**The Dissertation Committee for Mei-Yen Chen**
**Certifies that this is the approved version of the following dissertation:**

# The Development of Bias in Perceptual and Financial Decision-making

**Committee:**

Russell A. Poldrack, Supervisor

W. Todd Maddox, Co-Supervisor

Alexander C Huk

Jonathan Pillow

Inderjit S. Dhillon

# The Development of Bias in Perceptual and Financial Decision-making

**by**

**Mei-Yen Chen, B.S.; M.S.**

**Dissertation**

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

**Doctoral of Philosophy**

**The University of Texas at Austin**
**August, 2014**

## Acknowledgements

I am thankful to many people who contribute their time and efforts to my dissertation. I first would like to thank my two advisors, Drs. Russell Poldrack and Todd Maddox. I would not be able to carry out this intellectual journey without their guidance. I also appreciate my committee members, Drs. Jonathan Pillow, Alex Huk, and Inderjit S. Dhillon, for the insightful comments and helping develop mathematical models.

Additionally, I feel grateful to the support from the Poldracklab and the Imaging Research Center. I thank Drs. Koji Jimura, Corey N. White, Jeanette Mumford, Jeffery Luci, Sanmi Koyejo, Tom Schonberg, Sarah Helfinstein, Akram Bakkour, and Nick Malecek for the critical and inspiring discussions and technical supports throughout my research. Also, finishing the massive number of brain scans on time and on budget is a highlight of my dissertation process. I thank Albert J. Elumn, Nathan Giles, Sagar Parikh, Ungi Kim, Brenda Gregory, Ashleigh Hover, Jave Del Rosario, Grace Shearrer, Joy Smyrson, Yu Cai, and Adiel Carlo for their contribution in data collection.

Last but not least, I thank my family for placing their unyielding faith on my abilities. And I cherish all the memories and fun of being in the US together with all my friends.

# The Development of Bias in Perceptual and Financial Decision-making

Mei-Yen Chen, Ph.D.

The University of Texas at Austin, 2014

Supervisors: Russell A. Poldrack and W. Todd Maddox

Decisions are prone to bias. This can be seen in daily choices. For instance, when the markets are plunging, investors tend to sell stocks instead of purchasing them with lower prices because people in general are more sensitive to the potential losses than the potential gains, or loss averse, in making financial choices. This also can be seen in laboratory tests. When participants receive higher payoffs for successfully discriminating a visual stimulus as one choice against the other, they begin choosing this higher-rewarded option more often even though the objective evidence indicates the alternative. In my dissertation, I used mathematical models and functional magnetic resonance imaging (fMRI) to track the development of bias in perceptual and financial decision-making and presented evidence characterizing the experience-sensitive and domain-general decision-making process in the human brains. The first chapter showed that bias could be developed through associating decision contexts and reward feedback from trial to trial in perceptual decision-making. Although the surface task differed, this learning process involved the same prediction error driven mechanisms implemented in the dopaminergic system as in financial decision-making. Furthermore, the frontal cortex increased its strength of connection between visual and value systems that accounted for

the growth of perceptual bias. The second chapter extended this feedback-driven acquisition process to examine the influences of experience on loss aversion in financial decision-making. The results showed that people learned to make riskier or more conservative decisions according to the feedback that they had received in different decision contexts. This alternation in loss aversion was achieved through modulation of the value system's sensitivity toward the potential gains in evaluation. The frontal cortex mediated this change. The third chapter used a mathematical model to identify the changes in financial decision-making that occurred faster than the temporal resolution of fMRI. The results suggested that people might simplify financial information into some rules of thumb for making a choice. These findings not only integrated the knowledge in different domains of decision neuroscience but also shed lights onto how one may refine the decision-making process against experiences.

# Table of Contents

# List of Tables

# List of Figures

# Introduction

Decisions are susceptible to bias. This is commonly seen in choices that people make on daily basis. For instance, it has been shown that if investors could abide by their financial plan, they would have gained more returns from the markets. However, since the potential losses have greater psychological impact than the potential gains, people tend to deviate from their plan by purchasing stocks when the markets roar and selling them when the markets plunge. Biased decisions are also ubiquitously observed across species when tested in laboratory settings. One easy way to demonstrate how much choices can depart from the objective evidence presenting in front of a decision maker is using perceptual discrimination tasks. When participants receive higher payoffs for successfully discriminating a visual stimulus as one choice against the other, they begin choosing this higher-rewarded option more often even though the visual stimulus indicates the alternative.

For decades, research into the neural mechanisms that process objective decision evidence and those that contribute to biases in decision-making has proceeded in parallel domains with limited crosstalk. Specifically, the literature on perceptual decisions has focused on how an individual's choices are influenced by the quality of sensory evidence, such as brightness, motion strength, vibration…etc. On the other hand, the literature on economic decisions has emphasized how individuals develop preferential choices from previous choice outcomes (Glimcher 2011; Lee, Seo, and Jung 2012; Montague and

Berns 2002). Recently, converging evidence from animal physiology (Ding and Gold 2013) and human neuroimaging (Summerfield and Tsetsos 2012) has motivated a call to investigate behavior that links these two factors—evidence and experience—in order to identify the domain-general decision-making processes in the brain.

The goal of my dissertation is to offer an integrated view of decision-making process in the human brains by tracing the development of bias in different domains of decision-making. To achieve this goal, functional magnetic resonance imaging (fMRI) and mathematical models were used to synthesize data collected from seemly different behavioral tasks into a common framework. The details of these studies and findings were reported in three chapters. In Chapter 1, I used a perceptual discrimination task to test the hypothesis that the reward feedback from trial to trial contributed to bias in perceptual decisions and that this development involved the reinforcement learning mechanism implemented in dopaminergic systems as shown in economic decision-making. In Chapter 2, I examined the influence of choice feedback on a common bias in financial decision-making process--loss aversion. I tested the hypothesis that pairing decision contexts and feedback could alter loss aversion and that this adjustment was achieved by tuning the interaction between cognitive control and value systems. In Chapter 3, I developed a mathematical model to identify the experience-induced changes in financial decision-making in order to characterize those changes in the decision-making process that were beyond the limited temporal resolution of the functional brain

imaging technique. This complimented the findings in Chapter 2. Overall, these findings not only bridge the literature gap between different domains of decision neuroscience but also provide insights into how one may refine daily choices in the face of experiences.

# Chapter 1: Multiple brain networks contribute to the acquisition of bias in perceptual decision-making

## INTRODUCTION

Decisions are driven both by the objective evidence presented to an individual and by the outcomes that the individual has learned to expect from the past. For decades, research in the neural mechanisms that processes each of the two factors in decision-making has proceeded in parallel with very little crosstalk. The literature on perceptual decisions has focused on how an individual's choices are influenced by the quality of sensory evidence, whereas the literature on economic decisions has emphasized how an individual's choices are driven by the expected reward arising from previous choice outcomes (Glimcher 2011; Lee, Seo, and Jung 2012; Montague and Berns 2002). Recently, converging evidence from animal physiology (Ding and Gold 2013) and human neuroimaging (Summerfield and Tsetsos 2012) has motivated a call to investigate behavior that links these two factors in order to identify the general neural mechanisms of decision-making processes across domains. Reward-induced bias in perceptual decisions (Edwards 1965; Green and Swets 1966) is a phenomenon that sits exactly at this literature gap. Investigating the neural mechanisms underlying such bias hence will provide an integrated view of decision-making processes in the brain.

Expected reward has a profound influence on perceptual decisions. When prompted to classify sensory information as one of the two alternatives offering

asymmetric payoffs, both humans and animals tend to prefer the higher-rewarded alternative (Feng, Holmes, Rorie, and Newsome 2009; Fleming, Whiteley, Hulme, Sahani, and Dolan 2010; Liston and Stone 2008; Mulder, Wagenmakers, Ratcliff, Boekel, and Forstmann 2012; Rorie, Gao, McClelland, and Newsome 2010; Summerfield and Koechlin 2010; Whiteley and Sahani 2008). At the behavioral level, this choice preference can be identified from the sigmoidal relationship between the strength of sensory evidence and the probability of choosing one of the alternatives, or psychometric function. The amount of bias is quantified as the horizontal shift of the indecision point in the psychometric function (Gold and Ding 2013; Green and Swets 1966; Macmillan and Creelman 2004). The choice preference also can be characterized by reaction time since the biased choices usually are made faster (Mulder et al. 2012; Summerfield and Koechlin 2010). Drift-diffusion modeling of the choices and reaction times suggests that information about reward generally affects the early stage of decision process (Mulder et al. 2012; Summerfield and Koechlin 2010; but see Blank, Biele, Heekeren, and Philiastides 2013 for alternatives), so that less sensory evidence is required to be accumulated in order to support the more beneficial option.

In contrast to the well-established theoretical and empirical work at the behavioral level, the neural mechanisms by which reward information contributes to bias in perceptual decisions remain an open question. Specifically, it remains unclear how the expectation of reward is formed in a neural system during perceptual decisions and once

formed, how this reward information influences different neural components of the perceptual decision process. One possibility is that reward information is formed as a task-set (Summerfield, Egner, Mangels, and Hirsch 2006) in the higher level (i.e. frontal and parietal cortices) of the perceptual decision hierarchy. It then selectively influences the sensory and motor system to facilitate the choice of higher-rewarded options (Duncan 2001; Liston and Stone 2008). However, this account is only partially supported by evidence from human functional magnetic resonance imaging (fMRI). After human participants receive explicit instructions for the payoffs of each perceptual choice, the activation of fronto-parietal cortices positively correlates with the amount of bias in their perceptual decisions (Fleming et al. 2010; Mulder et al. 2012; Summerfield and Koechlin 2010). Yet, the correlation between perceptual bias and activation in sensory or motor cortices is inconsistent across studies (Fleming et al. 2010; Mulder et al. 2012; Summerfield and Koechlin 2010; Serences 2008).

Another possibility is that bias in perceptual decisions reflects the subjective value that is learned from the previous choice outcomes. Evidence shows that as macaque monkeys developed a bias toward the higher-rewarded options in a motion discrimination task, the firing-rate of midbrain dopaminergic neurons increased when the animals received greater reward for correctly identifying the motion direction, and decreased when reward was absent after an incorrect choice (Nomoto, Schultz, Watanabe, and Sakagami 2010). This activation pattern is analogous to a reinforcement learning signal

(Schultz 1998; Sutton and Barto 1998) in which the difference between the expected and actual choice outcomes (reward prediction error, RPE) is used to update the subjective value of each option, which is putatively encoded in the ventromedial prefrontal cortex (vmPFC) (Bartra, McGuire, and Kable 2013; Garrison, Erdeniz, and Done 2013; Glimcher 2011; Montague and Berns 2002). The value difference between each option hence determines which options are worth repeating in the future (Lee et al. 2012). This observation has inspired theoretical work that combines drift-diffusion and reinforcement learning models to simulate the potential interaction between cortical and basal ganglia in the development of perceptual bias (Bogacz and Larsan 2011; Rao 2010). Nevertheless, until now, no empirical study has directly investigated this reinforcement-learning mechanism and its influence on different decision-making networks at the level of the whole-brain when individuals acquire perceptual bias.

In the present study, we approach these issues by tracking how humans developed bias in perceptual decisions using computational models and functional brain imaging. During fMRI acquisition, participants performed a motion discrimination task (Britten, Shadlen, Newsome, and Movshon 1993; Newsome, Britten, and Movshon 1989; Shadlen and Newsome 2001) with pre-trial cues signaling one of two different reward contexts. Trial-wise reward feedback was delivered to participants so that a correct response to one of the motion directions was reinforced more strongly in each reward context. To maximize reward, the subject must combine information about the stimulus

and the potential reward, such that predicted reward exerts a greater effect on choices when the stimulus is weak. As the experiment proceeded, the indecision point of perceptual choices gradually shifted when it was measured in different reward contexts. This shift was quantified using a reinforcement-learning model that estimated the subjective value of each option according to the association of reward prediction error (RPE) and the reward contexts (Watkins and Dayan 1992). Consistent with a role for value learning mechanisms, the RPE signals associated with the acquisition of perceptual bias positively correlated with the activation of ventral striatum, dlPFC, and parietal cortex. As the bias grows, we find that the value signals integrate into perceptual decision network through increasing functional connectivity in networks involved in motor preparation (vmPFC-motor cortex), stimulus evaluation (frontal-vmPFC-visual cortex), and cognitive control (parietal-anterior cingulate cortex [ACC]). These results enhance our understanding of the neural mechanisms underlying bias acquisition, and provide a fundamental linkage between the perceptual and economic decision-making processes in the brain.

## MATERIALS AND METHODS

### Participants

Twenty-four human participants completed the behavioral paradigm in the MRI scanner (12 females, 12 males; age range: 18 - 30). One participant was excluded because of extreme parameter estimates in the fMRI data analysis. All participants were recruited through posted flyers and were prescreened. They were free of any self-reported

neurological or psychiatric diseases, had normal or corrected-to-normal visual acuity and normal color vision, and right-handed. They gave written informed consent for participation. The Institutional Review Board of the University of Texas at Austin approved all experimental procedures.

**Stimuli**

All stimuli were generated in Matlab version 7.10.0, using the Psychophysics Toolbox extension, version 3.0.10 (Brainard 1997; Pelli 1997). Each motion stimulus was composed of 150 white dots moving inside a donut-shaped display patch on a black background. The display patch was centered on the screen and extended from 4 to 8 degrees of visual-angle. Within the display patch, every dot moved at the speed of 8 degrees of visual-angle per second. Some dots moved coherently toward one direction while the others moved randomly. The percentage of coherently moving dots determined the motion strength (coherence level). The presentation of the dots was controlled to remove local motion signals (Britten et al. 1993; Newsome et al. 1989; Palmer, Huk, and Shadlen 2005; Shadlen and Newsome 2001). Upon stimulus onset, the dots were randomly located on the first three video frames. They were relocated after two subsequent frames, so that the dots in frame 1 were repositioned in frame 4, and the dots in frame 2 were repositioned in frame 5, etc. When repositioned, each dot was either randomly presented at the new location or aligned with the pre-determined motion direction (upward or downward), depending on the pre-determined motion strength on

that trial. Each stimulus was composed of 12 video frames with 60 Hz video frame refresh rates.

**Procedures and task**

Participants first performed a practice session in the laboratory to familiarize themselves with the random-dot motion discrimination task. In the practice session, a trial began with a red fixation cross that was presented at the center of the display screen for 1.5s. Then, a patch of moving dots was presented for 200ms. After the stimulus offset, participants had to decide whether the global motion direction of these dots was upward or downward by pressing the corresponding spatially congruent buttons within 600ms. Error feedback was presented for 1.5s for incorrect or slow responses; otherwise, the next trial continued immediately with presentation of the fixation cross. On each trial, the motion stimulus was a random sample from one of the 9 coherence levels (0%, ±6%, ±12%, ±64%, ±80%; positive sign: upward motion, negative sign: downward motion). The correct response for 0% coherence trials was decided using a random number generator, so that the probability of being either correct or incorrect on this trial type was equal over the entire experiment. The total of 540 trials (9 coherence levels x 60 repetition) was broken down into six 90-trial blocks. The participants could take a break after completing each block.

The fMRI scan was conducted no more than 7 days after the practice session. In the scanner, participants were asked to decide the motion direction of moving-dots

presented in two independent reward contexts with a goal to get as many reward points as possible over the experiment. The task structure and the timeline of events on a trial are illustrated in Figure 1. Before a trial started, a white fixation-cross presented at the center of the display screen during the jittered inter-trial interval (truncated exponential distribution; mean: 4s, range 2.7 - 12.7s). At the beginning of a trial, the color of the fixation-cross changed into either blue or yellow for 1s to signal the reward context of the trial. Then, a motion stimulus was presented for 200ms. The participant had up to 800 ms from stimulus onset to decide the motion direction by pressing the corresponding button. After this 800-ms response window, a number appeared on the screen for 1.5s to inform the participant how many reward points that they earned from their decisions. The payoff of the two possible motion direction choices was associated with the reward contexts. In one context correct upward motion choices led to more reward points, whereas in the other state correct downward motion choices led to more reward points (Figure 1). The total reward points were converted into US dollar as bonus at the end of the experiment. The participants were unaware of this payoff structure before they began the task. They were simply instructed to decide the motion direction on each trial in order to harvest the most reward points.

Several procedures were implemented in the experimental design in order to rule out other potential sources of decision bias rather than reward itself. First, each context was paired with equal numbers of trials in each motion direction and the same levels of

motion discriminability, which controlled the prior-induced bias (Green and Swets 1966). The motion stimulus was a random sample from one of 7 coherence levels (0%, ±4%, ±12%, ±64%) on each trial. Every coherence level repeated 20 times within each state. The total 280 trials (2 reward contexts x 7 coherence levels x 20 repetition) were equally distributed in the 5 scanning runs for each participant. Moreover, the presentation of reward contexts were independent from trial to trial, which rules out the potential confound that participants used the sequential pattern to guide their choices (Daw, Gershman, Seymour, Dayan, and Dolan 2011; Glascher, Daw, Dayan, and O'Doherty 2010). Finally, the response buttons and the colors of the fixation cross were counter-balanced across participants in order to remove other potential confounds. The participants' choice and the reaction time were recorded on each trial.

**Behavioral data analysis**

We applied the hierarchical logistic models to evaluate run-by-run changes of bias and discrimination, using the lme4 package (http://cran.r-project.org/web/packages/lme4/index.html) in R Version 3.0.0 (http://www.r-project.org/). The full model included five exploratory regressors: coherence-level, reward context, run-number, context-by-run, and coherence-by-run interaction. The intercept was taken as a random effect across participants. When testing the learning effects across the five scanning runs, we used a Chi-square test to compare the goodness-of-fit of this full model against the model in which either of the interaction terms was reduced.

Furthermore, we applied computational models to capture the cross-correlation between the feedback on the previous trial and the choice on the next trial. We implemented two reinforcement-learning models with different hypotheses regarding how participants might use their experiences about contexts, choices, and rewards to develop biases in perceptual decision-making. In each case, a logit function was used to generate the probability of binary choices on every trial based on the reward each participant had received so far and the motion stimulus that was presented on the particular trial. The models are described in detail in the next section. Auto-correlation functions were computed from the residuals of the best-fit learning model at the group level using the *acf* function in R (Box, Jenkins, and Reinsel 1994; Pinheiro and Bates 2009) to identify additional factors potentially missing in the model.

**Reinforcement learning models**

We assumed that on the $t^{th}$ trial, an individual chose probabilistically according to the value difference of each motion direction ($Q_t(m_i)$; a binary variable i, respectively indicating the upward and downward motion) and the perceived motion strength. This relationship can be described by a logit function with the linear combination of the value difference between each motion direction and the perceived motion strength (Green and Swets 1966; Macmillan and Creelman 2004). Since the perceived motion strength was monotonic with the physical stimulus (Britten et al. 1993), we used the coherence level on trial t $(S_t)$ to model the contribution of physical stimuli to the choice. The probability of choosing upward motion hence is:

$$P_t(m_1) = {}^{1}\Big/_{\{1 + \exp\,[\beta_0 * \left(Q_t(m_0) - \ Q_t(m_1)\right) + \ \beta_1 * \ S_t]\}} \quad (1)$$

where the $b_0$ and $b_1$ reflected how the choice probability was influenced by the subjective value difference and motion stimulus on the current trial respectively.

