity to losses in the anterior insula during economic choice scales with depression severity. *Psychol. Med* 7, 1–13.
- Eppinger B, Schuck NW, Nystrom LE, Cohen JD, 2013 Reduced striatal responses to reward prediction errors in older compared with younger adults. *J. Neurosci* 33 (24), 9905–9912. [PubMed: 23761885]
- Fisher RA, 1954 *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.
- Galvan A, 2010 Adolescent development of the reward system. *Front. Hum. Neurosci* 4, 6. [PubMed: 20179786]
- Garrison J, Erdeniz B, Done J, 2013 Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci. Biobehav. Rev* 37 (7), 1297–1310. [PubMed: 23567522]
- Gotlib IH, Hamilton JP, Cooney RE, Singh MK, Henry ML, Joormann J, 2010 Neural processing of reward and loss in girls at risk for major depression. *Arch. Gen. Psychiatr* 67 (4), 380–387. [PubMed: 20368513]
- Green DM, Swets JA, 1974 *Signal Detection Theory and Psychophysics*. R. E. Krieger Pub. Co, Huntington, N.Y.
- Hauser TU, Iannaccone R, Walitza S, Brandeis D, Brem S, 2015 Cognitive flexibility in adolescence: neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. *Neuroimage* 104, 347–354. [PubMed: 25234119]
- Helfinstein SM, Kirwan ML, Benson BE, Hardin MG, Pine DS, Ernst M, Fox NA, 2013 Validation of a child-friendly version of the monetary incentive delay task. *Soc. Cognit. Affect Neurosci* 8 (6), 720–726. [PubMed: 22569186]
- Henriques JB, Davidson RJ, 2000 Decreased responsiveness to reward in depression. *Cognit. Emot* 14 (5), 711–724.
- Herting MM, Gautam P, Chen Z, Mezher A, Vetter NC, 2017 Test-retest reliability of longitudinal task-based fMRI—implications for developmental studies. *Developmental Cognitive Neuroscience*.
- Hevey D, Thomas K, Laureano-Schelten S, Looney K, Booth R, 2017 Clinical depression and punishment sensitivity on the BART. *Front. Psychol* 8, 670. [PubMed: 28512440]
- Horowitz P, Hill W, 1980 *The Art of Electronics*. Cambridge University Press, Cambridge, Eng. ; New York.
- Knutson B, Westdorp A, Kaiser E, Hommer D, 2000 FMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage* 12 (1), 20–27. [PubMed: 10875899]
- Lahat A, Benson BE, Pine DS, Fox NA, Ernst M, 2016 Neural responses to reward in childhood: relations to early behavioral inhibition and social anxiety. *Soc. Cognit. Affect Neurosci*
- Lamm C, Benson BE, Guyer AE, Perez-Edgar K, Fox NA, Pine DS, Ernst M, 2014 Longitudinal study of striatal activation to reward and loss anticipation from mid-adolescence into late adolescence/early adulthood. *Brain Cognit.* 89 (Suppl. C), 51–60. [PubMed: 24485273]
- Liu X, Powell DK, Wang H, Gold BT, Corbly CR, Joseph JE, 2007 Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. *J. Neurosci* 27 (17), 4587–4597. [PubMed: 17460071]
- Luking KR, Pagliaccio D, Luby JL, Barch DM, 2015 Child gain approach and loss avoidance behavior: relationships with depression risk, negative mood, and anhedonia. *J. Am. Acad. Child Adolesc. Psychiatry* 54 (8), 643–651. [PubMed: 26210333]
- Maitra R, Roys SR, Gullapalli RP, 2002 Test-retest reliability estimation of functional MRI data. *Magn. Reson. Med* 48 (1), 62–70. [PubMed: 12111932]
- McGraw K, Wong SP, 1996 Forming Inferences about Some Intraclass Correlation Coefficients.
- Moutoussis M, Story GW, Dolan RJ, 2015 The computational psychiatry of reward: broken brains or misguided minds? *Front. Psychol* 6, 1445. [PubMed: 26483713]

