

Supporting Information

Supporting Information Corrected June 19, 2017

Suzuki et al. 10.1073/pnas.1600092113

SI Methods

Participants. We recruited 26 healthy participants (11 female participants; age range, 22–38 y; mean \pm SD, 29.35 ± 4.71) from the general population. All participants were preassessed to exclude those with any previous history of neurological/psychiatric illness, and they gave their informed written consent. Two participants were excluded from the subsequent behavioral and neuroimaging analyses because in the postexperiment questionnaire, they doubted that the other persons' choices were real. We therefore used the data from the remaining 24 participants (10 female participants; age range, 22–38 y; mean \pm SD, 29.50 ± 4.85).

Experimental Task. Our task contained three types of trials (Fig. 1A). In Self trials, participants made decisions under risk (see below for details); in Observe trials, participants observed decisions of another person ("observee") who had performed the Self trials; and in Predict trials, participants were asked to predict the observee's decisions.

The whole experiment consisted of five sessions, in which the three types of trials were presented to the participants in a block-wise manner (Fig. 1B). Sessions 2 and 4 involved all three trial types; and Sessions 1, 3, and 5 included only Self trials. Notably, each block began with the presentation of text and a picture indicating the trial type (Instruction phase; Fig. 1A). Because these instruction phases were presented only at the first trial in each block, the neural responses to this phase were unlikely to contaminate the decision-related responses of interest.

The observee for Observe/Predict trials was different between Sessions 2 and 4, and the two observees were distinguished by the images of two males presented in the instruction phase (see Fig. 1A and Fig. S1D). To minimize potential effects of their appearance, the images were taken from the back, and association between the images and the observees was randomized across participants. We instructed participants that the choices they observed in each session were made by a real person recorded from a previous experiment. In actuality, as in previous studies (14, 15, 17, 19, 20), the observees' choices were generated by computer algorithms. The two simulated observees had different risk-preferences (Fig. S1B and F): one for Session 2 was risk-averse ($\alpha = -0.013$ and $\beta = 5$ in mean-variance utility function; see *SI Methods, Behavioral Data Analyses*), and the other for Session 4 was risk-seeking ($\alpha = 0.03$ and $\beta = 5$) or vice versa (counterbalanced across participants). Furthermore, in the participant instructions, we emphasized that the observees did not have any further information about the task such as knowledge of the outcome of the gambles.

In Self trials, participants chose whether to accept or reject a gamble for themselves. If participants chose accept, they gambled for some amount of money; otherwise, they took a guaranteed \$10. Reward probability and magnitude of the gamble were varied in every trial (Fig. S1A), so that risk of the gamble (mathematical variance of reward) was decorrelated with the expected value of reward. At the beginning of each trial, participants were asked to make a choice between a gamble and the sure option by pressing a button with their right hand (index finger for the left option; and middle finger for the right option) within 4 s (Decision phase; Fig. 1A). Here, two options, accept and reject, were randomly positioned to the left or right of the screen in every trials, and information about the gamble was indicated by a pie chart (size of the green area indicates the probability and the digits denote the magnitude of reward). After making a response, the chosen option was highlighted by a yellow frame (Confirmation phase, 1 s), fol-

lowed by the uniformly jittered intertrial interval (ITI) (2–6 s). Notably, the outcome of the gamble was not revealed to the participants in each trial (one choice was randomly selected and actually implemented at the end of the experiment; see *SI Methods, Reward Payment* for details), to exclude potential effects of learning from reward feedback.

In Observe trials, participants observed a choice made by the observee. The trials were designed to minimize differences from the Self trials, and so the timeline of a trial and the set of gambles presented were the same between the two types of trials (Fig. 1A and Fig. S1A and B). Specifically, information about the gamble the observee would accept or reject was shown in the decision phase by a pie chart, and then the observee's actual choice was revealed to the participants in the confirmation phase (the observee's response time was 3–4 s), so that the participants could learn the observee's risk-preference. Furthermore, participants were also asked to press the third key after observing the observee's choice, to make motor-related neural responses in Observe trials comparable with those in the other types of trials.