The second part of the model was constructed to simulate how the trial-wise reward feedback was used to estimate the subjective value of each motion direction. Two Q-learning models (Watkins and Dayan 1992) were separately specified such that the reward prediction error (RPE) was used to update action value differently. The RPE was used to updated the action value pertained to the context in one model and the action value regardless of context in the other model.

On every trial, the context-dependent action value was adjusted by keeping track of the context in which a choice was made. The context-dependent RPE ($\delta_t'$) was computed as the difference between expected and the actual outcomes ($r_t$) followed by a choice made in the specific context. This error was scaled by a constant learning rate ($\alpha'$) and added into the value of each motion direction choice made in the same context before:

$$\delta_t = \ r_t - \ Q_t(m_t \mid c_t),$$

$$Q_{t+1}(m_t \mid c_t) = \ Q_t(m_t \mid c_t) + \ \alpha' * \ \delta_t' \quad (2)$$

where $c_t$ and $m_t$ takes binary values respectively for indicating the context on trial t and the motion direction choice made in this context.

A second model was constructed based on the assumption that context was ignored. In this case, the action value was adjusted by adding RPE ($\delta_t$), weighted by a constant learning rate ($\alpha$), to the value of each motion direction ($Q_t(m_t)$), regardless of the contexts where the choice was made:

$$\delta_t = r_t - Q_t(m_t)$$

$$Q_{t+1}(m_t) = Q_t(m_t) + \alpha * \delta_t \quad (3)$$

We used maximum likelihood method to obtain parameter estimates for each model. The likelihood that a sequence of choice and feedback (D) was generated by a set of free parameters ($\theta_M \in \{\beta_0, \beta_1, \alpha\}$) was the product of Equation 1 over the total trials (Daw, 2009):

$$P(\theta_M \mid D) \propto P(D \mid \theta_M) = \prod_t P_t(m_1 \mid \theta_M) \quad (4)$$

We fit each model's free parameters by minimizing the negative log of Equation 4 with nonlinear optimization function (fmin) in the Scipy toolbox for python (http://docs.scipy.org/doc/scipy-0.7.x/reference/optimize.html). We initialized the value of each motion direction choice with the participants' initial bias toward either motion direction, if they showed any in the practice session. Once the best-fit parameters were

determined for each participant, the indecision point on each trial was computed by solving the coherence level that yielded equal chance of choosing either motion direction given the rest of parameters in Equation 1. For model comparison, the Akaike Information Criterion (AIC) was computed by summing the negative log of Equation 4 over all participants and taking the number of parameters as a penalty term (Akaike 1974).

**MRI data acquisition**

Imaging data were collected using a Siemens Skyra 3T MR scanner. Functional data were collected using a T2*-weighted multi-band echo-planar imaging sequence (Moeller et al., 2010) with 60° flip-angle (TR: 2 s; TE: 30 ms; FOV: 256 mm, multi-band acceleration factor: 2, parallel acceleration factor: 2, matrix size: 128). Forty-eight oblique axial slices were collected in interleaved fashion with 2 mm isotropic resolution. To reduce dropout in orbito-frontal cortex, the slices were tilted at a 10-15° angle off of the anterior-commissure-posterior-commissure line and higher-order shimming was applied. T1-weighted anatomical images was collected using an MP-RAGE sequence with 9° flip angle (TR: 1.9 s; TE: 2.43 ms; FOV: 256 mm; Matrix size: 256 x 256, 192 slices; slice thickness: 1 mm).

**Image preprocessing and registration**

FMRI data preprocessing was carried out using FSL Version 5.0.1 (FMRIB's Software Library: www.fmrib.ox.ac.uk/). All image time series were aligned with the MCFLIRT tool, and the resulting motion parameters were used to compute frame-wise

displacement (FD) and temporal derivative of the root mean square variance over voxels (DVARS) to identify bad time points (FD>.5; DVARS>.5) (Power, Barnes, Snyder, Schlaggar, and Petersen 2012). The skull was removed from the functional images with the brain extraction tool (BET) and from the structural images using FreeSurfer (https://surfer.nmr.mgh.harvard.edu/). Spatial smoothing was applied using a Gaussian kernel of FWHM 5 mm. The grand-mean intensity was normalized over the entire 4D dataset by a single multiplicative factor, and a high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=50.0 s). This same high-pass filter was applied to the design matrix for analyzing the fMRI time-series. All functional images were registered to the high resolution structural image using Boundary-Based Registration (BBR) then the high resolution structural image to the MNI-152 2 mm template using the FLIRT linear registration (12 DOF) tool of FSL.

**fMRI analysis**

We used multi-stage general linear model (GLM) approach to analyze the brain imaging data, using FSL FEAT (FMRI Expert Analysis Tool) Version 6.00. The first-level model was estimated separately for each run and each participant. All five runs were combined within participant using a fixed-effects model. At the group level, the FLAME 1 mixed-effects model of FSL was applied to all participants (Worsley 2001). All the statistical maps were corrected by cluster-based random field theory using clusters determined by $Z > 2.3$ and a family-wise error corrected cluster significance threshold of

P=0.05 (Worsley 2001). The statistics maps of all analyses were projected onto the group-averaged brain from this study for visualization.

**GLM model**

The first level of GLM contained parametric modulated regressors to identify the brain mechanisms underlying the acquisition of bias in perceptual decisions as well as nuisance regressors to control for potential confound. The parametric modulated regressors included 1) the absolute value of the coherence level (duration between the onset and offset of the stimulus), 2) the trial-wise amount of bias (duration between trial onset and stimulus offset) and 3) the reward prediction error derived from both learning models (duration between the onset and offset of the reward feedback). The values of the bias were derived as the absolute value difference between the two choices from the best-fit reinforcement-learning model at the group-level for each participant. All the values of parametric modulated regressors were mean-centered before entering the GLM. Nuisance regressors in the model were 1) a boxcar regressor encoding trial-evoked activity (duration between the onset of the context and the next ITI), 2) a boxcar regressor between the stimulus onset and the time when a key press was detected to control the reaction time (RT), and 3) a confound file including all the motion correction parameters (estimated translation and rotation and their first derivatives, FD, and DVARS) together with single-time-point regressors for each time point that exceeded the FD/DVARS thresholds (which effectively performs "scrubbing" of those time points) (Power et al. 2012). All the regressors except the motion-correction regressors in the first-level model

were convolved with a double-gamma hemodynamic response function. Their temporal derivatives were also included in the model to accommodate for potential slice timing differences. Except for that the RT regressor that was orthogonalized relative to the regressor for the trial-evoked activity, all other regressors entered the GLM without orthogonalization.

**Psychophysiological interaction (PPI) analysis**

To examine the effect of the value system on the acquisition process of perceptual bias, we define a seed region (10-mm sphere around the vmPFC; MNI coordinates: X=-6, Y=39, Z=-8) according to previous results on value-based decisions (Tom, Fox, Trepel, and Poldrack 2007). Likewise, to examine the effect of the fronto-parietal system on the acquisition of bias, we defined seed regions for frontal cortex (MNI coordinates: X=-45, Y=21, Z=0) and parietal cortex (MNI coordinates: X=-36, Y=-39, Z=45) that have been replicated by previous study on perceptual bias (Fleming et al. 2010; Summerfield and Koechlin 2010). The BOLD activation of the seed regions was extracted from each participant's individual brain in each run. For each individual and each run, the neural signal of the seed region was estimated by deconvolving the BOLD signals using the deconvolution algorithm of SPM (Gitelman, Penny, Ashburner, and Friston 2003). The interaction between the seed region and the regressor modulated by trial-wise amount of bias was generated in the neural domain and then reconvolved with hemodynamic function. The first-level design matrix of the PPI analysis was the above-mentioned GLM design matrix with two additional regressors: 1) the raw time course extracted from the

seed, and 2) a PPI regressor (the interaction between the amount of bias and the mean BOLD response in the seed region) with duration between the trial onset and the stimulus offset.

## RESULTS

### Behavioral results

To visualize how the decisions changed as the experiment unfolded, the group average of the choice probability in each context is plotted against the coherence level across the five runs (Figure 2A). We first applied hierarchical logistic regression models to identify the factors driving the changes over the experiment. The intercept of each of these models was taken as a random effect across individual participants. The full model assumes that both the participants' ability to discriminate motion direction *and* their preference for one of the motion directions changes from run to run in the experiment. If either of these factors is not constant over the experiment, removing either term from the full model should significantly reduce the model fit to the data. We used a chi-square test for model comparison to evaluate whether the drop in goodness-of-fit between the full and the reduced models reaches significance. We find that the interaction between coherence level and run number can be eliminated from the full model ($\chi^2(4) = 6.15, p = .1881 > \alpha = .05$), suggesting that the participants' ability to discriminate motion direction did not change over the entire experiment. However, removing the interaction between the reward context and the run number from the full model

significantly reduces the model fit to the data ($\chi^2(4) = 72.22, p < .0001$), suggesting that the degree of bias did change across runs of the experiment.

We further examined how the participant acquired bias in perceptual decision-making over the experiment. The participant may utilize the previous reward that had been received in a particular context to guide the next choice that would be made in the same context along with the motion coherence of the present stimulus. Alternatively, the participant may simply adjust their perceptual choice according to the reward obtained on the previous trial (independent of context) as well as the strength of the motion stimulus. These two possibilities were evaluated through model comparison. We find that after controlling for individuals' abilities to discriminate motion direction, the reinforcement-learning model that associates previous reward and the next choice in the same context fits the data better than the learning model that ignores the context (Table 1).

One may suspect that the participant simply applied pre-existing knowledge about perceptual uncertainty and reward to make a choice rather than adjusting their choice from trial to trial (Whiteley and Sahani 2008). If this is the case, the above-mentioned hierarchical logistic regression model that treats each choice as an independent observation should fit the data better than the reinforcement-learning model that accounts for the cross-correlation between previous decision outcome and the next

choice made in the same context. The result of model comparison indicates that the reinforcement-learning model does provide a better fit to the data than the hierarchical logistic regression model (Table 1). Furthermore, using the reinforcement-learning model (Equation 1 and 2), we are able to track the trajectory of the indecision points from trial to trial (Figure 2B) and reproduce the choice probability up to the final trial of each run (solid lines in Figure 2A) at the group level.

One may also suspect that in addition to the context and reward association, the previous choice (Lau and Glimcher 2005) or even the sequential structure of stimulus types (Cho et al. 2002) alone could contribute to the observed bias. We calculated the average correlation between residuals that are lagged behind a certain number of trials (autocorrelation functions) after the context and reward association has been accounted for in the reinforcement-learning model. If the response-by-response or stimulus-by-stimulus structure was an additional source of bias in our experiment, we should observe that some of the autocorrelations in the residuals are significantly nonzero since these factors were ignored in the learning model. As illustrated in Figure 2C, we compared the autocorrelation functions estimated from the residuals against those estimated from an independent random process with the same number of trials. The dashed horizontal lines show the 95% confidence intervals for autocorrelations expected from an independent random process. All the autocorrelations estimated from the residuals fall well within this

confidence interval. This suggests that there is no appreciable temporal structure in the residuals after the context and reward association was taken into account.

## Neuroimaging results

### *The neural correlates of the acquired decision bias*

Through the analyses of the behavioral data, we found that the amount of perceptual bias on each trial can be modeled as the subjective value difference between the two motion directions at that particular time point of the reinforcement-learning process. If perceptual bias shares the same neural basis as economic decisions, we expect to see that the activation of the value-based decision network (Bartra et al. 2013) positively correlates with the amount of bias that has been acquired at the point of each trial of the experiment. Using the absolute value difference between the two motion directions in each context that is estimated from the best-fit reinforcement-learning model as a parametric modulated regressor, we find that the activation of vmPFC and midbrain dopaminergic areas is positively correlated with the bias acquired in each reward context (Figure 3; Table 2).

Moreover, we find that the bias acquired through reinforcement has a wider influence on the perceptual decision networks in the brain compared to that induced by explicit instructions about payoffs (Fleming et al. 2010; Summerfield et al. 2010; Mulder

et al. 2012) (Figure 3; Table 2). A consensus in human functional neuroimaging literature is that the perceptual bias induced by explicit instructions interferes with the intermediate stage of decision process in the fronto-parietal cortices. However, we find that the degree of acquired bias extends to a wider set of brain regions. In addition to the frontal and parietal regions, we also observe that the activation in the visual and motor cortices positively correlate with the amount of acquired bias from trial to trial. No regions showed negatively correlations with this regressor after whole-brain correction.

***Learning signals for the acquisition of decision bias***

Human fMRI studies have identified that when individuals use reward feedback to adjust the subjective value of taking each action, the signal pertaining to the difference between the expected and actual reward (reward prediction error, RPE) is represented in ventral striatum, a major target of midbrain dopaminergic neurons (Pagnoni, Zink, Montague, and Berns 2002). Recent findings further distinguish that when multiple reward contexts are involved in a task, the activation of fronto-parietal cortices in addition to the ventral striatum correlates with the PE signals (Daw et al. 2011; Glascher et al. 2010). Following the same framework, our reinforcement-learning model assumes that perceptual bias is acquired from adjusting the subjective value of each action (motion directions) in the context using prediction errors. Based on the above-mentioned findings, we suspect that our task may involve two types of RPE signals. One is derived from the action-outcome association; the other is derived from the action-outcome association that is contingent on a specific context. We used two learning models (see Methods section

for details) to generate each type of RPE as parametric modulated regressors in order to search for its neural correlates.

We found that in perceptual decisions, both the context-based and context-free RPE regressor reveals a similar pattern of brain activation. Activation in ventral striatum, fronto-parietal, visual, and motor cortices positively correlated with both RPE signals when each of the regressors was modeled separately in a GLM. This pattern is consistent with the meta-analysis result of RPE signals in the brain elicited by economic decisions (Garrison et al. 2013). However, distinct patterns were revealed when both of the regressors were presented together in the same GLM. The RPE derived from the contextual action-outcome association positively correlated with the ventral striatum, the fronto-parietal cortices, and the posterior and anterior cingulate gyrus (Figure 4, yellow-red; Table 2). In contrast, PE signals from the context-free RL model positively correlated with activity in the visual and motor cortices (Figure 4, blue-light blue; Table 2). Although interesting, the interpretation of these findings is made difficult by the very high degree of collinearity between the context-free and context-based prediction errors.

One may argue that the activation pattern that uniquely correlates with the context-based RPE simply reflects the level of surprise about the reward (Ding and Gold 2010; Gottlieb 2012). We therefore added the absolute value of this RPE into the above-mentioned model to adjust for the effect of surprise. The signed context-based RPE still

elicits the same activation pattern (whole-brain corrected) after the level of surprise is controlled in the statistical model. Furthermore, no brain regions show activation that correlates with this unsigned RPE after whole-brain correction for multiple comparisons. Thus, this result rules out the potential confound in the RPE signals owning to the surprise about the reward.

***The functional connectivity patterns underlying the growth of bias***

We further used psychophysiological interaction (PPI) analysis to examine how the perceptual and economic decision-making networks interact with each other as participants acquire bias in perceptual decisions. The reinforcement learning literature suggests that the value of each action is encoded in motor areas and sent to vmPFC to guide choices (Wunderlich, Rangel, and O'Doherty 2009). Using the same framework, our learning model assumes that perceptual bias at the behavioral level reflects the subjective value difference between each action (i.e. choosing one of the two motion directions) in the current context. If perceptual bias is guided by the value system, we should observe that the functional connectivity between vmPFC and motor region increases during the decision period as the bias grows. Using vmPFC (Tom et al. 2007) as the seed region in the PPI analysis, we find that the task-related interaction between vmPFC and motor cortex increases as the amount of bias increases over the experiment (Figure 5; Table 3).

We also evaluated the hypothesis that perceptual bias results from a task set encoded in the frontal and parietal cortices. According to this theory, bias arises from the mechanism that frontal and parietal regions provide top-down influence to facilitate the visual representation or motor response that favors the higher-valued option (Fleming et al. 2010; Summerfield et al. 2006). If this is the case, we should observe that the connectivity between the fronto-parietal and sensorimotor network gradually increase as individuals develop bias. To further distinguish the role of each individual brain region, we separated the frontal and parietal cortices into two individual seed regions and extract each of their activation as a regressor in two independent PPI analyses. We centered the seed regions of frontal and parietal cortices at the MNI coordinates in which the BOLD signals have been repetitively shown to correlate with the magnitudes of reward-induced bias in perceptual decisions (Fleming et al. 2010; Summerfield and Koechlin 2010). We find that these two seed regions yield distinct connectivity patterns. Using the left frontal cortex as a single seed region, its connectivity with vmPFC, ventral striatum, and visual cortex increased as bias grew (Figure 5; Table 3). On the other hand, the connectivity between left parietal and ACC increased as bias grew (Figure 5; Table 3).

## DISCUSSION

We demonstrated that bias in perceptual decision-making could be acquired through a reinforcement learning mechanism. In perceptual decisions, individuals constantly learn the values of the two potential perceptual choices (e.g. the two motion directions in our task) by associating the deviation between the expected and actual

outcome (RPE) with each reward context. When the context is clearly indicated, a perceptual choice reflects the sum of the present sensory information and the expected value difference between the two options that has been learned so far. At the neural level, the activation of ventral striatum, frontal, and parietal cortices positively correlate with the contextual RPE derived from modeling each individual's performance over the experiment. Furthermore, three distinct functional connectivity patterns echo the growth of such bias, suggesting that the value signals became increasingly integrated with neural systems involved in action, stimulus evaluation, and cognitive control. These results reveal the shared neural mechanism between perceptual and economic decisions, and highlight the involvement of multiple control networks during the development of bias.

There is increasing convergence across different neuroscience methods and across species regarding the involvement of reinforcement learning processes in perceptual decisions. Most of these findings focus on the improvement in individuals' ability to detect or discriminate sensory information, or perceptual learning (Kahnt, Grueschow, Speck, and Haynes 2011; Law and Gold 2008). Here, we extend this line of research by showing that bias in perceptual decision-making can be acquired through the same learning mechanism. Given the commonality between our findings and the findings in perceptual learning, one may suspect that the RPE signals that we report in ventral striatum, frontal, and parietal cortices may simply reflect perceptual learning (Kahnt et al. 2011) rather than the development of bias. However, we found that participants' ability to

discriminate the motion direction did not improve significantly over the experiment. Our experiment may be too short to elicit the effect of perceptual learning, which usually requires more than a thousand trials (Kahnt et al. 2011; Law and Gold 2008). This mismatched temporal dynamic between perceptual learning and the acquisition of bias further rules out this alternative interpretation.

Our findings also speak to the great interest in the literature regarding the integration of reward and sensory information in the brain during perceptual decisions. Using general linear model, it has been shown that reward information integrates into the intermediate level (frontal and parietal cortices) of perceptual decision hierarchy in the brain (Mulder et al. 2012; Summerfield and Koechlin 2010) that may further facilitate downstream visual and motor processes (Fleming et al. 2010; Liston and Stone 2008; Serences 2008). In this study, we further identified the neural locus for the integration of reward information into perceptual decision networks in the brain by tracking changes in functional connectivity patterns as bias developed. We found that functional connectivity between vmPFC-motor cortices, frontal-vmPFC-visual cortices, and parietal-ACC increases with the growth of perceptual bias. These distinct connectivity patterns reveal a more integrated view than previous findings and suggest that multiple mechanisms contribute to bias in perceptual decisions through integration of value processing with action, sensory, and control systems.