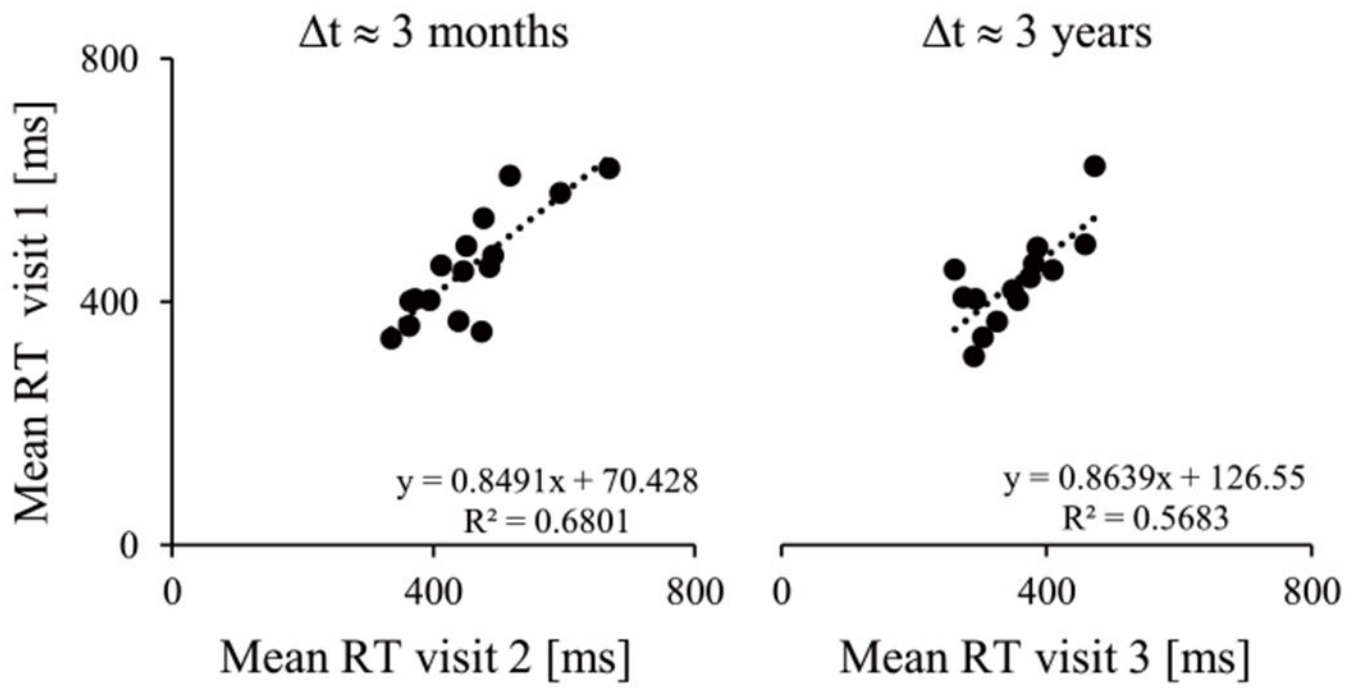


- Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey G, Jones PB, Bullmore ET, Robbins TW, Fletcher PC, 2008 Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol. Psychiatr* 13 (3), 267–276, 239.
- Neter J, Wasserman W, Kutner MH, 1996 *Applied Linear Statistical Models*. Irwin, Chicago.
- Olds J, Milner P, 1954 Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. Comp. Physiol. Psychol* 47 (6), 419–427. [PubMed: 13233369]
- Palminteri S, Justo D, Jauffret C, Pavlicek B, Dauta A, Delmaire C, Czernecki V, Karachi C, Capelle L, Durr A, Pessiglione M, 2012 Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. *Neuron* 76 (5), 998–1009. [PubMed: 23217747]
- Pan PM, Sato JR, Salum GA, Rohde LA, Gadelha A, Zugman A, Mari J, Jackowski A, Picon F, Miguel EC, Pine DS, Leibenluft E, Bressan RA, Stringaris A, 2017 Ventral striatum functional connectivity as a predictor of adolescent depressive disorder in a longitudinal community-based sample. *Am. J. Psychiatr* 174 (11), 1112–1119. [PubMed: 28946760]
- Pan WX, Schmidt R, Wickens JR, Hyland BI, 2005 Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci* 25 (26), 6235–6242. [PubMed: 15987953]
- Pizzagalli DA, Jahn AL, O’Shea JP, 2005 Toward an objective characterization of an anhedonic phenotype: a signal detection approach. *Biol. Psychiatr* 57 (4), 319–327.
- Radau J, Schmidt A, Borgwardt S, Heinz A, Schlagenhauf F, McGuire P, Fusar-Poli P, 2015 Ventral striatal activation during reward processing in psychosis: a neurofunctional meta-analysis. *JAMA Psychiatry* 72 (12), 1243–1251. [PubMed: 26558708]
- Raemaekers M, du Plessis S, Ramsey NF, Weusten JM, Vink M, 2012 Test-retest variability underlying fMRI measurements. *Neuroimage* 60 (1), 717–727. [PubMed: 22155027]
- Rolls ET, McCabe C, Redoute J, 2008 Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebr. Cortex* 18 (3), 652–663.
