Cognitive Strategies Regulate Fictive, but not Reward Prediction Error Signals in a Sequential Investment Task

Xiaosi Gu,[1,2] Ulrich Kirk,[3] Terry M. Lohrenz,[2] and P. Read Montague[1,2,4]*

[1]*Wellcome Trust Centre for Neuroimaging, University College London, London, United Kingdom*
[2]*Human Neuroimaging Laboratory, Virginia Tech Carilion Research Institute, Roanoke, Virginia*
[3]*Institute of Psychology, University of Southern Denmark, Odense, Denmark*
[4]*Department of Physics, Virginia Polytechnic Institute and State University, Blacksburg Virginia*

◆━━━━━━━━━━━━━━━━━━━━━━━━◆

Abstract: Computational models of reward processing suggest that foregone or fictive outcomes serve as important information sources for learning and augment those generated by experienced rewards (e.g. reward prediction errors). An outstanding question is how these learning signals interact with top-down cognitive influences, such as cognitive reappraisal strategies. Using a sequential investment task and functional magnetic resonance imaging, we show that the reappraisal strategy selectively attenuates the influence of fictive, but not reward prediction error signals on investment behavior; such behavioral effect is accompanied by changes in neural activity and connectivity in the anterior insular cortex, a brain region thought to integrate subjective feelings with high-order cognition. Furthermore, individuals differ in the extent to which their behaviors are driven by fictive errors versus reward prediction errors, and the reappraisal strategy interacts with such individual differences; a finding also accompanied by distinct underlying neural mechanisms. These findings suggest that the variable interaction of cognitive strategies with two important classes of computational learning signals (fictive, reward prediction error) represent one contributing substrate for the variable capacity of individuals to control their behavior based on foregone rewards. These findings also expose important possibilities for understanding the lack of control in addiction based on possibly foregone rewarding outcomes. *Hum Brain Mapp 00:000–000, 2013.* © 2013 Wiley Periodicals, Inc.

Key words: decision-making; reward prediction errors; fictive learning; emotion regulation; reappraisal; insula; fMRI

◆━━━━━━━━━━━━━━━━━━━━━━━━◆

INTRODUCTION

Recent computational models and experimental probes support the notion of multiple learning mechanisms in healthy individuals [Chiu et al., 2008; Daw et al., 2011; Glascher et al., 2010; Lohrenz et al., 2007; Montague et al., 2004, 2006; Pagnoni et al., 2002; Simon and Daw, 2011]. Reward prediction errors derived from ongoing differences between expected and actually experienced rewards (temporal difference (TD) errors) have a significant impact on choice behavior [Montague et al., 2004, 2006; Schultz et al., 1997]. However, these signals do not fully capture the complexity of decision-making processes. Recent evidence demonstrates that learning can also be driven by fictive errors derived from foregone outcomes ("what might have happened") [Chiu et al., 2008; Hayden et al., 2009; Lohrenz et al., 2007]. One central physical substrate supporting these mechanisms is dopaminergic signaling in the brain [Niv et al., 2005; Rangel et al., 2008]. In healthy individuals, both fictive [Lohrenz et al., 2007] and reward prediction errors [Montague et al., 2002] activate the striatum [Chiu et al., 2008; Montague et al., 2002], a dopaminoceptive structure that is commonly implicated in decision-making tasks and works closely with a network of brain regions such as the anterior insular cortex (AIC), orbitofrontal cortex (OFC), and the amygdala [Hsu et al., 2005; Li et al., 2011; Seymour et al., 2004].

An outstanding question is how these computational learning signals interact with top-down cognitive influences [Dayan et al., 2000; Montague et al., 2004]. In people with compromised top-down control (e.g. addicted individuals), fictive errors are computed in the brain, but fail to emerge as signals to guide choice behavior [Chiu et al., 2008], suggesting that fictive learning signals might interact with top-down cognitive input. On the other hand, cognitive strategies used to regulate emotions such as reappraisal have been shown to modulate neural activity related to reward anticipation [Delgado et al., 2008; Staudinger et al., 2011], loss aversion [Sokol-Hessner et al., 2013; Sokol-Hessner et al., 2009], and risky choices [Martin and Delgado, 2011] during decision-making. With the consideration of individual differences in learning and decision-making [Chiu et al., 2008; Daw et al., 2011; Glascher et al., 2010; Lohrenz et al., 2007; Montague et al., 2004, 2006; Pagnoni et al., 2002; Simon and Daw, 2011], it is therefore important to investigate the interplay between emotional regulation strategies and computational learning signals in individual decision-makers, which might open a window into intervention and treatment of psychiatric conditions with abnormal decision-making patterns.

In the current study, we investigated the impact of cognitive influences implemented through an emotion regulation strategy on fictive and reward prediction error signals in healthy adults. We employed a sequential investment task (Fig.1a; modified from [Chiu et al., 2008; Lohrenz et al., 2007]) and functional magnetic resonance imaging (fMRI), together with a cognitive reappraisal strategy, to frame subjects on the overall earnings based on their decisions ("Regulate"; see Materials and Methods and Supporting Information), in comparison to a control strategy focused on each local decision ("Attend"). Similar cognitive strategies have been proven successful in regulating emotions in both clinical [Kober et al., 2010; Volkow et al., 2010] and nonclinical settings [Delgado et al., 2008; Gross, 1998; Ochsner et al., 2012; Wager et al., 2008]. In the current sequential investment task, the Regulate strategy could modulate (1) neither fictive nor reward prediction errors, (2) both fictive and reward prediction errors, (3) fictive signals only, (4) reward prediction errors only. Our working hypothesis was that the Regulate scenario would be associated with a diminished impact of the fictive error only because of its status as a learning signal generated by foregone choices ("what might have been" had decisions been different) and its reported vulnerability to changes in psychophysiological states [Chiu et al., 2008]. Under this hypothesis, we should expect an attenuation of the weight of fictive errors on a subject's next bet and this attenuation should be accompanied by reduced neural responses to fictive errors but not reward prediction errors (the "experienced" errors in our description above). We also explored individual differences in fictive and reward prediction error learning signals.

MATERIAL AND METHODS

Participants

Seventy healthy adults were recruited from community populations in Virginia ($N = 63$) and London ($N = 7$). Seven participants recruited in Virginia were excluded due to excessive head motion (>4 mm), yielding a final sample of 63 healthy adults (age mean ± standard deviation (SD): 32 ± 13 years; 34 females; 56 from Virginia and 7 from London) with normal or adjusted to normal vision, no contraindication to MRI, and reported no previous or current psychiatric or neurological conditions. Subjects were informed of the study requirements and provided written consent prior to participation. The study was approved by the Institutional Review Board of Virginia Tech and the University College London ethics committee.

Stimulus and Procedure

Participants performed a sequential investment task (Fig. 1a) under two task conditions: a cognitive reappraisal task "Regulate," and a control task "Attend" (see Supporting Information for a full description). There were ten "Regulate" markets and ten "Attend" markets, presented in a randomized order for each participant. The Regulate scenario focused the subjects on their entire sequence of choices and the overall performance (e.g. "... Remind yourself that you are making many of these similar decisions. Do not keep a running total—simply approach these
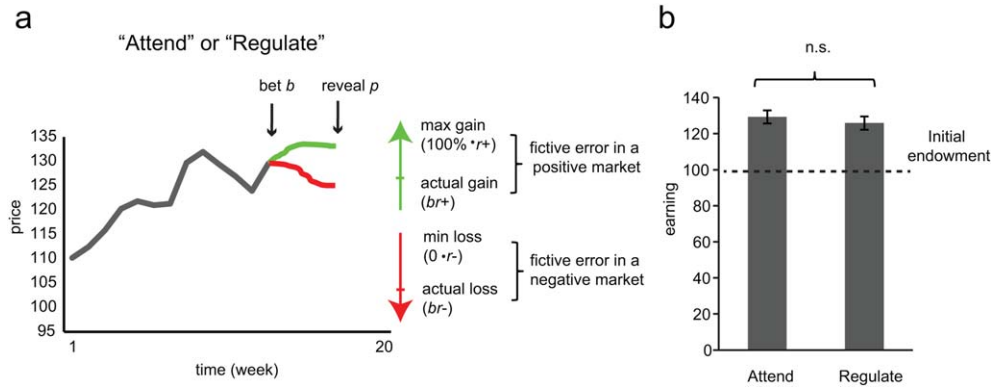
Figure 1.

