Base rate neglect and neural computations for subjective weight in decision under uncertainty

Yun-Yen Yang[a] and Shih-Wei Wu[b,c,1]

[a]Department of Psychology, Rutgers University, Newark, NJ 07102; [b]Institute of Neuroscience, National Yang-Ming University, Taipei 112, Taiwan; and [c]Brain Research Center, National Yang-Ming University, Taipei 112, Taiwan

Base rate neglect, an important bias in estimating probability of uncertain events, describes humans' tendency to underweight base rate (prior) relative to individuating information (likelihood). However, the neural mechanisms that give rise to this bias remain elusive. In this study, subjects chose between uncertain prospects where estimating reward probability was essential. We found that when the variability of prior and likelihood information about reward probability were systematically manipulated, prior variability significantly affected the degree to which subjects underweight the base rate of reward probability. Activity in the orbitofrontal cortex, medial prefrontal cortex, and putamen represented the relative subjective weight that reflected such bias. Further, sensitivity to likelihood relative to prior variability in the putamen correlated with individuals' overall tendency to underweight base rate. These findings suggest that in combining prior and likelihood, relative sensitivity to information variability and subjective-weight computations critically contribute to the individual heterogeneity in base rate neglect.

base rate neglect | probabilistic inference | orbitofrontal cortex | medial prefrontal cortex | putamen

Base rate neglect, the tendency to underweight base rate or prior information compared with current, individuating information when estimating probability of uncertain events, is an important bias in human probabilistic inference (1, 2). It highlights a long-lasting research program in experimental psychology that uses ideal Bayesian inference as a model for human performance and seeks to gain insights into how humans estimate probability through examining systematic deviations in probability estimates from the ideal model prediction (3). The neural mechanisms for how systematic biases, such as base rate neglect, arise, however, remain elusive. Investigating these biases at the computational and neural implementation levels are crucial to understanding a wide array of cognitive computations that require probabilistic inference.

Previous research investigated the neurocomputational substrates of probabilistic inference in a wide variety of tasks, an essential first step to understanding potential biases in estimation. These studies examined how people use cue reliability as prior information to guide perceptual decision making (4, 5), make financial decisions based on prior information about a partner's reputation (6), combine prior and likelihood information about reward probability (7, 8), infer latent causes (9) and other people's intentions (10), and make visuomotor decisions that require combining the uncertainty in prior and likelihood information (11). Together, they pointed to the role of neural systems critical to reward processing and decision making, including the medial prefrontal cortex (mPFC), orbitofrontal cortex (OFC), and putamen, in inference-related computations, from representing prior knowledge (4, 6, 8, 11), current observation or sensory evidence (7, 8), individuals' relative sensitivity to different sources of uncertainty (11), to integrating prior and current information (8–10).

While these studies provided crucial insights into probabilistic inference, they did not investigate how systematic biases in probabilistic inference arise in the brain. The goal of this study is to address this question by focusing on base rate neglect. However,

there are at least three major challenges in investigating the neural mechanisms for base rate neglect, which we outline below.

First, in task paradigms suitable for neurobiological investigations, the field has yet to see a paradigm that robustly reveals or replicates base rate neglect. Across the studies highlighted above, subjects either achieved near-optimal performance in combining prior and current information (8, 9), or it is unclear whether subjects achieved near-optimal performance (4–7, 10, 11). To address this challenge, it is critical to identify the statistical properties of prior and likelihood information that directly and robustly contribute to base rate neglect.

Second, there is little consensus on which behavioral metric one should use to quantitatively characterize base rate neglect and to identify brain regions involved in computing this metric. To address this challenge, we propose that subjective weight—how a decision maker weighs prior and likelihood information when combining them—should be the behavioral metric because base rate neglect is fundamentally about underweighting the prior information. However, how should one investigate the neural mechanisms for subjective-weight computations? What could be the starting neural hypothesis? Bayesian decision theory provides a key insight. That is, subjective weight assigned to prior and likelihood should be determined by the relative variability of these two sources of information. Therefore, a reasonable hypothesis would be that in computing subjective weight, the brain

takes into account prior and likelihood variability. To test this hypothesis, first it is important to independently manipulate the variability of both prior and likelihood so that we could estimate how subjective weight changes in response to information variability. Second, to quantitatively characterize the degree of bias—underweighting the base rate—we should compare individual participants' subjective weight with the ideal weight assigned by the Bayesian decision maker. At the neural implementation level, the relation between information variability and subjective weight also leads to an important hypothesis. That is, neural substrates for subjective-weight computations should be tightly linked with neural representations for information variability. Together, these observations point to three critical analyses that aim to 1) identify brain activity representing subjective weight, 2) identify neural correlates of information variability, and 3) characterize the connecting properties between neural substrates identified in analyses 1 and 2 at the time of prior-likelihood integration.

Finally, the third challenge is to dissociate inference from choice computations. It remains unclear whether probabilistic inference computations, combining prior and likelihood information for the purpose of estimating probability of potential outcomes, are dissociable from choice computations that use information about the outputs of probabilistic inference computations to make decisions (4, 11–15). To address this challenge, in our decision task we temporally separated inference from choice such that during inference, it would be difficult and therefore less likely for the subjects to engage in choice computations.

We found that the degree to which human subjects underweight base rate can be attributed to the variability of prior information but not the variability of likelihood information. Using blood-oxygen-level–dependent (BOLD) functional magnetic resonance imaging (fMRI), we identified neurocomputational substrates for the subjective weighting of likelihood information relative to the base rate and found that the relative sensitivity of brain activity to information variability correlated with an individual's overall tendency to underweight base rate. These results indicate that biases in probabilistic inference, pervasively observed in human probability judgments, arise from information-weighting computations and relative sensitivity to information variability in the brain.

Results

We designed two tasks to investigate the neural computations for probabilistic inference—the integration of prior and likelihood information—in which human subjects learned probability of reward associated with different visual symbols and were then asked to make choices between lotteries. In order to establish prior knowledge about reward probability through experience, subjects ($n = 28$) first completed a prior-learning session (session 1, a behavioral session). To study how subjects combined prior knowledge with likelihood information about reward probability, they came back the next day for a second session (session 2, prior-likelihood integration, an fMRI session).