Predict trials were introduced just to confirm that the participants learned the observee's behavioral tendency (i.e., risk-preference) through the observation of his choices in Observe trials, and therefore the number of trials was smaller relative to the other two trial types (Fig. 1B and Fig. S1A–C). Each trial began with the decision phase in which participants predicted whether the observee would accept or reject the gamble presented by a pie chart (Fig. 1A). The prediction was immediately highlighted by a yellow frame, initiating the confirmation phase. Here, the observee's actual choice was not revealed to the participants in each trial to prevent them from further learning (if the prediction was correct, the participant earned \$10 at the end of the experiment; see *SI Methods, Reward Payment* for details).

Postexperiment Questionnaire. After the experimental task, participants were asked to rate "How much do you think the first/second person is likely to accept a risky gamble?" (data shown in Fig. 3B), "How similar are you and the first/second person in terms of your preference for risky gambles?", and "How likely are you to accept a risky gamble?" in a six-point scale (from "not at all" to "very much"). We also asked another yes/no question, "Do you strongly doubt that the other persons' choices presented during the task were real?" (used for the exclusion of participants; *SI Methods, Participants*).

Reward Payment. Participants received a show-up fee of \$30. In addition, at the end of the experiment, the computer randomly selected one of the participant's choices from the Self trials, and the selected choice was actually implemented. The computer also selected one of the participant's predictions from the Predict trials at random. If the selected prediction was correct, the participant obtained \$10. The total earnings were on average $\$52 \pm \14 (mean \pm SD across participants). Importantly, because participants did not know which trial would be selected, they should have treated every trial as if it were the only one.

Additional Behavioral Experiment. Eleven healthy participants (five female participants; age range, 19–48 y; mean \pm SD, 29.36 ± 8.08) took part in the experiment. The experimental procedure was almost the same as the main experiment, but the participants observed/predicted choices of another person for one session and a computer agent for another session. In a session with the

computer agent, we instructed participants to observe/predict the computer's outputs.

The combination of the observees' type (human or computer) and their risk-preference (risk-averse or -seeking) was counter-balanced across participants. That is, there were four cases: (i) participants observed/predicted a risk-averse human in Session 2 and a risk-seeking computer in Session 4 ($n = 3$); (ii) a risk-seeking human in Session 2 and a risk-averse computer in Session 4 ($n = 3$); (iii) a risk-averse computer in Session 2 and a risk-seeking human in Session 4 ($n = 3$); and (iv) a risk-seeking computer in Session 2 and a risk-averse human in Session 4 ($n = 2$). Here, choices (outputs) of the risk-seeking and the -averse human/computer observees were generated by the same algorithms as the human observees in the main experiment (see *SI Methods, Experimental Task*).

Behavioral Data Analyses. We used three approaches to estimate participants' risk-preference for each session. Notably, the three approaches were found to provide highly consistent estimates at least in our task (Fig. S1 *E* and *F*).

One is a simple model-free approach based on the proportion of accepting gambles. That is, risk-preference is measured by $P(\text{gamble}) - P^*(\text{gamble})$, where $P(\text{gamble})$ denotes the participant's proportion of accepting gambles and $P^*(\text{gamble})$ indicates the proportion given the risk-neutral preference. That is, positive and negative values indicate risk-seeking and -averse preferences, respectively. In this approach, risk-neutrality can be defined by the **absolute value** of risk-preference (smaller values close to 0 indicate the more risk-neutral).

Another model is based on **mean-variance utility function**; the utility of each option is constructed by

$$U(X) = E(X) + \alpha V(X),$$

where $E(X)$ indicates **expected value**, whereas $V(X)$ denotes **risk of the option**. Let p be the reward probability and r be the reward magnitude of the option. Then, $E(X) = pr$ and $V(X) = r^2 p(1 - p)$. Here, risk-preference is characterized by α (negative values for risk-averse and positive for risk-seeking preference). The **utility** governs the participant's probability of accepting the gamble as follows:

$$q(\text{gamble}) = 1/[1 + \exp\{-\beta(U(\text{gamble}) - U(\text{sure}))\}],$$

where β is a parameter controlling the degree of stochasticity in the choice. We estimated the risk-preference α , as well as the parameter β , by fitting this model to the participant's choice data (i.e., **estimating the parameter values that maximize the likelihood function**).