Since perceptual decisions are usually analyzed under the framework of drift-diffusion process, one obvious question is how our finding relates to this tradition. In fact, many theoretical approaches have been proposed to account for the function of basal ganglia and its contribution to the learning of perceptual bias in terms of drift-diffusion process (Bogacz and Larsan 2011; Rao 2010). However, it is very challenging to empirically test these theories at the whole brain level. The difficulty is that bias induced by reward usually is weak in human studies using reaction time tasks (Mulder et al. 2012) because of the speed-and-accuracy trade-off (Maddox and Bohil 1998; Simen et al. 2009). In order to investigate the acquisition process of perceptual bias, we applied decision deadline to boost the effect of reward-induced bias since this type of bias usually occurs in fast choices and suggests the influence on the starting point of the drift-diffusion process (Mulder et al. 2012; Summerfield and Koechlin 2010). Using this manipulation, we present the first empirical evidence showing the role of basal ganglia in the acquisition of perceptual bias and three functional connectivity patterns pertaining to expressing the acquired bias in the human brain. However, the drawback of this manipulation is that it limits the interpretation of our finding to the speeded decisions and prevents us from applying the drift-diffusion model to the present data. Future study should focus on identifying the neural mechanisms by which bias presents as individuals freely adjust reaction time in order to maximize reward.

In conclusion, the present study shows that perceptual bias can arise from the neural mechanisms that learn the association between contexts and choice outcomes. The learning signals (contextual RPE) are observed in ventral striatum, frontal, and parietal cortices. The acquired bias mirrors the learned value difference between each perceptual choice at the behavioral level and correlates with multiple connectivity patterns suggesting that the information about value contributes to perceptual bias through interactions with multiple systems. Our results demonstrate the pervasive effects of reinforcement learning mechanisms on the whole-brain connectivity by which the subjective expectation of reward colors the interpretation of objective evidence in decision-making process.

# Chapter 2: Training of loss aversion modulates neural sensitivity toward potential gains

## INTRODUCTION

When facing uncertain outcomes, people tend to be more sensitive toward the potential losses than the potential gains. This tendency can lead to poor investment decisions that derail investors from their financial plans. Theories in psychology and economics attribute loss aversion to different sensitivities when transforming the possible gains and losses into subjective value (Kahneman & Tversky, 1979; 1984; Tversky & Kahneman, 1992). This asymmetrical response to potential outcomes has found its neural basis in the brain regions processing value (Tom, Fox, Trepel, & Poldrack, 2007) or emotional (Canessa et al., 2013; De Martino, Camerer, & Adolphs, 2010; Sokol-Hessner, Camerer, & Phelps, 2013) information.

Although multiple neural systems entail the loss aversion effect by responding more strongly to potential losses than to potential gains, the regulation of loss aversion has been focused majorly on its emotional component. Strategies for regulating negative emotional responses, such as reappraisal or taking different perspectives, has been shown effective in lessening the aversive emotion triggered by the potential losses, which correlates with the reduction of loss aversion in financial decisions (Sokol-Hessner et al., 2009). However, in addition to the emotion system (amygdala) in the brain, the value system (ventral striatum and vmPFC) is another neural basis of behavioral loss aversion (Tom et al., 2007). Several mechanisms have been shown to effectively alter the

responses of the value system in order to make optimal choices. However, the possibilities of utilizing these potential mechanisms to tackle the value system with the aim of regulating loss aversion in financial choices remain unexplored. Here we provide the first evidence showing a novel mechanism through which loss aversion can be dynamically adjusted in order to fulfill the demands of current decision contexts.

Economic decisions can be altered through modulating the value system, specifically the ventral medial prefrontal cortex (vmPFC). One of the mechanisms that can modify the response of vmPFC in value-based decision-making is reinforcement learning. Through trial-and-error, the difference between the expected and the actual choice outcomes (reward prediction errors) triggers the activation in ventral striatum that in term adjusts the subjective value encoded in the vmPFC (Bartra, McGuire, & Kable, 2013; Garrison, Erdeniz, & Done, 2013; Glimcher, 2011; Lee, Seo, & Jung, 2012; Schultz, 1998). This subjective value then determines which actions are worth repeating in the future. Since loss aversion can be explained by the asymmetrical responses of the value system (ventral striatum and vmPFC) toward the potential gains and the potential losses (Tom et al., 2007), this reinforcement learning mechanism implies that loss aversion can be either enhanced or reduced by providing reward feedback to encourage either type of decisions about uncertain financial outcomes.

Alternatively, the value system can also be modulated by the cognitive control system in order to achieve the desirable goal by following certain rules. Overcoming tempting but suboptimal choices according to the current decision context requires the cognitive control system (frontal and parietal cortices) to suppress the unwanted drives (Bunge, 2004; Miller & Cohen, 2001). This interaction between the cognitive control and value systems in a task can be measured as functional connectivity and has been shown to correlate with whether people can successfully inhibit their temptation of eating preferable but unhealthy food in order to follow their diet plans (Hare, Camerer, & Rangel, 2009). This mechanism implies that offering a rule for evaluating uncertain outcomes in a decision context should elicit the interaction between the cognitive control and value systems. This in turn may alter the level of loss aversion according to the present decision context.

We showed that the level of loss aversion is modifiable, and that this is associated with changes in activation of the value system (vmPFC). During functional brain imaging scans, participants were asked to decide whether to accept or reject a stock offering equal chance to either win or lose a certain amount of money. These stocks were presented in two different decision contexts. One context provided reward feedback to encourage riskier decisions (i.e. loss neutral) whereas the other offered reward feedback that reinforced more conservative decisions (i.e. more loss averse). In order to maximize the reward, the participants must use feedback from trial to trial to learn the optimal rule

for deciding in either context. At the end of this training, they were probed to apply what they had learned when evaluating uncertain outcomes in each context without trial-wise feedback.

We found that the level of loss aversion can be enhanced or reduced according to the decision context by respectively decreasing or increasing the sensitivities toward the potential gains but not the potential losses. The activation in cognitive control system (frontal, parietal, and middle temporal cortices) mirrored this behavioral pattern during the training. Furthermore, over the course of learning, the functional connectivity between the cognitive control and value system gradually decreased when the participants evaluated the potential gains in the loss neutral context. After this training, while participants' choices were modulated by the decision contexts, the value system becomes more sensitive toward the potential gains in the loss neutral context than that in the conservative context. The findings suggest that the level of loss aversion can be flexibly adjusted via releasing value system's inhibited response toward the potential gains from the cognitive control system. These highlight the important role of cognitive control systems in decision-making under risk and provide insights about how cognitive control can refine investment choices in order to achieve the desired financial goal.

## MATERIALS AND METHODS

### Participants

Sixty human participants completed the behavioral paradigm in the MRI scanner (31 females, 29 males; age range: 18 – 30 with mean 22.9-year-old). The sample size was determined by a power analysis using the fMRIpower software package (http://fmripower.org/; Mumford & Nichols, 2008) based on pilot imaging data from 8 participants. Two of the participants were discarded from the brain imaging analyses; one due to missing the anatomical image, and the other due to excessive head movement (more than one-third of the volumes were considered "bad time points" according to the motion correction procedures detailed in the Preprocessing section). All participants were recruited through posted flyers and were prescreened. They were free of any self-reported neurological or psychiatric diseases, had normal or corrected-to-normal visual acuity and normal color vision, and right-handed. They gave written informed consent for participation. The Institutional Review Board of the University of Texas at Austin approved all experimental procedures.

### Stimuli and Apparatus

We utilized a mixed gambles task for measuring individuals' sensitivity to potential gains and losses (Tom et al., 2007). To fit into the cover story of this experiment, which will be described in details in the Procedures and task section, we

called these stimuli “stocks”, and they had equal chance to gain or loss a certain amount of money. In the baseline and the probe session, the potential gains ranged from $10 to $40 with the increment of $4 whereas the potential losses ranged from $5 to $20 with the increment of $2 (Figure 6, top- and bottom-right). In the training session, each stock was a random sample from six levels of gain-to-loss ratio (.52–.79, .88–1.07, 1.18 – 1.8, 2.1–2.41, 2.59–3.17, 3.21–6.33) with a potential gain selected uniformly between $9 and $62 and a potential loss selected uniformly between $3 and $52 (Figure 6, middle-right). The number of trials at each level of the gain-to-loss ratio was adjusted according to the decision criterion of each decision context (1.12 and 2.5, respectively), so that there was the same number of the acceptance and rejection responses in each context. This ensured that the observed behavior was not due to response bias (Green & Swets, 1966). The overall expected value of the two stimulus sets were both positive (2.84, 8.56) but significantly different ($t(238) = -6.44$, $p<.0001$).

The experiment was programmed and delivered in python version 2.6 on Mac OS 10.6.8. The randomization procedure was conducted using numpy version 1.6.1 (http://www.numpy.org/), and the stimulus presentation and the response recording were controlled by pygame version 1.9.1 (http://www.pygame.org/).

**Procedures and task**

The experiment was presented to participants as a stock market investment scenario. As illustrated in the left panel of Figure 6, three phases were designed to

measure participants' own loss aversion tendency (baseline), to shift their loss aversion tendency using contextual feedback (training), and to test how this training was consolidated in each context without immediate feedback (probe). Participants performed all the phases during the fMRI acquisition. They did not receive endowment before they came for the experiment. They were told that the outcome that they received in the first phase would be their initial reward points to start the rest of the phases. The decision contexts and response buttons were counter-balanced across the group of participants to remove any potential stimulus or response confounds. The inter-trial interval was sampled from a truncated exponential distribution (mean = 3 sec., range = 2.5 – 8 sec.) and the randomized orders of trial types and stimuli presentation in all the three phases were the top sixty design matrices from 10,000 simulation with the highest efficiency for contrasting the gain variable between the two decision contexts and for contrasting the loss variable between the two decision contexts. The procedures of each phase are detailed below.

*Phase 1: Baseline*. As shown in Figure 6 (top-left), on each trial, participants saw a set of stocks offering 50-50 chances to either win or lose a certain amount of money. They had to press one of the four buttons to indicate whether they strongly accept, weekly accept, weekly reject, or strongly reject each stock into their own investment portfolio in 2.5 s. The trial would be aborted if no response made during this time window. At the end of this phase, one of the trials was randomly selected and

revealed the outcome. If the participant previously accepted this selected stock, a Bernoulli trial with probability of .5 would determine the payoff according to the amount of potential gains and potential losses offering by this stock. If the participant rejected this selected stock, the payoff would be zero. The outcome of this selected stock then became the initial number of reward points for the participants to continue collecting reward points in the next phase.

*Phase 2: Training* This phase was designed to use feedback for encouraging participants to use two different risk preferences as they decided whether to accept or reject a stock for one of two clients specified by a pre-trial cue. Participants were asked to review stocks with one of the two clients and suggest that client whether to accept or reject the stock into the client's investment portfolio. Figure 6 (middle-left) illustrates the events timeline of a trial. In the beginning of every trial, the participants saw for whom they would be providing investment suggestion for 1 s. Then, they saw a stock that has equal chance to win or lose certain amount of money. They had up to 2.5 s to decide whether they would suggest the client to accept or reject that stock by pressing buttons. Immediately after the participants responded, they saw the number of reward points that they earned or deducted from the client for 1.5 s. As shown in Figure 6, one of the clients was loss-neutral and would accept the stock as long as the potential gain is 1.12 times greater than the potential losses (riskier context). The other was loss averse and would accept the stock at least the potential gains were 2.5 times greater than the potential losses

(conservative context). If the participants' suggestion matched each client's investment preference, they would earn 10 reward points; otherwise, they would lose 10 reward points. Participants were unaware of the clients' preferences before they began the experiment. They simply were told to make suggestion that they thought would maximize their total reward points over the course of this phase.

*Phase 3: Probe*. This phase was designed to evaluate how the previous experiences in the two decision contexts would influence the participants' decision-making under risk without the guidance of immediate feedback. The events timeline were the same as in the previous phase, except for that no feedback was presented on every trial and that the stimuli in each decision context were the same set of stocks used in the first phase (Figure 6, bottom-left). The participants were requested to maximize their reward points that were accumulated in the background by applying what they had learned about each client's preference when making a suggestion. The decision criterion and payment structure in each context was exactly the same as in the previous training. At the end of this phase, the total reward points that have been collected by each participant over the three phases were convert into bonus and added into the compensation (500 reward points = 1 USD).

**Behavioral analysis**

*Baseline phase*. We followed the same analysis procedures as in the (Tom et al., 2007) to estimate each individuals' loss aversion tendency. We applied logistic

regression to estimate three parameters in the model from each individual's choice data: choice biases ($\beta_0$), sensitivity toward the potential gains ($\beta_{gains}$), and sensitivity toward the potential losses ($\beta_{losses}$). The loss aversion tendency for each individual was indicated as the ratio between the sensitivities toward the potential losses and the potential gains ($\lambda = -\beta_{losses}/\beta_{gains}$). We added the negative sign in this indicator because the absolute value of the potential losses was used in the model.

*Training phase.* The analyses were done at the two levels. At the first level, we applied logistic regression model to each of the five scanning runs to estimate choice biases, sensitivity toward the potential gains and losses for each participant. The loss aversion indicator ($\lambda$) was also derived for each individual from run to run in each condition.

At the second level, all these parameters were used as the dependent variables for repeated measure ANOVA test. We focused on testing whether there was any significant interaction between the two decision contexts and the five scanning runs. This would show how the sensitivities toward the gains and losses changed when people learned to become riskier or more conservative in evaluating the uncertain outcomes.

*Probe phase.* The above-mentioned analysis procedures were used in this phase as well. At the second level analysis, we used paired t-test to examine whether the

participants' loss aversion tendencies were significantly different across the two conditions. The same test was applied to identify the differences in either the sensitivities toward the potential gains or the potential losses across the two decision contexts.

All the first level analysis was done using the logistic regression function in the scikit-learn package version 0.1 (http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html) that empirically provided more robust parameter estimation against the problems of complete separation in binary data. The second level analysis was done using R with the nlme package (http://stat.ethz.ch/R-manual/R-devel/library/nlme/html/lme.html). We were aware that the two-level analysis could be done using the hierarchical logistic regression in the same package. However, the parameter estimation of this model did not converge, leading us to adopt the two-level analysis as stated.

**FMRI acquisition**

Imaging data were collected using a Siemens Skyra 3T MR scanner. Functional data were collected using a T2*-weighted multi-band echo-planar imaging sequence (TR=1.16 ms, TE=30 ms, flip angle=63 degrees, voxel size=2.4 X 2.4 X 2 mm, distance factor=20%, 96x96 matrix, FOV=230 mm, MB factor=4, 64 slics). To reduce dropout in orbito-frontal cortex, the slices were tilted at a 30° back from the anterior-commissure-posterior-commissure line and higher-order shimming was applied. T1-weighted

anatomical images was collected using an MP-RAGE sequence with 8° flip angle (256 sagittal slices, voxel size=.8 X .8 X .8 mm, TR=2.4 s, TE=1.94 ms, TI=1000 ms, PAT=2, FOV=205 mm).

**Image preprocessing and registration**

FMRI data preprocessing was carried out using FSL Version 5.0.1 (FMRIB's Software Library: www.fmrib.ox.ac.uk/). All image time series were aligned using the MCFLIRT tool, and the resulting motion parameters were used to compute frame-wise displacement (FD) and temporal derivative of the root mean square variance over voxels (DVARS) to identify bad time points (FD>.5; DVARS>.5) (Power et al., 2012). The skull was removed from the functional images with the brain extraction tool (BET) and from the structural images using FreeSurfer (https://surfer.nmr.mgh.harvard.edu/). Spatial smoothing was applied using a Gaussian kernel of FWHM 5 mm. The grand-mean intensity was normalized over the entire 4D dataset by a single multiplicative factor, and a high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=50.0 s). This same high-pass filter was applied to the design matrix for analyzing the fMRI time-series. All functional images were registered to the high resolution structural image using Boundary-Based Registration (BBR) then the high resolution structural image to the MNI-152 2 mm template using the FNIRT nonlinear registration (12 DOF; Wrap resolution: 10mm) tool of FSL.

**FMRI analysis**

We used a multi-stage general linear model (GLM) approach to analyze the brain imaging data with FSL FEAT (FMRI Expert Analysis Tool) Version 6.00. The first-level model was estimated separately for each run and each participant. All five runs in the training and two runs in the probe were separately combined within participant using a fixed-effects model. When testing the learning effect, a linear contrast was added into the second level design matrix. At the group level, the FLAME 1 mixed-effects model of FSL was applied to all participants (Worsley, 2001). All the statistical maps were corrected by cluster-based random field theory using clusters determined by $Z > 2.3$ and a family-wise error corrected cluster significance threshold of $P=0.05$ (Worsley, 2001). The statistical maps for all analyses were projected onto the group-averaged brain from this study for visualization.

**General Linear Model (GLM)**

The first level of GLM contained parametric modulated regressors to identify the brain activation correlating with the potential gains and the potential losses as well as nuisance regressors to control for potential confound. A confound file included all the motion correction parameters (estimated translation and rotation and their first derivatives, FD, and DVARS) together with single-time-point regressors for each time point that exceeded the FD/DVARS thresholds (which effectively performs "scrubbing" of those time points)(Power et al., 2012). All the regressors except the motion-

correction regressors in the first-level model were convolved with a double-gamma hemodynamic response function. Their temporal derivatives were also included in the model to accommodate for potential slice timing differences. All parametric modulated regressors were mean centered before entering the model. The regressors used in the design matrix of each phase were described in details below.

*Baseline phase*. There were five regressors in the design matrix. Two parametric modulated regressors encoded the amount of potential gains and the amount of potential losses. The duration of these regressors started from trial onset and lasted for the mean response time of each participant in this phase. Three nuisance regressors were used to model each trial, response time, and missing trials. The trial and missing trial regressors were boxcar regressors beginning at the trial onset and lasting the duration of the mean response time of each participant in this phase as well. Response time was modeled as an additional parametric regressor starting from the trial onset with 1 sec. duration and mean-centered response time as its amplitude. This procedure was used to model the effect of response time in the rest of the phases as well.

*Training and probe phase*. Nine regressors were included in the design matrix of the training phase. Four parametric modulated regressors were used to model the potential gains and losses in each of the two conditions. These regressors started from the trial onset with the duration of mean response time plus two seconds to accommodate the

presentation of the context cues and feedbacks. Another parametric regressor indicated the amplitude of feedback starting from the feedback onset with the duration of 1 sec. Four additional regressors were included: two for the onset of trials from each context, one for the missing trial, and the other one for the response time. These regressors were generated with the same procedures as described in the previous section, except for that the duration here was two seconds plus the mean of each participant's response time in each run. In the probe phase, the regressors were the same as in the training phase, except for that there was no regressor indicating the feedback since there was no feedback in this phase. The duration of the regressors was the participant's mean response time in each run plus 1 second to account for the presentation of the cue. Four nuisance regressors were included: two for the onset of trials from each context, one for the missing trial, and the other one for the response time. The regressor for the response time was parametrically modulated with 1 second duration and the mean-centered response time as its amplitude.

