Influence of expected reward on perceptual decision making

Abbreviated title: Reward influence on perceptual choice

Mohsen Rakhshan [1]*, Vivian Lee [1]*, Emily Chu [1]*, Lauren Harris [2], Lillian Laiks [3], Peyman Khorsand [1], Alireza Soltani [1]

[1]*Department of Psychological and Brain Sciences, Dartmouth College, NH*
[2]*Department of Psychology, Florida State University, Tallahassee, FL*
[3]*Department of Psychological and Brain Sciences, Institute for Neuroscience, Texas A&M University, TX*

* *Equal contribution*

Author contributions: PK and AS designed the study; VL, EC, LH, and LL performed the experiments; MR, PK, and AS analyzed the data; MR, VL, EC and AS wrote the manuscript; all authors contributed to editing the manuscript.

Corresponding author: AS, Department of Psychological and Brain Sciences, Dartmouth College, Hanover NH 03755, soltani@dartmouth.edu

Manuscript information: 9 figures and 1 Extended data figure; 29 pages
Abstract: 195 words
Introduction: 658 words
Results: 3680 words (including figure captions)
Discussion: 1424 words

Keywords: prior entry, response bias, temporal judgment, decision circuit, computational modeling

Abstract

Perceptual decision making is influenced by reward expected from alternative choice options or actions, but the underlying neural mechanisms are currently unknown. More specifically, it is debated whether reward effects are mediated through changes in sensory processing or later stages of decision making, or both. To address this question, we conducted two experiments in which male and female human subjects made saccades to what they perceived to be the first or second of two identical, but asynchronously presented targets with different expected reward. We found similar shifts in target selection between the two experiments indicating that expected reward did not influence sensory processing and, instead, caused response bias. Moreover, the observed response bias was suboptimal and independent of an individual's sensitivity to sensory signals. To elucidate underlying neural mechanisms, we simulated our experiments with a cortical network model based on alternative mechanisms for how reward could exert its influence. We found that our observations are most compatible with reward-dependent input to the output layer of the decision circuit. Together, our results suggest that during a temporal judgment task, reward information influences perceptual choice by changing later stages of decision making rather than early sensory processing.

Introduction

Perceptual decision making depends on both sensory information and reward expected from alternative outcomes (Sugrue, Corrado, & Newsome, 2005). Recently, many studies have begun to investigate the mechanisms by which reward information may influence perceptual decision making (Cicmil, Cumming, Parker, & Krug, 2015; Diederich, 2008; Diederich & Busemeyer, 2006; Farashahi, Ting, Kao, Wu, & Soltani, 2018; Gao, Tortell, & McClelland, 2011; Liston & Stone, 2008; Rajsic, Perera, & Pratt, 2017; Tosoni, Committeri, Calluso, & Galati, 2017; Voss, Rothermund, & Brandtstädter, 2008). Some studies suggest that reward information mainly affects perceptual choice by altering the starting or end point of decision-making processes and thus, results in response bias (Diederich, 2008; Diederich & Busemeyer, 2006; Feng, Holmes, Rorie, & Newsome, 2009; Gao et al., 2011; Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012; Rorie, Gao, McClelland, & Newsome, 2010; Summerfield & Koechlin, 2010). For example, Diederich and colleagues propose that value-based perceptual decision making follows a two-stage process in which the payoff of the alternative choices are evaluated first before sensory information is processed independently of the payoff information (Diederich, 2008; Diederich & Busemeyer, 2006). Others, however, argue that reward directly influences the processing of sensory information during perceptual decision making (Cicmil et al., 2015; Liston & Stone, 2008; Pleger, Blankenburg, Ruff, Driver, & Dolan, 2008; Voss et al., 2008). Understanding how additional information, such as reward, is incorporated into perceptual decision making can provide valuable insight into neural mechanisms underlying these processes (Christopoulos, Bonaiuto, & Andersen, 2015; Christopoulos & Schrater, 2015; Farashahi, Ting, et al., 2018; Gao et al., 2011; Rorie et al., 2010; Stanford, Shankar, Massoglia, Costello, & Salinas, 2010; Sugrue et al., 2005).

These competing hypotheses —influence of reward on sensory processing and influence on later stages of decision making —predict that unequal expected reward should result in perceptual bias or response bias, respectively (Liston & Stone, 2008). That is, by changing sensory processing, reward could alter perception, making the more rewarding target to appear brighter similar to the effect of selective attention (Carrasco & Barbot, 2019; Carrasco, Ling, & Read, 2004). On the other hand, modulation of later decision processes could bias response toward the more rewarding option/target without any changes in perception. Nonetheless, evidence

supporting either hypothesis mostly has been based on fitting performance and reaction time data using different models (e.g. drift diffusion model) and thus, is model dependent. Moreover, it is unclear whether positive and negative expected reward outcomes (e.g., gains and losses) influence perceptual decision making similarly or differently.

To address these questions and distinguish between the two alternative hypotheses, here we used two sets of experiments to directly measure the influence of unequal expected reward on perception and on choice. In these experiments, human subjects performed a temporal judgment task under different reward conditions. More specifically, in Experiments 1 and 2, subjects made free saccades to report what they perceived to be the first (saccade-first task) or second (saccade-second task) of the two identical targets that appeared on the computer screen with varying onset asynchrony, respectively. We manipulated the amount of gained/lost reward points upon correct/incorrect response on the two sides of the fixation cross under three reward conditions: neutral, gain, and loss. This design allowed us to estimate the shift in target selection and sensitivity of choice in response to different target onset asynchronies (TOAs) and how this shift and sensitivity were affected by reward manipulation.

We compared the observed shifts in target selection during Experiments 1 and 2 to determine whether reward influenced sensory processing and/or later stages of decision making. A bias in sensory processing (perceptual bias) would result in opposite shifts in target selection, whereas a response bias would cause similar shifts in our experiments. Additionally, we tested whether observed shifts in behavior were optimal considering individual's sensitivity to sensory information. Finally, we used a biophysically-plausible cortical network model to replicate the experimental data in order to reveal neural mechanisms underlying the influence of reward on perceptual choice.

Materials and Methods

Ethics statement. All participants gave informed consent to participate according to a protocol approved by the Dartmouth College Institutional Review Board. All subjects signed a written consent form before participating in the experiments.

Experimental design. Subjects performed a modified version of the paired-target task with two visual targets (gabor patches) appearing at varying time intervals with respect to each other (**Fig.**

1). All stimuli were presented on an FSI AM250 monitor, which has a refresh rate of 60 Hz and resolution of 1920 × 1080 pixels. Subjects were seated 60 cm from the computer screen. Eye movements were recorded using a video-based eye-tracking system (Eyelink 1000, SR Research Ltd, Mississauga, Canada). To minimize head movements and maintain distance from the screen, subjects were seated with their chin on a chin rest.

At the beginning of each trial, subjects fixated on a white cross at the center of the screen for at least 500 ms. Fixation was considered broken when the subject's eye position deviated 112.5 pixels (~2° visual angle) from the fixation cross before the fixation period ended. If the subject broke fixation, a new trial would begin after a 1000 ms pause. After successful fixation, the amounts of reward points to be earned/lost upon a correct/incorrect response to each target were signaled on the side of the fixation cross corresponding to that target. The amounts of reward points were presented close to the fixation cross and large enough to be read without breaking fixation (~2 ° visual angle from the fixation cross; see **Fig. 1** for the trial sequence). To allow for better discrimination, gain (reward points expected from a correct response) and loss (reward points expected from an incorrect response) points were presented in green and red, respectively. After reward information disappeared from the screen (~1000 ms), there was a variable interval between 500 to 1000 ms (uniform distribution) before the first target appeared on the screen. The second target then appeared after an interval selected from the following values: 0, 16.7, 33.3, 50, and 66.7 ms. We refer to these values as the target onset asynchrony (TOA). Targets were presented in equal distance from the fixation cross (~ 7° visual angle) and large enough to be observed without breaking fixation.