a) Experimental paradigm. Participants made investment choices under two task conditions: a cognitive reappraisal task "Regulate," and a control task "Attend." At each time point, the subject makes an investment decision b (0–100% of current portfolio). The market price p is then revealed with two possible directions of outcomes: increase or decrease. In a positive market where market return $r^+ > 0$, the best fictive outcome would be having invested 100%, therefore the fictive error in a positive market $f^+ = 100\% \cdot {}^+ - br^+$. Conversely, when the market return is negative ($r^- < 0$), the best fictive outcome would be having invested nothing, therefore the fictive error in a negative market $f^- = 0 \cdot {}^- - br^-$. **b)** There was no significant difference in overall earning between Attend and Regulate strategies ($N = 63$). Each subject was endowed with 100 money units to invest. n.s., not significant. Error bars represent standard error.

investment decisions keeping in mind their context"); it also requires the subjects to take the perspective of a trader (e.g. "…You take risks with money every day, for a living. All that matters is that you come out on top in the end - a loss or gain here or there will not matter in terms of your overall portfolio. In other words, you win some and you lose some"). Such strategies have been shown to reduce loss aversion in laboratory settings [Sokol-Hessner et al., 2009, 2013] and relate to trader performance in real-life investment scenarios [Fenton-O'Creevy et al., 2011]. In contrast, the Attend strategy (see [Sokol-Hessner et al., 2009, 2013]) clearly tells the subject that every decision counts and that she should take the perspective of her own.

We used twenty historic stock markets, similar to the ones used in previous studies (see [Chiu et al., 2008; Lohrenz et al., 2007]). There were 10 "Regulate" markets and 10 "Attend" markets, presented in a randomized order for each participant. Each market was considered a task block; there were 20 events (i.e. twenty investment decisions) in each market, yielding a total of 400 trials (200 trials for "Attend" and 200 trials for "Regulate"). The average task duration was approximately 35 min (ranging from 20 to 54 min; mean ± SD: 2,077 ± 460 s). Detailed task instructions were given to the subjects upon their arrival (outside of the scanner), and were repeated right before the actual scanning started (inside the scanner). During scanning, subjects saw a screen with the word "Attend" or "Regulate" which indicated the task for each market/block at the beginning of each market.

Participants were informed that they would have $100 US Dollars (Virginia) or £100 British Pounds (London) as their initial portfolios (i.e. total amount of money they

have at a given time point) at the beginning of the experiment, and were informed that their final payment would be scaled according to their score in the experiment. No payment was given to the subjects before the experiment. At each time point t, the subject used a two-button box to move a slide bar to make an investment decision b_t (0–100% of current portfolio) without a time constraint. Their mean response times ranged from 600 ms to 1,702 ms (mean ± SD: 1,075 ± 261 ms); 750 ms after they submitted their choices, the market price p_t was revealed and the fractional market price change and subjects' portfolio were updated. Market information for all previous segments then remained on the screen. The slide bar then changed from gray to red after another 750 ms, and subjects started to make investment decisions for the next market segment. There are two possible directions of outcomes: increase or decrease. Trials with increased or decreased market price compared with the previous time point were considered positive and negative markets, respectively. The market return r_t equals $(p_t - p_{t-1})/p_{t-1}$ and the gain is defined as $g_t = b_t \cdot r_t$. In a positive market where $p_t - p_{t-1} > 0$, $r_t^+ < 0$, and positive gain $g_t^+ = b_t \cdot r_t^+$, the best fictive outcome would be having invested 100%, therefore, the positive fictive error $f^+ = 100\% \cdot r_t^+ - b_t \cdot r_t^+$ (or $f^+ = r_t^+ - g_t^+$). Conversely, when the market return is negative ($p_t - p_{t-1} < 0$, $r_t^- < 0$, and negative gain $g_t^- = b_t \cdot r_t^-$), the best fictive outcome would be having invested nothing; therefore, the fictive error in a negative market $f^- = 0 \cdot r_t^- - b_t \cdot r_t^-$ (or $f^- = - g_t^-$). We will focus our main analyses on f^+ based on previous research suggesting that in healthy participants, investment behavior is mostly driven by f^+, but not f^- [Chiu et al., 2008; Lohrenz et al., 2007].

Behavioral Data Analysis

We regressed the subject's next bet (b_{t+1}) against previous bet (b_t), positive and negative market return (r_t^+ and r_t^- respectively), and the interaction terms ($b_t r_t^+$ and $b_t r_t^-$) simultaneously cross task conditions by coding Attend and Regulate as two indicator variables (*regstats* function in MATLAB, R2012a, The MathWorks, Inc, Natick, MA):

$$\tilde{b}_{t+1} = \beta_0 + \beta_1 \cdot (\text{Attend}) + \beta_2 \cdot (\text{Attend} \cdot \tilde{b}_t) + \beta_3 \cdot (\text{Attend} \cdot r_t^+) + \beta_4 \cdot (\text{Attend} \cdot r_t^-) + \beta_5 \cdot (\text{Attend} \cdot b_t r_t^+) + \beta_6 \cdot (\text{Attend} \cdot b_t r_t^-) + \beta_7 \cdot (\text{Regulate}) + \beta_8 \cdot (\text{Regulate} \cdot \tilde{b}_t) + \beta_9 \cdot (\text{Regulate} \cdot r_t^+) + \beta_{10} \cdot (\text{Regulate} \cdot r_t^-) + \beta_{11} \cdot (\text{Regulate} \cdot b_t r_t^+) + \beta_{12} \cdot (\text{Regulate} \cdot b_t r_t^-)$$

Here $\tilde{b}_t$ is the within-subject z-normalized bet. Because the data were pooled across subjects, and because we are primarily interested in the within-subject trial-to-trial fluctuations rather than the subject's general level of bet, we z-normalized the bets within subject so that they are comparable among subjects. Differences between regression coefficients of two task conditions were tested by performing linear hypothesis tests of the regression coefficients (F tests) using *linhyptest* in MATLAB.

We also assessed the influence of TD errors on subjects' next investment choice with a multiple regression model with the previous bet (b_t) and TD as regressors:

$$\tilde{b}_{t+1} = \beta_0 + \beta_1 \cdot (\text{Attend}) + \beta_2 \cdot (\text{Attend} \cdot \tilde{b}_t) + \beta_3 \cdot (\text{Attend} \cdot \text{TD}_t) + \beta_4 \cdot (\text{Regulate}) + \beta_5 \cdot (\text{Regulate} \cdot \tilde{b}_t) + \beta_6 \cdot (\text{Regulate} \cdot \text{TD}_t)$$

where $\text{TD}_t = \tilde{g}_t - \tilde{b}_t$, and $\tilde{g}_t$ is within-subject z-normalized gain/loss. In other words, TD represents the difference between the actual gain at time t and the expected gain at that time, which corresponds to the bet. This definition of TD requires the gain to be comparable in scale to the bet; therefore, we also z-transformed the gain term, considering we already z-scored the bet term. Statistical significance was determined at $P < 0.05$ two-tailed.

We also explored individual differences in the extent to which fictive error signals influence investment decisions both behaviorally and neurally in a post hoc analysis. As established previously, the fictive error f^+ is constructed as the difference between the maximum fictive gain $100\% \cdot r_t^+$ and the actual gain $b_t \cdot r_t^+$. Therefore, for f^+ to exert an impact on one's behavior, two criteria must be met: (1) the influence of r_t^+ on the next bet must be positive, and (2) the influence of $b_t r_t^+$ on the next bet must be negative. Considering that the impact of r_t^+ on the next bet is positive in most individuals (3 out of 63 individuals showed negative beta weights of r_t^+ and were excluded from this analysis), we identified different learning strategies by examining the direction of the impact of $b_t r_t^+$ on the next bet for each individual. We identified two groups of subjects groups as fictive learners (*f*-learners, $N = 31$) and temporal difference learners (TD-learners, $N = 29$; Fig. 4) respectively: *f*-learners' next bets were negatively driven by $b_t r_t^+$, and therefore significantly influenced by f^+ while TD-learners' behavior were positively driven by $b_t r_t^+$ (see Results for details).