- Rothkirch M, Tonn J, Kohler S, Sterzer P, 2017 Neural mechanisms of reinforcement learning in unmedicated patients with major depressive disorder. *Brain* 140 (4), 1147–1157. [PubMed: 28334960]
- Satterthwaite TD, Kable JW, Vandekar L, Katchmar N, Bassett DS, Baldassano CF, Ruparel K, Elliott MA, Sheline YI, Gur RC, Gur RE, Davatzikos C, Leibenluft E, Thase ME, Wolf DH, 2015 Common and dissociable dysfunction of the reward system in bipolar and unipolar depression. *Neuropsychopharmacology* 40 (9), 2258–2268. [PubMed: 25767910]
- Schmidt A, Palaniyappan L, Smieskova R, Simon A, Riecher-Rossler A, Lang UE, Fusar-Poli P, McGuire P, Borgwardt SJ, 2016 Dysfunctional insular connectivity during reward prediction in patients with first-episode psychosis. *J. Psychiatry Neurosci* 41 (6), 367–376. [PubMed: 26854756]
- Schultz W, 1998 Predictive Reward Signal of Dopamine Neurons. *The American Physiological Society*.
- Schultz W, 2006 Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol* 57, 87–115. [PubMed: 16318590]
- Schultz W, 2013 Updating dopamine reward signals. *Curr. Opin. Neurobiol* 23 (2), 229–238. [PubMed: 23267662]
- Schultz W, 2016 Dopamine reward prediction error coding. *Dialogues Clin. Neurosci* 18 (1), 23–32. [PubMed: 27069377]
- Schultz W, 2017 Reward prediction error. *Curr. Biol* 27 (10), R369–R371. [PubMed: 28535383]
- Schultz W, Apicella P, Ljungberg T, 1993 Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci* 13 (3), 900–913. [PubMed: 8441015]
- Schultz W, Stauffer WR, Lak A, 2017 The phasic dopamine signal maturing: from reward via behavioural activation to formal economic utility. *Curr. Opin. Neurobiol* 43, 139–148. [PubMed: 28390863]
- Shrout PE, Fleiss JL, 1979 Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull* 86 (2), 420–428. [PubMed: 18839484]