In the other approach based on an **exponential utility function**, the **utility of each option** is constructed by

$$U(X) = p r^\rho,$$

where ρ represents the participant's risk-preference. ρ is less and greater than 1 if she/he is risk-averse and risk-seeking, respectively. The value of ρ can be estimated by model fitting, as well as the case with the mean-variance utility function.

fMRI Data Acquisition. The fMRI images were collected using a 3T Siemens (Erlangen) Tim Trio scanner located at the Caltech Brain Imaging Center with a 32-channel radio frequency coil. The BOLD signal was measured using a one-shot T2*-weighted echo planar imaging sequence [volume repetition time (TR): 2,780 ms; echo time (TE): 30 ms; flip angle (FA): 80°]. Forty-four oblique slices (thickness, 3.0 mm; gap, 0 mm; field of view, 192 × 192 mm; matrix, 64 × 64) were acquired per volume. The slices were aligned 30° to the anterior-posterior commissure

(AC-PC) plane to reduce signal dropout in the orbitofrontal area (59). After the five functional sessions, high-resolution (1-mm³) anatomical images were acquired using a standard Magnetization Prepared Rapid Gradient Echo (MPRAGE) pulse sequence (TR: 1,500 ms; TE: 2.63 ms; FA: 10°).

fMRI Data Analyses.

Preprocessing. We used the Statistical Parametric Mapping 8 (SPM8) software (Wellcome Department of Imaging Neuroscience, Institute of Neurology) for image processing and statistical analysis. fMRI images for each participant were preprocessed using the standard procedure in SPM8: after slice timing correction, the images were realigned to the first volume to correct for participants' motion, spatially normalized, and spatially smoothed using an 8-mm full width at half maximum Gaussian kernel. High-pass temporal filtering (using a filter width of 128 s) was also applied to the data.

GLM I. A separate GLM was defined for each participant. The GLM contained parametric regressors representing risk (standard deviation of reward) and expected value (expected value of reward) of the gamble at the time of decision (Fig. 1A). Specifically, the participant-specific design matrices contained the following regressors: boxcar functions in the period of the decision phase (i.e., the duration is equal to the response time; see Fig. 1A) for Self, Observe, and Predict trials (regressors for Observe and Predict trials were included only in Sessions 2 and 4); boxcar functions in the period of the confirm phase (duration, 1 s) for Self, Observe, and Predict trials (regressors for Observe and Predict trials were included only in Sessions 2 and 4); a boxcar function in the period of the instruction phase (duration, 4 s); and a stick function at the timing of motor response (i.e., button press). Furthermore, trials on which participants failed to respond were modeled as a separate regressor. We also included parametric modulators at the period of the decision phase for Self, Observe, and Predict trials separately, representing the risk and the expected value of the gamble. All of the regressors were convolved with a canonical hemodynamic response function. In addition, six motion-correction parameters were included as regressors of no-interest to account for motion-related artifacts. In the model specification procedure, serial orthogonalization of parametric modulators was turned off. For each participant, the contrasts were estimated at every voxel of the whole brain and entered into a random-effects analysis.

GLM I-2. The GLM contained parametric regressors representing response time, decision-related response (1 for accept the gamble; 0 for reject), and motor-related response (1 for choosing the left option, 0 for the right; note that accept and reject were randomly positioned left or right in every trial) at the time of decision, in addition to the other regressors included in GLM I. **GLM II.** The GLM contained parametric regressors representing reward probability and reward magnitude of the gamble at the time of decision, instead of the risk and the expected value, in addition to the other regressors included in GLM I.

GLM III. The GLM contained a parametric regressor representing utility of the gamble, estimated based on mean-variance utility function, at the time of decision, instead of the risk and the expected value, in addition to the other regressors included in GLM I.

GLM IV. The GLM contained parametric regressors representing squared reward magnitude and reward uncertainty, defined as $p(1 - p)$, where p denotes the reward probability, at the time of decision, instead of the risk and the expected value, in addition to the other regressors included in GLM I.