Image Acquisition and Preprocessing

The anatomical and functional imaging was conducted on two identical 3.0 Tesla Siemens Trio scanners in Virginia and one 3.0 Tesla Siemens Trio scanner in London. Scanner and country were coded as nuisance variables at the second level fMRI analysis. High-resolution T1-weighted scans (1.0 mm $\times$ 1.0 mm $\times$ 1.0 mm) were acquired using an MP-RAGE sequence. Functional images were acquired using echo-planar imaging (EPI), and angled $30°$ with respect to the anterior-posterior commissural line. The detailed settings for the functional imaging were: repetition time (TR) = 2,000 ms; echo time (TE) = 25 ms; flip angle = $90°$; 37 slices; voxel size: 3.4 mm $\times$ 3.4 mm $\times$ 4.0 mm. The functional scans were adjusted for slice timing, realigned to the first volume, coregistered to the T1 image, normalized to a standard template (MNI, Montreal Neurological Institute), and spatially smoothed with an $8 \times 8 \times 8$ mm full-width-at-half-maximum (FWHM) Gaussian kernel.

fMRI Data Analysis

Event-related analyses of the fMRI data were conducted using statistical parametric mapping (SPM8; Wellcome Department of Imaging Neuroscience, London, UK). General linear modeling (GLM) [Friston et al., 1995] was conducted for the functional scans from each participant by modeling the observed event-related blood-oxygen-level dependent (BOLD) signals and regressors to identify the relationship between the task events and the hemodynamic response. Regressors of 0 s duration related to visual and motor events were created by convolving a train of delta functions representing the sequence of individual events with the default SPM basis function, which consists of a synthetic hemodynamic response function (HRF) composed of two gamma functions [Friston et al., 1998]. The regressors include: market type screen; initial market history screen; key press; Attend: reveal of market price of first round; Attend: reveal of rounds 2 to 19; Attend: reveal of market price of last round; Regulate: reveal of market price of first round; Regulate: reveal of rounds 2 to 19; Regulate: reveal of market price of last round. Six parameters generated during motion correction were entered as covariates. TD and fictive errors (f^+, f^-) were entered as parametric regressors at the onsets of revealing the market prices. Although the lack of time constraints in the decision period and the lack of jitter may compromise the efficiency of this task for fMRI, using TD and fictive

errors derived from model-based approaches as parametric modulation could potentially increase the efficiency compared with contrast-based approaches. Linear contrasts of the parameter estimates were made to identify the effects of temporal difference errors and fictive errors (TD, f^+, f^-) under Attend and Regulate, and their differences, for each participant. These images from all participants were then entered into a second-level group analysis conducted with a random-effects statistical model. One-sample t-tests were conducted for effects common to all participants and two-sample t-tests for group comparisons between the two types of learners. Significant activations related to the effects of TD errors under Attend are Regulate were identified at $P < 0.05$ level corrected for family-wise errors (FWE). All other activations were identified with a height threshold of P value exceeding 0.005 uncorrected in conjunction with an extend threshold of 10 voxels (resampled as $2 \times 2 \times 2$ mm) to maintain a balance between Type I and Type II errors (Lieberman and Cunningham, 2009]. Unbiased regions of interest (ROIs) were created using the MarsBaR toolbox (http://marsbar.sourceforge.net/) based on the main effect of fictive error averaged cross both task conditions. These ROIs include left anterior insular cortex (centered at $[-34, 16, -12]$), right lateral orbitofrontal coertex (LOFC; centered at $[44, 22, -8]$), right medial orbitofrontal cortex (MOFC; centered at $[-4, 32, -16]$), and left striatum (centered at $[-16, -4, 14]$), with 4 mm radius. Individual subject's parameter estimates were then extracted from each ROI for each task condition.

Psychophysiological Interaction (PPI) Analysis

We conducted PPI analysis [Friston et al., 1997; Gitelman et al., 2003] to explore such the functional connectivity between AIC and other regions under the modulation of the reappraisal strategy. The bilinear term in PPI represents the interaction between physiological activity and a psychological context input, which modulates the connectivity between the seed voxel of interest (VOI) and other brain regions, and has a directional character [Stephan et al., 2004]. The time series data of the first eigenvariate of the left AIC seed VOI derived from the ROI analysis were temporally filtered and mean corrected as in conventional SPM analysis. Bayesian estimation was used to deconvolve the time series of the BOLD signal to generate the time series of the neuronal signal for the VOI. The time series of the neuronal signal for responses to the events were created, resulting in one vector (the PPI regressor) representing the interaction between the reappraisal strategy and the AIC VOI (the psychophysiological interaction variable), a second vector (the P regressor) representing the contrast of fictive error-related Attend versus Regulate difference (the psychological variable), and a third vector (the Y regressor) representing the AIC VOI time course (the physiological variable). These regressors were forward-convolved with

the canonical HRF, and then entered into the regression model along with vectors for other events. Model estimation was performed and the resulting SPM showed areas with significant differential connectivity to the VOIs due to context manipulations. The PPI analysis was carried out for each subject and the resulting images of contrast estimates were entered into a random effects group analysis. The statistical significance was set at a height threshold of P value exceeding 0.005 uncorrected in conjunction with an extend threshold of 10 voxels.

RESULTS

Behavioral Modulatory Effect on Fictive Errors

The reappraisal strategy ("Regulate") did not significantly change overall earning compared with the control task ("Attend") ($P > 0.05$; Fig. 1b). Subjects' raw investment levels were higher in the Regulate condition (mean ± SD: $45 \pm 15\%$) compared with the Attend condition (mean ± SD: $42 \pm 16\%$; paired t-test $P = 0.014$), consistent with previous finding on reduced loss/risk aversion under a similar reappraisal strategy [Sokol-Hessner et al., 2009]. The normalized bets were not significantly different between conditions ($P > 0.9$). In the Attend condition, the previous bet b_t, positive and negative market return r_t^+ and r_t^-, and the positive interaction term $b_t r_t^+$ significantly predicted the next bet b_{t+1} (all $Ps < 0.001$; Table I). These results are consistent with previous results that fictive gain signals guided behavior in healthy adults [Chiu et al., 2008; Lohrenz et al., 2007].

Importantly, when participants were asked to perform the Regulate strategy, $b_t r_t^+$ no longer significantly predicted the next bet b_{t+1} ($P > 0.05$; Fig. 2a), while the regression coefficients of b_t, r_t^+, r_t^- (all $Ps < 0.001$) and the negative interaction term $b_t r_t^-$ ($P < 0.05$), were significant (Table I). We then tested the significance of differences between these regression coefficients under Attend and Regulate (Table II). Only the regression coefficients for $b_t r_t^+$ were significantly different between Attend and Regulate ($P = 0.007$; Fig. 2a). Direct comparison between Attend and Regulate using individual betas showed a similar pattern: only the betas of the $b_t r_t^+$ term were attenuated under Regulate (paired-t test $P = 0.06$; all other $Ps > 0.1$). Fictive errors were correlated with TD errors under the Attend ($r = 0.35$, $P < 0.01$), but not Regulate condition ($r = 0.11$, $P > 0.3$); the difference between these two correlation coefficients was not significant ($P > 0.1$). There was no significant effect of site for any of these behavioral coefficients (all $Ps > 0.5$). These results suggest the reappraisal strategy significantly attenuated the influence of fictive errors on investment behavior.