In the prior-learning session, subjects learned through feedback information about reward probability associated with two visual symbols. Each symbol represented a unique probability distribution on reward probability (beta distribution). The two distributions had the same mean (0.5 or 50% reward) but different variance (Fig. 1A). In each trial, subjects were presented with one symbol and asked to estimate the probability of receiving a reward that was randomly drawn from the probability distribution associated with the symbol (Fig. 1B). Subjects were then given feedback and rewarded based on how close his or her estimate was to the true probability of reward in that trial, an incentive compatible procedure developed to motivate subjects to learn probability distributions.

To investigate how and how well subjects combined prior and likelihood information, we estimated the weights they assigned

to likelihood information relative to prior, referred to as subjective weight, using a lottery decision task (session 2). In each trial, they were asked to choose between two lotteries, a symbol lottery and an alternative lottery, that differed only in the probability of winning a fixed monetary reward (Fig. 1C). For the symbol lottery, information about its reward probability was partially revealed through two pieces of information: prior and likelihood. In other words, subjects had to estimate its reward probability by using these two pieces of information. By contrast, information about reward probability of the alternative lottery was unambiguously revealed in numeric form so that there was no need to estimate its reward probability. Since reward magnitude was the same between the two options, subjects should always choose the option she or he believed to have the larger probability of reward. Therefore, through subjects' choices we can infer his or her estimates of reward probability associated with the symbol lottery on a trial-by-trial basis, which allowed us to estimate subjective weight.

To examine the effects of prior and likelihood variability on subjective weight, we independently manipulated the variability of prior and likelihood information about reward probability associated with the symbol lottery in a 2 × 2 factorial design (Fig. 1E). In each trial, the prior information was one of the two symbols subjects learned in the prior-learning session. The presented symbol served to inform subjects, in the current trial, which distribution the reward probability of the symbol lottery was drawn from. The likelihood information showed outcomes of repeated realizations of the symbol lottery (each red dot indicated a reward outcome, and each white dot indicated a no-reward outcome) such that if the sample size were infinitely large, the proportion of red dots would be equivalent to the reward probability of the symbol lottery in that trial. In other words, when the sample size (the number of times the lottery was executed) is small, likelihood information is unreliable in indicating probability of reward. As the sample size increases, likelihood information becomes more reliable because the proportion of red dots is more likely to reflect the true reward probability (Fig. 1D). Hence, by manipulating sample size we effectively manipulated the variability of likelihood information.

To dissociate probabilistic inference from choice, two important computations that can be highly correlated, each trial consisted of two stages: an inference stage followed by a choice stage. At the inference stage, prior and likelihood information associated with the symbol lottery was presented. Reward probability of the alternative lottery was not revealed until the choice stage (where subjects had to indicate his or her decision). By design, reward probability of the alternative lottery was determined randomly so that it would be difficult for the subjects, on a trial-by-trial basis, to predict the alternative lottery at the inference stage. Hence, brain activity identified to be associated with probabilistic inference at the inference stage would be less likely to be attributed to choice computations that involved comparing the reward probability between the two options.