**Psychophysiological Interaction (PPI)**

We used PPI analysis to examine the changes in functional connectivity between the vmPFC and the rest of brain as the participants become more conservative or risky in evaluating the uncertain outcomes. The vmPFC region was selected from the peak response to the potential gains in the probe phase (10-mm sphere around the vmPFC; MNI coordinates [x, y, z] = [-12, 54, -4]). The BOLD activation of the seed region was extracted from each participant's individual brain in each run. For each individual and

each run, the neural signal of the seed region was estimated by deconvolving the BOLD signals using the deconvolution algorithm of SPM (Gitelman, Penny, Ashburner, & Friston, 2003). The interaction between the seed region and the regressor parametrically modulated by the potential gains or the potential losses was generated in the neural domain and then reconvolved with hemodynamic function. The first-level design matrix of the PPI analysis was the above-mentioned GLM design matrix with five additional regressors: the raw time course extracted from the seed and four PPI regressors. The PPI regressors indicate the interaction between the mean BOLD response in the seed region and the amount of gain or losses in either of the two context, respectively.

## RESULTS

### Behavioral results

The participants followed the instructions to accumulate the reward points over the three phases of the experiment. On average, they earned -8.5 points in the first phase (standard deviation = 5.76) as their initial points, 1490 points over the five training runs (standard deviation = 182.12), and 972 points in the probe (standard deviation = 156.54). There was no significant effect of the initial points on the performance in the subsequent phases. The correlation between the number of initial points and the points that collected in the training (Pearson's correlation = -0.09, $p = .48 > \alpha = .05$), the probe (Pearson's correlation = -0.001, $p = .94 > \alpha = .05$), or the total points gained over these two subsequent phases (Pearson's correlation = -0.06, $p = .65 > \alpha = .05$).

Using our paradigm, we replicated previous findings that people are loss averse when facing uncertain outcomes. In the first phase of experiment, the participants were asked to decide whether to accept or reject a stock offering an equal chance to win or lose a certain amount of money, which later became their initial reward points to start the rest of the sessions. Using hierarchical logistic regression with individuals taken as random effects in the model, we found that the participants did not show any significant preference for either choice ($z = 1.608$, $p = 0.108 > \alpha = .05$). Instead, their choices were significantly dominated by the potential gains (averaged $\beta_{gain} = .27$, $z = 27.05$, $p<.0001$) and the potential losses shown on each stock ($\beta_{loss} = -.40$, $z = -23.53$, $p<.0001$). The negative value here is because the absolute value of the potential losses was used in the model. Each participant's level of loss aversion also was characterized by the loss aversion indicator ($\lambda$): the ratio between the absolute parameter estimates for the potential losses (Figure 7B, top panel) versus potential gains (Figure 7B, top panel) based on the logistic regression that was individually applied to everyone's choice data (Tom et at., 2007). The group average of this indicator ($\lambda$) is 1.4 (standard deviation= .54; median = 1.29, range = .63 – 3.58). This value reflected the indecision points shown in Figure 8A. When the ratio between the potential gains and the potential losses fell around 1.4, participants were equally likely to accept or reject those stocks (green grids in the left panel of Figure 8A) and spent longer time to decide (brown grids in the right panel of Figure 8A). The value of the loss aversion indicator was lower than the previous study ($\lambda$=1.93 [Tom et al., 2007]). A potential reason why participants in the present study

showed weaker loss aversion could be that they did not receive an endowment beforehand as in that study.

During training, we found that loss aversion can be either enhanced or reduced through feedback. As shown in Figure 6 (middle left panel), the same task was presented in two decision contexts with feedback to either enhance (conservative) or reduce (risky) loss aversion. In the risky context, the feedback encouraged the participants to accept the stocks as long as the potential gains were 1.12 times greater than the potential losses (loss neutral). In the conservative context, the feedback reinforced them to accept those as long as the potential gains were at least 2.5 times greater than the potential losses (more loss averse). Figure 7A shows that this procedure gradually changed the participants' levels of loss aversion. In the first run, the indecision points of accepting a stock in both decision contexts were around the third level of the gain-to-loss ratio (gray dashed lines in Figure 7A). The mean gain-to-loss ratio of this bin was 1.49 that was within one standard deviation of the original loss aversion indicator at the group level (mean $\lambda = 1.4$, standard deviation = .54). However, as the experiment proceeded, the indecision points shifted leftward when the participants were encouraged to use a risker criterion (Figure 7A; light blue - blue). On the contrary, this point shifted rightward when a more conservative criterion was reinforced (Figure 7A; light brown - brown).

We also used logistic regression model to identify how the participants' loss aversion changed in each decision context over the course of training. The ratio between the regression coefficients for the loss and gain variables (loss aversion indicator) estimated in each decision context was then plotted against the five training runs to illustrate these changes (Figure 7B, bottom panel). We found a significant interaction effect between the decision context and the training runs on the loss aversion indicator (repeated measured ANOVA, context x run interaction: $F(4, 531) = 9.6570$, $p < .0001$), suggesting that the participants gradually learn to become risky or conservative according to the decision context.

We further tested whether the changes in the sensitivity toward the potential gains or the potential losses contributed to the different loss aversion in the two decision contexts over the training. As shown in the top panel of Figure 7B, we found that the participants became more sensitive toward the potential gains when they were encouraged to become risker compared to when they were encouraged to become more conservative (repeated-measure ANOVA, context x run interaction: $F(4, 531) = 13.6393$; $p < .0001$). They also become more sensitive toward the potential losses in both conditions over this course (repeated-measure ANOVA, context x run interaction: $F(4, 531) = 4.1548$; $p = .0025 < \alpha = .05$). However, this incremental trend was very similar in both contexts (Figure 7B). This indicates that the sensitivity toward the potential gains largely mediated the training of loss aversion in the two decision contexts.

Finally, we found that two decision contexts continued influencing the participants' decisions about the uncertain outcomes even when no immediate feedback was presented. As shown in Figure 8B and 8C, the indecision points (green grids in the left panels) were close to the criteria that the participants were encouraged to used in the previous session (gain-to-loss ratio as 1.12 or 2.5). It also took longer for the participants to decide when the potential gains and potential losses pair fell near these ratios (brown grids in the right panels). Their loss aversion tendencies were significantly different between the two contexts (paired t-test: $t(59) = -12.3943$, $p < .0001$; Figure 7B bottom panel). This distinct pattern of loss aversion can be attributed to the different sensitivities toward the potential gains (paired t-test: $t(59) = 8.3373$, $p < .0001$; Figure 7B top panel) but not that toward the potential losses (paired t-test: $t(59) = -0.3781$, $p = 0.7067 > \alpha = .05$; Figure 7B top panel) in the two contexts.

**Brain imaging results**

In the baseline phase, we replicated previous findings that loss aversion correlated with the responses of the value system toward the potential gains and the potential losses. Behaviorally, we found that the participants would accept the stock at least the potential gains are 1.4 times greater than the potential losses, which decisions also took them longer to make (Figure 8A). We also found that the activation in the value system correlated with the participants' choices. As shown in Figure 9, the vmPFC (MNI coordinates [x, y, z] = [4, 38, -14]; cluster size = 603 voxels; $Z_{max} = 4.26$, $P = .0007$)

positively correlated with the potential gains after whole-brain corrected for multiple comparisons. It also de-activated as the potential losses increased, as shown previously (Tom et al., 2007); however, this effect was weaker in our sample (Figure 9, uncorrected), which may because our participants on average were less loss averse as previous study.

***Changing the sensitivity toward the potential gains relies on cognitive control***

We first used the general linear model to identify the brain areas whose activity was correlated with changes in loss aversion over the course of learning. The potential gains and losses entered the design matrix as parametric modulated regressors that were separated into two sets for modeling the effect of the two decision contexts. We searched the whole brain for areas showing different neural sensitivities toward the potential gains or the potential losses between the two decision contexts. We found that the value and cognitive control systems responded more strongly toward the potential gains in the risky than the conservative conditions over the training (Figure 10A, Table 4). To be qualified as the learning mechanisms, this difference should also increase from run to run, indicating the linear trend of learning as shown in the behavioral results (top panel in Figure 10B). We found that the activation of three brain regions in the cognitive control system matched this selection criterion when they responded to the potential gains (Figure 10B, red-yellow, Table 4 and Figure 10C; red shaded plots). These brain regions were right middle frontal gyrus (MNI coordinates [x, y, z] = [30, 36, 52]; cluster size = 653 voxels; $Z_{max} = 4.33$, P = .0002), right angular gyrus of the parietal cortex (MNI

coordinates [x, y, z] = [46, -54, 32]; cluster size = 664 voxels; $Z_{max}$ = 4.58, P = .0002), and right middle temporal gyrus (MNI coordinates [x, y, z] = [60, -8, -22]; cluster size = 373 voxels; $Z_{max}$ = 4.29, P = .0157). No activation in other regions showed this linear trend that survived the whole-brain correction for multiple comparisons.

We also searched for brain regions that gradually showed differentiable activation toward the potential losses between the two contexts. We found that the activation in the cognitive control and value systems (Figure 10A; Table 4) also responded more strongly toward the potential losses during training. However, only the activation in the right frontal pole (MNI coordinates [x, y, z] = [30, 60, 12]; cluster size = 417 voxels; $Z_{max}$ = 4.2, P = .0065) showed a linear trend such that it became increasingly stronger in the riskier context compared to the conservative context over training (Figure 10B, blue-light blue, Table 4, and Figure 10C; blue shaded plots).

***Cognitive control system releases supression from the value system during learning***

We further tested the hypothesis that the interaction between the value and cognitive control systems was necessary to execute preferable choices according to the decision context (Hare et al., 2009). We centered the seed region at the vmPFC (MNI coordinates [x, y, z] = [-12, 54, -4]) with 10mm radius covering the surrounding voxels. If using the decision contexts to adjust loss aversion requires the cognitive control system to modulate the value system, we should observe that the functional connectivity between the two is altered in the two contexts during the training. Furthermore, this change should

be more pronounced in response to the potential gains than to the potential losses based on our behavioral results. As shown in Figure 11A and Table 5, the interaction between vmPFC and the left middle frontal cortex (MNI coordinates [x, y, z] = [-35, 14, 48]; cluster size = 508 voxels; $Z_{max} = 4.15$, P = .0014) and the left supramarginal gyrus of the parietal cortex (MNI coordinates [x, y, z] = [-47, -46, 43]; cluster size = 401 voxels; $Z_{max} = 3.38$, P = .0076) became differentiable across the two contexts when responding to the potential gains over the training. As shown in the plots of Figure 11B, the functional connectivity between the vmPFC and frontal cortex gradually increased when the participants became less sensitive toward the potential gains in the context where conservative choices were preferred. However, this connectivity gradually decreased when the participants learned to be more sensitive toward the potential gains in the context where riskier choices were reinforced. On the other hand, the functional connectivity between vmPFC and the parietal cortex was overall greater in the conservative context than that in the risky context (Figure 11B). When responding to the potential losses, we found that the functional connectivity between the vmPFC and the superior parietal cortex (MNI coordinates [x, y, z] = [11, -52, 66]; cluster size = 1706 voxels; $Z_{max} = 4.2$, P < .0001), angular gyrus (MNI coordinates [x, y, z] = [54, -54, 31]; cluster size = 331 voxels; $Z_{max} = 3.53$, P = .016), and midbrain area (MNI coordinates [x, y, z] = [-2, -21, -36]; cluster size = 408 voxels; $Z_{max} = 4.35$, P = .0041) were greater in the conservative than that in the riskier context during the training (Figure 11A, blue-light blue; Table 5).

The functional connectivity between the vmPFC and the rest of the brain only occurred in the training. During the probe, no brain areas showed significant differences in functional connectivity between the vmPFC across the two decision contexts in responding to either the potential gains or losses. This is also the case after we restricted the search within the regions showing different connectivity patterns in the two contexts during the training.

***Value and cognitive control systems modulate loss aversion after training***

The above-mentioned finings together with the suggested function between the cognitive control and value system (Hare et al., 2009) implies that reducing loss aversion by being more sensitive toward the potential gains relies on the interaction of cognitive control systems with the value system. If this is the case, after this training, we should expect that the value system remained more sensitive toward the potential gains when probed in the context reinforcing riskier choices. Our data supported this hypothesis. We found that both the value (vmPFC, MNI coordinates [x, y, z] = [-12, 54, -4]; cluster size = 404 voxels; $Z_{max}$ = 3.81, P = .0154) and the cognitive control systems (middle frontal gyrus: MNI coordinates [x, y, z] = [28, 36, 40]; cluster size = 464 voxels; $Z_{max}$ = 3.83, P = .0065; angular gyrus: MNI coordinates [x, y, z] = [54, -48, 18]; cluster size = 337 voxels; $Z_{max}$ = 4.07, P = .0424) become more sensitive toward the potential gains in the context where the participants were trained to be riskier compared to the more conservative context (Figure 12, red-yellow; Table 4). Behaviorally, the

participants responded to the potential losses with equal sensitivities in the two conditions (Figure 7B, top panel); however, we observed that the parietal (supramarginal gyrus: MNI coordinates [x, y, z] = [56, -42, 28]; cluster size = 933 voxels; $Z_{max} = 3.98$, P < .0001) and supplementary motor area (MNI coordinates [x, y, z] = [4, -2, 60]; cluster size = 365 voxels; $Z_{max} = 3.65$, P = .0357) were more sensitive toward the potential losses in the riskier context than that in the more conservative context (Figure 12, blue – light blue; Table 4).

## DISCUSSION

We have illustrated that a cognitive control process can account for the adjustment of behavioral and neural loss aversion according to the environmental demands. We found that people become more sensitive toward the potential gains when learning to be loss neutral. This change correlates with the linearly increasing activation of the cognitive control system (frontal and parietal cortices) during the training. Meanwhile, the functional connectivity between the vmPFC and the frontal lobe decreased when the participants learned to become more sensitive toward potential gains in the risky context. After this training, the two decision contexts continued modulating the participants' loss aversion tendencies: they were more sensitive toward the potential gains in the loss neutral context while their sensitivities toward the potential losses were the same in both contexts. Underlying this behavioral change, we found that both the cognitive and the value systems activated more strongly toward the potential gains in the

loss neutral context than that in its conservative counterpart. These suggest that gradually reducing the influence of the cognitive control system is necessary for the value system to become more sensitive toward the potential gains, which then neutralizes the tendency toward loss aversion in decision-making under risk.

Some theories propose that regulating negative emotion could modulate loss aversion; however, our results suggest that this idea is relatively speculative. We did not observe any activation in brain areas involved in emotion process (amygdala) correlated with loss aversion tendencies in any phase of our experiment. Fundamentally, whether negative emotion can explain the overly sensitive toward the potential losses, which leads to behavioral loss aversion, is still debatable since the results from neural imaging studies are inconsistent (Canessa et al., 2013; De Martino et al., 2010; Sokol-Hessner et al., 2013; Tom et al., 2007). Specifically, the activation in amygdala, the brain region correlating with negative emotion process and used as an indicator of the emotional component of loss aversion, is not consistently observed in the literature. Even though this discrepancy may be due to the difference in experimental designs and regulating loss aversion did involve regulating negative emotion (Sokol-Hessner et al., 2009), the evidence supporting the role of cognitive control system in this regulation is indirect. In the recent finding, the cognitive control system (dlPFC) showed greater baseline responses during the decision period in the "emotion regulation" condition. However, this activity neither particularly correlates with the process of the potential losses nor links to

the activation in amygdala via functional connectivity (Sokol-Hessner et al., 2013). It is hence doubtful that loss aversion can be reduced by emotion regulation.

We present a novel mechanism that a direct relationship between the cognitive control and the value systems can regulate the loss aversion through increasing the sensitivity toward the potential gains. This finding extends the functional connectivity between cognitive control and value system in value-based decision-making. The cognitive control system (dlPFC) has been shown sending signals to value system (vmPFC) via a two-node functional connection to suppress the value of the unwanted feature, relative to the goal, of each decision target (Hare et al., 2009). Here, instead of this two-node connectivity between the cognitive control and the value systems, we report a direct functional connection and bidirectional effect on this suppression mechanism in modulating loss aversion. When more conservative choices were reinforced, the dlPFC gradually showed stronger functional connectivity with vmPFC in response to the potential gains. On the contrary, there was weaker functional connectivity between the two in response to the potential gains when loss neutral was promoted in a context. This suggests that relaxing the value system's response toward the potential gains from the cognitive control system is necessary for learning to take riskier choices. This account has been verified with our finding that the vmPFC became more sensitive toward the potential gains in the loss neutral context after the training. Our findings together suggest that future research on loss aversion should take a "loss neutral"

perspective. In addition to studying why people are more sensitive toward the potential losses, it also will be fruitful to examine why people are less sensitive toward the potential gains. Our finding implies that it may because the value system is normally inhibited by the cognitive control system.

Furthermore, our finding implies that cognitive control system may play extra role in decision-making under risk according to the current decision context. We found that during the probe, the activation in the supplementary motor and the parietal cortices was stronger toward the potential losses in the risky context than in the conservative context. Recent findings showed that the activation pattern in these brain regions could be used to predict whether a risky or a safe choice would be made subsequently after the risk level in a decision context increased (Helfinstein et al., 2014). Taken together, these imply that becoming more sensitive toward the potential losses in a risky decision context might facilitate the cognitive control system to execute a safe choice subsequently. Unfortunately, we could not examine this hypothesis further since our task did not require the participants to make sequential decisions about risk taking. We leave the relationship between processing potential losses and determining the time when a safe choice should be executed for future studies.

One may suspect that our findings may not actually speak to the core of loss aversion since in our task, the participants decided "for others" rather than "for

themselves". This seems to be a valid point at the first glance since it has been shown that people make different choices when the perspective shifts (Sokol-Hessner et al., 2009). If this account applies, we should observe that the participants' sensitivities toward the potential gains and the potential losses in the first run (when they just changed the "perspective") should be different from the baseline session when they decided for themselves. However, we found that the loss aversion was almost the same as in the baseline in either context of the first run (Figure 7B). This suggests that our participants reacted similarly as deciding for themselves even though the perspective changed. Our findings simply demonstrated the mechanism underlying daily situations regarding how decision-making under risk could be modulated by decision contexts. The question regarding how this experience influenced individuals' own loss aversion is beyond the scope of our study. This interesting question is awaited for future exploration.

In conclusion, we have shown that loss aversion can be neutralized through increasing the sensitivity toward the potential gains in response to the environmental demands. This can be achieved by releasing the value system from being suppressed by the cognitive control system when evaluating the potential gains during the training. As a result, when probed, the value system activates more strongly toward the potential gains in the loss neutral than in the conservative context. These point out the important but often overlooked interaction between cognitive control and evaluation in decision-

making under risk and offer insights onto how investors may refine their decision-making process to avoid being derailed from their financial plans owning to loss aversion.

# Chapter 3: Modeling the decision-making process of financial choices

## INTRODUCTION

Financial decisions can be made in very short period of time. For instance, a trader has to decide whether to buy or sell a share in response to a sudden change in the financial markets. Some of these fast decision making processes may not be captured by the BOLD signals measured with functional MRI. In order to better understand the decision making process, we used mathematical models to decompose the choice and response time data into important psychological factors underlying the decision-making processes in this chapter.