Neural Modulatory Effect on Fictive Errors

In parallel with the behavioral findings, we identified robust fictive error f^+ related activation in the striatum,

TABLE I. Behavioral regression results

	Regressor	Coefficient	SE	t value	P
Model 1	Attend	1.98	7.66	−0.26	0.79
	Attend · b	0.62	0.01	71.30	<0.0001
	Attend · r+	4.93	0.29	17.29	<0.0001
	Attend · r−	−3.29	0.28	−11.94	<0.0001
	Attend · br+	−2.96	0.58	−5.07	<0.0001
	Attend · br−	−0.15	0.55	−0.27	0.78
	Regulate	1.98	7.66	0.26	0.80
	Regulate · b	0.63	0.01	73.58	<0.0001
	Regulate · r+	4.23	0.31	13.83	<0.0001
	Regulate · r−	−3.64	0.31	−11.93	<0.0001
	Regulate · br+	−0.69	0.60	−1.14	0.25
	Regulate · br−	−1.42	0.60	−2.38	0.02
Model 2	Attend	2.08	7.76e+10	0.27	0.79
	Attend · b	0.81	0.01	79.87	0.01
	Attend · TD	0.23	0.008	30.12	0.01
	Regulate	2.08	7.76e+10	0.27	0.79
	Regulate · b	0.86	0.01	84.01	<0.0001
	Regulate · TD	0.25	0.008	32.65	<0.0001

Model 1: multiple regression model with next bet as dependent variable and the following regressors: b, investment decision; $r+$, positive market return; $r−$, negative market return; $br+$, actual gain in positive markets; $br−$, actual loss in negative markets. Model 2: multiple regression model with next bet as dependent variable and the following regressors: b; TD, temporal difference error. SE, standard error.

AIC, lateral, and medial orbitofrontal cortices (LOFC and MOFC), and midbrain nuclei (substantia nigra and red nucleus), in the Attend condition (Fig. 2b and Supporting Information Table S1; $P < 0.005$ uncorrected and $k > 10$). Importantly, these activations were attenuated in the Regulate condition (Fig. 2b and Supporting Information Table S2). A direct comparison between Attend and Regulate showed significant attenuation in f^+ related activity in left AIC (extending into putamen) and inferior frontal gyrus (IFG) including LOFC (Supporting Information Table S3). The reverse contrast did not yield any significant activation, suggesting that Regulate did not enhance f^+ related activity compared with Attend. There was no significant effect of site in the striatum, insula, or IFG/OFC.

We further conducted unbiased ROI analysis based on seeds selected from the main effect of fictive error averaged across both task conditions (Fig. 2c and Supporting Information Table S4), including AIC (centered at [−34, 16, −12]), LOFC (centered at [44, 22, −8]), MOFC (centered at [−4, 32, −16]), and striatum (centered at [−16, −4, 14]). Paired t-test suggests that left AIC activity showed a significant reduction under Regulate, compared with Attend ($P < 0.05$). Although striatum and other ROIs showed a similar pattern, the differences between Regulate and Attend did not reach statistical significance (all Ps > 0.05). These results suggest that the attenuation in the weight of fictive errors on behavior under the Regulate condition is predominantly accompanied by reduced activation in AIC.
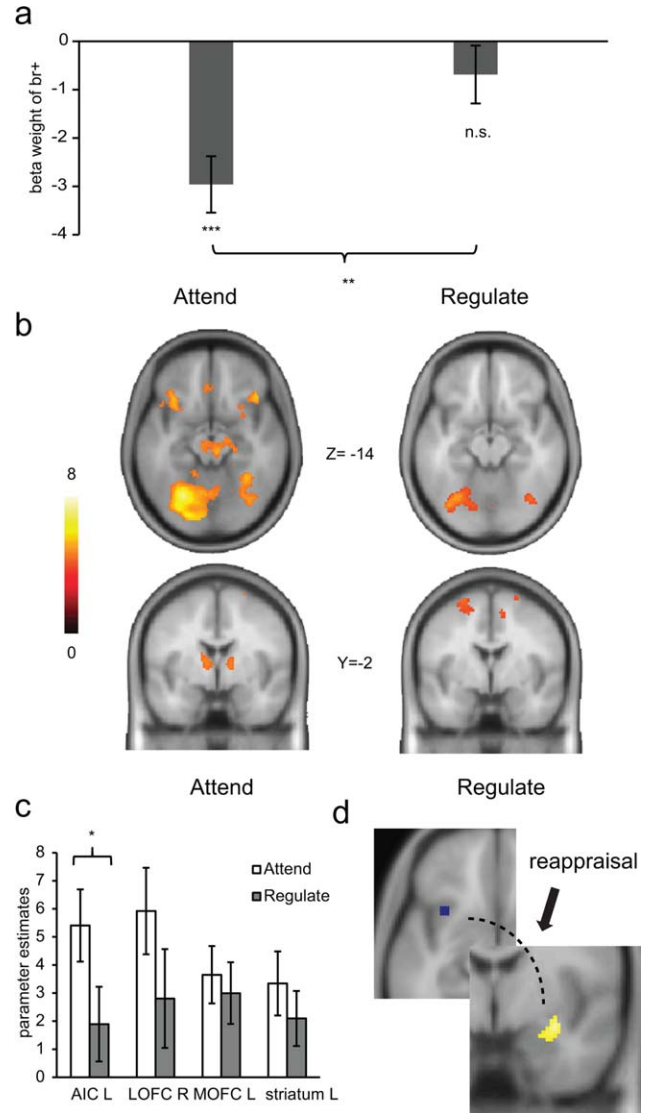
Figure 2.

Behavioral and neural modulatory effects on fictive errors ($N = 63$). **a)** The reappraisal strategy significantly attenuated the beta weights of br^+ on the next bet. **b)** Fictive errors f^+ activate the striatum, anterior insular cortex (AIC), lateral (LOFC) and medial orbitofrontal (MOFC) cortices, and midbrain nuclei, in the Attend, but not in the Regulate condition ($P < 0.005$ uncorrected). **c)** Region of interest analysis suggests that most robust f^+-related task difference is in left AIC. **d)** Functional connectivity analysis suggests that AIC-amygdala connectivity is decreased by the reappraisal strategy ($P < 0.005$ uncorrected). Blue dot indicate the seed voxels in the left AIC utilized in the PPI analysis. Decreased connectivity in the right amygdala is displayed on coronal slice. ***$P < 0.001$, ** $P < 0.01$, *$P < 0.05$; n.s., not significant. L, left. br^+: interaction term of bet (b) and positive market return (r^+). Error bars represent standard error.

TABLE II. Linear contrasts of regression coefficients

	Contrast	F value	P
Model 1	Attend · b vs. Regulate · b	0.99	0.32
	Attend · r+ vs. Regulate · r+	2.82	0.09
	Attend · r− vs. Regulate · r−	0.70	0.40
	Attend · br+ vs. Regulate · br+	7.37	0.007
	Attend · br− vs. Regulate · br−	2.44	0.12
Model 2	Attend · b vs. Regulate · b	8.50	0.004
	Attend · TD vs. Regulate · TD	3.67	0.06

Model 1: multiple regression model with next bet as dependent variable and the following regressors: b, investment decision; $r+$, positive market return; $r-$, negative market return; $br+$, actual gain in positive markets; $br-$, actual loss in negative markets. Model 2: multiple regression model with next bet as dependent variable and the following regressors: b; TD, temporal difference error. SE, standard error.

We then explored changes in functional connectivity parameters using the same AIC ROI as our seed region (centered at $[-34, 16, -12]$). Relative to Attend, the Regulate strategy significantly decreased the functional connectivity between AIC and the amygdala (Fig. 2d and Supporting Information Table S4). Regulate did not enhance the functional connectivity between AIC and other regions at the same threshold ($P < 0.005$ uncorrected, $k > 10$). The amygdala has been implicated in a wide range of processes involving aversive emotions [Delgado et al., 2008; LaBar et al., 1998; Sokol-Hessner et al., 2013]. Therefore, these results together suggest that negative feelings associated with fictive error signals were likely to be reduced in the Regulate condition.