GLM V. The GLM contained a parametric regressor representing "update of the belief about others' risk-preference" at the confirm phase in Observe trials, in addition to the regressors included in GLM I. Here, the update of the belief was defined as the D_{KL} between the posterior and the prior (26), derived from the best-fitting Bayesian learning model (*SI Text 4* and Table S2). Notably, in the model specification procedure, we concatenated

five sessions to capture the long-term learning process across sessions (note: constant regressors coding for each session were included).

GLM V-2. The GLM contained parametric regressors representing the observer's response time, decision-related response (1 for accept the gamble; 0 for reject), and motor-related response (1 for choosing the left option; 0 for the right) at the confirm phase in Observe trials, in addition to the regressors included in GLM V.

PPI analysis. The PPI analysis was performed by using the standard procedure in SPM8. We first extracted a BOLD signal from the dlPFC identified as tracking update of the belief about others' risk-preference (Fig. 3C; 8-mm sphere centered at the peak voxel) and then created a PPI regressor by forming an interaction of the BOLD signal (physiological factor) and the boxcar function regressor in the period of the decision phase for Self trials over the course of the five sessions (psychological factor). The GLM for the PPI analysis therefore contained the following regressors: (i) a physiological factor, the BOLD signal in the bilateral dlPFC; (ii) a psychological factor, the boxcar function regressor in the period of the decision phase for Self trials; and (iii) a PPI factor, the interaction term of the psychological and physiological factors, as well as the other regressors included in GLM V.

Whole-brain analysis. We set our significance threshold at $P < 0.05$ whole-brain FWE corrected for multiple comparisons at cluster level. The minimum spatial extent, $k = 63$, for the threshold was estimated based on the underlying voxel-wise P value, $P < 0.005$ uncorrected, by using the AlphaSim program in Analysis of Functional NeuroImages (AFNI) (60).

SI Text 1

To test whether participants became "rational" by observing the others' choices, we conducted two behavioral analyses based on different notions of rationality. The set of gambles used in this experiment involved two risk-free gambles (i.e., reward probability is one, and the magnitude is \$20 and \$30; Fig. S1A). Rational participants should prefer the risk-free gambles to the sure \$10, regardless of their risk-preference (called "monotonicity" in economics). One definition of rationality is therefore the act of choosing the risk-free gambles (sure \$20 and \$30) over the sure \$10. Given this, we examined whether the proportion of choosing the risk-free gambles was changed across sessions and found no significant effect (ANOVA, $P = 0.59$; Fig. S4A). Another possible definition of rationality is to make decisions based on a risk-neutral preference (i.e., relying only on expected value but not on risk). Note that we here define rationality in the sense it maximizes expected monetary reward, apart from the conventional definition (i.e., maximization of expected utility) used in economics. We checked whether participants became risk-neutral as the experiment proceeded and found no significant effect (ANOVA, $P = 0.18$; Fig. S4B). In sum, we conclude that there is no significant evidence that participants became more "rational" through observation of others' behavior.

One might argue that the shift in participants' risk-preference can be explained by "regression to the baseline" (e.g., each participant's preference in Session 1 or the mean over the participants). Additional behavioral analyses, however, provided no significant evidence that participants' risk-preference approached the baseline as the experiment proceeded (ANOVA: $P = 0.30$ for the regression to each participant's baseline; and $P = 0.47$ for the mean preference; Fig. S4C).

SI Text 2

Here, we aimed to show that the shift in participants' behavior is better captured by the change in their risk-preference rather than the change in their subjective judgment about probabilities. To this end, we constructed two computational models, one with varying risk-preference across sessions and the other with varying

probability-weighting across sessions and compared their goodness-of-fits to the participants' actual choices.

In both models, participants are assumed to convert the instructed objective probability to the subjective probability as suggested in the previous study (22):

$$\pi(p) = 1/\exp\{[\log(1/p)]^\gamma\},$$

where γ governs the probability-weighting. If $\gamma = 1$, the subjective probability is identical to the objective probability. Then, given the subjective probability, the participants compute the utility of each option based on mean-variance utility function (*SI Methods, Behavioral Data Analyses*).