The subject's task in Experiment 1 (respectively, Experiment 2) was to saccade to the first (respectively, second) target that appeared on the screen. Subjects were instructed to select the perceived correct target by making a saccade to the target side only after both targets were presented without any time pressure (i.e., the task was not a reaction time task). Following each response, subjects received reward feedback – a green (red) circle with the amount of reward points earned (lost). They were instructed to earn as many reward points as possible since the total reward points were exchanged for potential monetary reward at the end of each experiment. Subjects were compensated with a combination of money and "t-points," that are extra-credit points for classes within the department of Psychological and Brain Sciences at Dartmouth

College. More specifically, in addition to the base rate of $10/hour or 1 t-point/hour, subjects were compensated by up to $10/hour depending on the total reward points they collected in each experiment. The order of two experiments was randomized across subjects to avoid possible confounds related to ordering effects. The experiments were programmed using PsychToolbox in MATLAB (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

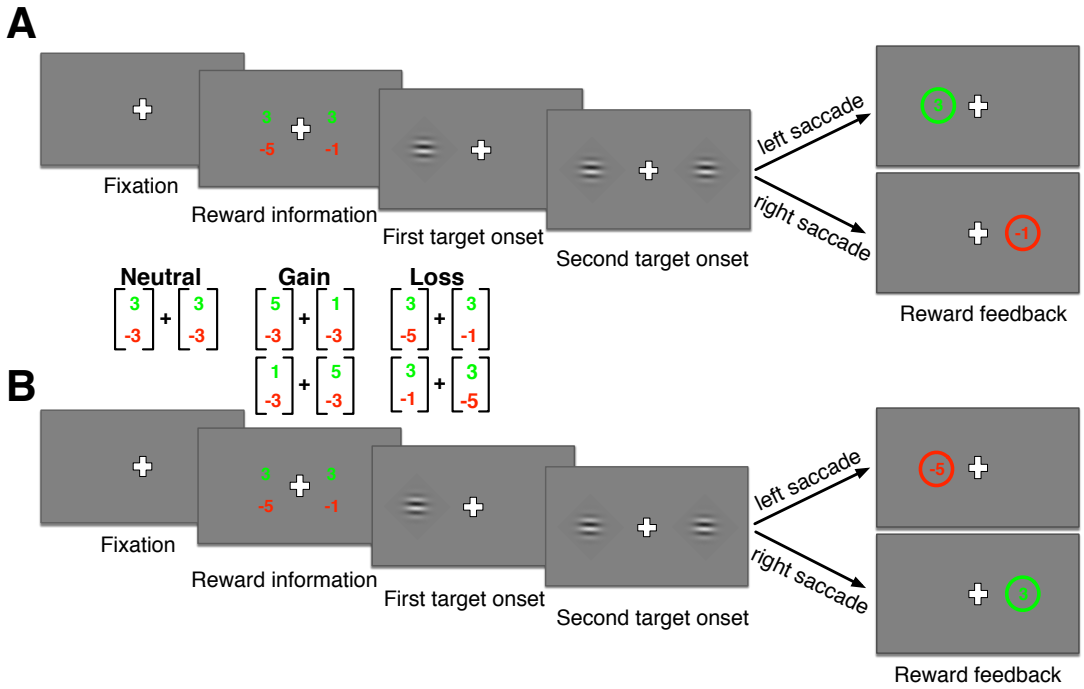

Figure 1. Schematic of the experimental paradigm. (**A**) Timeline of a trial during Experiment 1. Each trial began with a fixation cross followed by the presentation of reward points associated with a correct/incorrect response for each target on both sides of the fixation cross (in green for gain and red for loss). Information about expected reward following correct and incorrect response was provided on each trial according to the three reward manipulation conditions (neutral, gain, and loss). Following the presentation of the reward information, two identical targets (gabor patches) appeared on the screen asynchronously. The subjects' task was to report the first target that appeared on the screen by making a saccade to the target. Reward feedback was then given by a green circle for correct response or red circle for incorrect response around the selected choice with reward points in the center of the circle. (**B**) Timeline of a trial during Experiment 2. Stimulus and reward information presentation were similar to Experiment 1 except the correct response required making a saccade to the second target that appeared on the screen.

Reward conditions. There were three reward conditions in each experiment: neutral, gain, and loss. In the neutral condition, correct and incorrect responses on the two targets were associated with a gain of 3 points and loss of 3 points, respectively, expressed as (gain,loss) = (+3,-3). In the gain condition, reward points for saccade to one target were (+5,-3) whereas the other target had reward points of (+1,-3). In the loss condition, the targets were associated with reward points of (+3,-1) and (+3,-5). These values were selected based on our pilot study and to ensure similar differences in gain and loss values assuming an average loss-aversion factor (equal to 2). During the gain and loss conditions, the side associated with the target with higher expected reward was randomly assigned on each trial to minimize possible side bias. The order of gain and loss conditions was randomized across subjects. Each reward condition consists of 180 trials and was performed in a single session of the experiment without pause (lasting about 15-20 minutes). Therefore, each session corresponded to one of the three reward conditions.

Data Analysis. Overall, 29 subjects completed a total of 200 sessions of the experiment; 8 subjects completed either Experiment 1 or 2, and 21 subjects completed both Experiments 1 and 2. Outliers were determined on a session-by-session basis and defined as sessions in which the level of sensitivity to the sensory information (see Eq. 1) deviated from the overall mean of all sessions by more than three times the standard deviation, indicating that subjects did not properly attend to the main task parameter (TOA) in those sessions. This resulted in removal of 20 sessions. Data from the remaining 180 sessions were used for the main analyses reported here. Importantly, our conclusion did not depend on the exact exclusion criteria because we obtained qualitatively similar result using less conservative criteria.

Fitting choice behavior and estimated parameters. We fit choice data from each session (reward condition) of the experiment separately to estimate different aspects of target selection. In Experiment 1, we defined the psychometric function as the probability of choosing the left target as a function of the TOA favoring the left target, TOA = $t_{\text{right}} - t_{\text{left}}$, where t_{left} and t_{right} are the onset time of the left and right targets, respectively. In Experiment 2, the psychometric function was defined as the probability of choosing the left target as a function of the TOA favoring the right target (by flipping the sign for the TOA on each trial) in order to measure performance as a function of the proper signal. We then used the standard maximum likelihood

estimation method to estimate the best model parameters by minimizing the negative log-likelihood to fit the psychometric function in each session using the following equation:

$$p_L(\text{TOA}) = \frac{1}{1 + e^{-(\beta_{\text{side}} + \beta_s \times \text{TOA} + c \times \beta_r)}} \qquad \text{(Eq. 1)}$$

where β_{side} indicates the subject's preference for saccade to target on the left side of the fixation cross regardless of what was presented, β_s represents sensitivity to sensory information (i.e. how selection changes as a function of the TOA), β_r denotes reward bias (i.e. preference for saccade to the target with the larger expected reward). Finally, c is a dummy variable to indicate the more rewarding side (i.e., it is 1 if the left target is more rewarding, 0 if both targets are equally rewarding, and -1 if the right target is more rewarding) on each trial.

Optimality analysis. The optimal reward bias is defined as the amount of shift in target selection that maximizes the total reward earned in a given session. To determine the optimal reward bias, we first calculated the expected amount of reward earned assuming a given level of sensitivity to sensory information and loss aversion. The subject's sensitivity can predict the overall number of correct choices and thus, the number of reward points they can earn. Loss aversion causes individuals to respond more strongly to the loss than gain points associated with a target. Considering these factors, the expected payoff associated with a target can be calculated by the following equations:

$$\text{payoff}_A\,(p,\gamma) = (p(\beta,\text{TOA}) * \text{gain}_A) + (\gamma(1 - p(\beta,\text{TOA})) * \text{loss}_A)$$

$$\text{payoff}_B\,(p,\gamma) = ((1 - p(\beta,\text{TOA})) * \text{gain}_B) + (\gamma p(\beta,\text{TOA}) * \text{loss}_B) \qquad \text{(Eq. 2)}$$

where γ represents the loss-aversion factor and $p(\beta, TOA)$ is equal to the probability that the choice is the correct target for a given TOA:

$$p(\beta,\text{TOA}) = \frac{1}{1 + e^{-(\beta \times \text{TOA})}} \quad \text{(Eq. 3)}$$

where β represents sensitivity to sensory information. The total expected reward points for a given value of shift in target selection (due to biased reward information) can be calculated using the following equation:

$$\text{payoff}\,(\mu,\beta,p,\gamma) = P_{\text{cho}}(\text{TOA},\mu,\beta) * \text{payoff}_A + \left(1 - P_{\text{cho}}(\text{TOA},\mu,\beta)\right) * \text{payoff}_B \quad \text{(Eq. 4)}$$

where μ is the amount of shift or reward bias and $P_{\text{cho}}(\text{TOA}, \mu, \beta)$ represents the probability of choice for a given TOA, μ, and β computed as follows:

$$P_{\text{cho}}(\text{TOA}, \mu, \beta) = \frac{1}{1 + e^{-(\beta \times \text{TOA} + \mu)}} \quad \text{(Eq. 5)}$$

As a result, the total expected amount of reward earned for all values of p for given values of μ and β is equal to

$$R(\mu, \beta, \gamma) = \int_0^1 \text{payoff}(\mu, \beta, p, \gamma)\, dp. \quad \text{(Eq. 6)}$$

where the integral is computed by binning and summating over all values of p. Finally, the optimal reward bias can be computed by finding the maximum value of R for specific values of sensitivity, reward bias, and risk-aversion factor using Eq. 6.

Stepwise regression. In order to determine the model that best explains variances in the saccadic reaction time, we used a stepwise regression model. In the first step, we included the reward condition (neutral, gain, loss), TOA, trial's accuracy (correct vs. incorrect response), and a dummy variable indicating whether the chosen target had assigned a higher or lower expected reward outcomes. In the second step, we included interaction terms and tested whether they significantly increased adjusted R-squared of the regression model.