Behavioral and Neural Modulatory Effects on Reward Prediction Errors

We then examined the effects of reappraisal on reward prediction errors represented by temporal difference (TD_t), where TD_t is computed as the ongoing difference between the z-scored gained reward $\bar{g}_t$ and the expected reward $\bar{b}_t$ (i.e. $TD_t = \bar{g}_t - \bar{b}_t$, where $g_t = b_t \cdot r_t$). A behavioral regression model was carried out with normalized bet $\bar{b}_t$ and TD_t as predictors against the next bet $\bar{b}_{t+1}$. TD_t significantly predicted the next bet under both Attend and Regulate conditions (both $Ps < 0.001$; Table I). The beta coefficients of TD_t did not differ between Attend and Regulate ($P > 0.05$; Fig. 3a and Table II), suggesting that the reappraisal strategy did not significantly modulate the influence of reward prediction errors on choice behavior. Direct comparison between Attend and Regulate using individual betas showed a similar pattern: the betas of the TD_t term were not significantly different (paired t-test $P > 0.1$). There was no significant effect of site for these behavioral coefficients (all $Ps > 0.5$).

Neurally, both Attend and Regulate yielded significant activation in the striatum, OFC, and other related brain regions (Fig. 3b and Supporting Information Tables S6 and S7; $P < 0.05$ corrected for family-wise errors and $k > 5$). A direct comparison between conditions did not reveal significant differences in TD-related activation in the AIC, striatum or other regions of interest ($P < 0.005$ uncorrected and $k > 10$; Supporting Information Table S8), although middle temporal/hippocampal activation was increased under Regulate. Direct comparison of individual parameter estimates of ROIs (AIC, LOFC, MOFC, and striatum) confirmed the GLM results: there was no significant difference in TD-related neural activity between Attend and Regulate in our regions of interest (all $Ps > 0.2$). There was no significant effect of site in the striatum, insula, or IFG/OFC. Taken together, these behavioral and neural findings suggest that TD errors were not significantly modulated by the reappraisal strategy in the current investment task.

Individual Differences: f-Learners and TD-Learners

In a post hoc exploratory analysis on individual differences, we identified two groups of subjects groups as fictive learners (f-learners, $N = 31$) and temporal difference learners (TD-learners, $N = 29$; Fig. 4) respectively: f-learners' next bets were negatively driven by $b_t r_t^+$, and therefore significantly influenced by f^+ while TD-learners'
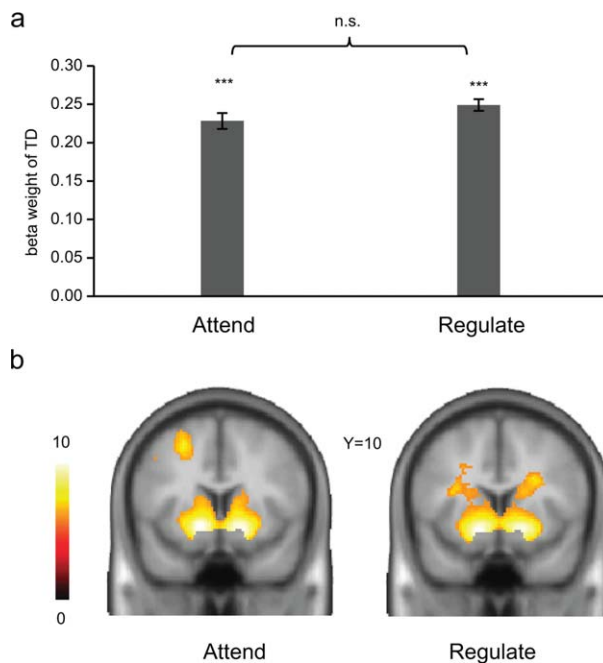

Figure 3.

Behavioral and neural modulatory effects on reward prediction errors ($N = 63$). **a)** The reappraisal strategy did not change the beta weights of temporal difference (TD) errors on the next bet. **b)** TD errors activate the striatum and other reward-related brain regions in both Attend and Regulate conditions ($P < 0.05$ corrected for family-wise error). ***$P < 0.001$; n.s.: not significant. Error bars represent standard error.
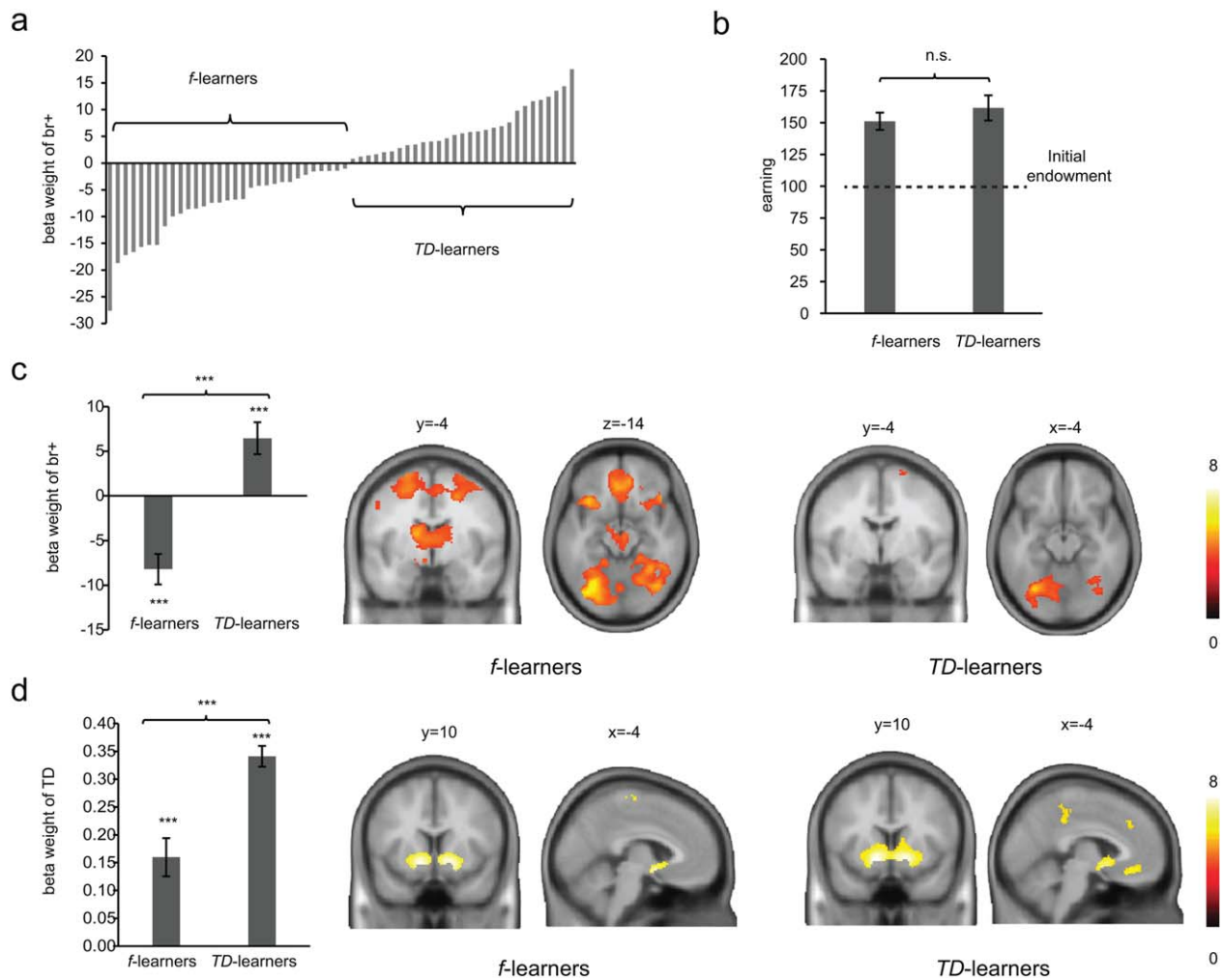
Figure 4.