Subjects Learned the Mean and Variance of Prior Distributions through Experience. We found that subjects learned the prior distributions well. Their trial-by-trial probability estimates in the prior-learning session (Fig. 2A, histogram in gray) captured the shape of the prior distributions (blue curve): the mode of probability estimates in response to both distributions was very close to 0.5; the estimates were symmetric around 0.5 and were more variable when the distribution had larger variance. Subjects' estimates of variability—the 90% interval estimates of reward probability provided at the end of each block of trials—also reflected the variability of the prior distributions (red in Fig. 2B). However, compared with the true 90% intervals (blue in Fig. 2B), subjects significantly underestimated the variability of
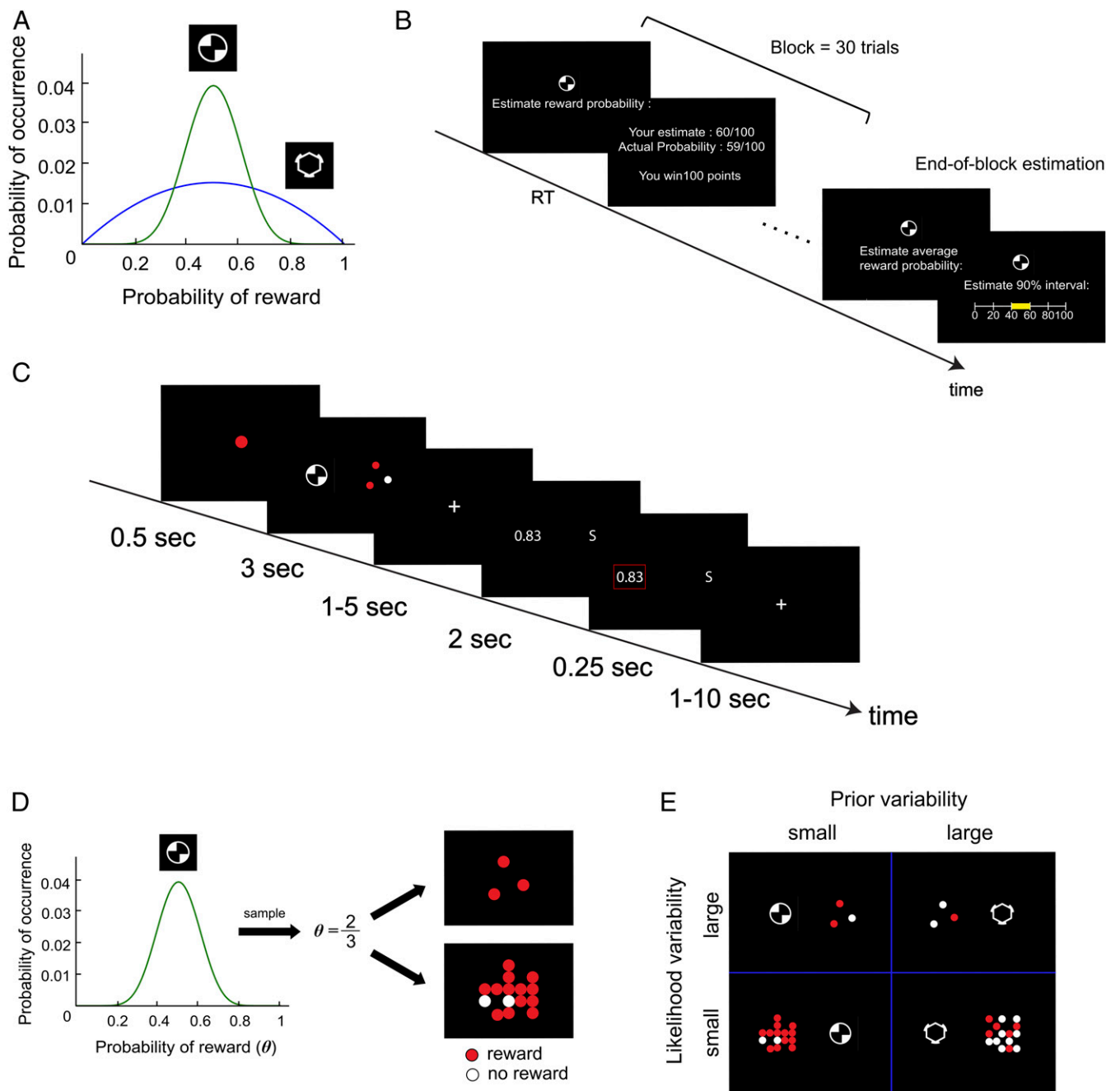
Fig. 1. Experimental design. (*A* and *B*) Prior-learning session. (*A*) Two beta distributions on reward probability that serve as prior information. Each distribution was represented by an abstract visual symbol. The two distributions had the same mean but different variance. The vertical axis represents the probability of occurrence when making reward probability (horizontal axis) discrete in steps of 0.01. (*B*) Prior-learning task. Subjects learned the two distributions through experience. In each trial, subjects were presented with a symbol and were asked to estimate the probability of receiving a monetary reward. After estimation, subjects received feedback and were rewarded based on how close his or her estimate was to the true reward probability. (*C–E*) Prior-likelihood integration session. (*C*) Lottery decision task. In each trial, subjects had to choose between the symbol lottery and the alternative lottery. Prior (abstract visual symbol on the left) and likelihood information (colored dots on the right) associated with the symbol lottery were first presented for 3 s. After a variable delay, reward probability associated with the alternative lottery was revealed explicitly in numeric form. Subjects had to indicate his or her decision within 2 s between the symbol lottery (represented by S on the right side of the screen in this example) and the alternative lottery (0.83 probability of winning a reward in this example) with a button press. Once a button was pressed, the identity of the chosen option was displayed (250 ms). No feedback on the reward outcome associated with the chosen option was given. (*D*) Illustrations on how likelihood information is generated in each trial. In this example, the probability of reward associated with the symbol lottery is 2/3. Likelihood information is the outcomes of repeated realization of this lottery, a 2/3 chance of winning a reward, for either 3 times (upper right screen containing 3 dots) or 15 times (bottom right screen containing 15 dots). Here a red dot indicates a reward outcome, and a white dot indicates a no-reward outcome. (*E*) Manipulation of prior and likelihood variability. In a 2 × 2 factorial design, we independently manipulated the prior variability (two prior distributions with different variance) and likelihood variability (two sample sizes: 3 dots or 15 dots).
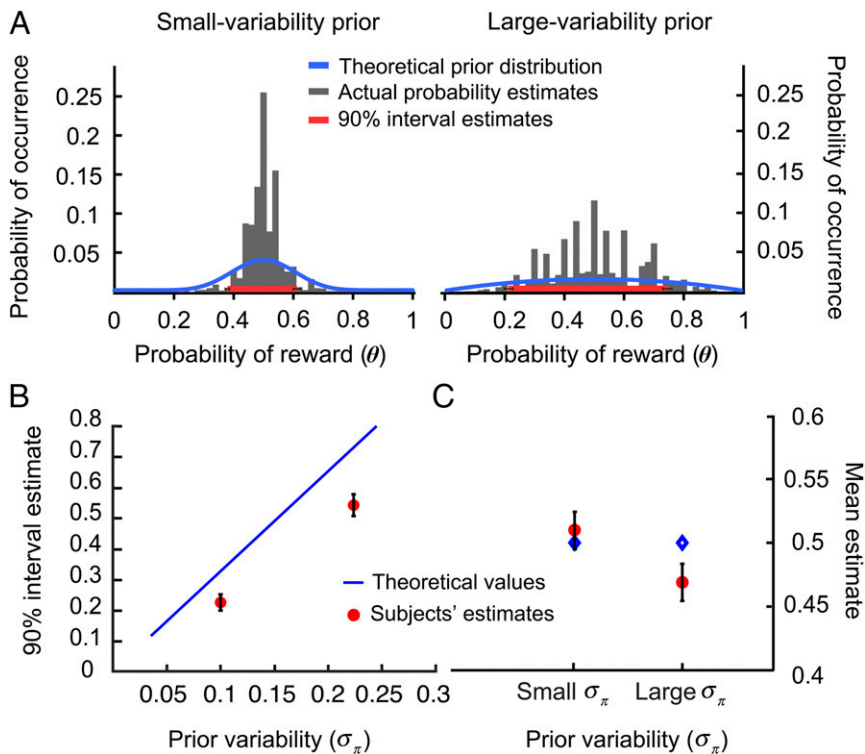
Fig. 2. Behavioral results: prior-learning session. (*A*) Comparison of probability estimates and prior distributions. Data from all subjects' trial-by-trial estimates of reward probability associated with the two prior distributions are summarized by the histograms in gray. The blue curves represent the prior distributions (the value on the vertical axis represents the probability of occurrence when making the reward probability discrete in steps of 0.01). (*B*) Variability estimates. Mean estimates (across subjects) of the 90% interval of reward probability (data points in red) associated with the two prior distributions are plotted against their corresponding SD (σ_π). The blue line represents the true 90% interval of prior distribution (mean = 0.5) with SD spanning from small (close to 0) to large (0.25). (*C*) Mean estimates (across subjects) of reward probability (data points in red) and the true mean of reward probability associated with the two prior distributions (0.5, in blue). Error bars represent ±1 SEM.

both distributions. For probability estimates, the mean estimates (across subjects) did not differ from the true mean (50% chance of reward) when prior variability was small but were significantly smaller than the true mean when prior variability was large [$t(27) = -2.14, P = 0.042$]. This could be due to increased task difficulty in estimating probability under large variability (16). Such difference, however, did not change the conclusion of the subjects' behavioral performance in the subsequent session (prior-likelihood integration session) in how they weighed prior and likelihood information.