Computational model. The basic model consisted of two cortical columns with two pools of excitatory neurons and one inhibitory pool of neurons in the superficial layer and two pools of excitatory neurons in the deep layer (Soltani, Noudoost, & Moore, 2013) (**Fig. 7A**). All neural pools in all the layers received a background input mimicking input from adjacent cortical neurons with different types of selectivity. The excitatory pools in the superficial layer also received visual input related to the presentation of targets on the screen. Specifically, the visual input to the two pools were similar except that they had different onset timing according to the TOA on each trial. Moreover, the two pools of excitatory neurons in the superficial layer were mutually inhibited using a shared pool of inhibitory interneurons. This mutual inhibition created a winner-take-all competition and caused activity in excitatory pools to diverge in both the superficial and deep layers. We used a mean-field approximation of a spiking network model of decision making to simulate the superficial layer (Wong & Wang, 2006). Each excitatory pool of neurons in the deep layer had weak self-excitatory recurrent connections. The deep layer then

projected its output to the brain stem or superior colliculus to direct a saccadic eye movement. We determined the choice of the network on each trial by identifying the first deep-layer excitatory pool whose activity passed 15 Hz (considered the winner pool). The full details of the basic model are described elsewhere (Soltani et al., 2013).

To simulate the observed effects of unequal reward information, we considered three alternative mechanisms. These mechanisms influenced different parts of the model, mimicking either modulations of sensory processing or later stages of decision-making processes. First, to simulate the effect of reward information on later stages of decision-making processes, we included a reward-based input to the excitatory pools in the deep layer (**Fig. 7B**; Mechanism 1). This input was independent of the amount and timing of the visual input (TOA) on a given trial. Second, to simulate the effect of reward information on sensory processing, we assumed two alternative mechanisms in which reward information could modulate the visual input to the decision circuit. In the first mechanism, the input evoked by targets with a larger and smaller expected reward was multiplied by $(1 + \lambda)$ and $(1 - \lambda)$, respectively, where λ is a constant that measures the modulation of visual input by reward information (**Fig. 7C**; Mechanism 2). This mechanism results in stronger input for the more rewarding target. In the second mechanism for modulating sensory processing, the input for target with larger expected reward was modulated through a shift in TOA in favor of the target with larger expected reward (**Fig. 7C**; Mechanism 3). This mechanism mimics faster processing of input related to the more rewarding target. In both Mechanisms 2 and 3, the reward cue presented to subjects modulates the processing of sensory information and thus, could result in perceptual bias. Finally, in order to simulate the saccade-second task (select the second target that appears on the screen) we assumed that the projections of the output in excitatory pools of the superficial layer is switched via a gating mechanism (green dashed lines in **Fig. 7B, C**) to allow the selection of the non-winner pool as the response.

Results

Influence of reward information on perceptual decision making. To study the influence of reward information on temporal judgment as one type of perceptual decision making, we used modified versions of the paired-target task. Two identical targets (gabor patches) appeared on each side of a computer screen with varying TOA. In Experiment 1 (Experiment 2), subjects reported what they perceived to be the first (second) of the two targets to appear on the screen by

making a saccade to the target followed by reward feedback. The amounts of reward points expected to be earned/lost upon correct/incorrect response for each target were presented on each side of the fixation cross at the beginning of each trial and were manipulated in three separate reward conditions (**Fig. 1**; see Materials and Methods).

To quantify the sensitivity to sensory information and possible shifts in choice due to unequal expected reward, we calculated the probability of choosing the left target as a function of the TOA, which we refer to as the psychometric function. We then fit individuals' psychometric function using a sigmoid function to estimate three parameters: sensitivity (β_s), side bias (β_{side}), and reward bias (β_r) (see Materials and Methods for more details). These parameters capture three features of target selection in our experiment: sensitivity measures the fidelity of target selection to sensory evidence (i.e. TOA); side bias measures an overall bias in choosing the left or right target independently of sensory evidence; and reward bias measures the shift in target selection toward the better target (i.e. target associated with a larger gain or smaller loss; **Fig. 2A-C**).

For both experiments, we investigated target selection in each of the three reward conditions (neutral, gain, and loss). We found that subjects were sensitive to the TOA in both experiments. The average values (±std) of sensitivity in Experiment 1 were equal to: 0.0432±0.0191, 0.0472±0.0222, and 0.0438±0.0179 (ms^{-1}) for the neutral, gain, and loss conditions, respectively. Notably, these average values were larger than zero (two-sided signed-rank test; neutral: $p = 5.18\times10^{-9}, d = 2.25$; gain: $p = 2.7\times10^{-5}, d = 2.29$; loss: $p = 2.7\times10^{-5}, d = 2.47$; **Fig. 2D**). Similarly, the average values of sensitivity in Experiment 2 were equal to: 0.0389±0.0195, 0.0432±0.0187, and 0.0378±0.0259 for the neutral, gain, and loss conditions, respectively. Again, these average values were larger than zero (two-sided signed-rank test; neutral: $p = 7.6\times10^{-9}, d = 2.08$; gain: $p = 2.7\times10^{-5}, d = 2.34$; loss: $p = 4\times10^{-5}, d = 1.76$; **Fig. 2G**).

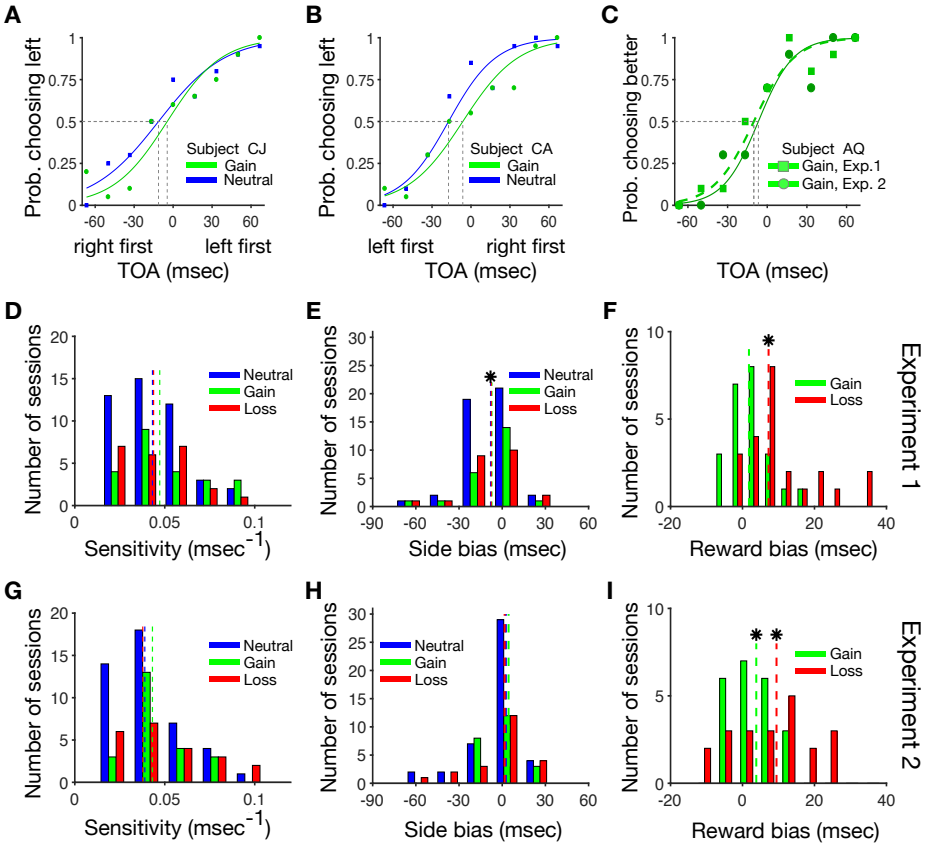

Figure 2. Influences of unequal expected reward on target selection. (**A-C**) Example psychometric functions in different conditions of Experiments 1 and 2. (**A**) Plotted is the probability of choosing the left target as a function of the TOA for an example subject in the gain and neutral conditions of Experiment 1. This subject exhibited slightly larger level of sensitivity in the gain compared to the neutral condition. Positive (negative) values of TOA correspond to the left (right) target to appear first. (**B**) The same as in (A) but for another subject who exhibited larger side bias in the neutral than gain conditions of Experiment 2. In these plots, side bias is equal to the TOA at which the probability of choosing the two targets are equal. In Experiment 2, negative (positive) values of TOA correspond to the left (right) target to appear first. (**C**) Psychometric functions of an example subject in the gain condition of Experiments 1 and 2. Plotted is the probability of choosing the better target as a function of the TOA for that target. For Experiment 1, negative (positive) values of TOA means the target with smaller (larger) expected reward has appeared first. In Experiment 2, negative (positive) values of TOA means the target with smaller (larger) expected reward has appeared second. In these plots, reward bias can be interpreted as the TOA at which the subject chooses the two targets equally. This subject exhibited similar reward bias in the two experiments. (**D-F**) Histograms plot the number of subjects with given values of sensitivity (**D**), side bias (**E**), and reward bias (**F**) under different reward conditions during Experiment 1. The dashed lines show the medians and each star indicates a significant difference from 0 (two-sided signed-rank test, $p < 0.05$). (**G-I**) The same as in D-F but for Experiment 2.

We also examined the side bias estimated from subjects' choice behavior. The side associated with the target with better reward outcomes had been randomly assigned on each trial to minimize such bias. In Experiment 1, we did not observe any significant side bias except in the neutral condition (two-sided signed-rank test; neutral: -8.11±17.21 ms, $p = 0.002$, $d = 0.47$; gain: -7.14 ±16.60 ms, $p = 0.083$, $d = 0.40$; loss: -7.60±22.72 ms, $p = 0.083$, $d = 0.36$; **Fig. 2E**). Furthermore, we did not find any significant side bias in any conditions of Experiment 2 (two-sided signed-rank test; neutral: 2.54±19.77 ms, $p = 0.55$, $d = 0.12$; gain: 4.41±14.63 ms, $p = 0.72$, $d = 0.06$; loss: 1.69±22.20 ms, $p = 0.88$, $d = 0.13$; **Fig. 2H**). Together, these results illustrate that subjects exhibited very small side bias in both experiments.