The decision-making process underlying binary choices can be modeled as a drift-diffusion process (Ratcliff, 1976). This model assumed that during the deliberation, a decision maker constantly accumulates noisy evidence that supports one of two possible choices. A decision is made once the accumulated evidence reaches a predetermined level, or decision bound. Three parameters control this evidence accumulation process. The drift rate determines the speed of evidence accumulation. The starting point indicates the decision maker's initial preference for either choice. The drift variance reflects the noise level of this accumulation process.

This model has been successfully applied to account for behavioral and neural activities data in different domains of decision-making. Especially, when applied to

value-based choices, it can capture the evaluation of costs and benefits (Basten, Biele, & Heekeren, 2010) and the attention weights that decision makers put on different attributes of the target as they browse at each alternative (Krajbich & Rangel, 2011; Krajbich, Armel, & Rangel, 2010; Krajbich, Lu, Camerer, & Rangel, 2012). Here, we applied the drift-diffusion model to identify which of the psychological factors in decision-making under risk can be shaped by experiences from trial to trial. To achieve this goal, we adapted a novel method for estimating parameters in drift-diffusion models in contrast to the traditional simulation-based fitting routine. First, the usage of analytical solutions for modeling the joint distribution of choice and reaction time were proposed and validated in this Chapter. This method then was applied to the empirical data in order to identify the psychological factors that account for the changes in decision-making under risk. Finally, we discuss the implementation of loss aversion in the drift-diffusion model framework and the potential extension of this modeling approach with learning algorithms that potentially could uncover the feedback-driven learning processes from trial to trial.

## METHODS

We applied the drift-diffusion modeling approach to the behavioral data obtained from the previous chapter. The details of this dataset, such as sample size, data collection, and the experimental procedures, were described in the Methods section in the

previous chapter. Here, we focused on our implementation of the drift-diffusion model to this dataset.

**Model assumptions and parameterization**

We first assumed that the evidence driving the choices in our task was the gains-to-losses ratio in our experiment. For each run and each context, we binned all the stimuli into 6 different levels according to the ratio between the gains and losses magnitudes (the range of the gain-to-loss ratio in each bin: .52–.79, .88–1.07, 1.18 – 1.8, 2.1–2.41, 2.59–3.17, 3.21–6.33). Each level hence provided different strengths of evidence to support either rejecting or accepting the potential outcomes. For each run, we used distinct drift rate ($\mu_{ij}$) to model the different speed of evidence accumulation driven by the gain-to-loss ratio at each level (i) in each decision context (j). In addition, we also tested the other assumption that the drift rate could be the subjective value that was the weighted sum of the potential gains ($x_{gains}$) and the potential losses ($-x_{loss}$) on each trial according to the Prospect Theory (Kahneman & Tversky, 1979; 1984; Tversky & Kahneman, 1992):

$$\mu_j = {x_{gain}}^{\alpha} - \lambda * (-x_{loss})^{\beta} \quad ,$$

where the $\lambda$ indicated the amount of loss aversion.

We further assumed that in each decision context within each run, this evidence accumulation process was perturbed by Gaussian random noise of constant magnitude ($\sigma_j$). Moreover, this process also was susceptible to initial preferences for either option in

each decision context ($x_{0j}$) that was independent of the evidence. The required amount of evidence, or decision bounds, were fixed as 1 or -1 respectively for the choice of acceptance or rejection. The non-decision time ($\tau_j$) associated with perceptual and motor processes also was included in the model for each decision context within each run.

**Joint distribution of choice and response time**

Given the Gaussian random noise assumption, the duration for accumulating evidence toward one of the decision bound, or reaction time (x), followed inverse Gaussian distribution:

$$f(x|a,\lambda) = \left[\frac{\lambda}{2\pi x^3}\right]^{1/2} * \exp\frac{-\lambda(x-a)^2}{2a^2x} \quad (5)$$

At the i[th] gain-to-loss level, the drift rate ($\mu_i$) and the distance between the starting point and one of the decision bounds ($|(\pm 1 - x_0)|$) determined the mean of the reaction time distribution:

$$a = \frac{abs(\pm 1 - x_0)}{\mu_i}. \quad (6)$$

The square of this distance together with the variance of the drift determined the spread of the reaction time distribution:

$$\lambda = \left(\frac{abs(\pm 1 - x_0)}{\sigma}\right)^2. \quad (7)$$

With the same assumption, the probability of reaching the lower bound was (Ratcliff, 1976):

$$\text{P(choice = reject)} = \left[e^{-(4*\mu_i/\sigma)} - e^{-(4*|-1-x_0|/\sigma)}\right] / \left[e^{-(4*\mu_i/\sigma)} - 1\right]. \quad (8)$$

We also assumed that the choice and the reaction time were independent. The product of the two distributions therefore became the joint distribution of choice and reaction time. Given the set of parameter $(x_0, \mu_i, \sigma, \tau)$, this joint distribution was:

$$p(choice, x \mid x_0, \mu_i, \sigma, \tau) = p(choice|x_0, \mu_i, \sigma, \tau) * f(x|x_0, \mu_i, \sigma, \tau). \quad (9)$$

**Parameter estimation**

Given a set of parameters ($\Theta=\{x_0, \mu_i, \sigma, \tau\}$), the likelihood function for getting the observed choice and response time over n trials (t: trial number) was:

$$L(\Theta|choice_t, x_t) = \prod_{t=1}^{n} p(choice_t, x_t|\Theta) \quad (10)$$

The maximum likelihood method was used for parameter estimation. We used the fmin function in the scipy optimize package (http://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.fmin.html#scipy.optimize.fmin) to search for the parameters that minimized the negative sum of the logarithm likelihood function for the obtained data.

**Model validation**

This analytical solution to choice and reaction time distribution was validated in two folds. First, the joint distribution from analytical solution was plotted against that derived from simulating 5000 Weiner (drift-diffusion) processes. Second, we used convergence test to evaluate how well our maximum likelihood fitting routine would recover the true parameters as the sample size increased. For conducting this test, we independently simulated 100, 500, 1000, 2500, and 5000 data points from the Weiner

processes. The parameters used for the simulation (true parameters) were entered as initial guess in the gradient descent parameter estimation algorithm in the scipy optimize package (fmin). This simulation-and-fit process was repeated 100 times for each sample size and four different sets of true paramters. The mean square errors between the true and the best-fit parameters were computed and then averaged over the 100 iterations as the indicator of convergence for each parameter set.

## RESULTS

### Analytical joint distribution can be used to fit drift-diffusion model

The visual comparison between the analytical and the empirical joint distributions are shown in the left panel of Figure 13. This comparison has been done using four different sets of parameters (Figure 13A – D, left panel). The empirical joint distribution (black and red dots) was derived from 5,000 data points simulated with each of the four sets of parameters. Each of the analytical joint distributions (black and red lines) was a result of entering the parameters into Equation 9. Generally speaking, the empirical and analytical joint distributions aligned well regardless of the different sets of parameters. However, they started to misalign when the sign of the starting point $(x_0)$ and the drift rate $(\mu_i)$ were different (Figure 13D, left panel). Specifically, the analytical distribution did not perfectly match the shape of the joint distribution for the processes ending at the lower bound. This may reflect the fact that the inverse Gaussian distribution is limited to the case when the drift rate and decision bound are both positive. When we

applied the inverse Gaussian distribution to the negative drift rate and decision bound, we forced them to be positive by taking absolute value. This procedure inevitably decoupled the correlation between processes ending in either bound when there was bias in the starting point.

Although this caveat existed, the maximum likelihood procedure combined with the analytical joint distributions still could recover the true parameters as the sample size increased. For each set of parameters, the convergence tests were conducted by comparing the best-fit parameters estimated from five different sample sizes (n = 100, 500, 1000, 2500, and 5000) using the mean square errors. As shown in the right panel of Figure 13, as the sample size increased, the errors between the estimated and the true parameters approached zero, indicating the convergence.

**Gain-to-loss ratio is the evidence driving a choice**

This drift diffusion model was applied to analyze the data in two different ways to test the how the potential gains and losses were transformed into the evidence during the decision-making process. The first model assumed that the ratio between the potential gains and losses might be the evidence determining the choice. The magnitudes of this evidence then were enhanced or reduced according to whether a risky or conservative choice was reinforced in a certain context. The second model assumed that the objective magnitudes of the potential gains and losses first transformed into subjective value (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), depending on each

individual's sensitivity toward gains and losses, which became the evidence driving a choice. These two assumptions were tested using model comparison. As shown in Table 5, we found that the model using the six levels of gain-to-loss ratio as the evidence for making a choice fit the obtained data better than the one calculating the subjective value. Overall, this best-fit model could account for 62% of the variance in the choice and 67% of the variance in response time over the experiment. This amount of variance in choices and response time that could be accounted for by this model was shown in each context and training runs in Figure 14.

**The effect of experience on the decision-making process**

The parameter estimation from the best-fit model is plotted according to the two different contexts over the five training runs to illustrate the effect of experiences (Figure 15). The drift rates showed the major differences between the two decision contexts over the five training runs. This interaction between the decision contexts and the training runs was significant at the levels where the gain-to-losses ratio was close to the acceptance criterion in each decision context (gray dashed lines). For instance, the drift rate of those originally unattractive (gain-to-loss ratio = .88–1.07, the 2$^{nd}$ level on the axis of Figure 15A) or indifference stocks (gain-to-loss ratio = 1.18 – 1.8, the 3$^{rd}$ level on the axis of Figure 15A) gradually became positive (more attractive to the participants) in the risky context, whereas the drift rate of the same stocks progressively turned negative in the conservative context (interaction effects on the 2$^{nd}$ level: $F(4,484) = 3.31$, $p=.01 < \alpha = .05$; on the 3$^{rd}$ level: $F(4, 484) = 14.31$, $p<.0001$). The training experience also

enhanced the attraction of an originally good financial offer (gain-to-loss ratio = 2.59–3.17, the 5th level o the axis of Figure 15A) in the risky context but reduced that in the conservative context (interaction effect on the 5th level: $F(4, 484) = 2.47$, $p = .04 < \alpha = .05$).

One might suspect that training may increase the signal-to-noise ratio by reducing the noises in the evidence accumulation process that can be quantified as the drift variance. However, as shown in Figure 15D, we found that the drift variance remained constant over the five training runs ($F(4, 484) = 1.9745$, $p = 0.10 > \alpha = .05$) and the two conditions ($F(1, 484) = 0.4497$, $p = .51 > \alpha = .05$). The interaction between decision context and run was also not significant ($F(4, 484) = 1.2609$, $p = .29 > \alpha = .05$).

We found that people tended to accept the stock regardless of the magnitudes of the potential gains and the potential losses. This can be seen from Figure 15B that the overall starting-point is greater than zero. There was also significant main effect on such bias across the two decision contexts ($F(1, 484) = 39.31$, $p<.0001$). The tendency to accept a stock was greater in the conservative context than that in the risky context. However, this tendency to accept the stock went down over the training runs (main effect on runs: $F(4, 484) = 2.78$, $p = .026 > \alpha = .05$). There was no significant interaction between the decision context and the training runs ($F(4, 484) = 1.728$, $p = .1425 > \alpha = .05$), suggesting that this bias was not a result of training.

Finally, the non-decision time became shorter over the training (Figure 15C). The main effect on run was significant ($F(4, 484) = 2.51$, $p = .04 < \alpha = .05$). There was no significant effect on the interaction between decision context and run ($F(4, 484) = .83$, $p = 0.5012 > \alpha = .05$) nor the main effect on the decision context ($F(1, 484) = 1.18$, $p = 0.2785 > \alpha = .05$).

## DISCUSSION

The influences of experience on the decision-making under risk were identified using the drift-diffusion model to complement the limited temporal resolution of the BOLD signals in functional neural imaging. We showed that the parameters in this model could be estimated more efficiently and flexibly with analytically derived joint distribution of choice and response time. The ratio between the potential gains and the potential losses rather than the subjective value computed from prospect theory better accounted for the evidence that drove decisions in our sample. Also, the magnitude of this decision evidence was modulated by the reinforcement in each decision context. This was shown as the effects of interaction between the decision contexts and the training runs on the drift rates. However, the drift variance remained constant over the training in both decision contexts, suggesting that the noises perturbing the evidence accumulation process remained the same. Finally, the non-decision time decreased over the training, indicating overall improvement of perceptual and motor processing in the task.

The finding that the evidence driving decisions about accepting or rejecting a potential financial outcome was better described as the gain-to-loss ratio suggests that financial decisions can be made by simplified the features of each option. It has been suggested that when choosing from multiple alternatives defined by two dimensions (e.g. considering the price and fuel economics when buying a car), people constantly switch attention across one of the dimensions and compare all the options alone this currently attended dimension at a time (Tversky, 1972; Usher & McClelland, 2004). For each option on this attended dimension, the negative difference across all the options would be weighted greater than the positive difference, indicating the loss aversion (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). All the different scores then were summed across all dimensions to derive the evidence supporting each alternative. Instead of switch between the dimensions for consideration, we found that people may use a linear combination to reduce the two dimensions into one and compare the options alone this new dimension (e.g. gain-to-loss ratio) in order to derive evidence supporting a choice. Furthermore, the weights for reducing two dimensions into one can be shaped by experiences in the current decision context. Our finding hence highlights the need to investigate the generation of heuristic strategies (e.g. dimension reduction) and its influences on the neural system in decision-making under risk.

The driving force underlying the changes in decision-making under risk is of great interest. This learning process may be similar to learning a rule to classify objects that required working memory to verbalize the rule (accepting that offer if the gains are x times greater than the losses) and test hypotheses. Such learning relies the frontal-striatum loop in the brain (Ashby & Maddox, 2005). However, this idea may not apply according to the results from the previous chapter: we only observed the activation in frontal lobe but not the striatum during the training and the probe. Alternatively, this learning process may be similar to the prediction-error-driven reinforcement learning mechanism (Schultz, 1998; Sutton & Barto, 1998). We have built the groundwork for testing this hypothesis. By showing that parameters in drift-diffusion model can be accurately estimated using analytical solutions, this approach can be combined with learning algorithms to estimate the cross-correlation between the feedback and the next choice. Unfortunately, deriving the detailed learning algorithms is beyond the timeline of the dissertation; hence, we left it to future study.

One may criticize that our modeling approach may ignore an important factor that participants may trade off speed and accuracy over the training. For instance, findings in lexical categorization (word vs. non-word) showed that over the training, the decision boundaries in the drift-diffusion process decreased. This suggested that experiences in two-alternative categorization made people became less cautious in making a choice (Dutilh, Krypotos, & Wagenmakers, 2011). However, in our study, we

fixed the height of the decision boundaries when applying the analytical solution to the drift-diffusion model in order to constrain the parameter estimation. This is an inevitable caveat of our approach.

In conclusion, we found that applying rules can better account for the decision-making process when people learned to make risker or more conservative financial choices over experiences. We also showed that parameters in drift-diffusion model could be estimated more efficiently with analytical solutions. This opened the potential to apply drift diffusion model to cases where trial-by-trial differences were important in decision-making.

# Conclusions

Through tracking how bias could be developed in perceptual and financial decisions, I presented evidence suggesting an integrated picture of a domain general decision-making system in the human brains. The core brain structures of this system were the frontal and parietal cortices. The major role of this system was constantly associating decision contexts and reward feedback in order to learn how objective evidence (e.g. visual stimulus in perceptual decisions, magnitudes of costs and benefits financial decisions) should be evaluated. Decision bias hence occurred when this system altered its strengths of link between the functionally specialized brain areas that processed the evidence under consideration based on the current decision context. For instance, the functional connectivity among the frontal, visual, and value systems increased as people developed the tendency to make a perceptual choice according to the reward magnitudes in the decision context rather than abide by the perceptual evidence presented in front of them. Moreover, when people were encouraged to make risky or conservative financial choices according to the decision context, the frontal cortex gauged its functional connectivity between the value systems to determine the weights of the potential gains in a financial choice. Results from analyzing the representation of this decision evidence with a mathematical model suggested that this system might further synthesize experiences in complicated choices (e.g. costs and benefits) into rules of thumb in each decision context. These findings highlight the importance of integrating domain-specific knowledge in

decision neuroscience into a whole and provide insights into how our decision-making processes could be refined in the face of experiences.

**Table 1.** Best-fitting parameter estimates in the study of Chapter 1

| Parameters | Learning Model I (contextual action-outcome association) | Learning Model II (action-outcome association) | Hierarchical logistic regression (context by run number interaction) |
|---|---|---|---|
| $b_0$ | .81 (.39, 1.23) | 1.03 (.01, 1.44) | -- |
| $\alpha$ | .02 (.009, .03) | .01 (.006, .02) | -- |
| $b_1$ | 4.01 (3.54, 4.80) | 3.54 (2.47, 4.11) | -- |
| AIC | **6672.9** | 7770.24 | 6962 |
| N(fit)/N(total) | 21/23 | 2/23 | -- |

Notes. The parameter values are shown as median and the interquartile range ($25^{th}$, $75^{th}$ percentile) across participants. Also shown are the proportions of participants whose data are better fit by each of the learning model. Bold fonts: the best-fit model. N: the number of participants, AIC: Akaike's Information Criterion.