Individual differences in fictive and reward prediction learning. **a)** Fictive learners (*f*-learners, $N = 31$) show negative beta coefficients of br^+ (interaction term of bet b and positive market return r^+), while temporal difference learners (TD-learners, $N = 29$) show positive beta weights. **b)** There is no difference in overall earnings between *f*-learners and TD-learners. **c)** In parallel with the behavioral difference between beta weights of br^+ of the two groups (left), *f*-learners, but not TD-learners, show fictive error f^+-related activation in the insula, striatum, orbitofrontal cortex, and other related brain regions ($P < 0.005$ uncorrected). **d)** TD-learners have greater beta coefficients of TD compared with *f*-learners, accompanied by TD-related brain activations in orbitofrontal cortex ($P < 0.05$ corrected for family wise error). ***$P < 0.001$. n.s.: not significant, Error bars represent standard error.

behavior were positively driven by $b_t r_t^+$ (both $Ps < 0.001$, two-sample t-test $P < 0.001$; Fig. 4c). Although both groups' behaviors were significantly driven by TD (both $Ps < 0.001$), TD-learners were more dependent on TD errors compared with *f*-learners (Fig. 4d; two-sample t-tests $Ps < 0.001$). The overall earning did not differ between these two types of learners (Fig. 4b).

In line with their behavioral differences, *f*-learners and TD-learners also showed different patterns of brain activations related to fictive and reward prediction errors (Fig. 4c,d, and Supporting Information Tables S9–S12). *f*-learners

showed robust f^+-related responses in the striatum, OFC and AIC, while these neural responses were absent in TD-learners at the same threshold (Fig. 4c; Supporting Information Table S9). Direct comparison between groups confirmed that *f*-leaners showed greater activation in the AIC and OFC than TD-learners (Supporting Information Table S10). On the other hand, while both *f*-learners and TD-learners showed robust TD-related activation in the striatum and OFC, TD-learners showed significantly stronger TD-related activation than *f*-learners in the OFC (Fig. 4d; Supporting Information Tables S11 and S12).
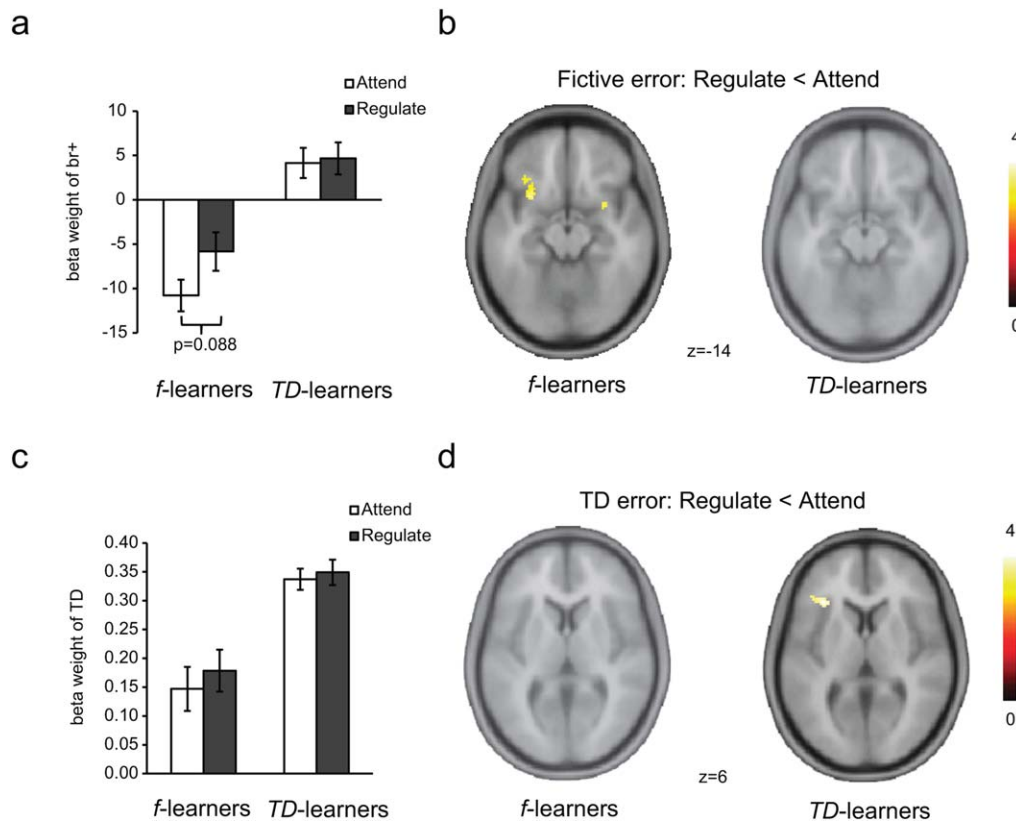
Figure 5.

Learner type and task interaction. **a)** The modulatory effect of the reappraisal strategy on the weight of fictive gain $br+$ over the next bet was mostly driven by f-learners ($P = 0.088$) but not temporal different (TD)-learners ($P > 0.8$). **b)** The Regulate condition reduced fictive error related activation in left anterior insula in f-learners, but not TD-learners ($P < 0.005$ uncorrected, $k > 10$). **c)** The reappraisal strategy did not influence the weight of TD on next bet in either f-learners or TD-learners ($Ps > 0.2$). **d)** The Regulate condition reduced TD error related activation in left anterior insula in TD-learners, but not f-learners ($P < 0.005$ uncorrected, $k > 10$). $br+$: interaction term of bet (b) and positive market return ($r+$). Error bars represent standard error.

We then explored the interaction between learner type and reappraisal (Fig. 5). Behaviorally, although the interaction between task and learner type was not significant ($P > 0.2$), planned comparison suggested a trend of reduced fictive learning in the Regulate condition in f-learners ($P = 0.088$), but not in TD-learners ($P > 0.8$; Fig. 5a). Regulate did not change the impact of TD on investment behavior in either group (Fig. 5c; $Ps > 0.2$). The Regulate condition reduced fictive error-related AIC and OFC activations only in f-learners, but not in TD-learners (Fig. 5b and Supporting Information Table S13). The Regulate condition did not alter TD-related brain activation in f-learners; however, it reduced TD-related AIC and frontal activations in TD-learners (Fig. 5d and Supporting Information Table S14). Together with the behavioral findings, these results suggest that different learning strategies in f-learners and TD-learners are subserved by distinct neural correlates and such individual differences in learning interact with the reappraisal strategy.

DISCUSSION

We provide human neuroimaging evidence demonstrating that fictive errors are more amenable to cognitive strategies such as reappraisal, when compared with reward prediction error signals; and that these learning signals and their interaction with cognitive influences vary among individual decision makers. These findings provide important insight into the dissociation in the nature of fictive and reward prediction error signals, the interaction between cognitive influences and these computational learning signals, and relevance to psychopathology and potential interventions.

Reappraisal Strategy Selectively Interacts With Computational Learning Signals

The primary finding of the current study is the selective coupling between the reappraisal strategy and fictive error

signals. In supervised actor-critic reinforcement learning [Rosenstein et al., 2004], TD errors derived from actually experienced rewards serve as the "critic" that is used to guide decision-making through the behavioral policies implemented by the "actor". Errors derived from fictive outcomes (ongoing differences between what might have been achieved and what was actually achieved) have been considered as an important learning signal that complements classic TD errors, and hence, constitute a second type of "critic" [Chiu et al., 2008; Lohrenz et al., 2007]. It remained unclear how the "actor" balances between these two sources of critics. In the current study, change in cognitive contexts modulates the behavioral and neural correlates of fictive, but not reward prediction errors. We speculate that to implement such modulation, cognitive input selectively biases the weight of fictive errors on behavior through the "actor." This also explains why fictive learning signals fail to guide decision-making in addicted individuals who lack behavioral control [Chiu et al., 2008]. An alternative account for the dissociation between TD learning and fictive learning would be that cognitive strategies modulate expected value signals, possibly at the level of the "actor." This possibility could be supported by increased level of raw bets under the Regulate condition (see Results), where subjects' bets can serve as a proxy of expected value in the current paradigm. It is also consistent with previous finding of modulation of expected rewards of conditioned stimuli [Delgado et al., 2008].