We then measured reward bias or the shift in target selection due to unequal expected reward. We found that during the gain and loss conditions of Experiment 1, subjects exhibited a significant positive bias toward the better target only in the loss condition (two-sided signed-rank test; gain: 1.79±5.92 ms, $p = 0.128$, $d = 0.36$; loss: 7.25±10.56 ms, $p = 0.0001$, $d = 1.0$; **Fig. 2F**). Furthermore, this shift in target selection was stronger during the loss than gain condition (two-sided Wilcoxon ranksum test; $\Delta = 5.45$ ms, $p = 0.002$, $d = 0.7$). Similarly, values of reward bias in the gain and loss conditions were significantly larger than zero in Experiment 2 (two-sided signed-rank test; gain: 3.82±10.40 ms, $p = 0.041$, $d = 0.43$; loss: 9.45±14.00, ms, $p = 0.006$, $d = 0.67$; **Fig. 2I**). We also found a trend for the shift in target selection to be larger in the loss than gain condition (two-sided Wilcoxon ranksum test; $\Delta = 5.62$, $p = 0.131$, $d = 0.67$). Nevertheless, there was no significant correlation between reward bias in the gain and loss conditions for subjects who performed both experiments successfully (i.e., validly completed all sessions) (Pearson correlation; Experiment 1: $r = -0.24, p = 0.32$, $N = 20$; Experiment 2: $r = 0.09, p = 0.72$, $N = 17$). Combining data from both experiments, we observed an overall larger bias in the loss than gain condition (two-sided signed-rank test; $\Delta = 6.00$±12.73, $p = 0.0085$, $d = 0.48$, $N = 37$; **Fig. 3** and **Fig. 3-1**). This result provides evidence for loss aversion in perceptual decision making with possible gains and losses.

The observed shift in target selection due to unequal expected reward could occur due to two different, but not exclusive, mechanisms: 1) differential enhancement of sensory processing, which can result in perceptual bias; and 2) increasing the tendency to make a saccade to the better target (response bias). In Experiment 1, a bias in sensory information processing would

result in a shift in target selection in the opposite direction of that in Experiment 2. In contrast, a response bias would result in similar shifts for the two experiments. To determine which mechanism best matches subjects' behavior, we then calculated the difference in reward bias observed in Experiments 1 and 2, separately for the gain and loss conditions. We found that for both gain and loss conditions, the differences in reward bias between the two experiments were not significantly different from zero (two-sided signed-rank test; gain: Δ = -1.27±13.20 ms, $p = 0.29$, $d = 0.26$; loss: Δ = 2.24±18.79 ms, $p = 0.74$, $d = 0.02$). This result indicates that the observed shift in target selection was most compatible with response bias.

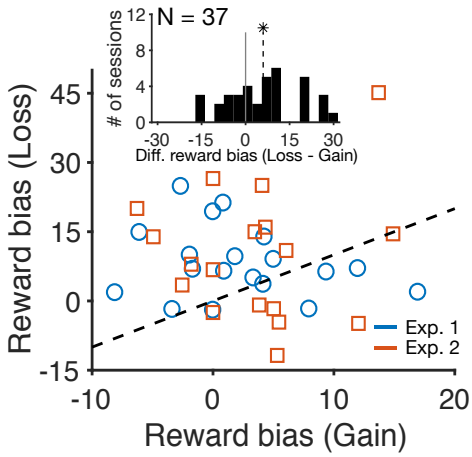

Figure 3. Larger shift in target selection during the loss than gain condition. Plotted is individuals' reward bias during the loss condition as a function of reward bias during the gain condition for Experiments 1 and 2. Here, we only included subjects who performed both experiments successfully (valid sessions). The inset plots the distribution of the difference between reward bias in the loss and gain conditions over both experiments. The dashed line is the median and the star indicates a significant difference from 0 (two-sided signed-rank test, $p < 0.05$).

Suboptimal shift in target selection. To test whether the shift in target selection due to unequal expected reward was optimal or not, we compared the observed shift with the predicted optimal shift based on individuals' sensitivity to sensory evidence (β_s) in Experiments 1 and 2. The optimal amount of shift in target selection (reward bias) should depend on individual's sensitivity to the TOA because for example, it is more advantageous for a subject who can discriminate the onset of the two targets well to mostly rely on sensory information, which in turn would result in a small reward bias. In contrast, a subject with poor discrimination could select the more rewarding target independently of the TOA to obtain more reward.

First, we calculated the optimal value of reward bias for given values of sensitivity and risk aversion separately in the gain and loss conditions (**Fig. 4A, D**). We found that in the gain condition, the optimal reward bias should decrease as loss aversion increases for a given level of sensitivity (**Fig. 4B**). Moreover, the optimal bias should decrease with larger sensitivity (**Fig. 4C**). In contrast, in the loss condition, the optimal reward bias should increase with more loss aversion (**Fig. 4E**) but decrease with larger sensitivity. These results illustrate that loss aversion should have opposite effects on reward bias in the gain and loss conditions and moreover, optimality requires that reward bias to be inversely correlated with the individual's' level of sensitivity to sensory evidence.

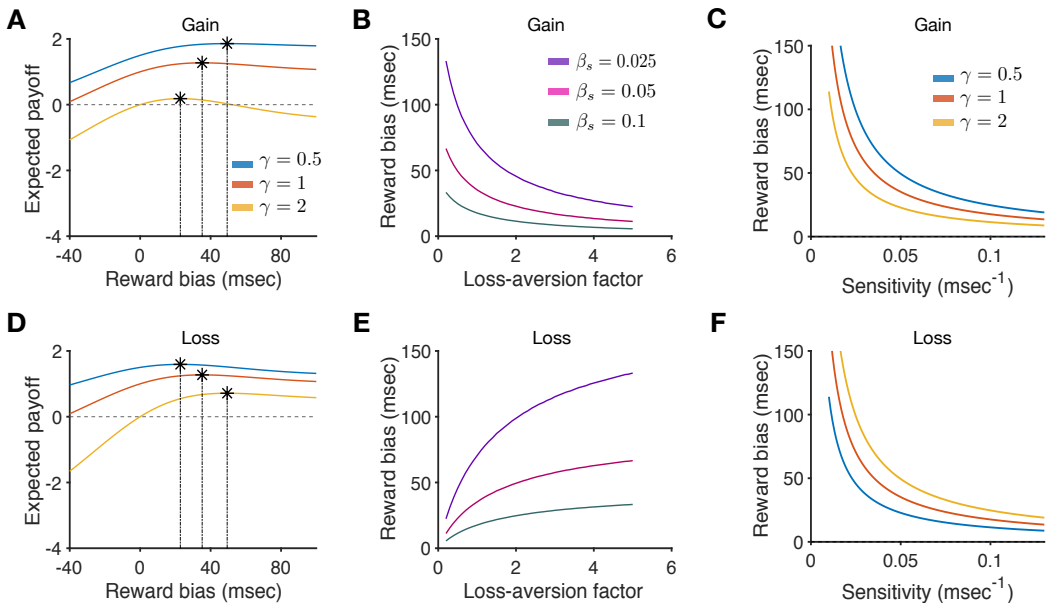

Figure 4. The optimal shift in target selection due to unequal expected reward in Experiments 1 and 2. (**A**) The expected payoff for a given level of sensitivity ($\beta_s = 0.05$ ms^{-1}) and reward bias during the gain condition. Each curve represents the expected payoff for a different level of risk aversion (risk aversion coefficient γ). The peak in each curve indicates the optimal reward bias. (**B**) The optimal reward bias as a function of the risk-aversion factor in the gain condition separately for three different values of the sensitivity. The optimal reward bias decreases as loss aversion increases in the gain condition. (**C**) The optimal reward bias as a function of the sensitivity to sensory evidence. The optimal reward bias diminishes as the sensitivity increases. (**D-F**) The same as in A-C but for the loss condition. The optimal reward bias increases as loss aversion increases in the loss condition. Nevertheless, the optimal reward bias decreases with larger values of the sensitivity similar to the gain condition.

We next compared the optimal and observed reward bias for different values of loss aversion in order to detect deviation from optimality. We used this approach because we did not have access to the degree of loss aversion in individual subjects. We found that in the gain condition of both experiments, subjects exhibited reward biases that were smaller than the predicted optimal biases based on loss neutrality (loss-aversion factor equal to 1) (two-sided signed-rank test; Δ = -39.58 ± 3.64 ms, $p = 3.5\times10^{-9}$, $d = 1.76$; **Fig. 5A-C**). This result is more compatible with very large loss aversion because the larger loss aversion gives rise to the smaller shift in the gain condition. At the same time, the observed reward biases in the loss condition were also smaller than the optimal values based on loss neutrality (two-sided signed-rank test; Δ = -33.29 ± 4.29 ms, $p = 6.3\times10^{-9}$, $d = 1.41$; **Fig. 5D-F**). Because the optimal shift increases with more loss aversion in the loss condition, the small shifts we observed point to loss-seeking as opposed to risk-aversive behavior as seen in the gain condition. These opposite patterns of results for the gain and loss conditions illustrate that observed values of reward bias were overall too small and thus, suboptimal. In addition, we did not find a significant correlation between reward bias and the sensitivity to sensory evidence as expected from any optimal shift in target selection (Pearson correlation; gain condition: Experiment 1: $r = -0.13, p = 0.55$; Experiment 2: $r = -0.34, p = 0.11$; loss condition: Experiment 1: $r = -0.17, p = 0.44$; Experiment 2: $r = -0.1, p = 0.64$). Together, these results illustrate that the amount of shift in target selection was determined heuristically and suboptimally.

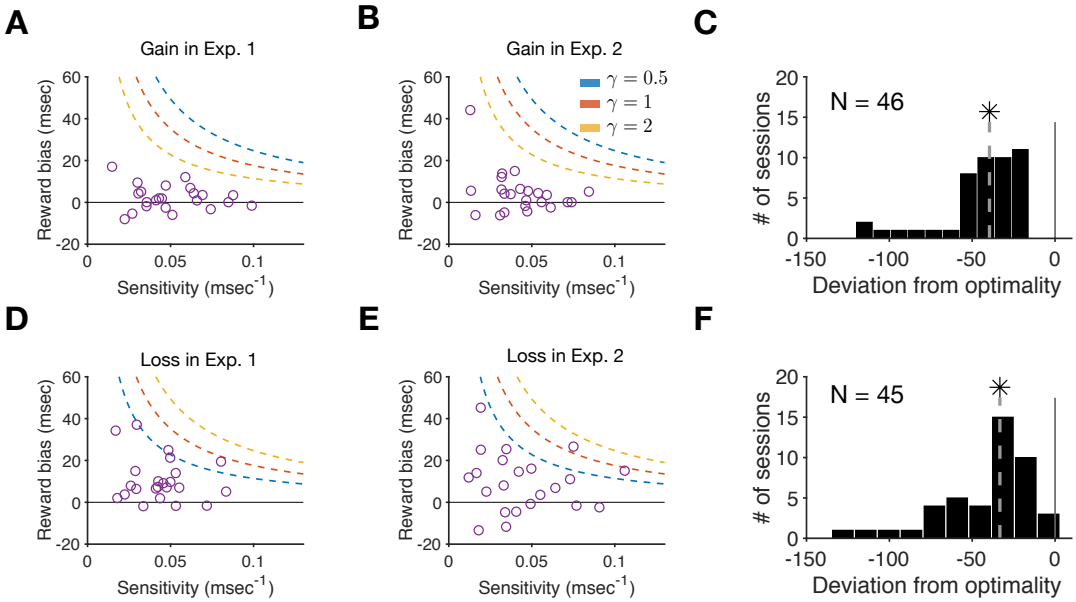

Figure 5. The amount of shift in target selection was suboptimal in the gain and loss conditions in both experiments. (**A**) Plot shows individuals' reward bias as a function of their sensitivity to the TOA. The blue, red, and yellow lines represent the optimal value of reward bias for different values of risk-aversion factor. (**B**) The same as in A but for Experiment 2. (**C**) Plotted is the distribution of the differences between observed and predicted optimal reward biases based on loss neutrality ($\gamma = 1$) across all subjects in the gain condition of Experiments 1 and 2. (**D-F**) The same as in A-C but for the loss condition in the two experiments. Note that in the loss condition a larger loss-aversion factor predicts larger reward bias (the opposite is true for the gain condition).

Saccadic reaction time reflects the effect of task parameters. In our experiments, the subjects were not instructed to saccade as quick as possible and thus, the tasks were not reaction time tasks. Nonetheless, we analyzed saccadic reaction time (SRT) to examine whether this measure reflects any task parameters. First, we observed that SRT increased as the absolute value of the TOA, reflecting the difficulty of temporal judgment, was increased (stepwise GLM; $F(5,32394) = 452, p = 10^{-273}$, adjusted $R^2 = 0.065$; TOA: $\beta = -1.11, p = 0.04$; **Fig. 6A**). Second, we found that SRT was on average smaller in the neutral condition compared to either gain or loss conditions ($\beta = -0.11, p = 1.03\times10^{-26}$; **Fig. 6A**). Third, we observed that SRT was significantly smaller on trials in which the response was correct compared to trials with incorrect response ($\beta = -0.076, p = 0.0005$; **Fig. 6B**). Finally, SRT was not significantly different on trials in which the choice was the target with higher expected reward compared to lower expected reward (saccade to better or worse targets; GLM; $F(4,316375) = 259, p = 2.18\times10^{-216}$, adjusted $R^2 = 0.059$; Expected reward: $\beta = -0.009, p = 0.19$), perhaps due to different effects of better and worse targets in the gain and loss conditions (**Fig. 6C**). The less pronounced effect of reward information on SRT could indicate different heuristics used to

process gain and loss information. Overall, these results show that saccadic reaction time was sensitive to the TOA and reward information, indicating that both types of information influenced perceptual choice.

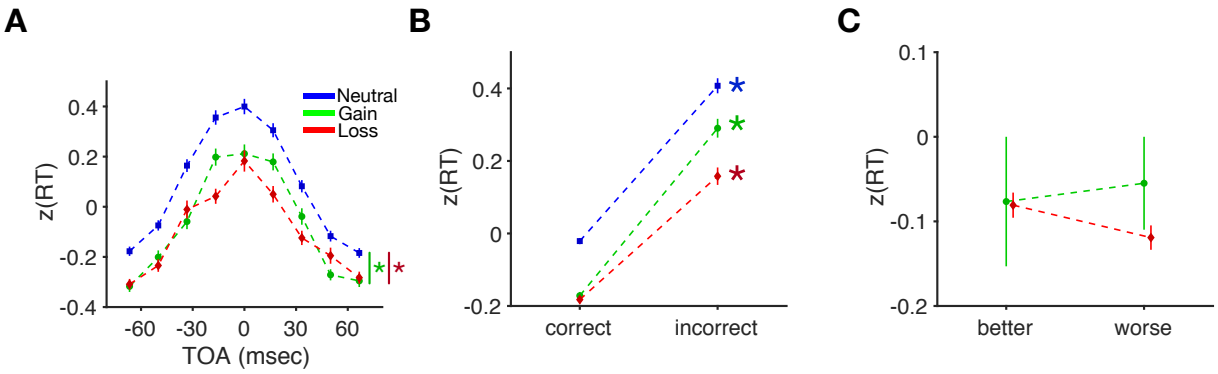

Figure 6. Saccadic reaction time was sensitive to TOA, trial's accuracy, and reward information. **(A)** Plotted are the z-scored saccadic reaction time (SRT) as a function of the TOA for neutral (blue squares), gain (green circles), and loss (red diamonds) conditions. An asterisk shows a significant difference between the average SRT in the neutral and gain or loss conditions. **(B)** Plot shows the average z-scored SRT on correct and incorrect trials. An asterisk shows a significant difference between the average SRT on correct and incorrect trials. (C) Plotted is the average SRT on trials in which the selected target was the target with higher or lower expected reward corresponding to the better and worse targets, respectively.

Plausible neural mechanisms for observed shifts in target selection. To reveal plausible neural mechanisms underlying the shifts in target selection, we simulated our experimental observations using a cortical network model that we have previously used to successfully simulate the paired-target task (Soltani et al., 2013). Specifically, we focused on capturing our two main experimental findings: 1) similar shifts in target selection during Experiments 1 and 2; and 2) non-significant but negative correlation between the shift in target selection and individuals' sensitivity to sensory evidence in both experiments.

The model consisted of two neural columns with two pools of excitatory neurons and one inhibitory pool of neurons in the superficial layer, and two pools of excitatory neurons in the deep layer. The two pools of excitatory neurons in the superficial layer were mutually inhibited using a shared pool of inhibitory interneurons. This mutual inhibition resulted in winner-take-all competition between the excitatory neurons in the superficial layer that subsequently drive the two excitatory pools in the deep layer. The deep layer then projects its output to the brain stem or

superior colliculus to direct a saccadic eye movement (**Fig. 7A**; see Materials and Methods for more details).

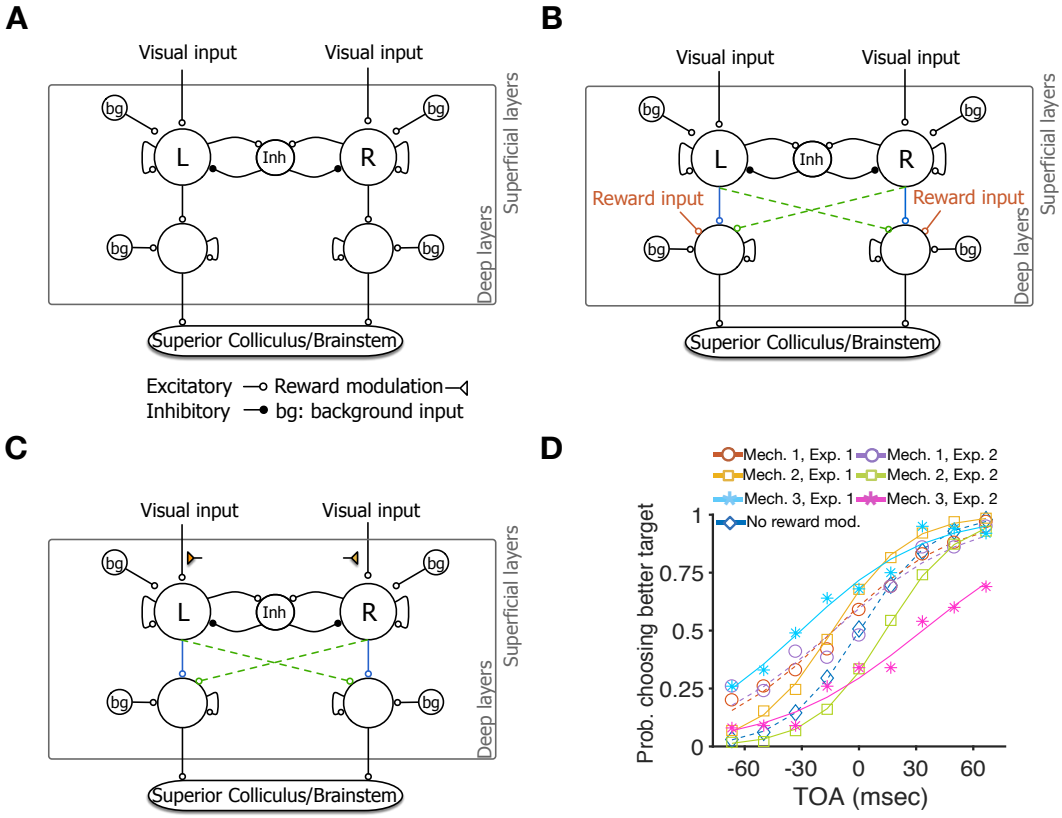

Figure 7. Alternative mechanisms for simulating the effects of unequal expected reward on perceptual decision making. (**A**) The basic architecture of the original network model used to simulate the paired-target task. (**B**) The extended model with independent reward input (Mechanism 1). In this model, excitatory pools in deep layer receive an additional reward-based input that was independent of the visual input. The blue solid (green dashed) lines show the output projection of excitatory pools of superficial layer to deep layer excitatory pools in Experiment 1 (Experiment 2). (**C**) The extended model with reward-dependent modulation of visual input. In this model, the visual input to the excitatory pools of superficial layer is modulated by reward information provided at the beginning of the trial. This modulation was performed via two different mechanisms. In the first mechanism, reward information strengthens (weakens) the visual input for the target with higher (lower) expected reward (Mechanism 2). In the second mechanism, unequal reward information results in faster processing (i.e. earlier onset) of the target with higher expected reward (Mechanism 3). (**D**) Probability of choosing the more rewarding target as a function of TOA (-TOA) in Experiment 1 (Experiment 2) using different mechanisms. The results with no reward modulation (diamond) are shown as the control. The model with independent reward input (Mechanism 1; circles) results in similar shift in target selection in the two experiments, whereas the model with reward-dependent modulation of visual input (Mechanisms 2 and 3; squares and asterisks, respectively) results in shifts that are opposite those observed in the two experiments.

To simulate the effect of reward information, we considered three alternative mechanisms that could influence different parts of the model to mimic different stages of decision-making processes (see Materials and Methods for more details). First, to simulate the effect of reward information on later stages of decision-making processes, we included a reward-based input to the excitatory pools in the deep layer (**Fig. 7B**; Mechanism 1). Importantly, this input was independent of the visual input on a given trial. Second, to simulate the effect of reward information on sensory processing, we assumed two different mechanisms. In the first mechanism, we assumed that reward cues could result in gain modulation of sensory input, giving rise to stronger input for the more rewarding target (**Fig. 7C**; Mechanism 2). In the second mechanism, we assumed that reward cues could cause a shift of TOA, advancing the onset of the target with higher expected reward (**Fig. 7C**; Mechanism 3).

Simulation results showed that that unequal expected reward results in significant shifts in target selection using all three mechanisms (**Fig. 7D**). These shifts, however, were similar for Experiments 1 and 2 only in the model based on Mechanism 1. We also examined the correlation between the shift in target selection and the sensitivity to visual input for a given set of model parameters. To generate target selection with different levels of sensitivity to sensory evidence, we changed the background noise in the input to the superficial layer. We found that the shifts in target selection and sensitivity to the TOA were negatively correlated in the model based on Mechanism 1, but these correlations were weak and not statistically significant (Pearson correlation; Experiment 1: $r = -0.14, p = 0.15$; Experiment 2: $r = -0.19, p = 0.054$; **Fig. 8A, D**). In contrast, in the model based on Mechanism 2, the shifts in target selection were negatively and positively correlated with sensitivity to the TOA in Experiments 1 and 2, respectively (Pearson correlation; Experiment 1: $r = -0.58, p = 1.75\times10^{-10}$; Experiment 2: $r = 0.56, p = 1.46\times10^{-9}$; **Fig. 8B, E**). We found qualitatively similar results for the model based on Mechanism 3, even though the shifts in target selection were not significantly correlated to sensitivity to the TOA (Pearson correlation; Experiment 1: $r = -0.09, p = 0.33$; Experiment 2: $r = 0.08, p = 0.4$; **Fig. 8C, F**). Overall, these results illustrate that the model based on Mechanism 1 is more compatible with our experimental data for two reasons: 1) exhibiting equal shifts in target selection in Experiments 1 and 2; and 2) non-significant but negative correlation between the shifts in target selection and sensitivity to the TOA in both experiments. These modeling results support the conclusion that the observed shift in subjects' behavior due to the

reward information is caused by changes in later decision-making processes and provide a plausible mechanism for how reward information influences perceptual choice.

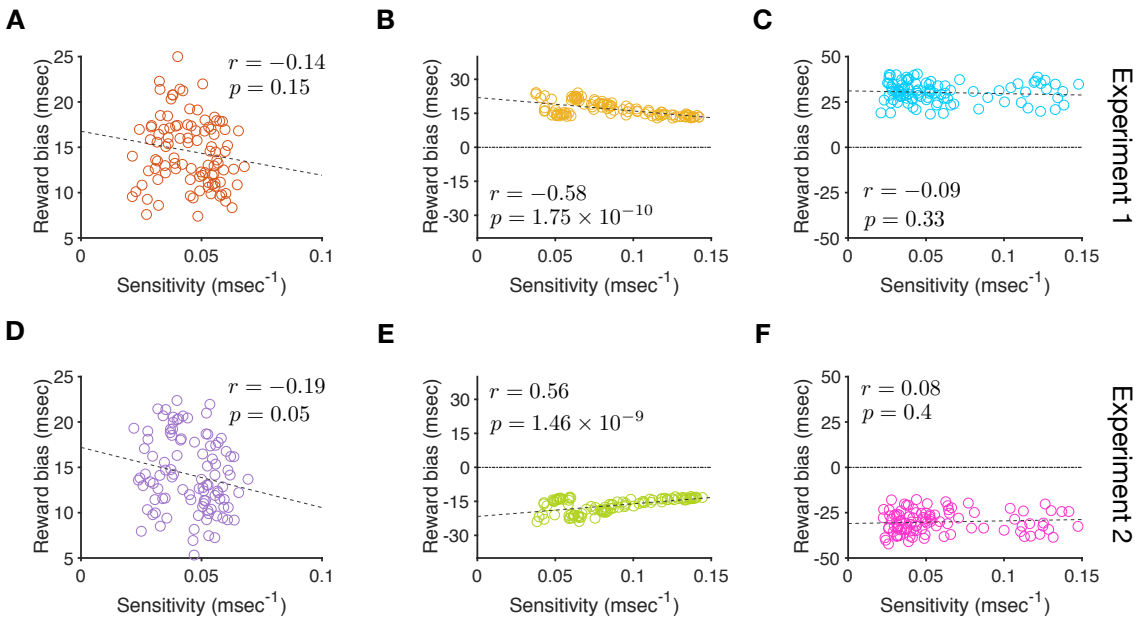

Figure 8. Shift in target selection due to unequal reward input was more compatible with Mechanism 1. (**A**) The shift in target selection is plotted as a function of the sensitivity to the TOA for Experiment 1 using the model based on Mechanism 1. Each point shows the replica of a subject. The grey dashed line shows the least squares line fitted on the simulated behavioral data. (**B–C**) The same as in A but using the model based on Mechanisms 2 and 3, respectively. (**D-F**) The same as in A-C but for Experiment 2. Only in the model based on Mechanism 1, shifts in target selection due to unequal reward input was negatively correlated with sensitivity to the TOA in both experiments.

Discussion

Series of studies in past decades have aimed to reveal neural mechanisms by which reward influences perceptual decision making. These studies have argued that unequal reward outcomes could either cause a response bias (i.e., increase the tendency to choose the target with larger expected reward) and/or result in a perceptual bias (i.e., differential processing of sensory information). To test these two alternative (but not necessarily exclusive) hypotheses for the influence of reward on perceptual decision making, in two sets of experiments, we asked subjects to saccade to the first or second target that appeared on the screen while we manipulated the amount of reward expected from the two alternative responses. Importantly, a bias in sensory

processing (perceptual bias) would result in opposite shifts in target selection in the two experiments, whereas a response bias would cause similar shifts. We found the same amounts of shift in the two experiments indicating that expected reward did not bias sensory processing and instead, resulted in response bias. These findings dovetail with results from recent studies that used modeling to determine the underlying process of perceptual decision making with unequal reward outcomes (Diederich, 2008; Diederich & Busemeyer, 2006; Gao et al., 2011).

Nonetheless, others have argued that reward can directly influence the processing of sensory information during perceptual choice (Cicmil et al., 2015; Liston & Stone, 2008; Pleger et al., 2008; Voss et al., 2008). A possible reason for the discrepancy between findings in these and our studies could be the difference in experimental paradigms. More specifically, here we used a temporal judgment task, which is a type of time-dependent perceptual decision making, and it is possible that reward exerts its influence differently during time-independent perceptual choice. However, there are studies (e.g., Diederich & Busemeyer, 2006) showing the effects of reward as response bias even in discrimination between the lengths of two lines (i.e., time-independent tasks). Future studies are required (using our approach) in other to test the generalizability of our findings to other types of perceptual decision making.

In addition to reward effects on perception, there is an extensive literature on whether attention influences perception by accelerating sensory processing (the 'prior entry' hypothesis) or inducing decision biases. Some of these studies have argued that attention enhances the speed of sensory processing (Hikosaka, Miyauchi, & Shimojo, 1993; Stelmach & Herdman, 1991) whereas other studies have maintained that observed effects are primarily due to attentional modifications of the decision mechanisms (Schneider & Bavelier, 2003). Most of these studies used the so-called temporal order judgment (TOJ) task to measure the point of subjective simultaneity (PSS) from attentional cueing. For example, in a study design similar to ours but concerned with measuring the effects of attention, Shore and colleagues (Shore, Spence, & Klein, 2001) asked subjects to report the first or second targets that appeared on the screen in order to separate changes in sensory processing from response biases. By comparing the shifts in psychometric functions in the two tasks, they found that attention mainly influences perception by accelerating sensory processing in addition to second-order response biases. In contrast, Schneider and Bavelier (Schneider & Bavelier, 2003) have argued that the shift in the PSS due to

attentional cueing in the TOJ task is not an adequate reason to accept the prior entry hypothesis. Instead, they suggest that one should compare shifts in the PSS in the TOJ task with those of in a simultaneity judgment task, in which the subjects report whether two stimuli appeared simultaneously or successively. By making this comparison, they showed that attentional cueing has little influence on accelerating sensory processing.

Here, rather than explicit attentional cueing, we used unequal reward information to bias subjects' processing of information and/or decision making. Importantly, this reward information could enhance the processing of the side with higher expected reward especially on harder trials (trials with zero or small TOA). We also provided reward feedback (correct or incorrect judgment) on each trial, which allowed subjects to correct their biases if desired so. The fact that we observed almost opposite results to those by Shore and colleagues (Shore et al., 2001) based on attentional cueing indicates that reward information influences perception very differently than how attention affects perception, and therefore, reward and attentional processes rely on different neural mechanisms to guide behavior. Nonetheless, attention has been shown to closely interact with processing reward information (Farashahi, Azab, Hayden, & Soltani, 2018; Serences, 2008; Soltani, Khorsand, Guo, Farashahi, & Liu, 2016; Spitmaan, Chu, & Soltani, 2018; Stănişor et al., 2013) and understanding their relationship is crucial for understanding both processes (Maunsell, 2004).

To reveal plausible neural mechanisms underlying our observations, we extended our biophysically-plausible cortical network model (Soltani et al., 2013) to simulate shifts in target selection due to unequal expected reward based on alternative mechanisms. Due to the nonlinear nature of the proposed network models, it is unclear which alternative model would be more compatible with our main experimental results *per se* (similar shifts in target selection, and lack of correlation between shift and sensitivity to the TOA in the two experiments). We found that our experimental results are more compatible with the influence of reward information on later stages of decision-making processes via biasing the activity in the output layer of the decision circuit toward the more rewarding target. Considering that lesions and reversible inactivation of the frontal eye field (FEF) cause similar shifts in target selection during the paired-target task (Schiller & Chou, 1998, 2000; Schiller & Tehovnik, 2003), it is conceivable that reward exerts its influence through modulations of the output layer of the FEF.

Importantly, we found that to reproduce experimental results, the input for biasing target selection should be independent of sensory evidence, which is more consistent with results from a few recent studies (Diederich, 2008; Diederich & Busemeyer, 2006; Gao et al., 2011). For example, using a task in which the subjects had to judge whether two lines are the same or different while manipulating the payoff for the two responses, Diederich and Busemeyer (Diederich & Busemeyer, 2006) have shown that the effect of unequal payoffs is more compatible with a two-stage processing of sensory and reward information. In their model, the decision maker first integrates reward information followed by the integration of sensory information (with no reward modulation) if no decision is made during the first stage. In another study, Gao and colleagues (Gao et al., 2011) used a leaky competing-accumulator model to show that reward information biases the initial state of the decision variable toward the target with higher expected reward. Both these studies illustrate that reward information does not interact with sensory evidence. The fact that all of the explored mechanisms can generate comparable shifts in target selection in Experiment 1 suggests that in order to distinguish the origin of reward effects, one needs to consider the appropriate task design in addition to the appropriate model.

Our results indicate that subjects determined the amount of response bias heuristically since the amounts of shift in target selection due to unequal expected reward were significantly smaller than the optimal values and moreover, were not correlated with individuals' sensitivity to sensory evidence. In addition, we observed a larger shift in target selection in the loss compared to gain condition in both experiments. This result resembles loss-aversion behavior during value-based choice (Tversky & Kahneman, 1992) and extends this phenomenon to perceptual decision making with different reward outcomes. Furthermore, our analyses of saccadic reaction time revealed how gain and loss information influenced choice differently. Future experiments with emphasize on speed (i.e., reaction time tasks) could provide useful information to test alternative models.

Similar but much weaker suboptimal behavior also has been observed for biased reward probabilities during perceptual decision making (Navalpakkam, Koch, & Perona, 2009; Voss et al., 2008). Interestingly, humans exhibit a closer-to-optimal criterion when they deal with unequal reward probabilities rather than unequal reward magnitudes on alternative options/actions (Maddox, 2002; Teichert & Ferrera, 2010). Our modeling results indicate that

shifts in target selection can be closer to optimal if reward information affects the processing of visual input instead of the final stages of decision making. Therefore, the difference in response to unequal reward probability and magnitude could be due to their influence on different stages of decision making. Finally, in environments that resemble more naturalistic settings, adjustments in choice and learning can occur in the absence of any optimization (Farashahi et al., 2017; Khorsand & Soltani, 2017). Future studies are required to test whether reward probability and magnitude exert their influence at different stages of decision making.

Acknowledgments

We would like to thank Daeyeol Lee, Shiva Farashahi, and Mehran Spitmaan for help in early stages of this work, and Patrick Cavanagh and Shih-Wei Wu for comments on the manuscript.

Competing interests

Authors declare that there are no competing financial, professional, or personal interests that might have affected the research described in this manuscript.

References

Brainard, D. H. (1997). The Psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Carrasco, M., & Barbot, A. (2019). Spatial attention alters visual appearance. *Current Opinion in Psychology*, *29*, 56–64. Retrieved from https://doi.org/10.1016/j.copsyc.2018.10.010

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*(3), 308–313. https://doi.org/10.1038/nn1194

Christopoulos, V., Bonaiuto, J., & Andersen, R. A. (2015). A biologically plausible computational theory for value integration and action selection in decisions with competing alternatives. *PLoS Computational Biology*, *11*(3), 1–31. https://doi.org/10.1371/journal.pcbi.1004104

Christopoulos, V., & Schrater, P. R. (2015). Dynamic integration of value information into a common probability currency as a theory for flexible decision making. *PLoS Computational Biology*, *11*(9), 1–26. https://doi.org/10.1371/journal.pcbi.1004402

Cicmil, N., Cumming, B. G., Parker, A. J., & Krug, K. (2015). Reward modulates the effect of visual cortical microstimulation on perceptual decisions. *ELife*, *4*(September), 1–25. https://doi.org/10.7554/eLife.07832

Diederich, A. (2008). A further test of sequential-sampling models that account for payoff effects on response bias in perceptual decision tasks. *Perception & Psychophysics*, *70*(2), 229–256. https://doi.org/10.3758/PP.70.2.229

Diederich, A., & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound-change, drift-rate-change, or two-stage-processing hypothesis. *Perception and Psychophysics*, *68*(2), 194–207. https://doi.org/10.3758/BF03193669

Farashahi, S., Azab, H., Hayden, B., & Soltani, A. (2018). On the flexibility of basic risk attitudes in monkeys. *The Journal of Neuroscience*, *38*(18), 4383–4398. https://doi.org/10.1523/JNEUROSCI.2260-17.2018

Farashahi, S., Donahue, C. H., Khorsand, P., Seo, H., Lee, D., & Soltani, A. (2017). Metaplasticity as a neural substrate for adaptive learning and choice under uncertainty. *Neuron*, *94*(2), 401–414.e6. https://doi.org/10.1016/j.neuron.2017.03.044