**Table 2.** GLM results in the Chapter 1

| Effect | Brain regions | Cluster size (voxels) | p-value | z-value | MNI coordinates | | |
|---|---|---|---|---|---|---|---|
| | | | | | X | Y | Z |
| Acquired bias | Frontal Orbital Cortex | 1336 | <.0001 | 4.13 | 46 | 20 | -12 |
| | Superior Frontal Gyrus | 949 | <.0001 | 3.37 | 0 | 34 | 50 |
| | Inferior Frontal Gyrus | 766 | <.0001 | 3.77 | -54 | 24 | 12 |
| | Posterior Supramarginal Gyrus | 729 | <.0001 | 3.98 | -52 | -46 | 10 |
| | Superior Lateral Occipital Cortex | 692 | <.0001 | 3.48 | -12 | -84 | 38 |
| | Posterior Middle Temporal Gyrus | 364 | <.0001 | 3.25 | 62 | -30 | -16 |
| | Postcentral Gyrus | 262 | <.0001 | 3.55 | 2 | -34 | 52 |
| | Brain-Stem | 237 | <.0001 | 3.89 | 14 | -22 | -14 |
| | Middle Frontal Gyrus | 167 | <.0001 | 3.13 | 38 | 0 | 64 |
| | Precentral Gyrus | 163 | <.0001 | 3.17 | 44 | 0 | 34 |
| | Frontal Medial Cortex | 105 | 0.005 | 2.95 | 0 | 46 | 2 |
| | Posterior Superior Temporal Gyrus | 100 | 0.007 | 3.16 | 58 | -34 | 4 |
| | Frontal Pole | 94 | 0.011 | 3.40 | -28 | 52 | 34 |
| | Temporal Pole | 91 | 0.014 | 3.26 | -54 | 4 | -24 |
| Contextual RPE | Superior Lateral Occipital Cortex | 3635 | <.0001 | 4.28 | 50 | -62 | 24 |
| | Posterior Supramarginal Gyrus | 1596 | <.0001 | 4.39 | -50 | -48 | 52 |
| | Superior Frontal Gyrus | 1395 | <.0001 | 4.26 | -16 | 18 | 52 |
| | Posterior Middle Temporal Gyrus | 1189 | <.0001 | 4.26 | -62 | -40 | -8 |
| | Middle Frontal Gyrus | 1134 | <.0001 | 3.69 | 38 | 26 | 52 |
| | Frontal Medial Cortex | 1065 | <.0001 | 4.43 | 2 | 44 | -4 |

**Table 2.** (Cont.)

| Effect | Brain regions | Cluster size (voxels) | p-value | z-value | MNI coordinates | | |
|---|---|---|---|---|---|---|---|
| | | | | | X | Y | Z |
| | Posterior Cingulate Gyrus | 1063 | <.0001 | 3.80 | 0 | -36 | 42 |
| | Frontal Pole | 443 | <.0001 | 3.58 | 50 | 44 | 20 |
| | Inferior Lateral Occipital Cortex | 308 | <.0001 | 3.15 | 42 | -78 | -2 |
| | Right Caudate | 270 | <.0001 | 4.42 | 12 | 12 | 4 |
| | Cerebellum | 234 | <.0001 | 4.15 | -40 | -68 | -38 |
| | Frontal Pole | 194 | <.0001 | 3.85 | -20 | 36 | -16 |
| | Left Caudate/Accumbens | 170 | <.0001 | 4.01 | -10 | 10 | 0 |
| | Frontal Orbital Cortex | 100 | 0.002 | 3.39 | -26 | 22 | -20 |
| | Anterior parahippocampal Gyrus, | 67 | 0.037 | 3.37 | 20 | -2 | -26 |
| Context-free RPE | Occipital Pole | 179 | <.0001 | 3.49 | 28 | -96 | 16 |
| | Pre-central Gyrus | 101 | 0.002 | 3.28 | -48 | -18 | 52 |

**Table 3.** PPI results in the Chapter 1

| Seed regions | Co-activating brain regions | Cluster size (voxels) | p-value | z-value | MNI coordinates | | |
|---|---|---|---|---|---|---|---|
| | | | | | X | Y | Z |
| vmPFC | Pre-central Gyrus | 56 | 0.037 | 3.44 | -36 | -30 | 68 |
| Frontal cortex | | | | | | | |
| (Left) | Occipital Pole | 325 | <.0001 | 3.27 | 22 | -92 | -2 |
| | Paracingulate Gyrus | 131 | <.0001 | 3.42 | 0 | 54 | 16 |
| | Frontal Medial Cortex | 117 | <.0001 | 3.24 | -4 | 38 | -20 |
| | Occipital Fusiform Gyrus | 90 | 0.003 | 3.10 | -48 | -70 | -26 |
| | Right Putamen/Caudate | 63 | 0.039 | 3.09 | 18 | 16 | -4 |
| Parietal cortex | | | | | | | |
| (Left) | Anterior Cingulate Cyrus | 98 | 0.001 | 3.07 | 10 | 38 | 6 |
| | Paracingulate Gyrus | 71 | 0.009 | 3.37 | 18 | 52 | 2 |

**Table 4.** FMRI results in Chapter 2: General Linear Model (Risky > Conservative).

| Variebles | Location | Cluster size (voxels) | Z statistics | P value | MNI (mm) X | Y | Z |
|---|---|---|---|---|---|---|---|
| **Training phase (Main effect)** | | | | | | | |
| Gains | Frontal Pole | 2699 | 4.12 | <.0001 | -22 | 42 | 52 |
| | Precuneous Cortex | 1761 | 4.27 | <.0001 | -2 | -58 | 30 |
| | Lateral Occipital Cortex | 1709 | 4.42 | <.0001 | -46 | -62 | 26 |
| | Frontal Pole | 566 | 4.3 | 0.0018 | 24 | 44 | 46 |
| | Lateral Occipital Cortex | 564 | 4.27 | 0.0018 | 52 | -60 | 26 |
| | Inferior Temporal Gyrus | 350 | 3.85 | 0.0371 | -68 | -30 | -24 |
| Losses | Lateral Occipital Cortex | 22321 | 6.45 | <.0001 | -46 | -62 | 24 |
| | Precuneous Cortex | 4262 | 6.06 | <.0001 | -4 | -58 | 34 |
| | Angular Gyrus | 2441 | 5.8 | <.0001 | 50 | -56 | 30 |
| | Middle Temporal Gyrus | 1411 | 4.93 | <.0001 | 64 | -10 | -20 |
| **Training phase (Linear trend)** | | | | | | | |
| Gains | Angular Gyrus (Parietal cortex) | 664 | 4.58 | .0002 | 46 | -54 | 32 |
| | Frontal Pole (Frontal cortex) | 653 | 4.33 | .0003 | 30 | 36 | 50 |
| | Middle Temporal Gyrus | 373 | 4.29 | .0157 | 60 | -8 | -22 |
| Losses | Frontal Pole (Frontal cortex) | 417 | 4.2 | .0065 | 30 | 60 | 12 |
| **Probe phase** | | | | | | | |
| Gains | Frontal Pole (Frontal cortex) | 464 | 3.83 | .0065 | 28 | 36 | 40 |
| | Frontal Medial Cortex (vmPFC) | 404 | 3.81 | .0154 | -12 | 54 | -4 |
| | Angular Gyrus (Parietal cortex) | 337 | 4.07 | .0424 | 54 | -48 | 18 |
| Losses | Supramarginal Gyrus (Parietal cortex) | 933 | 3.88 | <.0001 | 56 | -42 | 28 |
| | Supplementary Motor Cortex | 365 | 3.65 | .0357 | 4 | -2 | 60 |

**Table 5.** Functional connectivity with vmPFC (Conservative > Risky).

| Variables | Location | Cluster size (voxels) | Z statistics | P value | MNI (mm) | | |
|---|---|---|---|---|---|---|---|
| | | | | | X | Y | Z |
| **Training phase (Main effects)** | | | | | | | |
| Gains | Middle Frontal Gyrus (Frontal cortex) | 508 | 4.15 | 0.0014 | -48 | 12 | 38 |
| | Supramarginal Gyrus (Parietal cortex) | 401 | 3.38 | 0.0076 | -44 | -54 | 44 |
| **Training phase (Linear trend)** | | | | | | | |
| Losses | Superior Parietal Lobule (Parietal cortex) | 1706 | 4.2 | <.0001 | 0 | -50 | 68 |
| | Midbrain | 408 | 4.35 | 0.0041 | -6 | -16 | -24 |
| | Angular Gyrus (Parietal cortex) | 331 | 3.53 | 0.0160 | 44 | -50 | 30 |

**Table 6.** Testing different representations of evidence in financial choices.

| Model | Negative log likelihood | AIC | BIC | Proportion of fit among the sample |
|---|---|---|---|---|
| **Gain-to-loss ratio: 6 levels** | **60449** | **67019** | **73987** | **.91** |
| Subjective value function | 66818 | 73118 | 79872 | .95 |

*Notes.* The bolded font indicates the model that best accounts for the obtained data.

**Figure 1 Experimental paradigm used in Chapter 1.** Each trial is composed of a context, motion stimulus (illustrated as white dots), and reward points. The number of reward points that one can earn depends on the context and the choice. P: probability; Δt: duration; AVG: average; s: second.

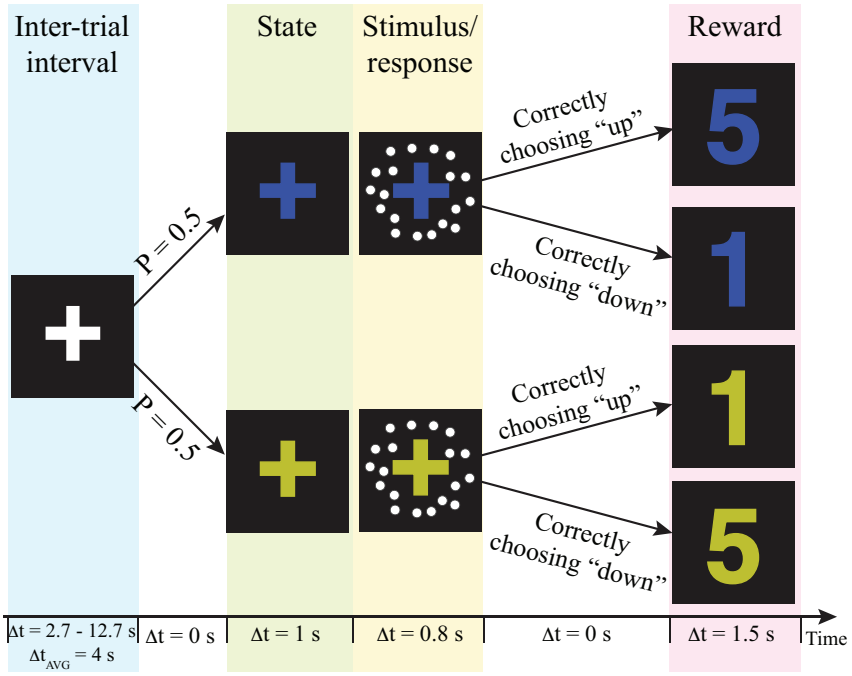


**Figure 2 Behavioral results in Chapter 1.** *A) Psychometric function*. The motion strength was plotted against the probability of choosing "up" in each context across the five runs (dots). This change is modeled with logit function in which its intercept reflects the value difference between the two motion directions that has been learned up to the end of each run (solid lines). The dashed lines are the indecision points. Motion strength: the percentage of coherent moving-dots; +/-: upward/downward motion. Error bars: ±1 s.e.m. *B) The bias acquisition process*. The trial numbers are plotted against the indecision points estimated by the reinforcement-learning model using individuals' data. Solid lines: group mean. Shaded areas: ±1 s.e.m. Dashed lines: the end of each run. Colors: corresponding to the reward context as illustrated in Figure 2A. C) *The autocorrelation functions*. The correlation estimates using the residuals from the context-based learning model is plotted against each lag. Solid lines: group average. Dashed lines: 95% confidence interval of the autocorrelation estimated from a random series with the same number of trials (a total of 280 trials).

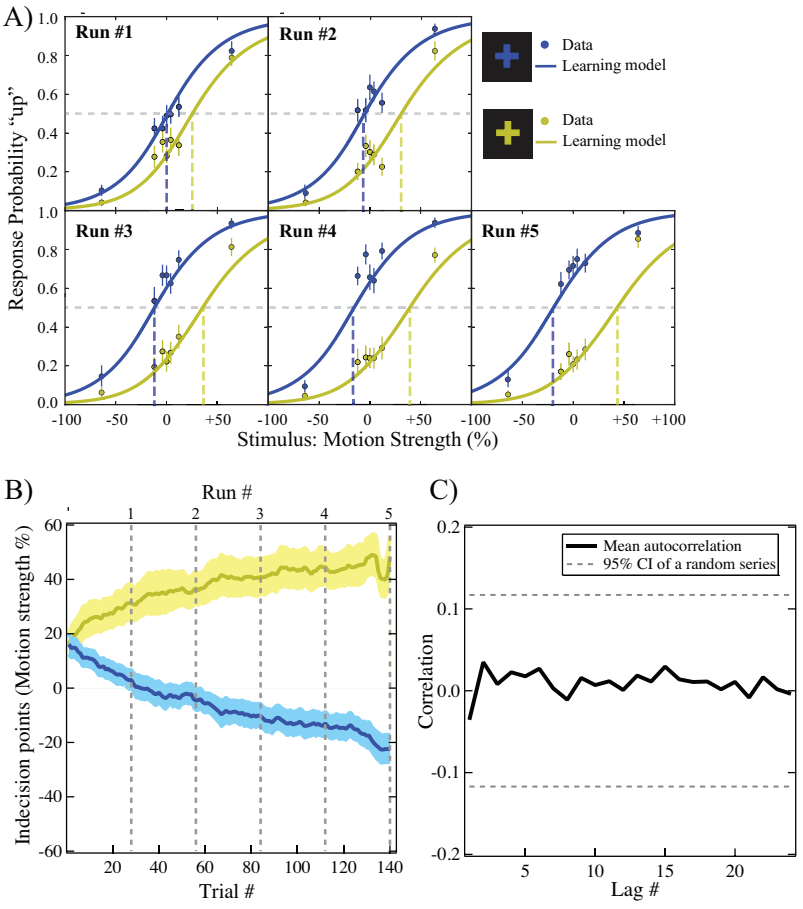


**Figure 3 The acquired perceptual bias in the bran.** The maps show the brain areas whose activation positively correlates with the amount of acquired bias on each trial. No brain areas negatively correlate with this signal after the whole-brain correction of multiple comparisons. All maps are presented at p < .05 whole-brain corrected using cluster-based Gaussian random field and overlaying on the mean anatomical images from the group of participants. R: right hemisphere; L: left hemisphere; Z: the MNI coordinate of the axial slice.

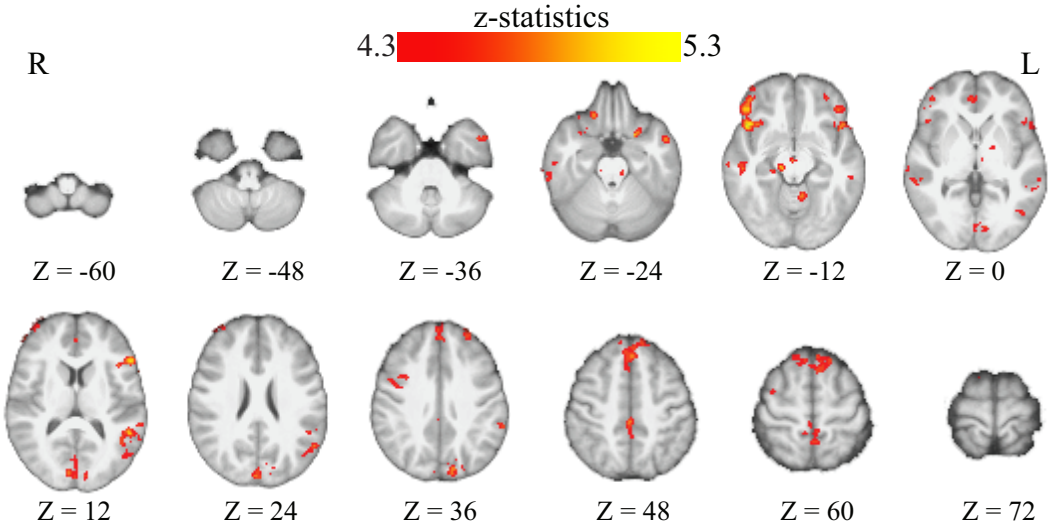


**Figure 4 The reward prediction errors.** The brain maps show the regions whose activation positively correlates with the two types of reward prediction error signals in perceptual decisions after adjusting one type over the other. All maps are presented at $p < .05$ whole-brain corrected using cluster-based Gaussian random field and overlaying on the mean anatomical images from the group of participants. Red-Yellow: context-based RPE; Blue-Light-blue: context-free RPE; R: right hemisphere; L: left hemisphere; Z: the MNI coordinate of the axial slice.

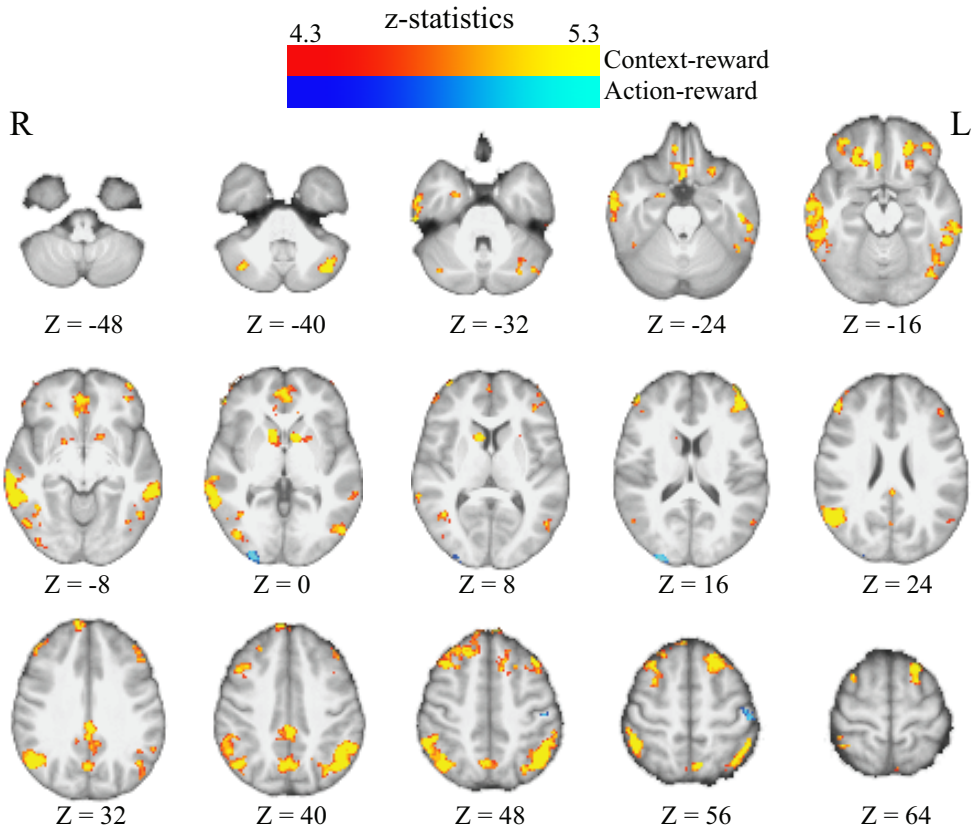


**Figure 5 Three functional connectivity patterns underlying the growth of perceptual bias.** The left panel shows the seed regions that were used in the psychophysiological interaction analyses. These seed regions were selected according to the literature and centered at the MNI coordinates (vmPFC: [-6, 39, -8]; left-frontal: [-45, 21, 0]; left-parietal: [-36, -39, 45]) with the radius of 10 mm. The brain maps on the right panel show the areas that positive correlate with the interaction between each of the seed regions and the amount of acquired bias on each trial. The statistical maps are corrected for multiple comparisons at the whole-brain level using cluster-based Gaussian random field correction at $P < 0.05$ and overlaying on the mean anatomical images from the group of participants. R: right hemisphere; L: left hemisphere; X, Y, Z: the MNI coordinates.