Suboptimal Integration: Subjects Underweight the Base Rate of Reward Probability. We found that subjects significantly underweight the base rate of reward probability and that the degree of base rate neglect can be primarily attributed to prior variability but not likelihood variability. We estimated how the subjects weighed likelihood of reward (indicated by the proportion of red dots) relative to the base rate of reward probability (the mean of the prior distributions at 0.5), referred to as subjective weight (subjective ϖ), and compared it with the weights assigned by the ideal Bayesian decision maker (ideal ϖ). To estimate subjective weight, for each subject and each condition (a combination of prior and likelihood variability) separately, we performed a logistic regression analysis on choice (*SI Appendix, SI Methods: Behavioral Analysis 1: Estimating Subjective Weight*). If the subjects completely ignored likelihood information (only considering base rate), subjective weight would be 0. By contrast, if the subjects only considered likelihood information (completely ignoring base rate), subjective weight would be 1.

The computation of the ideal weight is illustrated in Fig. 3*A*. The example on the left indicates a situation where variability of the prior distribution is relatively smaller than the variability of the likelihood function. In this case, the ideal decision maker would "trust" the prior more by assigning smaller ϖ (ideal $\varpi = 0.11$). By contrast, the example on the right illustrates a situation where the ideal decision maker would weigh likelihood information more

heavily than the prior (ideal $\varpi = 0.79$) because variability of the likelihood function is relatively smaller than the variability of the prior distribution. In principle, the ideal weight changes as a function of both the variability of the prior distribution and the sample size of the likelihood information (Fig. 3*B*).

We found that the participants clearly adjusted subjective weight in response to likelihood variability: as likelihood variability increased, subjective weight decreased (Fig. 3*C*). They also adjusted subjective weight in response to prior variability: as prior variability increased, subjective weight increased. These results were qualitatively consistent with the direction predicted by the ideal weight (blue lines in Fig. 3*C*). Notably, we found that compared with the likelihood variability, subjects showed smaller adjustment in subjective weight in response to changes in prior variability [$t(27) = -2.14, P = .042$].

When we compared the subjective weight with the ideal weight (*SI Appendix, SI Methods: Behavioral Analysis 2: Ideal Decision Maker Analysis*), we found both near-optimal and suboptimal performance (Fig. 3*C*). When prior variability was large, mean subjective weight (across subjects) did not differ from the ideal weight regardless of likelihood variability. By contrast, when prior variability was small, mean subjective weight was significantly larger than the ideal weight, indicating that subjects underweight the base rate. This pattern was observed regardless of likelihood variability. Further, this suboptimal behavior—underweighting of base rate—cannot be attributed to individual differences in the variability estimates of the prior distributions in the prior-learning session: subjective weight was not correlated with the 90% interval estimates (Fig. 3*D*) the participants provided in the prior-learning session. It was also not correlated with the SD of the subjects' probability estimates (Fig. 3*E*). Together with the result showing that subjects had relatively accurate knowledge about prior variability (Fig. 2*B*), these results indicate that weighting the base rate and knowledge about prior variability might be dissociable.
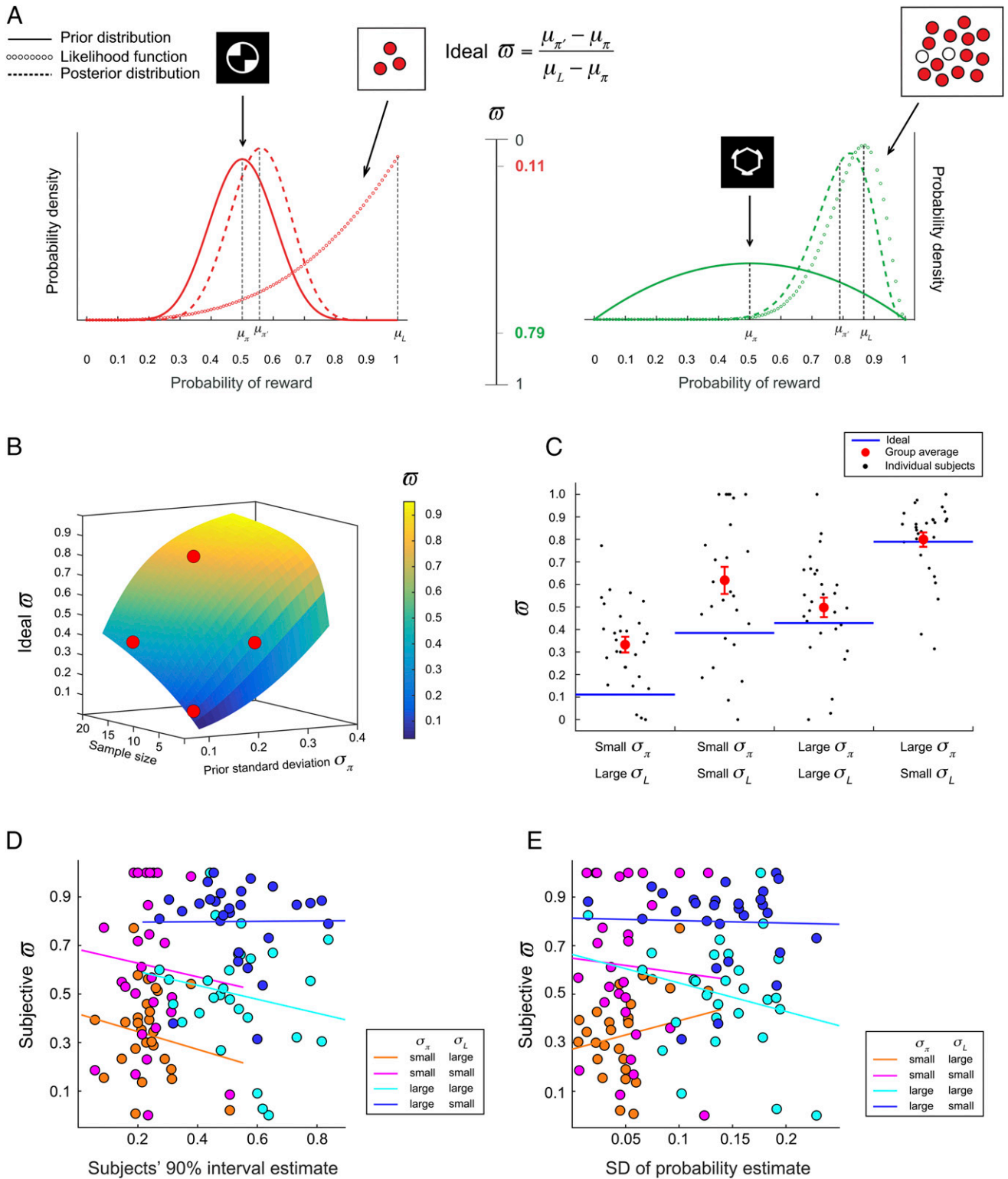
Fig. 3. Behavioral results: prior-likelihood integration session. (*A*) Two examples illustrating relative-weight computation of the ideal Bayesian decision maker based on prior and likelihood information about probability of reward (referred to as the ideal weight or ideal ϖ). (*B*) Landscape of the ideal weight plotted as a function of the variability of prior information (SD of prior distribution, σ_π and variability of likelihood information (sample size). The four red dots indicate the combinations of prior and likelihood variability used in this study and their corresponding ideal weights. (*C*) Estimated subjective weight compared with the ideal weight (blue lines). Data points in black indicate individual participants' subjective weight; data points in red indicate the mean of subjective weight averaged across subjects. Error bars represent ±1 SEM. (*D*) Subjective weight plotted against subjects' estimated 90% interval of reward probability from the prior-learning session. (*E*) Subjective weight plotted against the SD of subjects' probability estimates from the prior-learning session.