Farashahi, S., Ting, C. C., Kao, C. H., Wu, S. W., & Soltani, A. (2018). Dynamic combination of sensory and reward information under time pressure. *PLoS Computational Biology*, *14*(3), 1–26. https://doi.org/10.1371/journal.pcbi.1006070

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*(2). https://doi.org/10.1371/journal.pcbi.1000284

Gao, J., Tortell, R., & McClelland, J. L. (2011). Dynamic integration of reward and stimulus information in perceptual decision-making. *PLoS ONE*, *6*(3), 5–7. https://doi.org/10.1371/journal.pone.0016749

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research*, *33*(9), 1219–1240. https://doi.org/10.1016/0042-6989(93)90210-N

Khorsand, P., & Soltani, A. (2017). Optimal structure of metaplasticity for adaptive learning. *PLoS Computational Biology*, *13*(6), 1–22. https://doi.org/10.1371/journal.pcbi.1005630

Kleiner, M., Brainard, D. H., Pelli, D. G., Broussard, C., Wolf, T., & Niehorster, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, S14. https://doi.org/10.1068/v070821

Liston, D. B., & Stone, L. S. (2008). Effects of prior information and reward on oculomotor and perceptual choices. *Journal of Neuroscience*, *28*(51), 13866–13875. https://doi.org/10.1523/JNEUROSCI.3120-08.2008

Maddox, W. T. (2002). Toward a unified theory of decision criterion learning in perceptual categorization. *Journal of the Experimental Analysis of Behavior*, *78*(3), 567–595. https://doi.org/10.1901/jeab.2002.78-567

Maunsell, J. H. R. (2004). Neuronal representations of cognitive state: reward or attention? *Trends in Cognitive Sciences*, *8*(6), 261–265. https://doi.org/10.1016/j.tics.2004.04.003

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*(7), 2335–2343. https://doi.org/10.1523/JNEUROSCI.4156-11.2012

Navalpakkam, V., Koch, C., & Perona, P. (2009). Homo economicus in visual search. *Journal of Vision*, *9(1)*(31), 1–16. https://doi.org/10.1167/9.1.31.Introduction

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*. https://doi.org/10.1163/156856897X00366

Pleger, B., Blankenburg, F., Ruff, C. C., Driver, J., & Dolan, R. J. (2008). Reward facilitates

tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *Journal of Neuroscience*, *28*(33), 8161–8168. https://doi.org/10.1523/JNEUROSCI.1093-08.2008

Rajsic, J., Perera, H., & Pratt, J. (2017). Learned value and object perception: Accelerated perception or biased decisions? *Attention, Perception, and Psychophysics*, *79*(2), 603–613. https://doi.org/10.3758/s13414-016-1242-0

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*(2). https://doi.org/10.1371/journal.pone.0009308

Schiller, P. H., & Chou, I. H. (1998). The effects of frontal eye field and dorsomedial frontal cortex lesions on visually guided eye movements. *Nat Neurosci*, *1*(3), 248–253. https://doi.org/10.1038/693

Schiller, P. H., & Chou, I. H. (2000). The effects of anterior arcuate and dorsomedial frontal cortex lesions on visually guided eye movements: 2. Paired and multiple targets. *Vision Research*, *40*(10–12), 1627–1638. https://doi.org/10.1016/S0042-6989(00)00058-4

Schiller, P. H., & Tehovnik, E. J. (2003). Cortical inhibitory circuits in eye-movement generation. *European Journal of Neuroscience*, *18*(11), 3127–3133. https://doi.org/10.1111/j.1460-9568.2003.03036.x

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology*, *47*(4), 333–366. https://doi.org/10.1016/S0010-0285(03)00035-5

Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, *60*(6), 1169–1181. https://doi.org/10.1016/j.neuron.2008.10.051

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science*, *12*(3), 205–212. Retrieved from http://www.ijrap.net/admin/php/uploads/1308_pdf.pdf

Soltani, A., Khorsand, P., Guo, C., Farashahi, S., & Liu, J. (2016). Neural substrates of cognitive biases during probabilistic inference. *Nature Communications*, *7*, 11393. https://doi.org/10.1038/ncomms11393

Spitmaan, M., Chu, E., & Soltani, A. (2018). Salience-driven value construction for adaptive choice under risk. *bioRxiv*

Soltani, A., Noudoost, B., & Moore, T. (2013). Dissociable dopaminergic control of saccadic target selection and its implications for reward modulation. *Proceedings of the National Academy of Sciences*, *110*(9), 3579–3584. https://doi.org/10.1073/pnas.1221236110

Stanford, T. R., Shankar, S., Massoglia, D. P., Costello, M. G., & Salinas, E. (2010). Perceptual decision making in less than 30 milliseconds. *Nature Neuroscience*, *13*(3), 379–385. https://doi.org/10.1038/nn.2485

Stănișor, L., van der Togt, C., Pennartz, C. M. A., Roelfsema, P. R., Stanisor, L., van der Togt, C., … Roelfsema, P. R. (2013). A unified selection signal for attention and reward in primary visual cortex. *Proceedings of the National Academy of Sciences*, *110*(22), 9136–9141. https://doi.org/10.1073/pnas.1300117110

Stelmach, L. B., & Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(2), 539–550. https://doi.org/10.1037/0096-1523.17.2.539

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews Neuroscience*, *6*(5), 363–375. https://doi.org/10.1038/nrn1666

Summerfield, C., & Koechlin, E. (2010). Economic value biases uncertain perceptual choices in the parietal and prefrontal portices. *Frontiers in Human Neuroscience*, *4*(November), 1–12. https://doi.org/10.3389/fnhum.2010.00208

Teichert, T., & Ferrera, V. P. (2010). Suboptimal integration of reward magnitude and prior reward likelihood in categorical decisions by monkeys. *Frontiers in Neuroscience*, *4*(NOV), 1–13. https://doi.org/10.3389/fnins.2010.00186

Tosoni, A., Committeri, G., Calluso, C., & Galati, G. (2017). The effect of reward expectation on the time course of perceptual decisions. *European Journal of Neuroscience*, *45*(9), 1152–1164. https://doi.org/10.1111/ejn.13555

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323. https://doi.org/10.1016/j.actao.2004.05.005

Voss, A., Rothermund, K., & Brandtstädter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, *44*(4), 1048–1056. https://doi.org/10.1016/j.jesp.2007.10.009

Wong, K.-F., & Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, *26*(4), 1314–1328. https://doi.org/10.1523/JNEUROSCI.3733-05.2006

Extended data

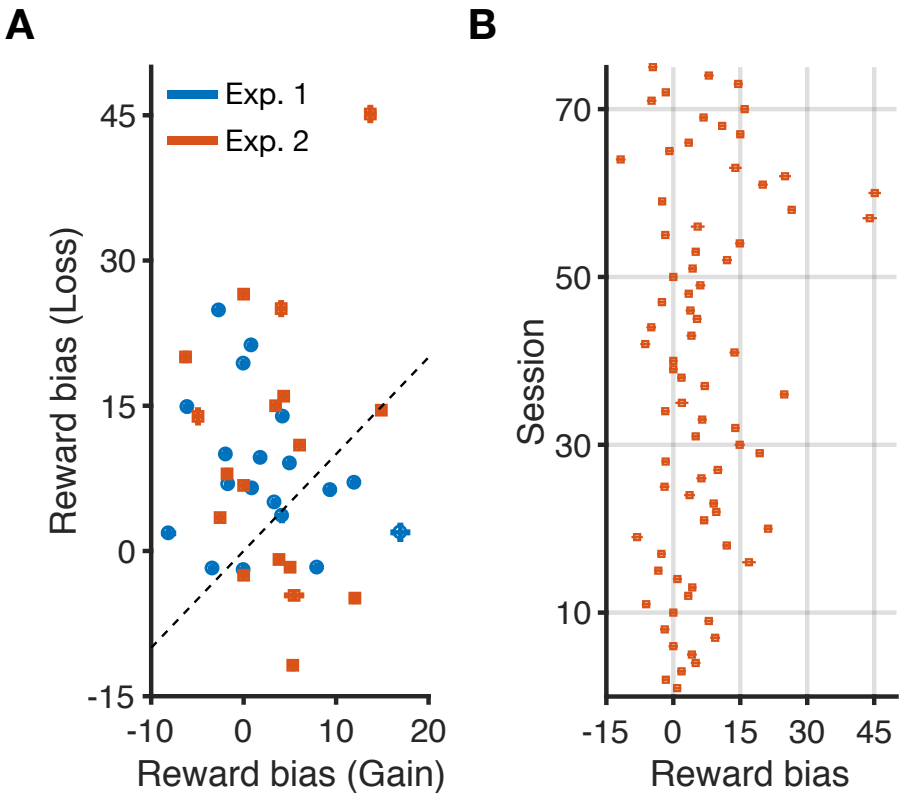

Figure 3-1. Robustness of the estimated reward biases. **(A)** Plotted are reward biases in the loss condition as a function of the reward biases in the gain condition with inclusion of estimation error for both Experiments 1 (blue) and 2 (orange). In most cases, the error bars are smaller than the symbol. (B) Plotted are estimation of reward biases and their estimation error for all sessions (for both Experiments 1 and 2 in both loss and gain conditions).