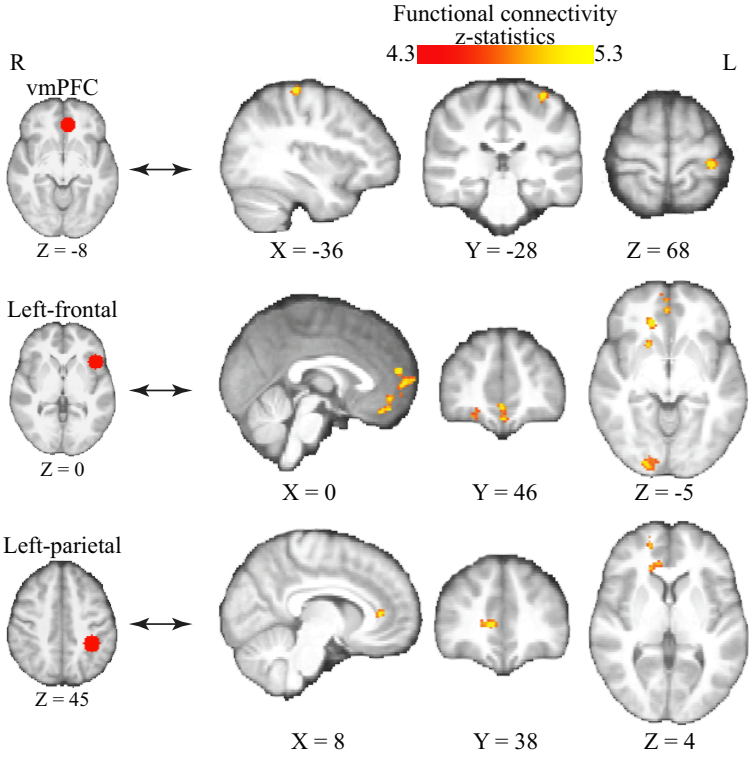


**Figure 6 Behavioral paradigms used in Chapter 2.** Left panel: The timeline of each event in the three phases of the experiment. Right panel: The ranges of the potential gains and the potential losses. In the baseline and probe phase, all the tested pairs of the potential gains and losses are shown as the grids. In the training phase, the tested pairs in each decision context are shown as the open circles. The blue and the yellow shaded areas show those pairs that should be accepted according to the decision criterion of each context. Those pairs fall in the gray shaded area should be rejected. ΔT: duration; s: seconds.

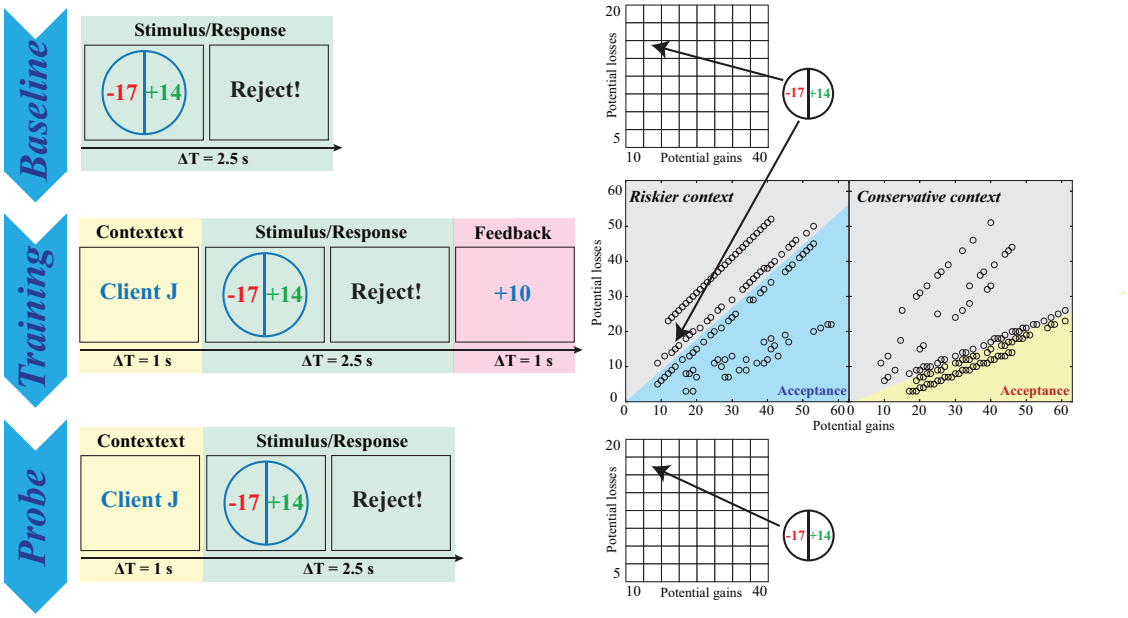


**Figure 7 Changes in behavioral loss aversion.** A) The probability of accepting a stock at each level of gain-to-loss ratio is gradually modulated by the decision contexts. The progression of training runs is shown as the gradient from light to dark colors. The ranges of the gain-to-loss ratio at each level, from the level-1 to the level-6, are .52–.79, .88–1.07, 1.18 – 1.8, 2.1–2.41, 2.59–3.17, 3.21–6.33. Dashed gray line: indecision points. B) The measurements of loss aversion over the three phases of the experiment. Top panel: The sensitivities toward the potential gains and the potential losses in the two decision contexts. Bottom panel: The loss aversion indicators in the two decision contexts. Black dashed line: the decision criterion in each context; Solid dots: group average; Error bars: ±1 standard error of the group average.

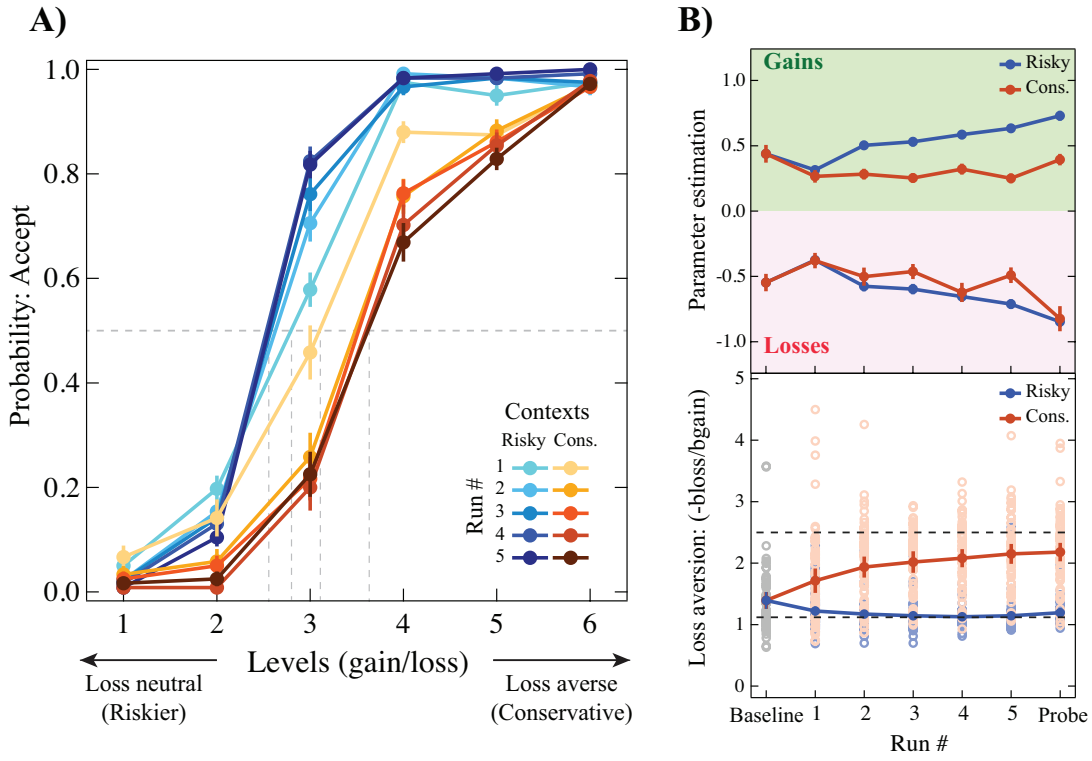


**Figure 8 Choice and response time in the baseline and probe phases.** A) Baseline. B) The loss neutral (risky) context of the probe. People should accept the stocks as long as the potential gains are 1.12 times greater than the potential losses. C) The loss averse (conservative) context of the probe. People should accept the stocks when the potential gains are at least 2.5 times greater than the potential losses. P(accept): Probability of acceptance.

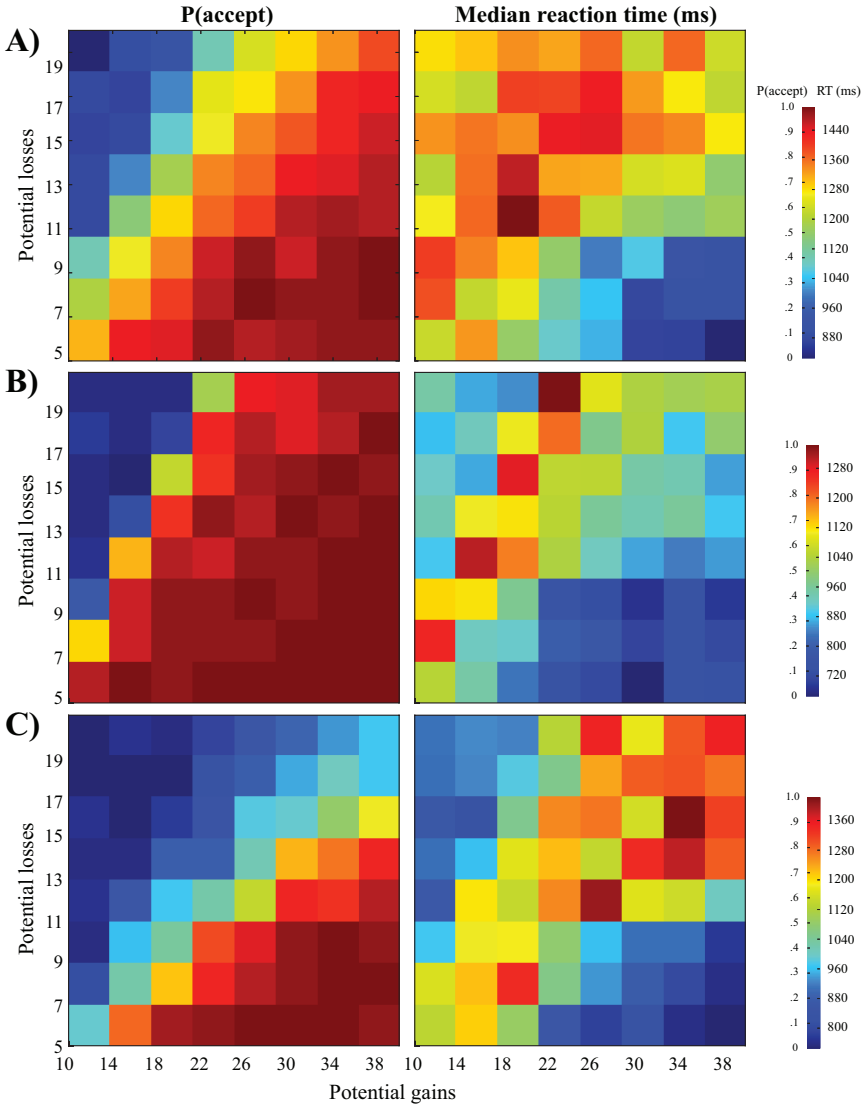


**Figure 9 FMRI results in the baseline phase.** Brain regions positively correlated with the potential gains but negatively correlated with the potential losses during the baseline phase. Notice that the real values of the potential losses were used in the design matrix; therefore, positive sign of the z-value indicates negative correlation. The brain map of the potential gains is presented at $p < .05$ whole-brain corrected using cluster-based Gaussian random field. The brain map of the potential losses is uncorrected. L: left hemisphere; Y: the MNI coordinates.

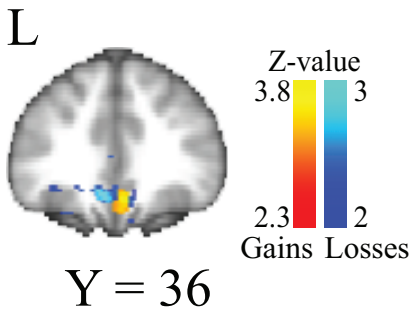


**Figure 10 Training effect on brain responses (Risky – Conservative).** A) Main effect. Brain areas showed greater responses toward the potential outcomes in the risky than in the conservative context. B) Linear trends. Brain areas gradually show greater responses toward the potential outcomes in the risky than that in the conservative context. C) The plots of the peak response areas shown in (B) against the five training runs. Red shaded plots: responses toward the potential gains; blue shaded plot: responses toward the potential losses; red solid lines and dots: group averaged response in the conservative context; blue solid lines and dots: group averaged in the risky context; solid dots: group average; error bars: ±1 standard errors. Left: left hemisphere; Right: right hemisphere; Y, Z: the MNI coordinates. All the brain maps are presented at $p < .05$ whole-brain corrected using cluster-based Gaussian random field.

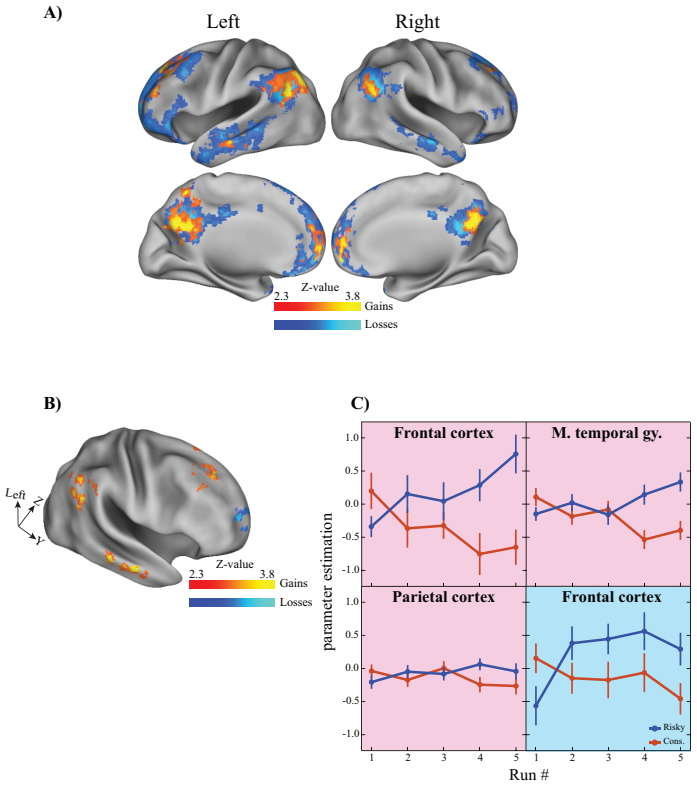


**Figure 11 Training effect on the functional connectivity with the vmPFC (Conservative - risky).** A) Brain areas show greater functional connectivity between the vmPFC over the training when participants evaluated the potential gains (red-yellow) or the potential losses (blue-light blue). The map is presented at $p < .05$ whole-brain corrected using cluster-based Gaussian random field. Left or right: left or right hemisphere. B) The plots of the responses of the peak voxel within each cluster in (A) against the five training runs. Red shaded plots: the responses toward the potential gains; blue shaded plot: the responses toward the potential losses; red solid lines and dots: measured in the conservative context; blue solid lines and dots: measured in the risky context; solid dots: group average; error bars: ±1 of the standard errors.

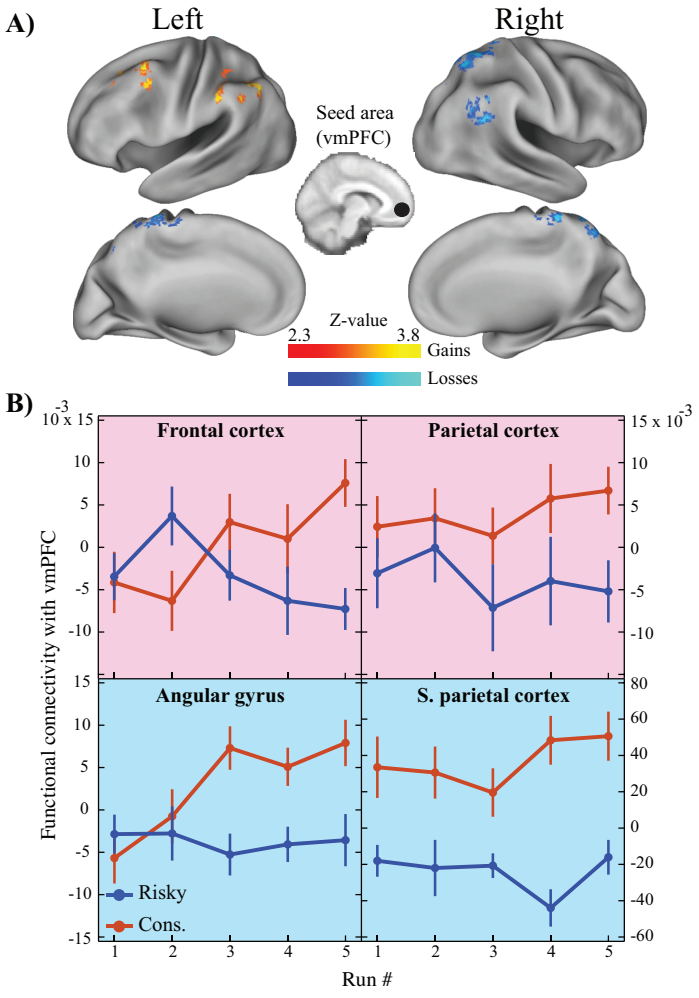


**Figure 12 Brain activation in the probe phase (Risky - conservative).** Brain areas show greater responses toward the potential gain (red-yellow) or the potential losses (blue-light blue) in the risky than that in the conservative context when probed after the training. All the maps are presented at $p < .05$ whole-brain corrected using cluster-based Gaussian random field. Y, Z: the MNI coordinates; Left or right: left or right hemisphere.

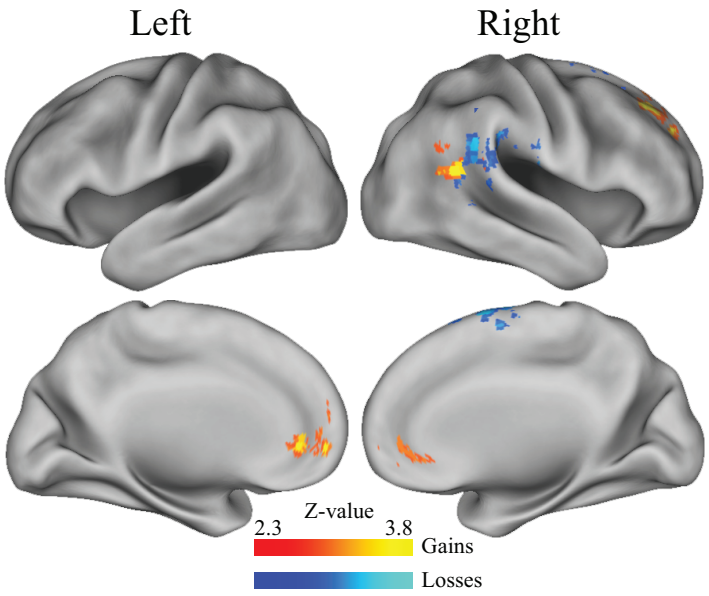


**Figure 13. Convergence test results.** A) – D): Results from four different parameter sets. The left panel shows the comparison of choice and response time joint distribution between the simulation and analytical solution. Red: response time ending at the upper bound; black: response time ending at the lower bound; dots: simulation results; lines: analytical solutions. The right panel plots the sample size used in the simulation against the mean square errors between the true and best-fit parameters. $\sigma^2$: drift variance; $\mu$: drift rate; $x_0$: starting point; $\tau$: non-decision time.