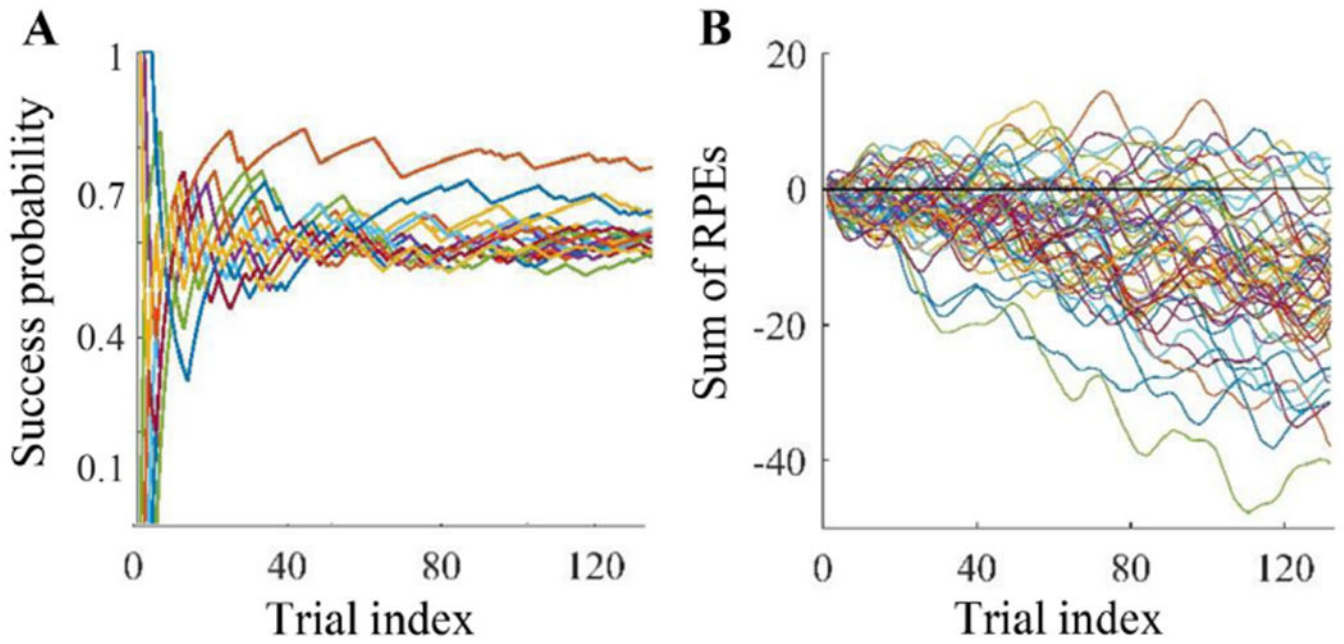
- Singer JD, Willett JB, 2003 *Applied Longitudinal Data Analysis : Modeling Change and Event Occurrence*. Oxford University Press, Oxford ; New York.
- Smith AR, Steinberg L, Chein J, 2014 The role of the anterior insula in adolescent decision making. *Dev. Neurosci* 36 (3–4), 196–209. [PubMed: 24853135]
- Somerville LH, Casey BJ, 2010 Developmental neurobiology of cognitive control and motivational systems. *Curr. Opin. Neurobiol* 20 (2), 236–241. [PubMed: 20167473]
- Staudinger MR, Erk S, Abler B, Walter H, 2009 Cognitive reappraisal modulates expected value and prediction error encoding in the ventral striatum. *Neuroimage* 47 (2), 713–721. [PubMed: 19442745]
- Sutton RS, Barto AG, 1998 *Introduction to Reinforcement Learning*. MIT Press.
- Tanner JM, 1986 Normal growth and techniques of growth assessment. *Clin. Endocrinol. Metabol* 15 (3), 411–451.
- Tavares JVT, Clark L, Furey ML, Williams GB, Sahakian BJ, Drevets WC, 2008 Neural basis of abnormal response to negative feedback in unmedicated mood disorders. *Neuroimage* 42 (3), 1118–1126. [PubMed: 18586109]
- Thomason ME, Marusak HA, 2017 Within-subject neural reactivity to reward and threat is inverted in young adolescents. *Psychol. Med* 47 (9), 1549–1560. [PubMed: 28238294]
- Ubl B, Kuehner C, Kirsch P, Ruttorf M, Diener C, Flor H, 2015 Altered neural reward and loss processing and prediction error signalling in depression. *Soc. Cognit. Affect Neurosci* 10 (8), 1102–1112. [PubMed: 25567763]
- Van Leijenhorst L, Zanolie K, Van Meel CS, Westenberg PM, Rombouts SA, Crone EA, 2010 What motivates the adolescent? Brain regions mediating reward sensitivity across adolescence. *Cerebr. Cortex* 20 (1), 61–69.
- Vetter NC, Steding J, Jurk S, Ripke S, Mennigen E, Smolka MN, 2017 Reliability in adolescent fMRI within two years - a comparison of three tasks. *Sci. Rep* 7 (1), 2287. [PubMed: 28536420]
- Vrieze E, Pizzagalli DA, Demyttenaere K, Hompes T, Sienaert P, de Boer P, Schmidt M, Claes S, 2013 Reduced reward learning predicts outcome in major depressive disorder. *Biol. Psychiatr* 73 (7), 639–645.
- Welvaert M, Rosseel Y, 2013 On the definition of signal-to-noise ratio and contrast-to-noise ratio for FMRI data. *PLoS One* 8 (11), e77089. [PubMed: 24223118]
- White SF, Geraci M, Lewis E, Leshin J, Teng C, Averbek B, Meffert H, Ernst M, Blair JR, Grillon C, Blair KS, 2017 Prediction error representation in individuals with generalized anxiety disorder during passive avoidance. *Am. J. Psychiatr* 174 (2), 110–117. [PubMed: 27631963]
- Whitmer AJ, Frank MJ, Gotlib IH, 2012 Sensitivity to reward and punishment in major depressive disorder: effects of rumination and of single versus multiple experiences. *Cognit. Emot* 26 (8), 1475–1485. [PubMed: 22716241]