Reward prediction errors, on the other hand, show robust resilience to the reappraisal strategy in the current study. This is in line with a previous finding that the impact of TD errors on behavior is not susceptible to addictive state and remains largely intact in chronic smokers [Chiu et al., 2008]. Although one previous study suggests that emotion regulation strategies can also modulate reward prediction errors [Staudinger et al., 2009], the strategy was different from our reappraisal strategy in that subjects were instructed to "distance" themselves from reward outcomes. Based on the selective coupling between the reappraisal strategy and fictive learning signals, we speculate that while both fictive and reward prediction errors guide decision-making, fictive signals might be more important in allowing behavioral flexibility while reward prediction errors are more robust and resilient to external modulatory factors. Such flexibility associated with fictive learning is important for developing potential intervention and treatment of psychopathology based on cognitive strategies.

level autonomic and interoceptive processes [Craig, 2009; Critchley et al., 2004], to high-level processes such as emotion [Fan et al., 2011; Lamm and Singer, 2010], empathy [Gu et al., 2010, 2012, 2013b; Singer et al., 2004], fairness [Kirk et al., 2011; Sanfey et al., 2003], risk and uncertainty [Bossaerts, 2010; Preuschoff et al., 2008; Ullsperger et al., 2010], trust and cooperation [King-Casas et al., 2008], norm violations [Montague and Lohrenz, 2007; Xiang et al., 2013], and cognitive control [Eckert et al., 2009; Menon and Uddin, 2010]. Therefore, AIC is considered a critical neural substrate in integrating bodily signals with top-down control [Craig, 2009; Singer et al., 2009), potentially in a Bayesian optimal fashion [Gu et al., 2013a; Seth, 2013]. The amygdala is a critical structure in general emotional processing [Pessoa and Adolphs, 2010; Phelps, 2006] as well as Pavlovian learning [Li et al., 2011; Rangel et al., 2008]. Recent work has singled out a computational role of the amygdala in encoding economic uncertainty [Coricelli et al., 2005; De Martino et al., 2006; Hsu et al., 2005]. Importantly, a recent study suggests that economic risk-related amygdala activation in a gambling task is modulated by an emotion regulation strategy that is identical to the one used in the current study [Sokol-Hessner et al., 2013]. The AIC also has dense reciprocal connections with almost all subnuclei of the amygdaloid complex [Mufson et al., 1981], which provides the neuroanatomical basis for functional connectivity between these two structures. Both the AIC and amygdala are known to be involved in rapid information integration, especially when stimuli are salient and relevant [Eckert et al., 2009; Kuo et al., 2009; LeDoux, 2000].Therefore, we speculate that there are at least two aspects of the top-down modulatory effect on the activity and connectivity of the AIC in the current paradigm. Firstly, the AIC could compute the quantity of fictive errors per se; by selectively acting on AIC activity and connectivity, the reappraisal strategy is able to exert influence on the weight of fictive signals on behavior directly. Alternatively, the AIC could encode subjective feelings associated with fictive errors (e.g. regret); by modulating AIC activity and AIC-amygdala connectivity, the reappraisal strategy then modulates subjective feelings associated with fictive signals. These two aspects are possibly intertwined during economic decision-making both in our experimental setting as well as in real-life decision-making; and their dissociation remains to be examined by future studies.

A role of the Anterior Insular Cortex and its Related Networks in Encoding Fictive Errors

In parallel with the behavioral findings, we find that the modulatory effect of cognitive strategies on fictive errors is primarily associated with reduced activity in the AIC and its associated brain regions such as the OFC and the striatum; as well as decreased AIC-amygdala connectivity. The AIC participates in a wide range of functions from low-

Individual Differences in Fictive and Reward Prediction Error Signals

In a post hoc analysis, we also identified individuals with different types of learning mechanisms, namely fictive learners and temporal difference learners; both types of learners have distinct behavioral and neural response patterns. Fictive learners' decisions are significantly driven by fictive outcomes and display robust neural activation

related to fictive errors; the reappraisal strategy modulate these fictive learning signals only in *f*-learners. TD-learners' behaviors are under greater influence of TD errors and show TD-related activation in the striatum and OFC, compared with fictive learners; their neural responses to TD errors are also modulated by the reappraisal strategy although the behavioral modulatory effect is not significant. While these results should be interpreted with caution due to the post hoc nature of the analysis, they could contribute to the existing literature on different learning mechanisms among individuals, such as model-based vs. model-free learning [Daw et al., 2011; Glascher et al., 2010] by demonstrating individual differences in learning from fictive and reward prediction errors under the modified actor-critic model as proposed earlier in the Discussion and as elsewhere [Chiu et al., 2008; Lohrenz et al., 2007]. Under such framework, it is not surprising that individual decision-makers exhibit varying capacities to make decisions based on fictive rewards, and that the same cognitive strategies interact with individual behavioral and neural responses to fictive outcomes differently. These findings could be informative for cognitive intervention and treatment programs of psychopathology by suggesting that individual differences in decision-making and learning should be taken into account.

CONCLUSION

Taken together, our results support the hypothesis that top-down cognitive strategies such as reappraisal can impact learning signals known to guide valuation and choice. Our findings are mainly twofold. First, while both fictive and reward prediction errors serve as important learning signals, only fictive signals are susceptible to cognitive strategies both behaviorally and neurally. Second, we present the first report of individual differences in fictive errors and its interaction with cognitive modulation. Overall, these findings suggest that the variable coupling of cognitive strategies to two important classes of learning signals (fictive, reward prediction error) represent one contributing substrate for the variable capacity of individuals to control their behavior based on foregone rewards. These findings also expose important possibilities for understanding the control or lack of control in addiction based on possibly foregone rewarding outcomes.

REFERENCES

Bossaerts P (2010): Risk and risk prediction error signals in anterior insula. Brain Struct Funct 214:645−653.

Chiu PH, Lohrenz TM, Montague PR (2008): Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. Nat Neurosci 11:514–520.

Coricelli G, Critchley HD, Joffily M, O'Doherty JP, Sirigu A, Dolan RJ (2005): Regret and its avoidance: A neuroimaging study of choice behavior. Nat Neurosci 8:1255–1262.

Craig AD (2009): How do you feel–now? The anterior insula and human awareness. Nat Rev Neurosci 10:59–70.

Critchley HD, Wiens S, Rotshtein P, Ohman A, Dolan RJ (2004): Neural systems supporting interoceptive awareness. Nat Neurosci 7:189–195.

Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011): Model-based influences on humans' choices and striatal prediction errors. Neuron 69:1204–1215.

Dayan P, Kakade S, Montague PR (2000): Learning and selective attention. Nat Neurosci 3(Suppl):1218–1223.

De Martino B, Kumaran D, Seymour B, Dolan RJ (2006): Frames, biases, and rational decision-making in the human brain. Science 313:684–687.

Delgado MR, Gillis MM, Phelps EA (2008): Regulating the expectation of reward via cognitive strategies. Nat Neurosci 11:880–881.

Eckert MA, Menon V, Walczak A, Ahlstrom J, Denslow S, Horwitz A, Dubno JR (2009): At the heart of the ventral attention system: the right anterior insula. Hum Brain Mapp 30: 2530–2541.

Fan J, Gu X, Liu X, Guise KG, Park Y, Martin L, de Marchena A, Tang CY, Minzenberg MJ, Hof PR (2011): Involvement of the anterior cingulate and frontoinsular cortices in rapid processing of salient facial emotional information. Neuroimage 54: 2539–2546.

Fenton-O'Creevy, M, Soane E, Nicholson N, Willman P (2011): Thinking, feeling and deciding: The influence of emotions on the decision making and performance of traders. J Org Behav 32:1044–1061.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997): Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6:218–229.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998): Event-related fMRI: characterizing differential responses. Neuroimage 7:30–40.

Friston KJ, Frith CD, Turner R, Frackowiak RSJ (1995): Characterizing evoked hemodynamics with FMRI. Neuroimage 2:157–165.

Gitelman DR, Penny WD, Ashburner J, Friston KJ (2003): Modeling regional and psychophysiologic interactions in fMRI: The importance of hemodynamic deconvolution. Neuroimage 19: 200–207.

Glascher J, Daw N, Dayan P, O'Doherty JP (2010): States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron 66:585–595.

Gross JJ (1998): Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. J Pers Soc Psychol 74:224–237.

Gu X, Gao Z, Wang X, Liu X, Knight RT, Hof PR, Fan J (2012): Anterior insular cortex is necessary for empathetic pain perception. Brain 135:2726–2735.

Gu X, Hof PR, Friston KJ, Fan J (2013a): Anterior insular cortex and emotional awareness. J Comp Neurol 521:3371–3388.

Gu X, Liu X, Guise KG, Naidich TP, Hof PR, Fan J (2010): Functional dissociation of the frontoinsular and anterior cingulate cortices in empathy for pain. J Neurosci 30:3739–3744.

Gu X, Liu X, Van Dam NT, Hof PR, Fan J (2013b): Cognition-emotion integration in the anterior insular cortex. Cereb Cortex 23: 20–27.

Hayden BY, Pearson JM, Platt ML (2009): Fictive reward signals in the anterior cingulate cortex. Science 324:948–950.

Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer CF (2005): Neural systems responding to degrees of uncertainty in human decision-making. Science 310:1680–1683.

King-Casas B, Sharp C, Lomax-Bream L, Lohrenz T, Fonagy P, Montague PR (2008): The rupture and repair of cooperation in borderline personality disorder. Science 321:806–810.

Kirk U, Harvey A, Montague PR (2011): Domain expertise insulates against judgment bias by monetary favors through a modulation of ventromedial prefrontal cortex. Proc Natl Acad Sci USA 108:10332–10336.

Kober H, Mende-Siedlecki P, Kross EF, Weber J, Mischel W, Hart CL, Ochsner KN (2010): Prefrontal-striatal pathway underlies cognitive regulation of craving. Proc Natl Acad Sci USA 107:14811–14816.

Kuo WJ, Sjostrom T, Chen YP, Wang YH, Huang CY (2009): Intuition and deliberation: Two systems for strategizing in the brain. Science 324:519–522.

LaBar KS, Gatenby JC, Gore JC, LeDoux JE, Phelps EA (1998): Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. Neuron 20:937–945.

Lamm C, Singer T (2010): The role of anterior insular cortex in social emotions. Brain Struct Funct 214:579–591.

LeDoux JE (2000): Emotion circuits in the brain. Annu Rev Neurosci 23:155–184.

Li J, Schiller D, Schoenbaum G, Phelps EA, Daw ND (2011): Differential roles of human striatum and amygdala in associative learning. Nat Neurosci 14:1250–1252.

Lieberman MD, Cunningham WA (2009): Type I and Type II error concerns in fMRI research: Re-balancing the scale. Soc Cogn Affect Neurosci 4:423–428.

Lohrenz T, McCabe K, Camerer CF, Montague PR (2007): Neural signature of fictive learning signals in a sequential investment task. Proc Natl Acad Sci USA 104:9493–9498.

Martin LN, Delgado MR (2011): The influence of emotion regulation on decision-making under risk. J Cogn Neurosci 23:2569–2581.

Menon V, Uddin LQ (2010): Saliency, switching, attention and control: A network model of insula function. Brain Struct Funct 214:655–667.

Montague PR, Berns GS, Cohen JD, McClure SM, Pagnoni G, Dhamala M, Wiest MC, Karpov I, King RD, Apple N, Fisher RE (2002): Hyperscanning: Simultaneous fMRI during linked social interactions. Neuroimage 16:1159–1164.

Montague PR, Hyman SE, Cohen JD (2004): Computational roles for dopamine in behavioural control. Nature 431:760–767.

Montague PR, King-Casas B, Cohen JD (2006): Imaging valuation models in human choice. Annu Rev Neurosci 29:417–448.

Montague PR, Lohrenz T (2007): To detect and correct: norm violations and their enforcement. Neuron 56:14–18.

Mufson EJ, Mesulam MM, Pandya DN (1981): Insular interconnections with the amygdala in the rhesus monkey. Neuroscience 6:1231–1248.

Niv Y, Duff MO, Dayan P (2005): Dopamine, uncertainty and TD learning. Behav Brain Funct 1:6.

Ochsner KN, Silvers JA, Buhle JT (2012): Functional imaging studies of emotion regulation: A synthetic review and evolving model of the cognitive control of emotion. Ann NY Acad Sci 1251:E1–E24.

Pagnoni G, Zink CF, Montague PR, Berns GS (2002): Activity in human ventral striatum locked to errors of reward prediction. Nat Neurosci 5:97–98.

Pessoa L, Adolphs R (2010): Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance. Nat Rev Neurosci 11:773–783.

Phelps EA (2006): Emotion and cognition: Insights from studies of the human amygdala. Annu Rev Psychol 57:27–53.

Preuschoff K, Quartz SR, Bossaerts P (2008): Human insula activation reflects risk prediction errors as well as risk. J Neurosci 28:2745–2752.

Rangel A, Camerer C, Montague PR (2008): A framework for studying the neurobiology of value-based decision making. Nat Rev Neurosci 9:545–556.

Rosenstein MT, Barto AG, Si J, Barto A, Powell W, Wunsch D (2004): Supervised Actor-Critic Reinforcement Learning. Handbook of Learning and Approximate Dynamic Programming. John Wiley & Sons, Inc. Hoboken, NJ, United States, pp 359–380.

Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003): The neural basis of economic decision-making in the Ultimatum Game. Science 300:1755–1758.

Schultz W, Dayan P, Montague PR (1997): A neural substrate of prediction and reward. Science 275:1593–1599.

Seth AK (2013): Interoceptive inference, emotion, and the embodied self. Trends Cogn Sci 17:565–573.

Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004): Temporal difference models describe higher-order learning in humans. Nature 429:664–667.

Simon DA, Daw ND (2011): Neural correlates of forward planning in a spatial decision task in humans. J Neurosci 31:5526–5539.

Singer T, Critchley HD, Preuschoff K (2009): A common role of insula in feelings, empathy and uncertainty. Trends Cogn Sci 13:334–340.

Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD (2004): Empathy for pain involves the affective but not sensory components of pain. Science 303:1157–1162.

Sokol-Hessner P, Camerer CF, Phelps EA (2013): Emotion regulation reduces loss aversion and decreases amygdala responses to losses. Soc Cogn Affect Neurosci 8:341–350.

Sokol-Hessner P, Hsu M, Curley NG, Delgado MR, Camerer CF, Phelps EA (2009): Thinking like a trader selectively reduces individuals' loss aversion. Proc Natl Acad Sci USA 106:5035–5040.

Staudinger, M.R., Erk, S., Abler, B., Walter, H. (2009): Cognitive reappraisal modulates expected value and prediction error encoding in the ventral striatum. Neuroimage, 47:713–721.

Staudinger MR, Erk S, Walter H (2011): Dorsolateral prefrontal cortex modulates striatal reward encoding during reappraisal of reward anticipation. Cereb Cortex 21:2578–2588.

Stephan KE, Harrison LM, Penny WD, Friston KJ (2004): Biophysical models of fMRI responses. Curr Opin Neurobiol 14:629–635.

Ullsperger M, Harsay HA, Wessel JR, Ridderinkhof KR (2010): Conscious perception of errors and its relation to the anterior insula. Brain Struct Funct 214:629–643.

Volkow ND, Fowler JS, Wang GJ, Telang F, Logan J, Jayne M, Ma Y, Pradhan K, Wong C, Swanson JM (2010): Cognitive control of drug craving inhibits brain reward regions in cocaine abusers. Neuroimage 49:2536–2543.

Wager TD, Davidson ML, Hughes BL, Lindquist MA, Ochsner KN (2008): Prefrontal-subcortical pathways mediating successful emotion regulation. Neuron 59:1037–1050.

Xiang T, Lohrenz T, Montague PR (2013): Computational substrates of norms and their violations during social exchange. J Neurosci 33:1099–1089.