An alternative interpretation of the base rate neglect observed here is that the subjects misrepresented prior variability because she or he forgot about the prior distributions, in particular, the variability of the distributions, learned the day before in the prior-learning session. To rule out this possibility, we performed a behavioral control experiment (experiment 2, $n = 28$ subjects) in which the time gap between the prior-learning session and prior-likelihood integration session was shortened to just 1 h and replicated the results of the original experiment (experiment 1). The subjects showed the same pattern of base rate neglect: they underweight base rate when prior variability was small but were closer to ideal Bayesian when the prior variability was large (Fig. 4A). Also consistent with experiment 1 was that subjects on average gave smaller prior-variability estimates when the prior variability was small than when the prior variability was large (Fig. 4 B and C). Subjective weight also did not significantly correlate with subjects' prior-variability estimates (Fig. 4B). However, subjective weight tended to be positively correlated with another measure of prior variability, the SD of subjects' trial-by-trial probability estimates in the prior-learning session (Fig. 4C), indicating that subjects who showed more variation in probability estimates in the process of learning the prior distributions tended to give less weight to the base rate. Compared with end-of-block variability estimates (Fig. 4B), these results suggest that trial-to-trial variation in probability estimates from the prior-learning session might be a more sensitive and reliable measure of the subjects' beliefs about the prior variability than the end-of-block variability estimates. Although subjects still showed the same pattern of base rate neglect, results from experiment 2 indicate the possibility that shortening the gap between sessions might promote the use of knowledge about prior variability in the prior-likelihood integration session.

In summary, the participants did change subjective weight in response to prior and likelihood variability in the direction consistent with Bayesian integration. However, subjects showed robust suboptimal integration by underweighting the base rate. Such underweighting was clearly seen when prior variability was small (when prior information should be trusted more because of its small variation) but not when prior variability was large. In other words, the variability of prior information, but not likelihood variability, significantly affected the degree to which the subjects underweight base rate. When prior variability was small, the subjects significantly underweight base rate. By contrast, when prior variability was large, subjects were closer to the ideal Bayesian.

Neural Representations for Subjective Weight. We found regions including the lateral OFC (lOFC), mPFC, and dorsal anterior cingulate cortex significantly correlated with subjective weight (Fig. 5A and *SI Appendix, SI Methods: Group-Level Covariate Analysis 1 under GLM-1* and Tables S1 and S2). To identify regions that represent subjective weight, we fit a general linear model (GLM) to BOLD response (*SI Appendix, SI Methods: GLM-1*) where for each subject separately, the average activity (across trials) in response to each condition (a combination of prior and likelihood variability) was estimated. A group-level covariate analysis was then performed to examine the correlation between the subject- and condition-specific subjective weights and their corresponding average brain activity. In a different analysis that examined both the effects of the average subjective weight (across subjects; computed for each condition separately) and the individual differences in subjective weight (the deviation from the group average of subjective weight separately computed for each condition), we found that the posterior parietal cortex and dorsolateral prefrontal cortex represented the group average subjective weight, while mPFC and lOFC represented the individual differences in subjective weight (Fig. 5B and *SI Appendix, SI Methods: Group-Level Covariate Analysis 2 under GLM-1* and Tables S3 and S4).

Further, leave-one-subject-out (LOSO) region-of-interest (ROI) analysis confirmed that mPFC and lOFC positively correlated with the individual differences in subjective weight across different conditions (Fig. 5 D and E) and that lOFC represented both the group average (Fig. 5C) and individual differences (Fig. 5D) in subjective weight. Since base rate neglect was identified based on subjective weight, these results indicate that lOFC and mPFC contribute to base rate neglect through subjective-weight computations.