**Fig. 1.**  
The task trial time course (adapted from Lahat et al., 2016).

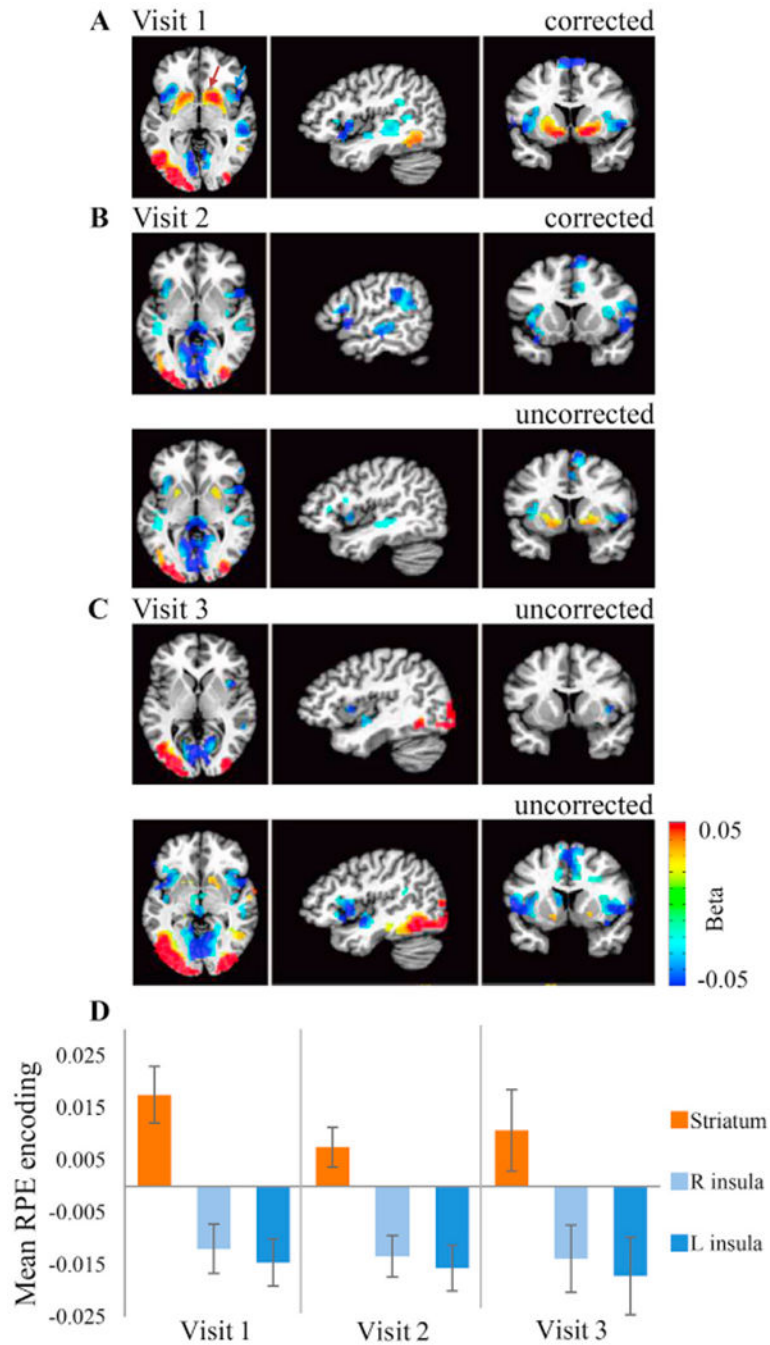


**Fig. 2.**  
Reliability of response time across the two time intervals.



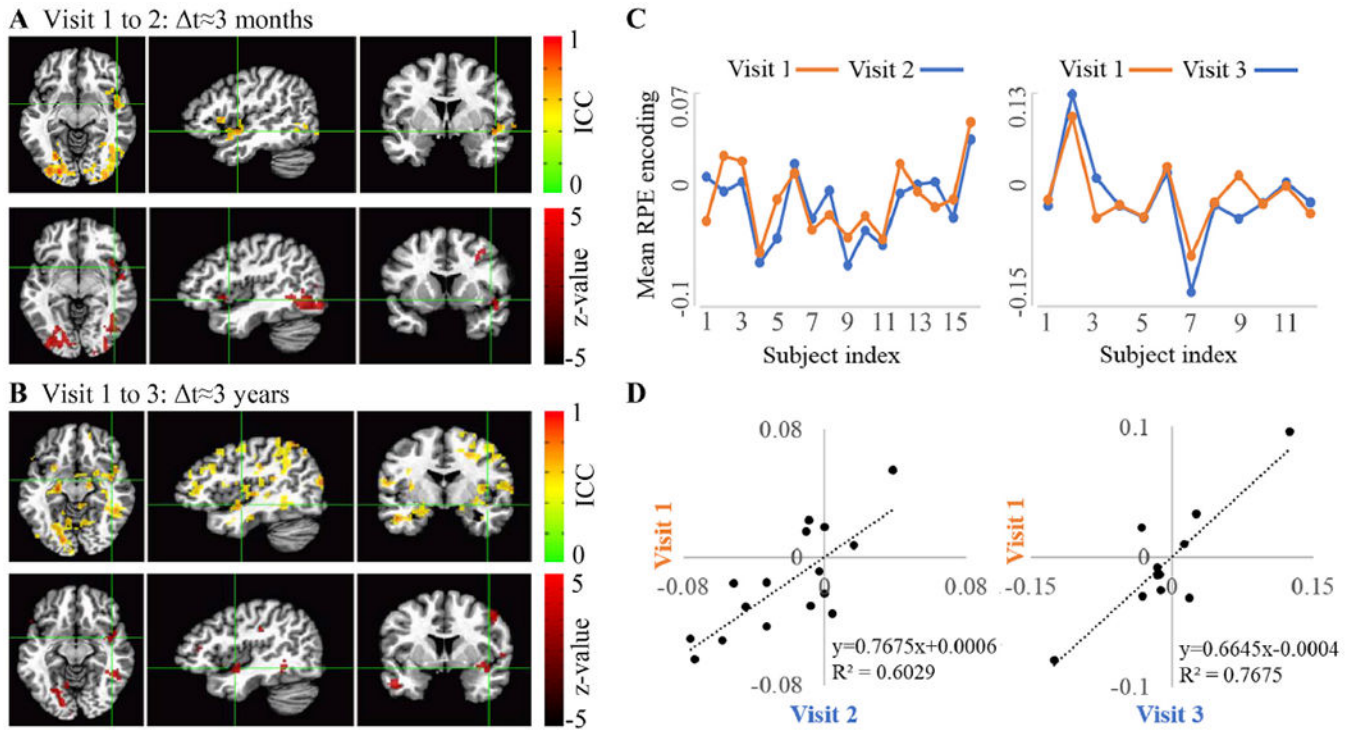
**Fig. 3.**

(A) Success probability values, exemplified for visit 1, each subject is represented by a single curve (probability per trial is the mean across all preceding trials). (B) RPE values are mostly negative, as demonstrated by the cumulative sum of RPE values (calculated per trial using all preceding trials values); each curve represents a single subject, and all visits are shown.