In the first model, we assign different risk-preference parameters into the five sessions and a single common probability-weighting parameter into the sessions. In contrast, in the second model, different probability-weighting parameters are assigned into the five sessions and one common risk-preference parameter is assigned. Fitting these models to each participant's behavior separately, we found the first model provided the better fit than the second model did [(Exceedance probability in Bayesian Model Selection (61) is 0.94]. Here, it is worth noting that each model's goodness of fit was assessed using the log model evidence, approximated by the Bayesian information criterion (BIC) (62), which penalizes models with more free parameters compared with those with fewer free parameters.

The same result was obtained when we used an exponential utility function (*SI Methods, Behavioral Data Analyses*) instead of a mean-variance utility. That is, the model with varying risk-preference provided a better fit than the model with varying probability-weighting (Exceedance probability favoring the first model is 0.99). These findings suggest that, regardless of the specific forms of utility function, the session-by-session behavioral shift we observed reflects a change in the participants' risk-preference rather than probability-weighting.

SI Text 3

We aimed to confirm that the shift in participants' behavior was better captured by the change in their risk-preference rather than the change in the simple bias for/against utility of gambling options (16). For this purpose, as in *SI Text 2*, we constructed two computational models and compare their goodness-of-fits. In both of the two models, a bias for/against utility of gambling options is incorporated. That is, under mean-variance utility function (*SI Methods, Behavioral Data Analyses*), the utility of a gambling option is computed by

$$U(X) = E(X) + \alpha V(X) + \delta,$$

where δ denotes the bias. The bias term can be similarly introduced into exponential utility function (*SI Methods, Behavioral Data Analyses*): $U(X) = p r^p + \delta$, which is equivalent to the other-conferred utility (OCU) model introduced by ref. 16.

As in *SI Text 2*, the first model has different risk-preference parameters for the five sessions and a single common bias parameter, δ , whereas the second model has different bias parameters and a single risk-preference parameter. The model comparison revealed that the first model provided the better fit than the second model. Exceedance probability (61) favoring the first model is 0.99 in the case with mean-variance utility function and 1.00 in the case with exponential utility function. These results support our claim that the shift in participants' behavior was driven by the change in their risk-preference.

SI Text 4

Here, we aimed to demonstrate that the shift in participants' behavior is better captured by a change in risk-preference rather

than a change in a simple bias for/against the choice probability of gambling options (24). Following the procedure in *SI Text 2*, we constructed two computational models and compared their goodness-of-fits. In both of the two models, a bias for/against choice probability of gambling options is incorporated. That is, after computing the utility of the gambling and sure options (based on either mean-variance or exponential utility function; *SI Methods, Behavioral Data Analyses*), participants make a choice based on the choice probability:

$$q(\text{gamble}) = \delta + (1 - \delta) / [1 + \exp\{-\beta(U(\text{gamble}) - U(\text{sure}))\}] \text{ if } \delta \geq 0,$$

$$q(\text{gamble}) = (1 + \delta) / [1 + \exp\{-\beta(U(\text{gamble}) - U(\text{sure}))\}] \text{ if } \delta < 0$$

where $\delta \in [-1, 1]$ denotes the bias to the choice probability. If $\delta = 1$, the participant always choose the gambling option independent of the utility; if $\delta = -1$, she/he always choose the sure \$10; and if $\delta = 0$, she/he makes decisions based only on the utility. Note that the model with the biased choice probability is referred to as parametric approach-avoidance decision model in the previous study (24).

As in *SI Text 2*, the first model has different risk-preference parameters for the five sessions and a single bias parameter, δ , whereas the second model has different bias parameters and a single risk-preference parameter. The model comparison revealed that the first model was better fit than the second model. Exceedance probability (61) favoring the first model is 0.98 in the case with the mean-variance utility function and 1.00 in the case with the exponential utility function. These results suggest that the shift in participants' behavior is captured by a change in risk-preference rather than a change in the bias to choice probability.

SI Text 5

To capture the computational process of participants' learning about the observees' risk-preference, we constructed a family of computational models and fit those models to the participants' actual prediction behaviors.

Bayesian Learning Models. In this family of models, an observee's risk-preference is inferred by a simple Bayesian learning algorithm. The Bayesian learner updates the estimation of an observee's risk-preference, when she/he observes the observee's choice (at the time of confirmation in Observe trials). Using the estimated risk-preference, she/he predicts the observee's choice in Predict trials.

Given the mean-variance utility function, the learner estimates and updates an observee's risk-preference, α_o , as follows:

$$p'(\alpha_o) \propto p(\alpha_o) p_o(C_o|\alpha_o),$$

where C_o and p_o denote the observee's choice and the choice probability, respectively. Here, the observee's choice probability can be derived by

$$p_o(C_o = \text{gamble}|\alpha_o) = 1 / \{1 + \exp[-\beta_o(U(\text{gamble}) - U(\text{sure}))]\} \\ = 1 / \{1 + \exp[-\beta_o(E(X) + \alpha_o V(X) - 10)]\},$$

where U denotes utility of each option, $E(X)$ and $V(X)$ represent expected value and risk of the gamble, and β_o controls the degree of stochasticity in the observee's choice (treated as a free parameter in the model fitting).

Based on the estimated risk-preference, the learner computes the observee's choice probability as follows:

$$p_o(C_o) = \int p_o(C_o|\alpha_o)p(\alpha_o)d\alpha_o;$$

and finally predicts the observee's choice:

$$p_s(C_o = \text{gamble}) = 1 / \{1 + \exp[-\beta_s(p_o(C_o = \text{gamble}) - p_o(C_o = \text{sure}))]\},$$

where p_s represents the probability that the learner predicts the observee's choice of the gamble or sure option, and β_s controls the degree of stochasticity in the prediction (a free parameter). At the beginning of each session, we assume that the Bayesian learner has a normally distributed prior belief about the observee's risk-preference, α_o .

Each model included in this family is different with respect to shape of the prior belief. In the Bayesian (simple) model, both mean and variance of the normal prior belief are free parameters, in addition to β_o and β_s . On the other hand, in the Bayesian (self-projection) model, the mean of the prior is extracted from the participant's own risk-preference in the last session (i.e., Session 1 for Session 2; and Session 3 for Session 4), whereas the variance is a free parameter. For Session 4, we constructed additional two models: a Bayesian (carry-over) and a Bayesian (carry-over + self-projection) model. In the Bayesian (carry-over) model, the mean of the prior is equal to the mean of the posterior in Session 2 (the variance is a free parameter). That is, learning in Session 2 carries over to Session 4, which would capture the fact that the participants' prediction performance at the beginning of Session 4 was below the chance-level (Fig. S6, *Right*). In the Bayesian (carry-over + self-projection) model, the mean prior is a weighted sum of the participant's own risk-preference and the mean of the posterior in Session 2, where the weight is a free parameter.

Null Models. In addition to the Bayesian learning models, we constructed two null models. One is a Random Prediction model that predicts observees' choice randomly. The other is a Fixed Probability model that predicts that observees accept a gambling option with a fixed probability (independently of the reward probability and magnitude).

Model Comparison. We fit each model to the aggregated data from all of the participants together because of the small number of Predict trials (i.e., we used a fixed-effects analysis). Each model's goodness of fit was then assessed using the BIC, which penalizes models with more free parameters compared with those with fewer free parameters. The model comparison process revealed that, for Session 2, the Bayesian (self-projection) model provided the best fit to the participants' actual prediction behaviors and that, for Session 4, the Bayesian (carry-over + self-projection) model best fit the data (Table S2). These results together suggest that the participants learned the observees' risk-preference in a manner consistent with a Bayesian learning algorithm featuring the use of their own risk-preference as a prior.

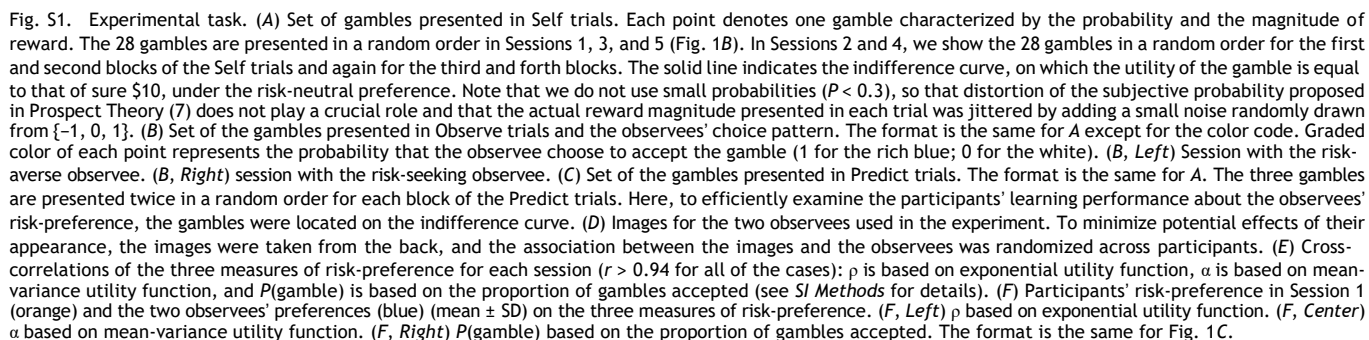
SI Text 6

We confirmed that the correlation (Fig. 3D) remained significant ($P < 0.05$), even when (i) controlling for individual differences in the strength of the neural representation of a belief-updating signal in the dlPFC (i.e., using partial correlation); (ii) removing outliers based on a Jackknife procedure (63); or (iii) analyzing the data separately for the left and the right dlPFC ($P < 0.05$ for both regions).

SI Text 7

The shift in risk-preference we observed here is unlikely to be explained by a simple imitation of the others' choices (i.e., by merely copying a stimulus-response association). In our experiment, on

that the observees did not have any further information about the task such as knowing the outcome of the gambles. Finally, our experimental task is fundamentally different from others used in previous studies, in that there is no effect of learning from reward feedback (10), emotion regulation (64), task contexts (6, 65), explicit incentive (66), or priming with a financial boom/bust (67) on the changes of risk-preferences.



6 of 9

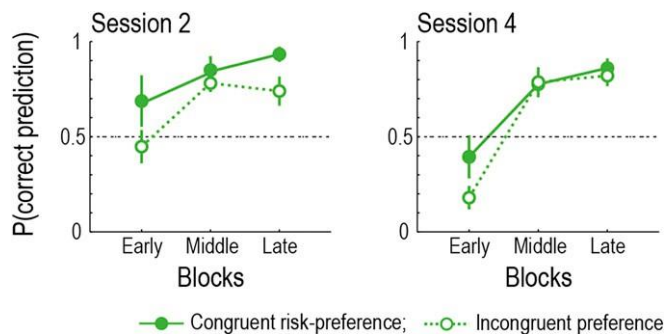


Fig. S6. Learning about the observees' risk-preference. Proportions of the correct predictions on Predict trials are plotted as a function of time. (Left) Session 2. (Right) Session 4. Filled points denote the case where participants' own risk-preference was congruent with the observee's preference. Open points represent the case where participants' risk-preference was incongruent with the observee's.

Table S1. Areas exhibiting significant changes in BOLD associated with the risk and the expected value

Region	Hemi	BA	x	y	z	t static	P value	Voxels
Risk								
Caudate (head)	Bi		9	20	7	6.26	0.000	189
Caudate (tail)	L		-27	-37	4	4.21	0.000	94
Expected value								
Visual cortex	Bi	17/18/19	-6	-70	16	7.25	0.000	4,151
PCC/TPJ/IPL	Bi	23/39/40	9	-40	37	6.93	0.000	1,487
mPFC	Bi	8/910/32	-12	41	40	5.27	0.000	1,491
Insula	L	13	-33	11	-8	4.78	0.000	73
Precentral gyrus	R	4	51	-19	46	4.45	0.000	154
IFG	R	47	27	29	-11	4.40	0.000	63
IPL	L	40	-30	-43	40	4.12	0.000	67

Activated clusters observed in the whole-brain analysis ($P < 0.05$, FWE corrected at cluster level) for risk and expected value of the gamble at the time of self-decision. The stereotaxic coordinates are in accordance with MNI space. t statistics and uncorrected P values at the peak of each locus are shown. In the far right column, the number of voxels in each cluster is shown. BA, Brodmann area; Bi, bilateral; Hemi, hemisphere; IFG, inferior frontal gyrus; L, left; PCC, posterior cingulate cortex; R, right; TPJ, temporoparietal junction.