Medial Superior Frontal Cortex Represents Variability of Prior and Likelihood Information. We found that many brain regions correlated, either positively or negatively, with likelihood variability (Fig. 6A and *SI Appendix, SI Methods: GLM-2 and Tables S5 and S6*). By contrast, only the occipital cortex positively correlated with prior variability (Fig. 6A and *SI Appendix, SI Methods: GLM-2 and Table S5*). However, it is challenging to interpret these findings because both positive and negative correlations carry potentially important information relevant to base rate neglect. To tackle this problem, we used subjective weight as a constraint when examining these correlations. From behavior, we know how subjective weight changed in response to prior and
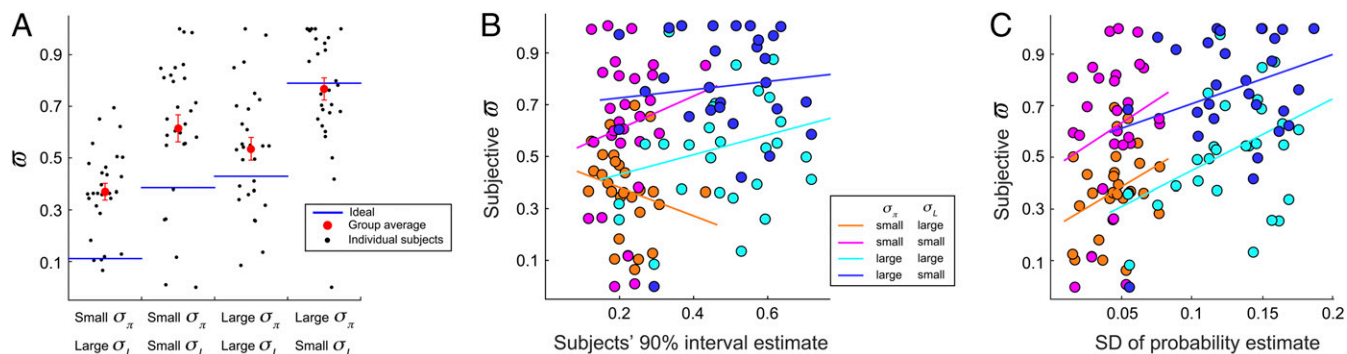
Fig. 4. Behavioral control experiment (experiment 2). Different from the fMRI experiment (experiment 1), in this experiment the time gap between prior-learning session and prior-likelihood integration session was shortened to 1 h. Conventions are the same as in Fig. 3 C–E. (A) Estimated subjective weight compared with the ideal weight (blue lines). Data points in black indicate individual participants' subjective weight; data points in red indicate mean subjective weight (across subjects). Error bars represent ±1 SEM. (B) Subjective weight plotted against subjects' end-of-block 90% interval estimate of reward probability from the prior-learning session. (C) Subjective weight plotted against the SD of subjects' trial-by-trial probability estimates from the prior-learning session. Individual subjects' SD was significantly correlated with the subjective weight in the condition where both prior variability and likelihood variability were large (cyan; $r = 0.41$, $P = 0.03$). The correlation was not significant in the other three conditions.
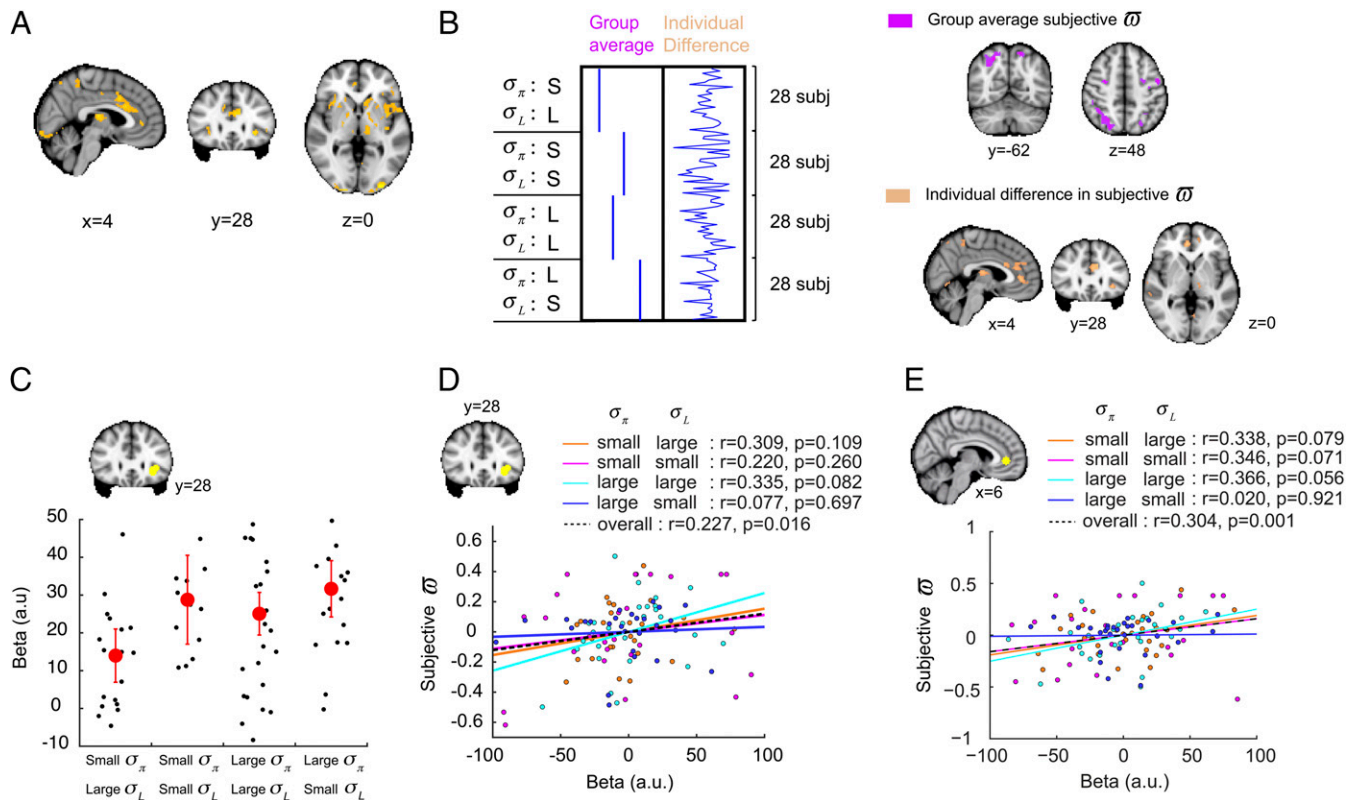
Yang and Wu

Fig. 5. Neural representations for subjective weight. (*A* and *B*) Whole-brain results. (*A*) Activity in regions that correlated with subjective weight. (*B*) Representations of mean and individual variations in subjective weight. On the left, we illustrate the group-level covariate analysis that implemented both the average subjective weight (across subjects for each condition) and individual differences in subjective weight (individual subjects' deviation from the condition average) as parametric regressors. Regions in the posterior parietal cortex and dorsolateral prefrontal cortex correlated with the group average subjective weight, while regions including lOFC and mPFC correlated with individual differences in subjective weight. (*C*) Independent LOSO ROI analysis on lOFC. Mean activity (beta; average across subjects) of each condition is plotted (data points in red; error bars represent ±1 SEM). The activity pattern closely resembles the behavioral result shown in Fig. 3*C*. Each data point in black represents an individual subject in a condition. (*D* and *E*) lOFC and mPFC ROIs represented individual participants' subjective weight. We plot subjective weight against brain activity in lOFC (5D) and mPFC (5E). Each data point represents a single subject's data in a condition. Data points from different conditions are coded by different colors. The correlation between brain activity and subjective weight is shown for each condition separately. The overall correlation indicates the correlation computed using all data points.

likelihood variability (Fig. 3*C*): prior variability positively correlated with subjective weight, while likelihood variability negatively correlated with it (left graph in Fig. 6*B*). This is indicated by subjective weight becoming larger when prior variability increased and when likelihood variability decreased (sample size from 3 dots to 15 dots). Hence, we examined regions that positively correlated with prior variability but negatively correlated with likelihood variability. A LOSO ROI analysis was performed on all regions that negatively correlated with likelihood variability at the whole-brain level (blue in Fig. 6*A*) to examine whether they also positively correlated with prior variability (*SI Appendix*, Table S7). We identified a region in the left medial superior frontal cortex (mSFC; anterior portion of BA 8) that fit this criterion (*SI Appendix, SI Methods: Independent ROI Analysis*). That is, this region positively correlated with prior variability but negatively correlated with likelihood variability (right graph in Fig. 6*B*). However, it should be noted that although mSFC showed positive correlation with prior variability, the significant difference in activity between prior and likelihood variability was primarily driven by the strong negative likelihood-variability coding and that after Bonferroni correction for the number of ROIs tested ($n = 46$), the positive correlation with prior variability was no longer statistically significant.

A Functional Network for Prior-Likelihood Integration on Reward Probability. We found that a potential mechanism involving information-selective (prior instead of likelihood) and variability-