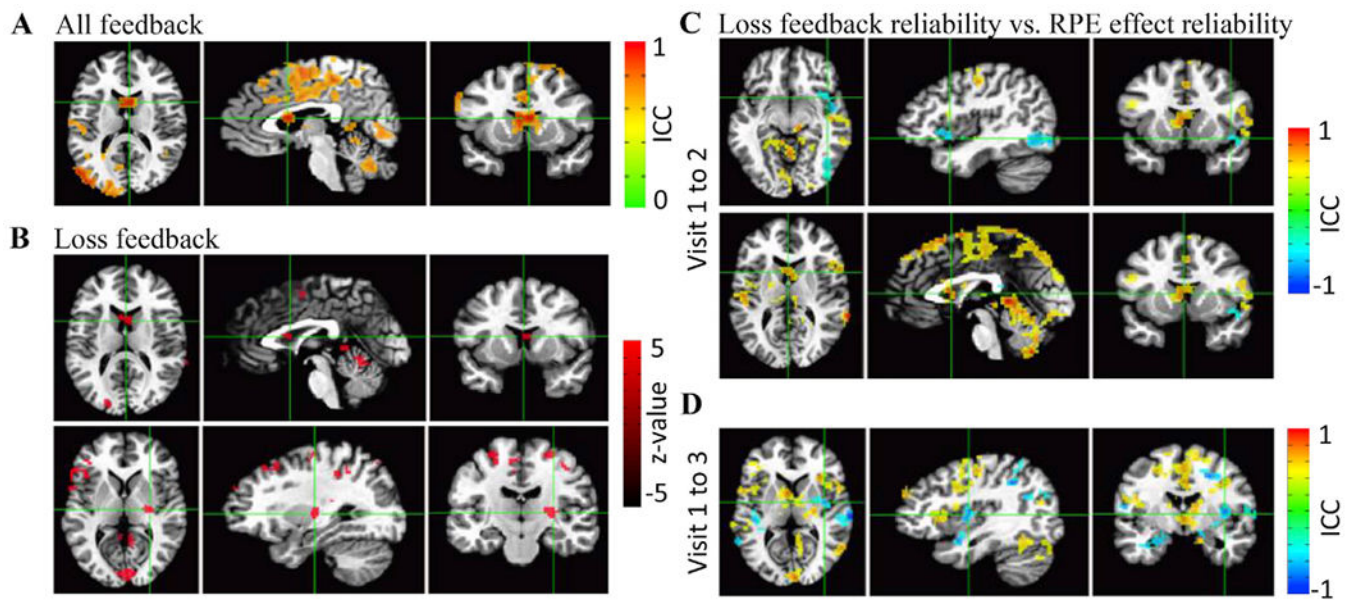
**Fig. 4.** RPE encoding. Voxel-wise group effect maps presenting the parametric modulation values of: (A) Visit 1, corrected for FWE with a threshold of  $p = 0.005$  and a minimal cluster size of 102 voxels; (B) Visit 2, upper panel is corrected similarly and the lower panel is with a non-corrected  $p = 0.005$ ; (C) Visit 3, using an uncorrected  $p = 0.005$  (upper panel), or an uncorrected  $p = 0.05$  (lower panel). (D) Mean ROI RPE encoding levels, across all voxels per region, averaged across all subject per visit (error bars depict the standard error). The effect size for each ROI (Cohen's  $d$ , peak beta value): right insula: visit 1 ( $-1.17$ ,  $-0.09$ ),

visit 2 (-2.03, -0.12), visit 3 (-1.24, -0.08); left insula: visit 1 (-1.41, -6.62), visit 2 (-3.11, -0.12), visit 3 (-1.25, -0.083); striatum: visit 1 (1.49, 0.074), visit 2 (1.81, 0.046), visit 3 (1.44, 0.05).



**Fig. 5.** Reliability of RPE signals. (A, B) ICC maps (top panels), thresholded for FWE and  $ICC > 0.45$  and the respective maps of significant ( $p < 0.05$ ) ICC z scores (lower panels). (C) Mean RPE encoding over reliable insular voxels across visits, and (D) shows the within-subject comparison of RPE encoding (where each data point is mean RPE signal value, of one subject, in two of the visits).





**Fig. 6.**

(A) Reliability of general feedback activation, with no RPE modulation, across the three time points. (B) Reliability of only loss trials feedback, transformed to z-score ( $p = 0.01$ ) across visit 1 to 2 (upper panel) and visit 1 to 3 (lower panel). Under this more stringent threshold only striatum remains reliable. (C) Contrast images of loss feedback reliability versus RPE encoding reliability, visit 1 to 2 and for visit 1 to 3 in panel (D).

**Table 1**

Group characteristics and demographics.

	Visit 1	Visit 2	Visit 3
Number of subjects	22	16	14
Mean Age (SD)	10.61 (0.32)	10.78 (0.37)	13.61 (0.56)
Gender, n females	13	11	9

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

RPE values with a fixed success probability model.

<b>Outcome</b>	<b>Reward Prediction Error (RPE)</b>
Winning reward of 0 stars	$0 - (0 \times 0.66) = 0$
Winning reward of 1 stars	$1 - (1 \times 0.66) = 0.34$
Winning reward of 2 stars	$2 - (2 \times 0.66) = 0.68$
Winning reward of 4 stars	$4 - (4 \times 0.66) = 1.36$
Missing reward of 0 stars	$0 - (0 \times 0.66) = 0$
Missing reward of 1 stars	$0 - (1 \times 0.66) = -0.66$
Missing reward of 2 stars	$0 - (2 \times 0.66) = -1.32$
Missing reward of 4 stars	$0 - (4 \times 0.66) = -2.64$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript