Disentangling reward processes underlying payoff maximization from individual differences in gain frequency bias and reinforcement learning

Pragathi Priyadharsini Balasubramani[1*], Juan Diaz-Delgado[1], Gillian Grennan[1], Fahad Alim[1], Mariam Zafar-Khan[1], Vojislav Maric[1], Dhakshin Ramanathan[1,2], Jyoti Mishra[1*]

[1]Neural Engineering and Translation Labs, Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA

[2]Department of Mental Health, VA San Diego Medical Center, San Diego, CA

[*]Correspondence should be addressed to:
Pragathi Priyadharsini Balasubramani & Jyoti Mishra
University of California, San Diego
Neural Engineering & Translation Labs (NEAT Labs)
9500 Gilman Drive Mail Code 0875
La Jolla, CA 92037
pbalasubramani@health.ucsd.edu & jymishra@health.ucsd.edu

Total Figures: 3

Total Words: 2941, Introduction (431), Results (1606), Discussion (904).

Abstract

Humans make choices based on both reward magnitude and reward frequency. Probabilistic decision making is popularly tested using multi-choice gambling paradigms that require participants to maximize task payoff. However, research shows that performance in such paradigms suffers from individual bias towards the frequency of gains as well as individual differences that mediate reinforcement learning, including attention to stimuli, sensitivity to rewards and risks, learning rate, and exploration vs. exploitation based executive policies. Here, we developed a two-choice reward task, implemented in 186 healthy human subjects across the adult lifespan, to understand the cognitive and neural basis of payoff-based performance. We controlled for individual gain frequency biases using experimental block manipulations and modeled individual differences in reinforcement learning parameters. Simultaneously recorded electroencephalography (EEG)-based cortical activations showed that diminished theta activity in the right rostral anterior cingulate cortex (ACC) as well as diminished beta activity in the right parsorbitalis region of the inferior frontal cortex (IFC) during cumulative reward presentation correspond to better payoff performance. These neural activations further associated with specific symptom self-reports for depression (greater ACC theta) and inattention (greater IFC beta), suggestive of reward processing markers of clinical utility.

Keywords: reward, risk sensitivity, reinforcement learning, anterior cingulate cortex, inferior frontal cortex, depression, inattention

Introduction

Cognitive and neural responses to reward and risk are quintessential to understanding human behavior. Gambling tasks predominantly form the experimental test beds for measuring reward and risk processing abilities in humans[1,2]. However, it is widely debated how well these tasks can separate decision-making based on frequency of gains/losses versus expected value, i.e. payoff of different choice-sets[1,3,3–12].

For instance, studies that have controlled for gain frequency in the Iowa Gambling Task (IGT) show that subject choices reflect their gain frequency preferences, which drive relatively immediate reinforcement based choice behavior, rather than expected values that benefit payoff in the long-term[13–17]. Additionally, many studies suggest that performance measures from gambling tasks are influenced by individual differences that are revealed within a reinforcement learning framework, e.g. sensitivity to rewards and risks, learning rate, and behavioral execution strategies[18–22]; prior neural studies of decision payoff have not fully accounted for these differences.

In this study, we aimed to address the shortcomings of existing reward processing assessments in two ways. First, we accommodated for individual gain frequency bias while assessing advantageous payoff-based decision making. Specifically, we designed a two-choice paradigm that implements two separate blocks – a $\Delta_0$payoff block (baseline block) where two reward choice-options have equal payoffs and reward variance suitable for measuring the immediate gain frequency bias, and a Δpayoff block (difference block) where the two-choice options have unequal payoffs suitable for measuring payoff influences. We thereby, teased apart measurements of immediate gain frequency biased response from long-term payoff (i.e. expected value) based response to understand the distinct cognitive and neural mechanisms underlying payoff decisions.

Second, we accounted for individual differences in reinforcement learning (RL) including attention to stimuli, sensitivity to rewards and risks, learning rate, and exploration vs. exploitation based executive policies, for understanding reward functions. In prior work, we have shown that reward and risk based processing can be explained within RL model frameworks[23–27]. RL models provide the ability to derive the underlying learning parameters forming the basis for individual differences in performance, including subjective sensitivity to risky or probabilistic outcomes[24], time scale of reward prediction or outcome discounting rate contributing to reward sensitivity, exploration versus exploitation in individual responses[28], influence of repeated choices on learning[29] and levels of attention[30,31]. Uniquely, in this study, we model and account for these RL parameters while estimating the data-driven neural correlates for payoff relevant decisions. Finally, we predict individual variations in self-reported mental health based on the cognitive and neural correlates of payoff relevant responses. We show that payoff relevant neural markers are sensitive to specific neuropsychiatric symptoms and thereby serve future clinical utility.

Results

Reward paradigm disentangles payoff-based performance from gain frequency bias.

Healthy adult subjects (N = 186, ages 18-80 years, 115 females) performed a two-choice gambling task, *Lucky Door*, which implemented two distinct blocks of choices; the $\Delta_0$payoff block delivered choice-sets with different gain frequencies but no differences in payoff, while the Δpayoff block delivered choice-sets with same gain frequencies as the earlier $\Delta_0$payoff block yet with long-term payoff (i.e. expected value) differences (**Figure 1A**). Specifically, the $\Delta_0$payoff block only varied the gain frequency associated with the choice doors, with one door leading to 70% positive reward outcomes (Rare Loss or RareL door) while the other resulting in 70% negative reward outcomes (Rare Gain or RareG door), yet maintaining the same reward average or long-term expected value/ payoff (**Supplementary Table 1**). The Δpayoff block was presented in a random sequence order relative to the $\Delta_0$payoff block across subjects. It had the same gain and loss frequency setup as the $\Delta_0$payoff block (choices randomly positioned on the left or right side of the screen, **Figure 1A**), but the rewards associated with the RareG door resulted in a larger long-term payoff than the RareL door. Participants executed 40 trials per block. The gain frequency bias (*Bias*) was computed from the $\Delta_0$payoff block as the difference between the proportion of RareL vs. RareG selections. Thus higher the preference for RareL to RareG door in the $\Delta_0$payoff block, higher is the gain frequency *Bias*. On the other hand, the payoff-based response (*Perf*) was computed as the difference between the proportion of RareG selections on the Δpayoff vs. $\Delta_0$payoff block. Therefore higher the preference for RareG door in Δpayoff block to RareG door in the $\Delta_0$payoff block, higher is the *Perf*. The RareG door was designed with greater payoff as choosing this door could selectively suggest payoff-based decision processing in subjects as opposed to simply choosing based on gain frequency in which case RareL should be preferred.

Given our specific focus on understanding advantageous long-term decisions measured by payoff-based responses (*Perf*), we first wanted to understand which behavioral predictors significantly predicted *Perf*. We implemented a multivariate regression model of *Perf* using demographic (age, gender, race, ethnicity, socio-economic status (SES) and mental health (anxiety, depression, inattention and hyperactivity) predictors as per **Table 1**. Mental health was assessed using self-reporting questionnaires detailed in our **Methods** section. The model also accounted for individual gain frequency *Bias* and order of block presentation. The overall *Perf* model was significant (R^2=0.43, p<0.0001). Interestingly the only variable that significantly predicted *Perf* was Gain frequency *Bias* (**Figure 1B**, β=0.37 ±0.04, t(151)=9.18, p<0.0001, f^2=0.58), showing that *Bias* can confound the understanding of payoff-based performance. Hereafter, we control for *Bias* in all payoff-relevant analyses below.

Reinforcement Learning models capture task-payoff related performance and suggest individually differentiating learning parameters.

We developed a reinforcement learning model for each subject that converged to their actual payoff-related performance (**Figure 2**). The motivation for building the RL model sprung from many folds of reasoning. First, it explains the converged behavioral dynamics in each participant sans experimental trial limitations. This is important since each choice in the task was associated with a probabilistic reward and subjects were presented with finite number of trials (40 per block), yet we wanted to characterize the stabilized and representative decision-making dynamics for each subject. Second, it characterizes the individual differences in each participant in terms of computational parameters that manifest in learning and executive control. We particularly focused on five key parameters controlling reward processing, risk taking and decision making, which included risk sensitivity factor, temporal reward discount factor, stimulus learning eligibility factor, inverse exploration index, and stimulus noise index (**Figure 2**, **Table 1**). Prior modeling studies have suggested that task behavior can be explained with greater accuracy when both the expected value (long-term average reward or payoff) and expected risk (reward variance) are considered for computing the utility that regulates the decisions. The risk sensitivity factor depicts how a subject trades off the accounting of reward variance from the reward average, higher factor values indicate risk aversiveness and lower levels indicate risk seeking nature of the subject (**Table 1**, model α). The temporal reward discount factor represents impulse control, with higher factor values depicting greater control (**Table 1**, model γ). Stimulus learning eligibility depicts the effects of repetitive choice on learning the underlying task's reward structure and learning the link between the stimuli and their associated behavioral responses, lower factor values indicate increased decay associated with the infrequent choices (**Table 1**, model λ). The inverse exploration index marks the exploration/exploitation tradeoff with lower values indicative of greater exploration (**Table 1**, model β). Finally, the stimulus noise factor represents the noise in stimulus representation due to inadequate attention, greater factor values represent greater noise (**Table 1**, σ). We found that the individual subject RL models fit the actual behavior data significantly well (**Figure 2B**, Sum of squares cost optimization on the number of selections of each choice-door option in each block of the experiment, Spearman correlation between model predicted and actual *Perf*, $\rho(185)=0.92$, $p<0.0001$).

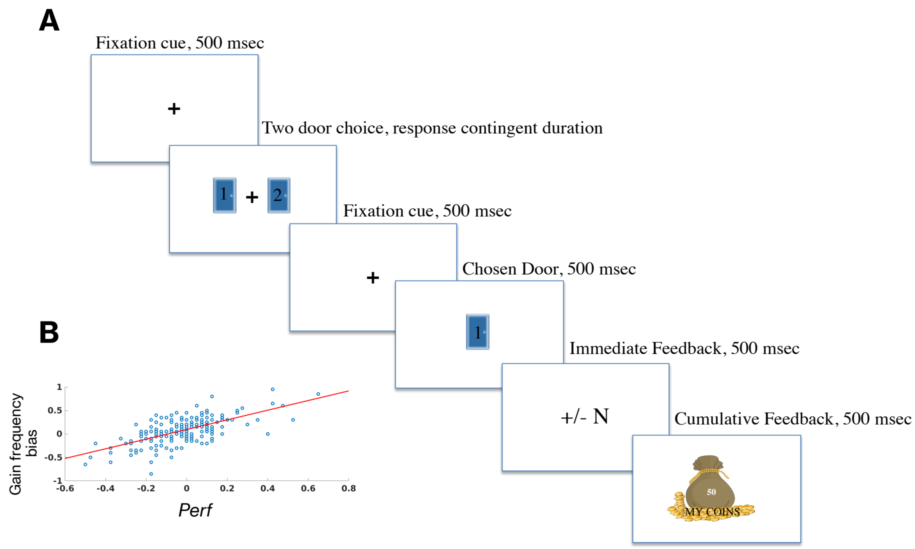


Figure 1. A) Schematic of the *Lucky Door* Task. Participants fixated for 0.5 sec, then chose from one of two choice doors. Post-response, fixation reappeared for 0.5 sec, followed by presentation of the chosen door for 0.5 sec, then immediate gain or loss feedback provided for 0.5 sec, and finally, cumulative feedback of all gains/losses up to the present trial shown for 0.5 sec. **B) Significant predictors of payoff-based performance, *Perf*.** Gain frequency *Bias* (top panel, β=0.36 ±0.04, p<0.0001) significantly predict *Perf*.

Demographics		**Median ± MAD**
Age		25.00 ± 14.87
Gender n (%)		
	Male	71 (38.17)
	Female	115 (61.83)
Ethnicity n (%)		
	Caucasian	116 (62.37)
	Black or African American	4 (2.15)
	Native Hawaiian, Pacific Islander	0 (0)
	Asian	37 (19.89)
	American Indian, Alaska Native	4 (2.15)
	Multi-racial	12 (6.45)
	Others	12 (6.45)
Race n (%)		
	Hispanic or Latino	25 (13.51)
	Not Hispanic or Latino	155 (83.78)
	Unknown	5 (2.70)
SES		5.00 ± 1.34
Mental Health		**Median ± MAD**
Anxiety		3 ± 2.88
Depression		3 ± 2.76
Inattention		4 ± 3.99
Hyperactivity		3 ± 2.96
Behavior		**Median ± MAD**
Model α		0.87 ± 0.13

Model γ	0.90 ± 0.11
Model λ	0.54 ± 0.21
Model β	22.94 ± 11.63
Model σ	1.64±0.23
Perf	-0.02±0.13
Bias	0.10±0.21

Table 1. Subject characteristics. Median ± MAD for subjects demographics variables, mental health self-report scores, and parameters from the reinforcement learning models. MAD: median absolute deviation, SES: socioeconomic status score. The normal threshold cut-off score for mental health symptoms is 5. For the reinforcement learning model, the parameter range was $\alpha \in (-1\ 1)$, $\gamma \in (0\ 1)$, $\lambda \in (0\ 1)$, $\beta \in (0\ 50]$, $\sigma \in [0.5\ 3]$.

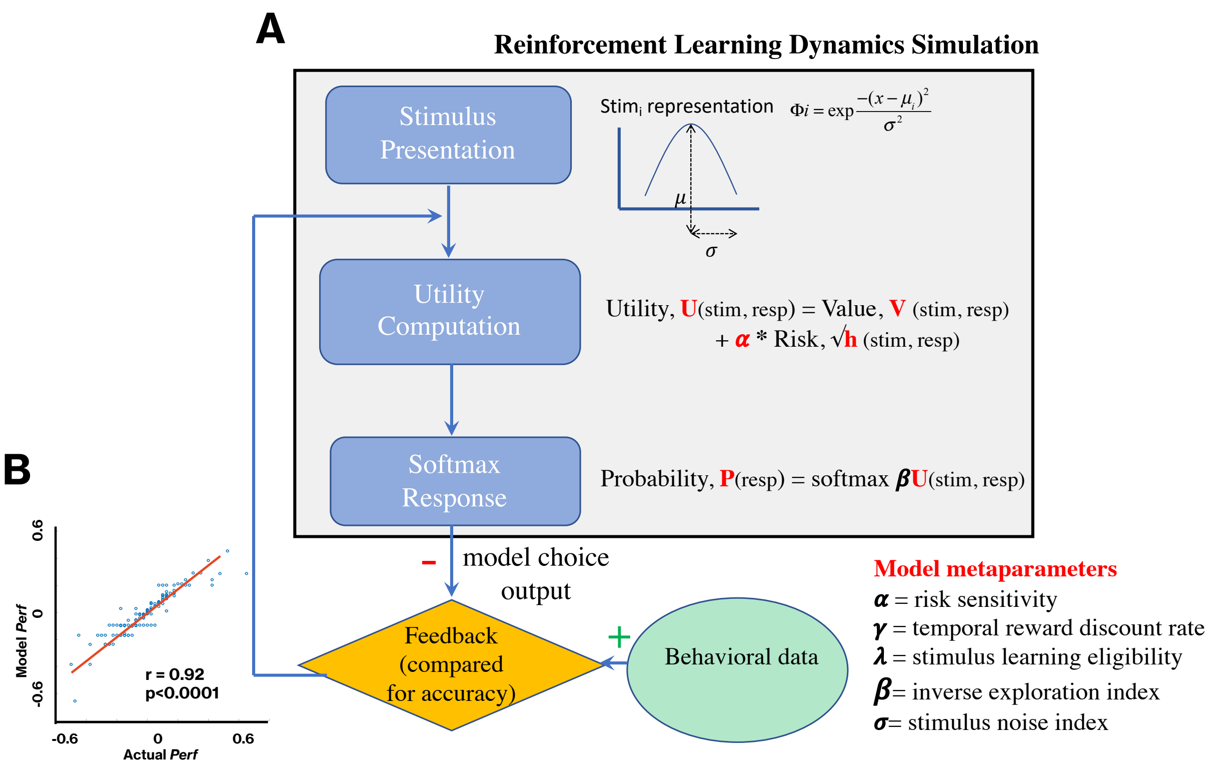


Figure 2. A) Reinforcement Learning model schematic representing the stimulus, value function and choice selection modules. The model results for number of selections associated with each of the choice door stimuli in each task block are compared against the actual selections made by each subject, for purposes of model optimization. The model uses the utility, U, associated with each choice response for making the decision, where the utility is a function of reward average and reward variance associated with choices. The decision in the model is taken using the SoftMax probability, P, of making the choices. Model parameters are highlighted as α (model agent's sensitivity to outcome reward variance), γ (temporal reward discounting), λ (influence of repeated choice on learning and decision-making), β (agent's exploration and exploitation index), and σ (noisy stimulus, x, representation from mean, μ, due to inadequate attention). **B) Model outcomes fit the actual behavior data for *Perf*** (payoff-based decisions, Spearman $\rho(185) = 0.92$, $p<0.0001$).

Right rostral anterior cingulate cortex and inferior frontal cortex code for decision making payoff.

Participants performed the reward task with simultaneous EEG that was analyzed in the theta (3-7 Hz), alpha (8-12 Hz), and beta (13-30 Hz) frequency bands in cortical source space parcellated as per the Desikan-Killiany regions of interest[32]. To identify the neural correlates underlying expected value i.e. payoff-based performance (*Perf*), we modeled these as predictors of *Perf* using robust multivariate linear regression accounting for gain frequency *Bias* that determined *Perf* (**Figure 1**), and the five RL model parameters (α, γ, λ, β and σ; **Figure 2**). Neural activations from three relevant trial periods were investigated: immediately post-presentation of selected choice but prior to reward (0-500 ms selected choice period), during presentation of trial reward (0-500 ms reward period), and during presentation of the cumulative reward up to that trial in the trial sequence (0-500 ms cumulative reward period); neural activations were the relative difference in activity on Δpayoff vs. $\Delta_0$payoff block RareG trials. Taking the relative block difference allowed non-task related individual EEG differences to cancel out, and relative responses to the RareG door were important for analysis because this door choice resulted in a larger long-term payoff than the other (RareL) door in the Δpayoff block. Family-wise error-rate corrections were applied for multiple comparisons.

Independently observed significant neural correlates of *Perf* are shown in **Supplementary Table 2**. We further accounted for these multiple independently significant neural predictors within a unified multivariate model for *Perf* that also included the significant *Bias*, and RL model parameter covariates. The results of this multivariate model showed theta activity in the right rostral anterior cingulate cortex (ACC) during the cumulative reward period (β=-43.74 ± 13.49, t(173)=-3.24, p=0.001 f^2=0.01) and beta activity in the right parsorbitalis region of the inferior frontal cortex (IFC) also during the cumulative reward period (β=-74.88 ± 22.97, t(174)=-3.26, p=0.001, f^2=0.03) as the most significant independent predictors of payoff-based decisions (**Figure 3A-B**); activity in the selected choice presentation period and the immediate reward period did not survive multiple comparisons. It is noteworthy that activity that occurs during the cumulative reward feedback is linked most strongly with performance linked with expected value.

Additionally, we checked whether these specific neural activity regressions showed any interactions with age and gender, but no interactions were found (all p>0.55). The scalp topographic activations corresponding to these neural source profiles are shown in **Supplementary Figure 1**.

Separately, we also investigated the neural correlates for gain *Bias* as differential activity towards frequent gains versus frequent losses on the $\Delta_0$payoff block. We controlled for RL parameters in these analyses as well and found *Bias* activations to be unique from *Perf*. Predictors of *Bias* included left superior frontal theta (β=-48.53 ± 15.09, t(171)=-3.21, p=0.001, f^2=0.02), and right rostral ACC alpha (β=-685.81 ± 201.26, t(170)=-3.41, p=0.0008, f^2=0.06) during the cumulative reward feedback period **(Supplementary Figure 2)**.

Neural correlates of payoff decision processes predict subjective mental health.

We next investigated whether the neural correlates of payoff decisions are relevant to subjective mental health by modeling anxiety, depression, inattention and hyperactivity self-report scores as the dependent variables in robust multivariate regression models. All demographic variables (age, gender, race, ethnicity, SES), all RL model parameters (α, γ, λ, β and σ), task performance variables of *Perf* and *Bias* as well as the two cumulative reward-processing neural correlates (right rostral ACC theta; right IFC beta) were included in each model as independent predictors.

The overall model for each symptom score was significant with false discovery rate (fdr, $p<0.05$) correction applied for multiple comparisons: anxiety ($R^2=0.24$, p=0.001), depression ($R^2=0.20$, p=0.012), inattention ($R^2=0.19$, p=0.03), and hyperactivity ($R^2=0.18$, p=0.03).

Amongst demographics, age negatively predicted anxiety (β=-0.02±0.004, t(149)=-5.04, p<0.0001, $f^2=0.17$), depression (β=-0.01±0.004, t(149)=-3.52, p=0.0006, $f^2=0.08$) and inattention (β=-0.01±0.004, t(142)=-3.69, p=0.0003, $f^2=0.10$). Moreover, a multiracial origin negatively predicted inattention in subjects (β=-0.65±0.32, t(142)=-2.06, p=0.04, f^2 all races=0.03). No other demographics were significant predictors of mental health symptoms. Amongst the RL model variables, reward discount factor (γ), which represents impulse control, negatively predicted inattention (β=-0.96±0.48, t(142)=-2.02, p=0.04, $f^2=0.03$) as well as hyperactivity scores (β=-0.97±0.44, t(149)=-2.20, p=0.03, $f^2=0.04$), and the stimulus decay factor positively predicted hyperactivity (β=0.99±0.50, t(149)=1.98, p=0.049, $f^2=0.03$). Notably, neural correlates of payoff performance also significantly predicted symptom scores: cumulative reward related rostral ACC theta activity positively predicted depression (β=304.83±153.92, t(149)=1.98, p=0.049, $f^2=0.07$); and right IFC beta activity positively predicted inattention (β=697.84±324.08, t(142)=2.15, p=0.03, $f^2=0.03$, **Figure 3 C-D**). Finally, the overall regression models for depression and inattention symptom scores were improved when taking the significant neural correlates into account vs. not (regression models compared with and without neural parameters for Depression: Fstat=3.92, p=0.049; Inattention: Fstat=4.64, p=0.03).

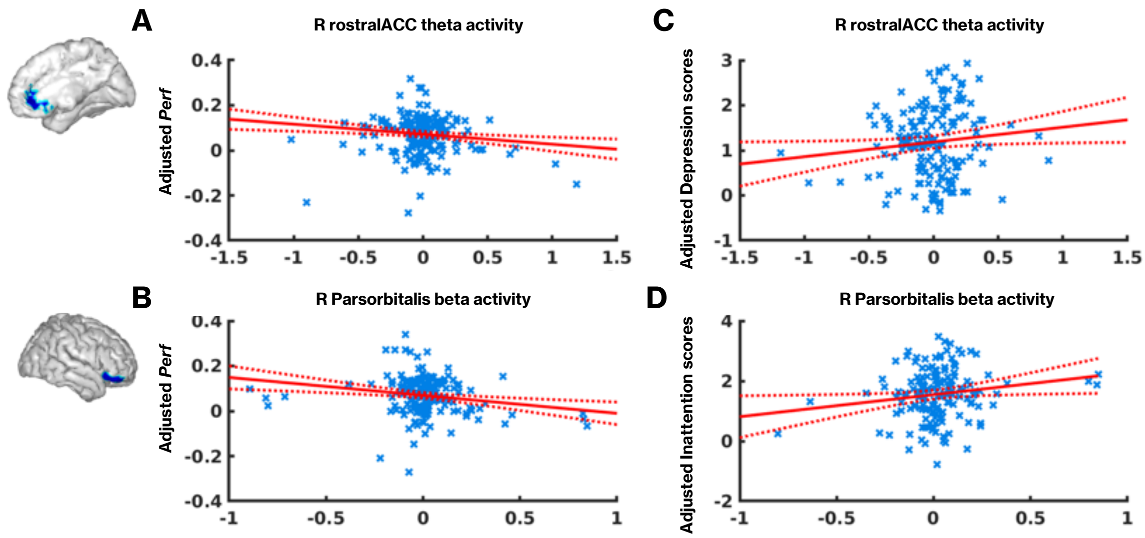


Figure 3: Neural correlates of payoff-based decision making in humans. Payoff based performance is negatively predicted by (A) right rostral anterior cingulate cortex (ACC) theta activity in the cumulative reward period, and by (B) right parsorbitalis (IFC) beta activity in the cumulative reward period. (C) Payoff performance related rostral ACC theta positively predicted depression and (D) parsorbitalis beta predicted inattention symptoms. The scatters are presented on an adjusted axis as obtained from the multivariate robust regression models. The x-axes in all cases are 10^{-3} source activity units.

Discussion

Reinforcement learning models suggest human choices ideally tend to maximize long-term beneficial outcomes[23,24,26,33,34]. However, many existing neuropsychological measures of decision-making that optimize for long-term payoffs don't reliably estimate the participant's ability to integrate rewards and make foresighted decisions, and instead suffer from biases to immediate outcome frequencies[16,19,35,36]. In our paradigm, we can disentangle gain frequency biases from payoff-based decision-making by introducing a $\Delta_0$payoff block with no payoff difference between choice options wherein decisions are purely based on gain frequency[15]. Comparing choices within the Δpayoff experimental block, designed to have similar reward distribution structure as the $\Delta_0$payoff block but differing only on the long-term outcome between options, allows measurement of individual long-term payoff sensitivity. Therefore, our study by its very design is able to perform this important distinction to tease apart individual payoff-based performance from bias towards gain frequency, and further leverage these measures to inform mental health behaviors.

The behavioral outcomes of our experiment varied based on individual subject characteristics. Payoff-based performance was significantly related to individual bias for observed frequency of gains; this is in line with prior studies of decision-making but wherein gain frequency decisions are often conflated with expected value[1,15]. We further modeled subjective differences using reinforcement learning (RL) models, and extracted sensitivity to rewards and risks in the outcomes (α), reward discounting rates through time (γ), attention levels (σ), explorative tendency during behavior (β), and repetitive behavior effects on general learning dynamics (λ) to explain each subject's behavior. Using the RL model outputs, we were able to account for individual differences when identifying the payoff performance related neural correlates.

We focused on three different time periods of the task, the first associated with processing of the selected choice, the ensuing reward presentation period and the cumulative reward period to understand how neural dynamics in these periods affect payoff-based performance. The selected choice period captures the processing associated with presentation of the chosen door after the actual decision period. We did not analyze the actual decision period since two different choice options are shown on the screen during this period and the signal associated with every choice option was difficult to explicitly assess. Accounting for demographics, gain frequency *Bias* and RL-model informed differences in learning and behavioral execution, the significant neural correlates of payoff sensitive performance were found in critical frontal executive regions of the right rostral anterior cingulate cortex (ACC) and inferior frontal cortex (IFC) during the cumulative reward period. Relatedly, most earlier studies on probabilistic reward processing have suggested medial prefrontal cortex (mPFC) to be a core region mediating decision performance[1–3,37–39].

More specifically, analyses showed theta activity in right rostral ACC negatively correlated with payoff-based performance. This finding is aligned with prior evidence for reward-based theta processing[40–42] and its widely studied relationship to long-term risk or uncertainty. This may be one reason why we observe a negative relationship between rostral ACC theta during cumulative reward presentation and effective payoff, *Perf*, whose magnitude inversely relates to uncertainty but positively to choice utility[43–46].

Similarly, right IFC (parsorbitalis) negatively correlated with the payoff performance, specifically in the beta spectral band during the cumulative reward feedback period. Many inhibitory control studies suggest that increased beta activity in this region during action stopping plays an important role in behavioral inhibition. Our results at the least suggest that decisions for advantageous choices may significantly interact with the stopping circuit for successful behavior[47–50]. Moreover, our findings observed during the cumulative reward feedback period may suggest facilitation of accurate response with optimal speed and inhibitory control in the next-trial[51].

In our study we also found that neural correlates of payoff-based performance are distinct from correlates of gain frequency bias. We found the medial prefrontal cortex, especially, left superior frontal theta and right rostral ACC alpha coded for bias-related activity. These areas had been previously suggested to play significant roles in probabilistic decision making[52–59] and our study by its ability to tease apart payoff-performance and bias, is able to also distinguish the underlying neural correlates.

Translational neuroscience studies show that reward based decision processing deficits are found in depression and in attention disorders, leading to difficulty in reward integration and foresighted choice-behaviors[23,60–65]. It is precisely such individuals who then focus on the immediate reward outcome in the short-term, characterized by a prolonged attenuation of temporal discounting of rewards[66,67]. Interestingly, rostral ACC activity in our paradigm during the cumulative reward feedback period that is a negative correlate of payoff performance, was a positive predictor depression scores—a link that is also supported by prior work[68,69]. Related research also suggests that rostral ACC theta activity is a significant pretreatment marker for depression treatment outcomes[70]. Additionally, right IFC beta activity during the cumulative reward feedback period showed a positive relationship to inattention scores, suggesting that decreased activity enables pursuit for maximal payoff with high attention, possibly by enabling flexibility in stopping impulsive decisions. These results are still limited by our study in healthy adults and need to be replicated in clinical populations.

Altogether, our study presents the importance of controlling for biases for immediate reward frequency and individual differences in learning while assessing advantageous, i.e. foresighted decision-making ability in humans. Our findings of payoff-relevant right rostral ACC theta activity and right IFC beta activity, both in the cumulative reward-related feedback period, are important for clinical translational application, especially for depression, and attention problems, and also suggests plausible neural targets for reward processing based interventions.

Methods

Participants. 198 adult human subjects (age mean ± standard deviation 35.44 ± 20.30 years, range 18-80 years, 115 females) participated in the study. All participants provided written informed consent for the study protocol approved by the University of California San Diego institutional review board (UCSD IRB #180140). Twelve of these participants were excluded from the study as they had a current diagnosis for a psychiatric disorder and current/recent history of psychotropic medications for a final sample of 186 healthy adult participants. All participants reported normal/corrected-to-normal vision and hearing and no participant reported color blindness. For older adults >60 years of age, participants were confirmed to have a Mini-Mental State Examination (MMSE) score >26 to verify absence of apparent cognitive impairment (Arevalo-Rodriguez et al. 2015). All data was collected prior to COVID-19 period of restricted research.

Surveys. All participants provided demographic information by self-report including age, gender, race (in a scale of 1 to 7: Caucasian; Black/African American; Native Hawaiian / Other Pacific Islander; Asian; American Indian / Alaska Native; More than one race; Unknown or not reported) and ethnicity; socio-economic status (SES) was measured on the Family Affluence Scale from 1 to 9 (Boudreau and Poulin, 2008), and any current/past history of clinical diagnoses and medications were reported. For older adults >60 years of age, participants completed the Mini-Mental State Examination (MMSE) and scored >26 to verify absence of apparent cognitive impairment[71]. All participants completed subjective mental health self-reports using standard instruments, ratings of inattention and hyperactivity obtained on the Adult ADHD Rating Scale (New York University and Massachusetts General Hospital. Adult ADHD-RS-IV with Adult Prompts. 2003; : 9–10), Generalized Anxiety Disorder 7-item scale GAD-7[72] and depression symptoms reported on the 9-item Patient Health Questionnaire, PHQ-9[73]. Symptoms for these psychiatric conditions were measured because they have been related to changes in reward processing[74–76].

Task Design. We investigated a two-choice decision-making task that enabled a rapid assessment and was easy to understand across the adult lifespan. In this task that we refer to as *Lucky Door*, participants chose between one of two doors, either a rare gain door (RareG, probability for gains p=0.3, for losses p=0.7) or a rare loss door (RareL, probability for losses p=0.3, for gains p=0.7). Participants used the left and right arrow keys on the keyboard to make their door choice. Door choice was monitored throughout the task. Additionally, in two separate blocks, we investigated whether the overall expected value (payoff) of the choice door can influence individual behavior. In the baseline block with $\Delta_0$payoff (no-payoff difference), the two choice doors did not differ in payoff (RareL door, p=0.3 for -70 coins and p=0.7 for +30 coins, payoff=0; RareG door, p=0.3 for +70 coins and p=0.7 for -30 coins, payoff=0). In the experimental difference block with Δpayoff (payoff difference), expected value or payoff was greater for the RareG door (p=0.3 for +60 coins, p=0.7 for -20 coins, payoff=+40) than for the RareL door (p=0.3 for -60 coins, p=0.7 for +20 coins; payoff=-40). Manipulation of payoff, with greater expected value tied to the RareG door, allowed for investigating individual propensities to prioritize long-term (or cumulative) vs. short-term (or immediate) rewards. The RareG door was assigned greater payoff because choosing this door could selectively suggest payoff-based decision processing in subjects as opposed to simply choosing based on gain frequency in which case RareL should be preferred. 40 trials were presented per block and block order was randomized across participants; two practice trials preceded the main Δpayoff / $\Delta_0$payoff blocks. **Figure 1A** shows a schematic of the task stimulus sequence and **Figure 1B** shows the reward distribution that was shuffled and updated after every 10 trials had been sampled from that set.

The *Lucky Door* task was deployed in Unity as part of the assessment suite on the *BrainE* (short for Brain Engagement) platform[77]. The Lab Streaming Layer (LSL[78]) protocol was used to time-stamp each stimulus/response event during the task. Study participants engaged with the assessment on a Windows 10 laptop sitting at a comfortable viewing distance.

Electroencephalography (EEG). EEG data was collected simultaneous to the *Lucky Door* task using a 24-channel SMARTING device with a semi-dry and wireless electrode layout (Next EEG—new human interface, MBT). Data were acquired at 500 Hz sampling frequency at 24-bit resolution. Cognitive event markers were integrated using LSL and data files were stored in xdf format.

Behavioral analyses. Task speeds were calculated as log(1/RT), where RT is response time in seconds. We computed the PayOff sensitive performance response (Perf) as the difference in proportion selection of the RareG door between the Δpayoff and the $\Delta_0$payoff blocks; RareG vs. RareL EVs differed only in the Δpayoff block. We computed Gain frequency bias (Bias) as the difference in proportion selection between the RareG and RareL doors in the $\Delta_0$payoff block where the payoff for both the doors was the same. While Perf is indicative of subjective payoff based selection of advantageous choices, Bias is indicative of inherent valence based selection of choices. For N fraction of responses in each block, we calculated:

$$\begin{aligned} Perf &= N_{exptRareG} - N_{baseRareG} \\ Bias &= N_{baseRareG} - N_{baseRareL} \end{aligned} \tag{1}$$

Reinforcement Learning (RL) Model. We simulated a RL model[24,29] to estimate 5 different parameters for each participant, including, risk sensitivity (α); reward discount factor (γ); stimulus learning eligibility (λ); exploration index (β); stimulus noise (σ).

The risk sensitivity parameter (α) measures how much the expected uncertainty associated with the door is accounted for in the computation of utility for decision making, the smaller the parameter value $\alpha \in (-1\ 1)$, $\alpha \rightarrow -1$ the higher is risk seeking, while a larger value, $\alpha \rightarrow 1$ indicates high risk aversiveness.

The reward discounting factor (γ) represents discounting of rewards through time in subjects, lower values $\gamma \in (0\ 1)$, $\gamma \rightarrow 0$ suggest impulsiveness in decisions accounting for immediate rewards alone for decision processing while higher values, $\gamma \rightarrow 1$ suggest long term integration of rewards for decisions.

The learning eligibility trace factor (λ) represents the extent of decay associated with the infrequent presentation of stimulus that affects the learning updates, lower values $\lambda \in (0\ 1)$, $\lambda \rightarrow 0$ suggest heavy loss of stimulus information for not being presented for a trial thereby causing lossy learning update, while higher values, $\lambda \rightarrow 1$ suggest conservation of stimulus information for learning in any trial.

The exploration index (β) measures the randomness associated with choice selection policy, lower values $\beta \in (0\ 50]$, $\beta \rightarrow 0$ suggest exploration while higher values, $\beta \rightarrow \infty$ suggest exploitation based on utility for decision making.

The stimulus noise (σ) measures the randomness associated with stimulus representation as attended and perceived by the subject, lower values, $\sigma \in [0.5\ 3]$, $\sigma \rightarrow 0$ suggest sharper representation while higher values, $\sigma \rightarrow \infty$ suggest noisier representation.

The simulation agent had reward distributions as in the real experiment but scaled down by multiplying with a factor 0.1, and varying with blocks (Δpayoff, $\Delta_0$payoff) that were randomly ordered. There were as high as 10000 trials in each block for letting model performance converge.

The agent has to choose between two doors each of which (stimulus, *s*) was represented by a radial basis function (Φi) as below:

$$\Phi s = \exp\frac{-(x-\mu_s)^2}{\sigma^2} \tag{2}$$

Here, the μ_s and σ denotes the mean ($s \in [1\ 2]$; door1 = 1; door2 = 2) and standard deviation (stimulus noise), respectively.

The door stimulus is multiplied with the weight matrix *wv* for computing its value function, *Q*, and *wr* for constructing its risk function, $\sqrt{h}$.

Utility associated with any state at a trial, *t*, is the combination of value and risk function[79,80], where the risk function is modulated by a risk sensitivity factor α. Higher the α, higher the risk aversiveness of the subject.

$$U(s,t) = Q(s,t) - \alpha\sqrt{h(s,t)} \text{ where}$$

$$\begin{aligned} Q(s,t) &= wv(s,t)\Phi(s) \\ h(s,t) &= wr(s,t)\Phi(s) \end{aligned} \tag{3}$$

The door choice selection is performed using the SoftMax principle defined as below. According to SoftMax, the probability for choosing a door at trial, *t*, is *P*(s,t):

$$P(s,t) = \frac{exp(\beta U(s,t))}{\sum_{i=1}^{n} exp(\beta U(i,t))} \tag{4}$$

Here, *n* is the total number of doors available, and *β* is the exploration index. Values of *β* tending to 0 make the choices almost equiprobable and is more exploratory whereas the *β* tending to ∞ makes the choice selection identical to exploitative choice selection.

After choice selection, the weight functions are updated using principles below. The choice value function *Q* at trial *t*+1 for door, *s*, may be expressed as,

$$Q_{t+1}(s,t) = Q_t(s,t) + \eta_Q \delta\Phi' e \tag{5}$$

where ηQ is the learning rate of the value function ($0 < \eta Q < 1$) for the stimulus variable, Φ'(s) =

$\Phi(s)e$, i.e., scaled by current eligibility trace, e, associated with the stimulus. The variable e is decayed to $e*\lambda*\gamma$ at every trial, where λ is the inverse decay parameter, and γ is the reward discount factor, for all the stimuli, but for the currently chosen e, it is further facilitated by an addition of 1, $e(t+1) = e(t) + 1$.

Δ is the temporal difference error represented as

$$\delta = r + \gamma\, max_{s'} Q(s',t) - Q_t(s,t) \qquad (6)$$

where r is the reward associated with taking an action, a, for stimulus, s, at time, t, and γ is the reward discount factor. Similar to the value function, the risk function h has an incremental update as defined by the below equation. Optimizing the risk function in addition to the value function is shown to capture human behavior well in a variety of cognitive tasks involving reward-punishment sensitivity, risk sensitivity, and time scale of reward prediction[24,81].

$$h_{t+1}(s,t) = h_t(s,t) + \eta_h \xi \Phi' e \qquad (7)$$

where ηh is the learning rate of the risk function ($0 < \eta h < 1$), and ξ is the risk prediction error expressed by the below equation.

$$\xi = \delta^2 - h_t(s,a,t) \qquad (8)$$

For simplicity, we model as $\eta_h = \eta_Q = 0.1$ as an initial optimization for our subjects for η provided a median of 0.1. The weights wv and wr are set to a small random number from set [-0.0005 0.0005] at trial = 1. The weights are normalized by dividing by their norm.

The cost function optimizes the frequency of selections of rare gain and rare loss options in $\Delta_0$payoff and Δpayoff blocks for every subject after running the simulation agent for 10 instances of twenty thousand trials each, and inferring the optimal parameters for every participant in our study using *fmincon* function in MATLAB. Cost function = sum of squares of the difference actual (Proportion# $RareG_{expt}$ + Proportion# $RareL_{expt}$ + Proportion# $RareG_{base}$ + Proportion# $RareL_{base}$) – simulated actual (Proportion# $RareG_{expt}$ + Proportion# $RareL_{expt}$ + Proportion# $RareG_{base}$ + Proportion# $RareL_{base}$). Optimization is carried out for the 5 parameters, risk sensitivity (α); reward discount factor (γ); stimulus decay (λ); exploration index (β); stimulus noise (σ), using *fmincon*(). We ran fmincon() 100 times to choose the parameter set with least cost for any subject. The simulated Perf measures correlated with the actual values significantly (Spearman correlation, $\rho(185) = 0.92$, $p<0.0001$, **Figure 2B**).

We also performed sensitivity analyses for our models, by systematically varying one of the five parameters at a time for about 200 different initial points linearly spaced within the boundary specific for each parameter boundary(α) = [-0.99 0.99], boundary(γ) = [0.01 0.99], boundary(λ) = [0.01 0.999], boundary(β) = [0.001 5], boundary(σ) = [0.5 3]. The other 4 parameters were kept at the median computed over the set of subjective parameter values each optimized using *fmincon*()

as mentioned above (**Supplementary Figure 3**). These analyses confirmed *Perf* and *Bias* measures to be sensitive to the range of the five core RL model parameters.

Neural data processing.

We applied a uniform processing pipeline to all EEG data acquired simultaneous to the *Lucky Door* task. This included: 1) data pre-processing, 2) computing event related spectral perturbations (ERSP) for all channels, and 3) cortical source localization of the EEG data filtered within relevant theta, alpha and beta frequency bands.

1) Data preprocessing was conducted using the EEGLAB toolbox in MATLAB[82]. EEG data was resampled at 250 Hz, and filtered in the 1-45 Hz range to exclude ultraslow DC drifts at <1Hz and high-frequency noise produced by muscle movements and external electrical sources at >45Hz. EEG data were average referenced and epoched to the chosen door presentation during the task, in the -.5 sec to +1.5 sec time window (**Figure 1**). Any missing channel data (one channel each in 6 participants) was spherically interpolated to nearest neighbors. Epoched data were cleaned using the autorej function in EEGLAB to remove noisy trials (>5sd outliers rejected over max 8 iterations; 0.91± 2.65% of trials rejected per participant). EEG data were further cleaned by excluding signals estimated to be originating from non-brain sources, such as electrooculographic, electromyographic or unknown sources, using the Sparse Bayesian learning (SBL) algorithm[83,84], https://github.com/aojeda/PEB) explained below in the cortical source localization section.

2) For ERSP calculations, we performed time-frequency decomposition of the epoched data using the continuous wavelet transform (cwt) function in MATLAB's signal processing toolbox. Baseline time-frequency (TF) data in the -250 ms to -50 ms time window prior to chosen door presentation were subtracted from the epoched trials (at each frequency) to observe the event-related synchronization (ERS) and event-related desynchronization (ERD) modulations[85]. Time-frequency decompositions of the chosen door evoked neural activity showed that most electrodes had significant ERS and ERD signatures at the channel level, with ERS predominant in the theta/alpha frequencies and ERD predominant in the beta frequency range (**Supplementary Figure 1**).

3) Cortical source localization was performed to map the underlying neural source activations for the ERSPs using the block-Sparse Bayesian learning (SBL) algorithm[83,84] implemented in a recursive fashion. This is a two-step algorithm in which the first-step is equivalent to low-resolution electromagnetic tomography (LORETA)[86]. LORETA estimates sources subject to smoothness constraints, i.e. nearby sources tend to be co-activated, which may produce source estimates with a high number of false positives that are not biologically plausible. To guard against this, SBL applies sparsity constraints in the second step wherein blocks of irrelevant sources are pruned. Source space activity signals were estimated and then their root mean squares were partitioned into (1) regions of interest (ROIs) based on the standard 68 brain region Desikan-Killiany atlas[32] shown in **Supplementary Figure 4**, using the Colin-27 head model[87] and (2) artifact sources contributing to EEG noise from non-brain sources such as electrooculographic, electromyographic or unknown sources; activations from non-brain sources were removed to clean the EEG data. Cleaned subject-wise trial-averaged EEG data were then specifically filtered in theta (3-7 Hz), alpha (8-12 Hz), and beta (13-30 Hz) bands and separately source localized in each of these bands to estimate their cortical ROI source signals. The source signal envelopes were

computed in MatLab (envelop function) by a spline interpolation over the local maxima separated by at least one time sample; we used this spectral amplitude signal for all neural analyses presented here. We focused on selected choice based processes in the 0-500 ms period in all frequency bands, reward processes during 0-500 ms after immediate reward presentation, and 0-500 ms after cumulative reward presentation (**Figure 1**).

Statistical Analyses

We fit robust multivariate linear regression models in MATLAB to investigate the behavioral relationships between the *Perf* measure and demographic variables – age, sex, race, ethnicity and SES, after controlling for *Bias* and order of block presentation. The response variable was log-transformed for normality and we identified significant factors contributing to the main effects. Similarly, we investigated the relationships between each of the mental health factors – anxiety, depression, inattention, hyperactivity and *Perf* measure after controlling for *Bias* and demographic variables. For regression models, we report the overall model R^2 and p-value, and individual variable β coefficients, t-statistic, degrees of freedom, and p-values. Effect sizes (Selya et al. 2012), 0.02 small, 0.15 medium, 0.35 large were calculated as $f^2 = (R^2_{FullModel} - R^2_{RestrictedModel})/(1 - R^2_{FullModel})$.

Channel-wise theta, alpha, beta ERS and ERD modulations for significant spectral activity were computed relative to baseline by first processing for any outliers; any activations greater than 5MAD from the median were removed from further analyses. The significant average activity across all trials were found by performing t-tests ($p<0.05$) across subjects, followed by false discovery rate (FDR, alpha = 0.05) corrections applied across the three dimensions of time, frequency, and channels[88].

For computing source level activity correlates of the behavioral *Perf* measure, we first found the difference in RareG door specific neural activations between Δpayoff and $\Delta_0$payoff blocks in three frequency bands – theta, alpha and beta and in three trial periods – selected choice (before reward), reward and cumulative reward time periods. We again used robust linear regression fits for identifying individual ROIs that relate to the *Perf* measure accounting for *Bias* that significantly affect the payoff decisions, as well as the five RL model parameters (α, γ, λ, β and σ). The results were family-wise error rate corrected for multiple comparisons for 3 trial periods and 3 frequency bands (FWER correction, $p<0.0055$). The independently identified ROIs (**Supplementary Table 2**) were further factored in a unified multivariate linear regression model to account for comparisons across ROIs; significant ROIs in this multivariate model were reported ($p<0.05$) after controlling for *Bias*, and the five RL model parameters. Similar steps were used for computing the source level activity correlates for the gain *Bias* measure. We first found the difference in RareL and RareG door specific neural activations in the $\Delta_0$payoff blocks in three frequency bands – theta, alpha and beta and in the three trial periods corresponding to selected choice, reward and cumulative reward. We then used robust linear regression fits for identifying individual ROI activations that relate to the *Bias* measure accounting for *Perf* that significantly affect the payoff decisions, as well as the five RL model parameters (α, γ, λ, β and σ). The results were family-wise error rate corrected for multiple comparisons for 3 trial periods and 3 frequency bands (FWER correction, $p<0.0055$). The independently identified ROIs were further factored in a unified multivariate linear regression model to account for comparisons across ROIs; significant ROIs in this multivariate model were reported ($p<0.05$) after controlling for *Perf*, and the five RL model parameters (**Supplementary Figure 2**).

Additionally, we used robust multivariate linear regressions to model the self-reported symptoms of Anxiety, Depression, Inattention, Hyperactivity using predictors of demographic variables, *Perf*, *Bias*, RL model parameters along with the identified neural correlates of *Perf* above. Adjusted responses from robust multivariate models were plotted using the plotAdjustedResponse function in MATLAB.

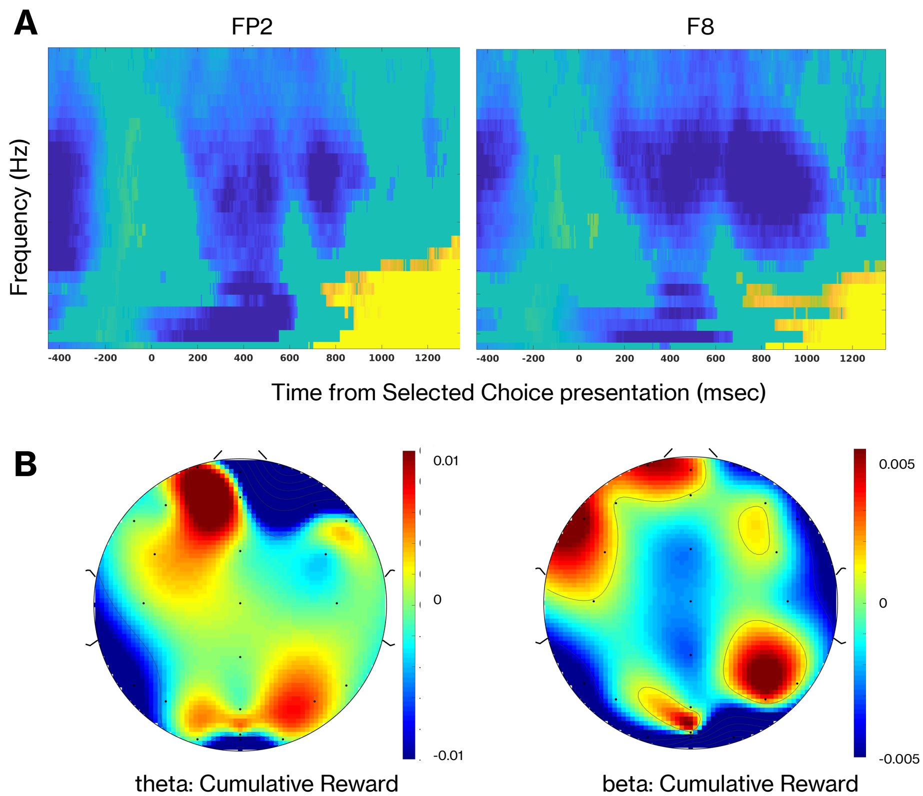
Supplementary material

	Δpayoff		$\Delta_0$payoff	
	RareL	**RareG**	**RareL**	**RareG**
1	20	-20	30	-30
2	20	-20	30	-30
3	20	-20	30	-30
4	20	-20	30	-30
5	20	-20	30	-30
6	20	-20	30	-30
7	20	-20	30	-30
8	-60	60	-70	70
9	-60	60	-70	70
10	-60	60	-70	70
sum	-40	40	0	0
average	-4	4	0	0
variance	1493.33	1493.33	2333.33	2333.33

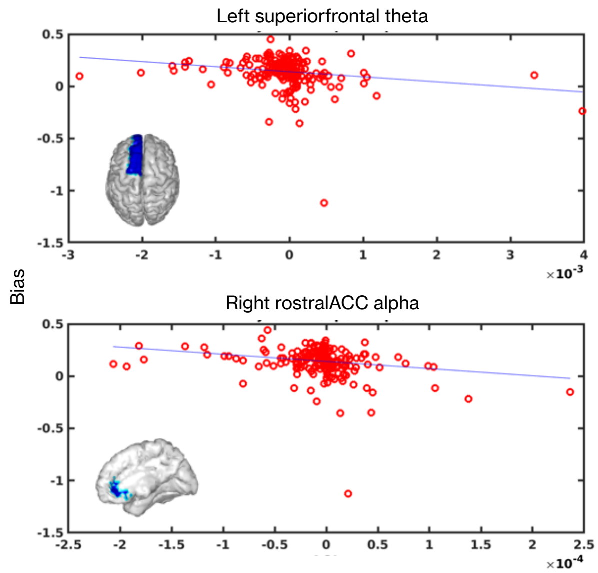
Supplementary Table 1. **Reward distributions for the door choices in the Δpayoff and $\Delta_0$payoff blocks**. The two door choices in either block were RareG (rare gains and frequent losses) and RareL (rare losses and frequent gains). RareG/RareL distributions had the same sum, average and variance in the $\Delta_0$payoff block, and different sum and averages but same variance in the Δpayoff block. Payoff (expected value) for RareG/RareL were the same in the $\Delta_0$payoff block and greater for RareG relative to RareL in the Δpayoff block.

Epoch	Freq	ROI	β	β SE	tstat	pValue
Selected choice	alpha	caudalanteriorcingulate R	-184.58	60.32	-3.06	2.58E-03
Selected choice	beta	lingual L	-29.87	10.34	-2.89	4.35E-03
Reward	beta	superiorfrontal L	-80.41	28.57	-2.81	5.48E-03
Reward	beta	superiorfrontal R	-115.88	34.43	-3.37	9.46E-04
CumuReward	theta	rostralanteriorcingulate R	-44.41	13.71	-3.24	1.44E-03
CumuReward	beta	parsorbitalis R	-79.67	23.05	-3.46	6.96E-04
CumuReward	beta	superiorfrontal L	-95.63	32.76	-2.92	4.00E-03
CumuReward	beta	superiorfrontal R	-105.19	33.21	-3.17	1.83E-03
CumuReward	beta	supramarginal R	-38.85	7.46	-5.21	5.44E-07

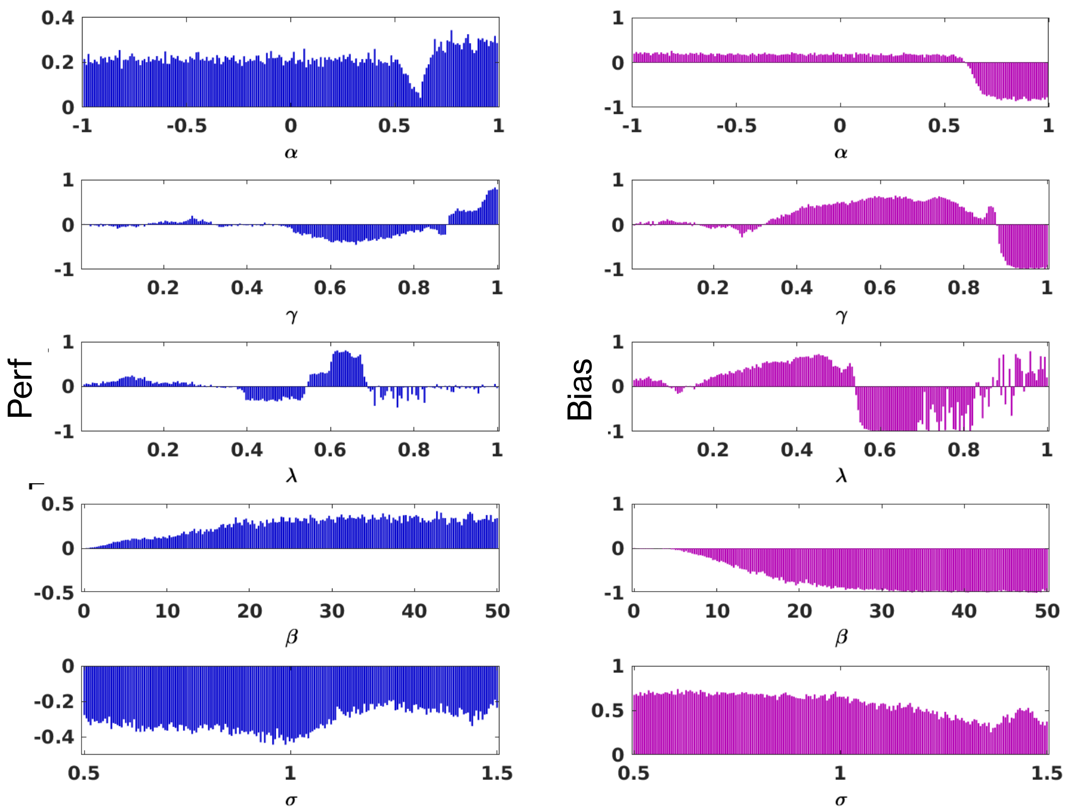
Supplementary Table 2: ROIs sensitive to Payoff-based decision-making performance (*Perf*) during selected choice presentation: 0-500 ms selected choice period; Reward: 0-500 ms immediate reward period and CumuReward: 0-500 cumulative reward period.



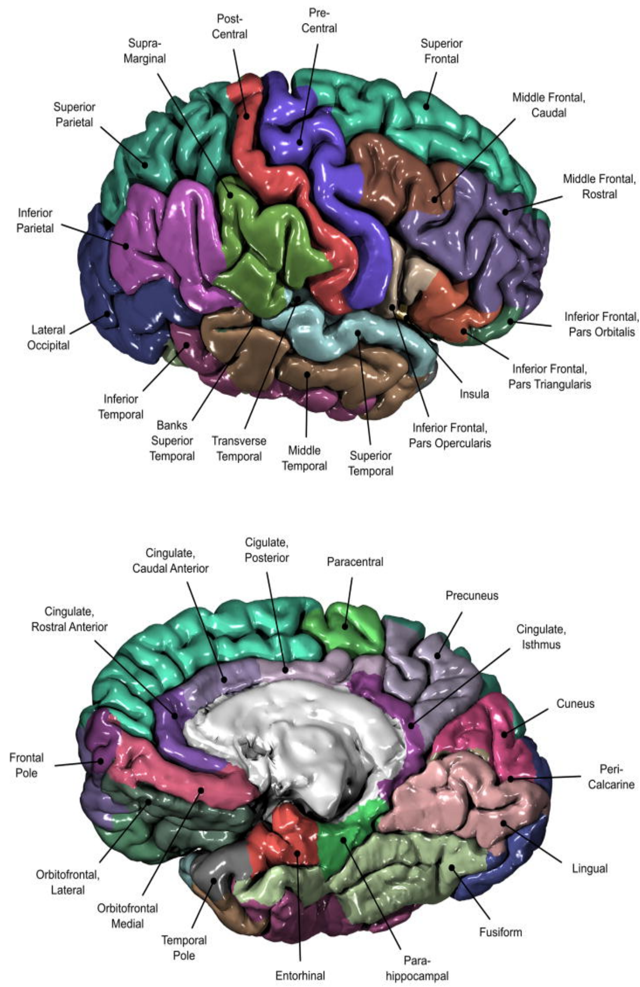
Supplementary Figure 1: (A) Grand-averaged spectral perturbations at scalp channels FP2 and F8 representing the event-related synchronizations and desynchronizations, closest in proximity to the identified *Perf* activations in the right rostral anterior cingulate cortex, and right pars opercularis, respectively. (B) Corresponding *Perf* scalp topographies, as differences in RareG trial activations between the Δpayoff and the $\Delta_0$payoff blocks for the relevant (500 ms-averaged) theta and beta activations are shown.



Supplementary Figure 2: Neural correlates of Gain frequency bias-based decision making. Bias based performance was negatively predicted by left superior frontal cortex theta activity in the cumulative reward period, and by right rostral anterior cingulate cortex (ACC) alpha activity in the cumulative reward period.



Supplementary Figure 3. Sensitivity analysis results for each model meta-parameter, α, γ, λ, β, σ, presented in rows and their simulation-derived model *Perf* and gain frequency *Bias* outcomes presented as corresponding columns. Only one parameter was varied at a time while the other parameters were held to their population median.



Supplementary Figure 4. Cortical source regions as per the Desikan-Killiany atlas (Desikan *et al.*, 2006).

Acknowledgements

This work was supported by University of California San Diego (UCSD) lab start-up funds (DR, JM), the Interdisciplinary Research Fellowship in NeuroAIDS (PB: R25MH081482), the Brain & Behavior Research Fund (PB), the Kavli Foundation (PB, JM), and the Sanford Institute for Empathy and Compassion (JM, PB). We thank Alankar Misra for software development of the *BrainE* software and several UCSD undergraduate students who assisted with data collection. The *BrainE* software is copyrighted for commercial use (Regents of the University of California Copyright #SD2018-816) and free for research and educational purposes. We thank Sabyasachi Shivkumar and Vignesh Muralidharan for their helpful feedback on the study analysis.

Data Availability

A part of the dataset used in this study is available on the open-access repository link: 10.5281/zenodo.4088951

Conflict of Interest.

The authors declare no conflict of interest.

References

1. Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. Deciding advantageously before knowing the advantageous strategy. *Science* **275**, 1293–1295 (1997).
2. Brevers, D., Bechara, A., Cleeremans, A. & Noel, X. Iowa Gambling Task (IGT): twenty years after – gambling disorder and IGT. *Front. Psychol.* **4**, (2013).
3. Bechara, A., Tranel, D. & Damasio, H. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* **123**, 2189–2202 (2000).
4. Bembich, S. *et al.* Differences in time course activation of dorsolateral prefrontal cortex associated with low or high risk choices in a gambling task. *Frontiers in human neuroscience* **8**, 464 (2014).
5. Christakou, A., Brammer, M., Giampietro, V. & Rubia, K. Right ventromedial and dorsolateral prefrontal cortices mediate adaptive decisions under ambiguity by integrating choice utility and outcome evaluation. *Journal of Neuroscience* **29**, 11020–11028 (2009).

6. Garon, N., Moore, C. & Waschbusch, D. A. Decision Making in Children With ADHD Only, ADHD-Anxious/Depressed, and Control Children Using a Child Version of the Iowa Gambling Task. *J Atten Disord* **9**, 607–619 (2006).
7. Gupta, R., Koscik, T. R., Bechara, A. & Tranel, D. The amygdala and decision-making. *Neuropsychologia* **49**, 760–766 (2011).
8. Mueller, E. M., Nguyen, J., Ray, W. J. & Borkovec, T. D. Future-oriented decision-making in Generalized Anxiety Disorder is evident across different versions of the Iowa Gambling Task. *Journal of Behavior Therapy and Experimental Psychiatry* **41**, 165–171 (2010).
9. Must, A., Horvath, S., Nemeth, V. L. & Janka, Z. The Iowa GamblingTask in depression – what have we learned about sub-optimal decision-making strategies? *Frontiers in psychology* **4**, 1–6 (2013).
10. Oberg, S. A., Christie, G. J. & Tata, M. S. Problem gamblers exhibit reward hypersensitivity in medial frontal cortex during gambling. *Neuropsychologia* **49**, 3768–3775 (2011).
11. Roca, M. *et al.* Executive Functions in Pathologic Gamblers Selected in an Ecologic Setting. *Cognitive and Behavioral Neurology* **21**, 1–4 (2008).
12. Woodrow, A. *et al.* Decision-making ability in psychosis: a systematic review and meta-analysis of the magnitude, specificity and correlates of impaired performance on the Iowa and Cambridge Gambling Tasks. *Psychological Medicine* **49**, 32–48 (2019).
13. Chiu, Y.-C. & Lin, C.-H. Is deck C an advantageous deck in the Iowa Gambling Task? *Behavioral and Brain Functions* **3**, 37 (2007).
14. Chiu, Y.-C. *et al.* Immediate gain is long-term loss: Are there foresighted decision makers in the Iowa Gambling Task? *Behavioral and Brain Functions* **4**, 13 (2008).

15. Lin, C., Chiu, Y. & Huang, J. Gain-loss frequency and final outcome in the Soochow Gambling Task: A Reassessment. *Behavioral and Brain Functions* **5**, 1–9 (2009).
16. Napoli, A. & Fum, D. Rewards and punishments in iterated decision making: An explanation for the frequency of the contingent event effect. in *10th International Conference on Cognitive Modeling. Philadelphia, PA* (Citeseer, 2010).
17. Singh, V. & Khan, A. Decision making in the reward and punishment variants of the Iowa gambling task: evidence of "foresight" or "framing"? *Frontiers in neuroscience* **6**, 107 (2012).
18. Franken, I. H. & Muris, P. Individual differences in decision-making. *Personality and Individual Differences* **39**, 991–998 (2005).
19. Furl, B. A. The influence of individual differences on the Iowa Gambling Task and real-world decision making. (Wake Forest University, 2010).
20. Harman, J. L. Individual differences in need for cognition and decision making in the Iowa Gambling Task. *Personality and individual differences* **51**, 112–116 (2011).
21. Newman, L. I., Polk, T. A. & Preston, S. D. Revealing individual differences in the Iowa Gambling Task. in 1067–1072 (Cognitive Science Society Austin, TX, 2008).
22. Weller, J. A., Levin, I. P. & Bechara, A. Do individual differences in Iowa Gambling Task performance predict adaptive decision making for risky gains and losses? *Journal of Clinical and Experimental Neuropsychology* **32**, 141–150 (2010).
23. Balasubramani, P. P. & Chakravarthy, V. S. Bipolar oscillations between positive and negative mood states in a computational model of Basal Ganglia. *Cogn Neurodyn* (2019) doi:10.1007/s11571-019-09564-7.

24. Balasubramani, P. P., Chakravarthy, S., Ravindran, B. & Moustafa, A. A. An extended reinforcement learning model of basal ganglia to understand the contributions of serotonin and dopamine in risk-based decision making, reward prediction, and punishment learning. *Frontiers in Computational Neuroscience* **8**, 47 (2014).
25. Chakravarthy, V. S., Balasubramani, P. P., Mandali, A., Jahanshahi, M. & Moustafa, A. A. The many facets of dopamine: Toward an integrative theory of the role of dopamine in managing the body's energy resources. *Physiology & behavior* (2018).
26. Gupta, A., Balasubramani, P. P. & Chakravarthy, S. Computational model of precision grip in Parkinson's disease: A Utility based approach. *Frontiers in Computational Neuroscience* **7**, (2013).
27. Muralidharan, V., Balasubramani, P. P., Chakravarthy, V. S., Lewis, S. J. & Moustafa, A. A. A computational model of altered gait patterns in parkinson's disease patients negotiating narrow doorways. *Frontiers in computational neuroscience* **7**, 190 (2014).
28. Doya, K. Metalearning and neuromodulation. *Neural networks* **15**, 495–506 (2002).
29. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction. Adaptive Computations and Machine Learning*. (MIT Press/Bradford, 1998).
30. Denison, R. N., Adler, W. T., Carrasco, M. & Ma, W. J. Humans incorporate attention-dependent uncertainty into perceptual decisions and confidence. *PNAS* **115**, 11090–11095 (2018).
31. Roelfsema, P. R. & Ooyen, A. van. Attention-gated reinforcement learning of internal representations for classification. *Neural computation* **17**, 2176–2214 (2005).
32. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980 (2006).

33. Balasubramani, P. P., Chakravarthy, V. S., Ali, M., Ravindran, B. & Moustafa, A. A. Identifying the Basal Ganglia Network Model Markers for Medication-Induced Impulsivity in Parkinson's Disease Patients. *PLoS One* **10**, e0127542 (2015).
34. Balasubramani, P. P., Moreno-Bote, R. & Hayden, B. Y. Using a Simple Neural Network to Delineate Some Principles of Distributed Economic Choice. *Frontiers in Computational Neuroscience* **12**, 22 (2018).
35. Gansler, D. A., Jerram, M. W., Vannorsdall, T. D. & Schretlen, D. J. Comparing alternative metrics to assess performance on the Iowa Gambling Task. *Journal of Clinical and Experimental Neuropsychology* **33**, 1040–1048 (2011).
36. Stocco, A., Fum, D. & Napoli, A. Dissociable processes underlying decisions in the Iowa Gambling Task: a new integrative framework. *Behavioral and Brain Functions* **5**, 1–12 (2009).
37. Balasubramani, P. P. & Hayden, B. Overlapping neural processes for stopping and economic choice in orbitofrontal cortex. *bioRxiv* 304709 (2018).
38. Kennerley, S. W., Behrens, T. E. & Wallis, J. D. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature neuroscience* **14**, 1581 (2011).
39. Moccia, L. *et al.* Neural correlates of cognitive control in gambling disorder: a systematic review of fMRI studies. *Neuroscience and Biobehavioral Reviews* **78**, 104–116 (2017).
40. Cavanagh, J. F., Frank, M. J., Klein, T. J. & Allen, J. J. Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage* **49**, 3198–3209 (2010).
41. Christie, G. J. & Tata, M. S. Right frontal cortex generates reward-related theta-band oscillatory activity. *Neuroimage* **48**, 415–422 (2009).

42. Zavala, B. *et al.* Cognitive control involves theta power within trials and beta power across trials in the prefrontal-subthalamic network. *Brain* **141**, 3361–3376 (2018).

43. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nature Neuroscience* **10**, 1214–1221 (2007).

44. Krain, A. L., Wilson, A. M., Arbuckle, R., Castellanos, F. X. & Milham, M. P. Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making. *NeuroImage* **32**, 477–484 (2006).

45. Paulus, M. P. & Frank, L. R. Anterior cingulate activity modulates nonlinear decision weight function of uncertain prospects. *Neuroimage* **30**, 668–677 (2006).

46. Preuschoff, K., Bossaerts, P. & Quartz, S. R. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* **51**, 381–90 (2006).

47. Aron, A. R. *et al.* Converging Evidence for a Fronto-Basal-Ganglia Network for Inhibitory Control of Action and Cognition. *Journal of Neuroscience* **27**, 11860–11864 (2007).

48. Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J. & Owen, A. M. The role of the right inferior frontal gyrus: inhibition and attentional control. *NeuroImage* **50**, (2010).

49. Muralidharan, V., Yu, X., Cohen, M. X. & Aron, A. R. Preparing to stop action increases beta band power in contralateral sensorimotor cortex. *Journal of cognitive neuroscience* **31**, 657–668 (2019).

50. Picazio, S. *et al.* Prefrontal Control over Motor Cortex Cycles at Beta Frequency during Movement Inhibition. *Curr Biol* **24**, 2940–2945 (2014).

51. Tops, M. & Boksem, M. A. S. *A potential role of the inferior frontal gyrus and anterior insula in cognitive control, brain rhythms, and event-related potentials. Front. Psychol. 2 (330)*. (2011).

52. Gehring, W. J. & Willoughby, A. R. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* **295**, 2279–2282 (2002).
53. Janssen, D. J. C., Poljac, E. & Bekkering, H. Binary sensitivity of theta activity for gain and loss when monitoring parametric prediction errors. *Social Cognitive and Affective Neuroscience* **11**, 1280–1289 (2016).
54. van Noordt, S. & Segalowitz, S. J. Performance monitoring and the medial prefrontal cortex: a review of individual differences and context effects as a window on self-regulation. *Front. Hum. Neurosci.* **6**, (2012).
55. Rogers, R. D. *et al.* Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological psychiatry* **55**, 594–602 (2004).
56. Sallet, J. *et al.* Expectations, gains, and losses in the anterior cingulate cortex. *Cognitive, Affective, & Behavioral Neuroscience* **7**, 327–336 (2007).
57. Schutter, D. J. L. G., de Haan, E. H. F. & van Honk, J. Anterior asymmetrical alpha activity predicts Iowa gambling performance: distinctly but reversed. *Neuropsychologia* **42**, 939–943 (2004).
58. Shenhav, A., Botvinick, M. M. & Cohen, J. D. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217–240 (2013).
59. Telpaz, A. & Yechiam, E. Contrasting losses and gains increases the predictability of behavior by frontal EEG asymmetry. *Front. Behav. Neurosci.* **8**, (2014).
60. Balasubramani, P. P., Chakravarthy, V. S., Ali, M., Ravindran, B. & Moustafa, A. A. Identifying the Basal Ganglia Network Model Markers for Medication-Induced Impulsivity in Parkinson's Disease Patients. *PLoS One* **10**, (2015).

61. Gradin, V. B. *et al.* Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* **134**, 1751–1764 (2011).
62. Groen, Y., Gaastra, G. F., Lewis-Evans, B. & Tucha, O. Risky behavior in gambling tasks in individuals with ADHD–a systematic literature review. *PLoS One* **8**, e74909 (2013).
63. Miller, W. R. & Seligman, M. E. Depression and the perception of reinforcement. *Journal of Abnormal Psychology* **82**, 62–73 (1973).
64. Silvetti, M., Wiersema, J. R., Sonuga-Barke, E. & Verguts, T. Deficient reinforcement learning in medial frontal cortex as a model of dopamine-related motivational deficits in ADHD. *Neural networks* **46**, 199–209 (2013).
65. Ziegler, S., Pedersen, M. L., Mowinckel, A. M. & Biele, G. Modelling ADHD: A review of ADHD theories through their predictions for computational models of decision-making and reinforcement learning. *Neuroscience & Biobehavioral Reviews* **71**, 633–656 (2016).
66. Eshel, N. & Roiser, J. P. Reward and punishment processing in depression. *Biological psychiatry* **68**, 118–124 (2010).
67. Pizzagalli, D. A., Sherwood, R. J., Henriques, J. B. & Davidson, R. J. Frontal brain asymmetry and reward responsiveness: a source-localization study. *Psychological Science* **16**, 805–813 (2005).
68. Boes, A. D., McCormick, L. M., Coryell, W. H. & Nopoulos, P. Rostral Anterior Cingulate Cortex Volume Correlates with Depressed Mood in Normal Healthy Children. *Biological Psychiatry* **63**, 391–397 (2008).
69. Yoshimura, S. *et al.* Rostral anterior cingulate cortex activity mediates the relationship between the depressive symptoms and the medial prefrontal cortex activity. *Journal of Affective Disorders* **122**, 76–85 (2010).

70. Pizzagalli, D. A. *et al.* Pretreatment Rostral Anterior Cingulate Cortex Theta Activity in Relation to Symptom Improvement in Depression: A Randomized Clinical Trial. *JAMA Psychiatry* **75**, 547–554 (2018).

71. Arevalo-Rodriguez, I. *et al.* Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews* (2015).

72. Spitzer, R. L., Kroenke, K., Williams, J. B. W. & Loewe, B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of internal medicine* **166**, 1092–1097 (2006).

73. Kroenke, K., Spitzer, R. L. & Williams, J. B. W. The PHQ-9: Validity of a Brief Depression Severity Measure. *Journal of general internal medicine* **16**, 606–613 (2001).

74. Admon, R. & Pizzagalli, D. A. Dysfunctional reward processing in depression. *Current Opinion in Psychology* **4**, 114–118 (2015).

75. Dillon, D. G. *et al.* Peril and pleasure: An RDOC-inspired examination of threat responses and reward processing in anxiety and depression. *Depression and anxiety* **31**, 233–249 (2014).

76. Luman, M., Tripp, G. & Scheres, A. Identifying the neurobiology of altered reinforcement sensitivity in ADHD: A review and research agenda. *Neuroscience & Biobehavioral Reviews* **34**, 744–754 (2010).

77. Balasubramani, P. P. *et al.* Mapping Cognitive Brain Functions at Scale. *NeuroImage* 117641 (2020).

78. Kothe, C., Medine, D., Boulay, C., Grivich, M. & Stenner, T. *'Lab Streaming Layer' Copyright*. (2019).

79. Bell, D. E. Risk, return, and utility. *Management science* **41**, 23–30 (1995).

80. d'Acremont, M., Lu, Z.-L., Li, X., Van der Linden, M. & Bechara, A. Neural correlates of risk prediction error during reinforcement learning in humans. *Neuroimage* **47**, 1929–1939 (2009).

81. Balasubramani, P. P., Chakravarthy, S., Ravindran, B. & Moustafa, A. A. A network model of basal ganglia for understanding the roles of dopamine and serotonin in reward-punishment-risk based decision making. *Name: Frontiers in Computational Neuroscience* **9**, 76 (2015).

82. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* **134**, 9–21 (2004).

83. Ojeda, A., Kreutz-Delgado, K. & Mullen, T. Fast and robust Block-Sparse Bayesian learning for EEG source imaging. *Neuroimage* **174**, 449–462 (2018).

84. Ojeda, A., Klug, M., Kreutz-Delgado, K., Gramann, K. & Mishra, J. A Bayesian framework for unifying data cleaning, source separation and imaging of electroencephalographic signals. *bioRxiv* 559450 (2019) doi:10.1101/559450.

85. Pfurtscheller, G. EEG event - related desynchronization (ERD) and event - releated synchronization (ERS). *Electroencephalography: Basic Principles, Clinical Applications and Releated Fields* 958–967 (1999).

86. Pascual-Marqui, R. D., Michel, C. M. & Lehmann, D. Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology* **18**, 49–65 (1994).

87. Holmes, C. J. *et al.* Enhancement of MR Images Using Registration for Signal Averaging. *Journal of Computer Assisted Tomography* **22**, 324–333 (1998).

88. Genovese, C. R., Lazar, N. A. & Nichols, T. Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. *NeuroImage* **15**, 870–878 (2002).