Table S2. Behavioral model fits to the participants' prediction data

Model	-LL	#prms	BIC
Session 2			
Random	297	0	595
Fixed Probability	296	1	599
Bayesian	232	4	488
Bayesian (self-projection)	229	3	477
Session 4			
Random	295	0	591
Fixed Probability	295	1	597
Bayesian	206	4	437
Bayesian (self-projection)	214	3	446
Bayesian (carry-over)	206	3	429
Bayesian (self-projection + carry-over)	191	4	407

Comparison of each model's goodness-of-fit to the participants' prediction behaviors. For BIC, smaller values indicate better fit. -LL, negative log likelihood (smaller values indicate better fit); #prms, no. of free parameters. The best model in each session is shown in bold.

Table S3. Areas exhibiting significant changes in BOLD associated with the belief-updating signal (D_{KL}) between the posterior and the prior

Region	Hemi	BA	x	y	z	t static	P value	Voxels
dIPFC	R	9	54	17	34	6.70	0.000	568
	L	9	-45	8	34	5.16	0.000	481
IPL	L	40	-36	-52	40	5.94	0.000	686
	R	40	42	-49	43	4.18	0.000	233
dmPFC	R	8	-3	35	49	5.38	0.000	804
Insula	L	13	-30	20	-5	5.27	0.000	117
V1	L	17	-12	-76	13	4.31	0.000	121

Activated clusters observed in the whole-brain analysis ($P < 0.05$, FWE corrected at cluster level) for belief-updating at the time of confirmation in Observe trials. The format is the same for Table S1. BA, Brodmann area; Hemi, hemisphere; L, left; R, right; V1, primary visual cortex.

Table S4. Literature survey about the neural correlates of decision-risk

Study	Contrast	Activated regions
Instructed-risk without feedback		
Hsu et al. (2005) (40)	Decision with risk > ambiguity	Caudate, culmen, lingual gyrus, cuneus, precuneus, precentral gyrus, precuneus, angular gyrus
Labudda et al. (2008) (43)	Decision with risk > without information incentive	Anterior cingulate gyrus, middle frontal gyrus, inferior parietal lobe, lingual gyrus, precuneus
Weber and Huettel (2008) (41)	Decision with risk > delayed reward	Lingual gyrus, middle occipital gyrus, superior parietal lobule, precuneus, middle frontal gyrus, inferior frontal gyrus, anterior cingulate cortex, orbitofrontal cortex, hippocampus, superior frontal sulcus, anterior insula, inferior temporal gyrus
	Choice of riskier option > safer option	Insula, posterior cingulate, anterior cingulate, medial frontal gyrus, caudate, postcentral gyrus, superior temporal gyrus, inferior frontal gyrus
Christopoulos et al. (2009) (37)	Decision with high risk > low risk	Dorsal anterior cingulate cortex
Symmonds et al. (2011) (42)	Variance of reward (parametric)	Posterior parietal cortex
Instructed-risk with feedback		
Paulus et al. (2003) (68)		
Huettel et al. (2006) (69)		
van Leijenhorst et al. (2006) (70)		
Lee et al. (2008) (71)		
Engelman and Tamir (2009) (72)		
Smith et al. (2009) (73)		
Xue et al. (2009) (74)		
Learned risk		
Paulus et al. (2001) (75)		
Volz et al. (2003) (76)		
Matthews et al. (2004) (77)		
Volz et al. (2004) (78)		
Cohen et al. (2005) (79)		
Huettel et al. (2005) (80)		
Huettel (2006) (81)		
Mohr et al. (2010) (82)		
Bach et al. (2011) (83)		
Payzan-LeNestour et al. (2013) (84)		

The top set of studies are those controlling for effects of learning from reward feedback: as in the present study, information about risk was provided by instruction, and the outcome of each choice was not revealed. See refs. 38 and 39 for the discrimination of instructed and learned value information. The middle set of studies are those not controlling for effects of learning: information about risk was provided by instruction, but the outcome of each choice was revealed to the participants (i.e., valuations can be influenced by the history of previous outcomes). Activated regions and the contrast were omitted. The bottom set of studies are those not controlling for effects of learning: information about risk was acquired through learning (i.e., the information was not instructed explicitly). Activated regions and the contrast were omitted. A region of interest discussed in the main text, caudate, is shown in bold.