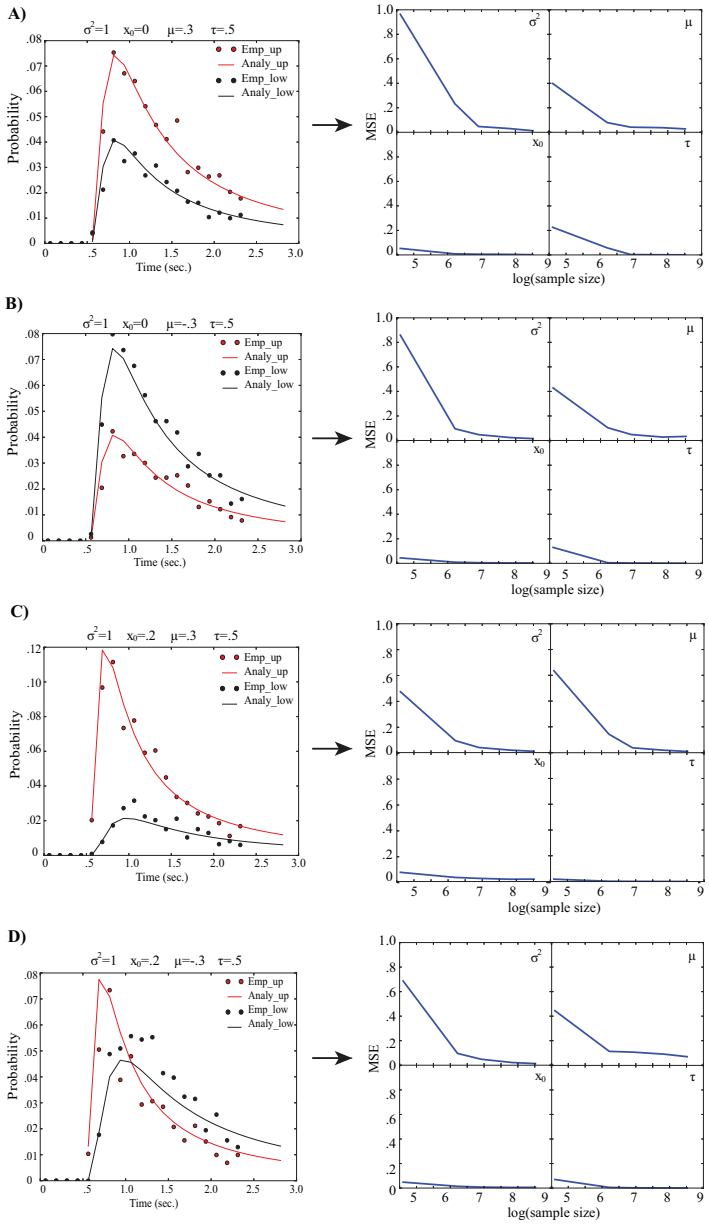


**Figure 14. The comparison between the model and the data.** The empirical and model derived choice accuracy and response time are plotted against each other together with the corresponding Spearman correlation and $R^2$ indicating the goodness of fit in each context and training run. The black line shows the cases of perfect model fit. Blue dots and font color: risky context; red dots and font color: conservative context.

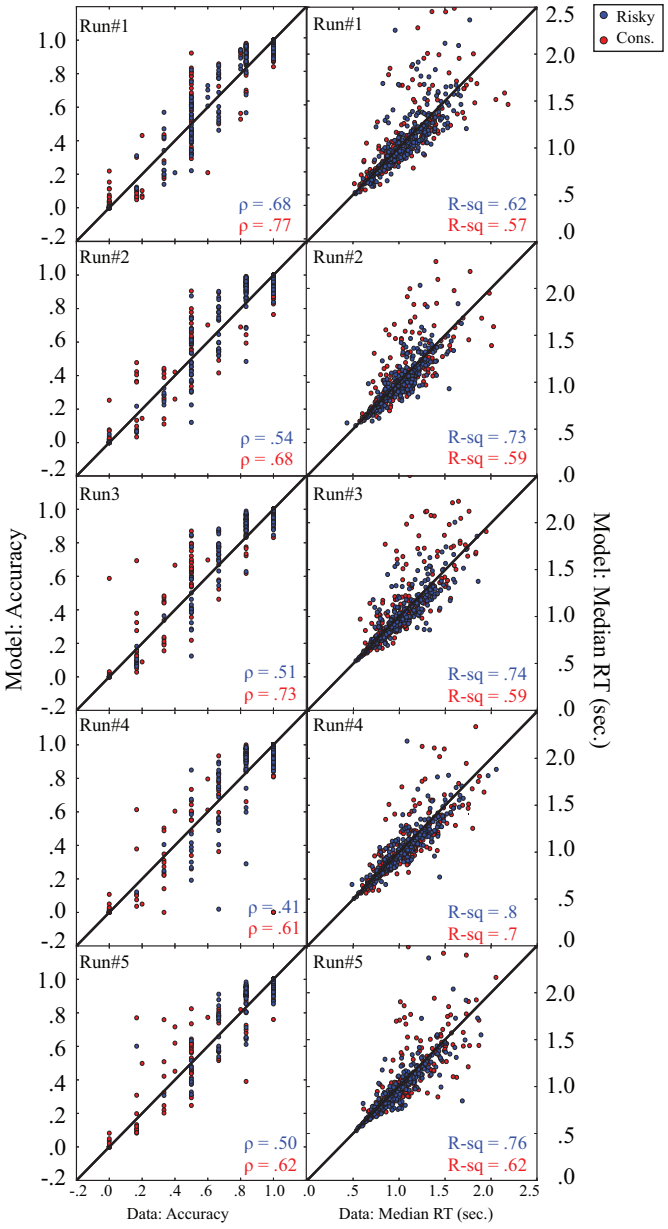
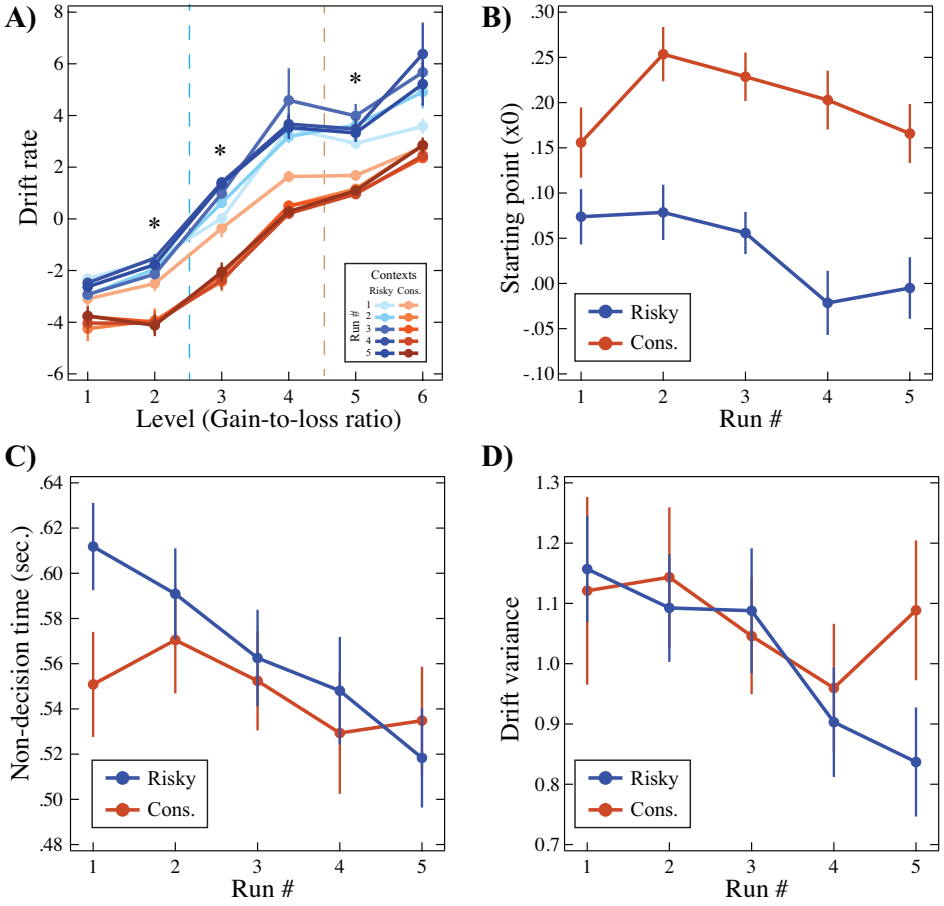


**Figure 15. Parameter estimation from the drift-diffusion model.** A) The drift rate estimated in the six different levels of gain-to-loss ratio. Blue gradients: risky context; brown gradients: conservative context; dashed line: criteria for acceptance in the risky (blue) and conservative (brown) context; stars: significant context-by-run interaction effects. B) The starting points. C) The non-decision time. D) The drift variance. Dots: group average; error bars: ± 1 standard error.



# References

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 19: 716–723.

Ashby, F. G., & Maddox, W. T. (2005). Human Category Learning. Annual Review of Psychology, 56(1), 149–178.

Basten, U., Biele, G., & Heekeren, H. R. (2010). How the brain integrates costs and benefits during decision making. Pnas, 1–6.

Bartra, O., McGuire, J. T., Kable, J. W. 2013. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage. 76: 412–427.

Blank, H., Biele, G., Heekeren, H. R., Philiastides, M. G. 2013. Temporal Characteristics of the Influence of Punishment on Perceptual Decision Making in the Human Brain. Journal of Neuroscience 33: 3939–3952.

Bogacz, R., Larsan, T. 2011. Integration of Reinforcement Learning and Optimal Decision-Making Theories of the Basal Ganglia. Neural Computation. 23: 817–851.

Bunge, S. A. (2004). How we use rules to select actions: A review of evidence from cognitive neuroscience. Cognitive, Affective, & Behavioral Neuroscience, 4(4), 564–579.

Box, G., Jenkins, G. M., Reinsel, G. 1994. Time Series Analysis: Forecasting and Control (3rd Edition). USA (NJ): Wiley.

Brainard, D. H. 1997. The Psychophysics Toolbox. Spatial Vision. 10: 433–436.

Britten, K. H., Shadlen, M. N., Newsome, W. T., Movshon, J. A. 1993. Responses of neurons in macaque MT to stochastic motion signals. Vision Neuroscience. 10: 1157–1169.

Canessa, N., Crespi, C., Motterlini, M., Baud-Bovy, G., Chierchia, G., Pantaleo, G., et al. (2013). The Functional and Structural Neural Basis of Individual Differences in Loss Aversion. Journal of Neuroscience, 33(36), 14307–14317.

Cho, R. Y., Nystorm, L. E., Brown, E. T., Jones, A. D., Braver, T. S., Holmes, P., Cohen, J. D. 2002. Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. Cognitive, Affective, & Behavioral Neuroscience. 4: 283–299.

Daw, N. D. (2009). Trial-by-trial data analysis using computational models. In Affect, learning, and decision making.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. Neuron. 69: 1204–1215.

De Martino, B., Camerer, C. F., & Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. Proceedings of the National Academy of Sciences, 107(8), 3788–3792.

Dutilh, G., Krypotos, A.-M., & Wagenmakers, E.-J. (2011). Task-Related Versus Stimulus-Specific Practice. Experimental Psychology (Formerly Zeitschrift Für Experimentelle Psychologie), 58(6), 434–442.

Ding, L., Gold, J. I. 2010. Caudate encodes multiple computations for perceptual decisions. Journal of Neuroscience. 30: 15747–15759.

Ding, L., Gold, J. I. (2013). The Basal Ganglia's Contributions to Perceptual Decision Making. Neuron. 79: 640–649.

Duncan, J. 2001. An adaptive coding model of neural function in prefrontal cortex. Nature Reviews Neuroscience. 2: 820–829.

Edwards, W. 1965. Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. Journal of Mathematical Psychology. 2: 312–329.

Feng, S., Holmes, P., Rorie, A., Newsome, W. T. 2009. Can Monkeys Choose Optimally When Faced with Noisy Stimuli and Unequal Rewards? PLoS Computational Biology. 5: e1000284.

Fleming, S. M., Whiteley, L., Hulme, O. J., Sahani, M., Dolan, R. J. 2010. Effects of Category-Specific Costs on Neural Systems for Perceptual Decision-Making. Journal of Neurophysiology. 103: 3238–3247.

Garrison, J., Erdeniz, B., Done, J. 2013. Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. Neuroscience and Biobehavioral Reviews. 37:1297–1310.

Gitelman, D. R., Penny, W. D., Ashburner, J., Friston, K. J. 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, J. P. 2010. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron. 66: 585–595.

Glimcher, P. W. 2011. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. PNAS. 108:15647–15654.

Gold, J. I., Ding, L. 2013. How mechanisms of perceptual decision-making affect the psychometric function. Progress in Neurobiology. 103: 98–114.

Gottlieb, J. 2012. Attention, Learning, and the Value of Information. Neuron. 76: 281–295.

Green, D. M., Swets, J. A. 1966. Signal detection theory and psychophysics. USA(NY): Wiley. Reprinted by Krieger, Huntington, 1974.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. Science, 324(5927), 646–648.

Helfinstein, S. M., Schonberg, T., Congdon, E., Karlsgodt, K. H., Mumford, J. A., Sabb, F. W., et al. (2014). Predicting risky choices from brain activity patterns. *Proceedings of the National Academy of Sciences*, *111*(7), 2470–2475.

Kahnt, T., Grueschow, M., Speck, O., Haynes, J.-D. 2011. Perceptual Learning and Decision-Making in Human Medial Frontal Cortex. Neuron. 70: 549–559.

Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. Econometrica, 47(2), 263–291.

Kahneman, D., & Tversky, A. (1984). Choice, Values, and Frames. American

Psychologist, 39(4), 341–350.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. Pnas, 108(33), 13852–13857.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. Nature Neuroscience, 13(10), 1292–1298.

Krajbich, I., Lu, D., Camerer, C. F., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. Frontiers in Psychology, 3, 1–18.

Lau, B., Glimcher, P. W. 2005. Dynamic Response-by-Response Models of Matching Behavior in Rhesus Monkeys. Journal of the Experimental Analysis of Behavior. 84: 555–579.

Law, C.-T., Gold, J. I. 2008. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. Nature Neuroscience. 11: 505–513.

Lee, D., Seo, H., Jung, M. W. 2012. Neural basis of reinforcement learning and decision making. Annual Review Neuroscience. 35: 287–308.

Liston, D. B., Stone, L. S. 2008. Effects of Prior Information and Reward on Oculomotor and Perceptual Choices. Journal of Neuroscience. 28:13866–13875.

Macmillan, N. A., Creelman, D. C. 2004. Detection Theory: A User's Guide (2nd ed.). USA: Lawrence Erlbaum.

Maddox, W. T., Bohil, C. J. 1998. Base-rate and payoff effects in multidimensional perceptual categorization. Journal of Experimantal Psychology: Learning, Memory,

and Cognition. 24: 1459–1482.

Miller, E. K., & Cohen, J. D. (2001). AN INTEGRATIVE THEORY OF PREFRONTAL CORTEX FUNCTION. Annu. Rev. Neurosci., 24, 167–202.

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., Uurbil, K. 2010. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. Magnetic Resonance in Medicine. 63: 1144–1153.

Montague, P. R., Berns, G. S. 2002. Neural economics and the biological substrates of valuation. Neuron. 36: 265–284.

Mumford, J. A., & Nichols, T. E. (2008). Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation. NeuroImage, 39(1), 261–268.

Mulder, M. J., Wagenmakers, E. J., Ratcliff, R., Boekel, W., Forstmann, B. U. 2012. Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff. Journal of Neuroscience. 32: 2335–2343.

Newsome, W. T., Britten, K. H., Movshon, J. A. 1989. Neuronal correlates of a perceptual decision. Nature. 341: 52–54.

Nomoto, K., Schultz, W., Watanabe, T., Sakagami, M. 2010. Temporally Extended Dopamine Responses to Perceptually Demanding Reward-Predictive Stimuli. Journal of Neuroscience. 30: 10692–10702.

Pagnoni, G., Zink, C. F., Montague, P. R., Berns, G. S. 2002. Activity in human ventral

striatum locked to errors of reward prediction. Nature Neuroscience.5: 97–98.

Palmer, J., Huk, A. C., Shadlen, M. N. 2005. The effect of stimulus strength on the speed and accuracy of a perceptual decision. Journal of Vision. 5: 376–404.

Pelli, D. G. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial Vision. 10: 437–442.

Pinheiro, J. E., Bates, D. 2009. Mixed-Effects Models in S and S-PLUS (Statistics and Computing) (1st ed.). USA: Springer.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., Petersen, S. E. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage. 59: 2142–2154.

Ratcliff, R. (1976). Retrieval process in recognition memory. Psychological Review, 83(3), 190–214.

Rao, R. P. N. 2010. Decision Making Under Uncertainty: A Neural Model Based on Partially Observable Markov Decision Processes. Frontiers in Computational Neuroscience. 4: .

Rorie, A. E., Gao, J., McClelland, J. L., Newsome, W. T. 2010. Integration of Sensory and Reward Information during Perceptual Decision-Making in Lateral Intraparietal Cortex (LIP) of the Macaque Monkey. PLoS ONE. 5: e9308.

Schultz, W. 1998. Predictive reward signal of dopamine neurons. Journal of Neurophysiology. 80: 1–27.

Serences, J. T. 2008. Value-Based Modulations in Human Visual Cortex. Neuron. 60:

1169–1181.

Shadlen, M. N., Newsome, W. T. 2001. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. Journal of Neurophysiology. 86: 1916–1936.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., Cohen, J. D. 2009. Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. Journal of Experimental Psychology: Human Perception Performance. 35: 1865–1897.

Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. Social Cognitive and Affective Neuroscience, 8(3), 341–350.

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. Pnas, 106(13), 5053–5040.

Summerfield, C., Koechlin, E. 2010. Economic Value Biases Uncertain Perceptual Choices in the Parietal and Prefrontal Cortices. Frontiers in Human Neuroscience. 4.

Summerfield, C., Tsetsos, K. 2012. Building Bridges between Perceptual and Economic Decision-Making: Neural and Computational Mechanisms. Frontiers in Human Neuroscience. 6: 1–20.

Summerfield, C., Egner, T., Mangels, J., Hirsch, J. 2006. Mistaking a house for a face: neural correlates of misperception in healthy humans. Cerebral Cortex. 16: 500–508.

Sutton, R. S., Barto, A. G. 1998. Reinforcement Learning: Introduction. USA (MA): MIT press.

Tom, S. M., Fox, C. R., Trepel, C., Poldrack, R. A. 2007. The neural basis of loss aversion in decision-making under risk. Science. 315: 515–518.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative representation of Uncertainty. Journal of Risk and Uncertainty, 5, 297–323.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157), 1124–1131.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*(4), 281–299.

Usher, M., & McClelland, J. L. (2004). Loss Aversion and Inhibition in Dynamical Models of Multialternative Choice. *Psychological Review*, *111*(3), 757–769.

Watkins, C. J. C. H., Dayan, P. 1992. Q-Learning. Machine Learning. 8: 279–292.

Whiteley, L., Sahani, M. 2008. Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. Journal of Vision. 8: 2.1–215

Worsley, K. J. 2001. Statistical Analysis of Activation Images.

Wunderlich, K., Rangel, A., O'Doherty, J. P. 2009. Neural computations underlying action-based decision making in the human brain. PNAS. 106: 17199–17204.