dependent (prior variability) functional connectivity between mSFC (variability coding) and mPFC (subjective-weight coding) might be at play to contribute to base rate neglect. We hypothesized that in the context of our task, mSFC, mPFC, and lOFC are part of a functional network that computes subjective weight by using information about prior and likelihood variability. To test this hypothesis, we performed two psychophysiologic interaction (PPI) analyses (17) using mSFC, shown to positively correlate with prior variability and negatively correlate with likelihood variability, as the seed region. In the first PPI analysis, we found that mPFC showed an overall increase in functional connectivity with mSFC at the inference stage of the trial (Fig. 6*C*; lOFC was marginally significant at $P = 0.059$; *SI Appendix, SI Methods: PPI Model 1* and Tables S8 and S9). In the second PPI analysis, we found that the strength in mSFC connectivity with both mPFC and lOFC was dependent on prior variability but not on likelihood variability (Fig. 6*D* and *SI Appendix, SI Methods: PPI Model 2* and Tables S10 and S11). This is closely related to the subjects' behavior in that the degree of base rate neglect was mainly affected by prior variability but not likelihood variability (Fig. 3*C*): the subjects significantly underweight base rate only in trials where prior variability was small but was statistically indistinguishable from ideal Bayesian when prior variability was large.
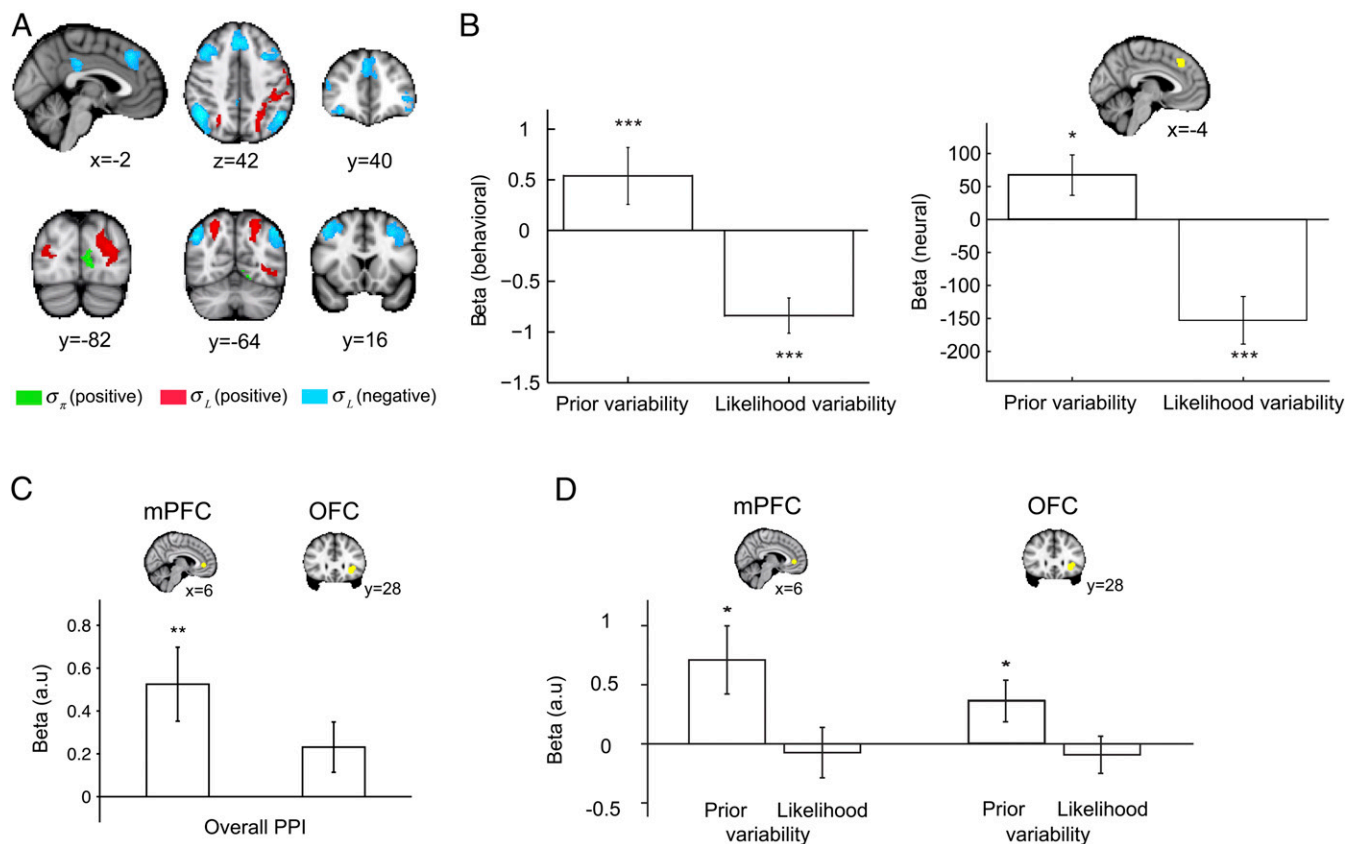
Fig. 6. mSFC represented the variability of prior and likelihood information. (A) Whole-brain results on regions that represented prior variability (in green; positive correlation) and likelihood variability (positive correlation in red; negative correlation in light blue). (B) (Left) The behavioral regression results showing that subjective weight positively correlated with prior variability [$t(27) = 4.233$, $P = 2.384\times10^{-4}$] and negatively correlated with likelihood variability [$t(27) = -8.349$, $P = 5.855\times10^{-9}$]. (Right) Results from LOSO ROI analysis in mSFC that positively correlated with prior variability [$t(27) = 2.194$, $P = 0.037$] and negatively correlated with likelihood variability [$t(27) = -4.237$, $P = 2\times10^{-4}$]. (C and D) Two PPI analyses using mSFC as the seed region. (C) PPI-1. LOSO ROI analysis examining functional connectivity between the mSFC and the ROIs that represented subjective weight (lOFC and mPFC). Beta value (vertical axis) represents the PPI contrast (labeled as the overall PPI on the horizontal axis) indicating whether there was an increase in functional connectivity (positive value) between mSFC and the subjective-weight ROIs at the inference stage of the trial. mPFC, $t(27) = 3.042$, $P = 0.005$; lOFC, $t(27) = 1.97$, $P = 0.059$. (D) PPI-2. LOSO ROI analysis examining whether functional connectivity, also at the inference stage of the trial, between the mSFC and subjective-weight ROIs was dependent on the variability of prior and likelihood information. Beta values correspond to PPI contrasts indicating the degree to which functional connectivity correlated with the prior variability [mPFC, $t(27) = 2.465$, $P = 0.02$; lOFC, $t(27) = 2.0582$, $P = 0.0493$] and likelihood variability [mPFC, $(t(27) = -0.345$, $P = 0.73$; lOFC, $t(27) = -0.589$, $P = 0.561$]. Positive values indicate positive correlations. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

Relative Sensitivity to Likelihood Variability in Putamen Is Associated with Overall Tendency to Underweight Base Rate. We found that the putamen contributes to individual subjects' overall tendency to underweight base rate through representing the relative sensitivity to information variability. To further explore the asymmetry in coding likelihood variability and prior variability, with many more regions representing likelihood variability than prior variability, we examined whether the relative sensitivity of brain activity in response to prior and likelihood variability is associated with individual subjects' overall tendency to underweight base rate. The relative sensitivity measure of brain activity, $\kappa = \beta_{\sigma_L} - \beta_{\sigma_\pi}$, reflects the difference in regression coefficients between likelihood variability (β_{σ_L}) and prior variability (β_{σ_π}). For each subject separately, we obtained a whole-brain map of κ. Also for each subject separately we obtained a behavioral measure of the overall tendency to underweight base rate (δ), defined as the sum of deviation of subjective weight from the ideal weight over the four conditions (2 prior variability × 2 likelihood variability), $\delta = \sum_{i=1}^{4}(subjective\ \varpi_i - ideal\ \varpi_i)$, where i represents condition. Larger values of δ indicate greater tendency (over all conditions) to underweight the base rate or equivalently, to

overweight the likelihood of reward. At the whole-brain level, we did not find any brain region whose relative sensitivity κ significantly correlated with δ. We subsequently performed the same analysis on the subjective-weight ROIs in mPFC and lOFC and found that they also did not represent δ.

We then examined putamen, previously shown to represent individual subjects' sensitivity to prior variability in a visuomotor decision task (11), and found that subjects who tended to underweight base rate more (across different conditions) showed greater sensitivity to likelihood variability relative to prior variability in putamen activity. Using anatomically defined putamen as ROI (Harvard–Oxford subcortical structural atlas), we found that its sensitivity to likelihood variability relative to prior variability (κ) positively correlated with δ (Fig. 7 A and B). Further, we performed the subjective-weight analysis (as in Fig. 5 D and E) and the information-variability analysis (as in Fig. 6B) on the putamen ROIs and found that it significantly represented individual differences in subjective weight but not information variability (*SI Appendix*, Fig. S5).

We also found that the putamen exhibited the same functional connectivity profiles with mSFC as mPFC and lOFC did with mSFC. Similar to what we found in Fig. 6D, functional connectivity
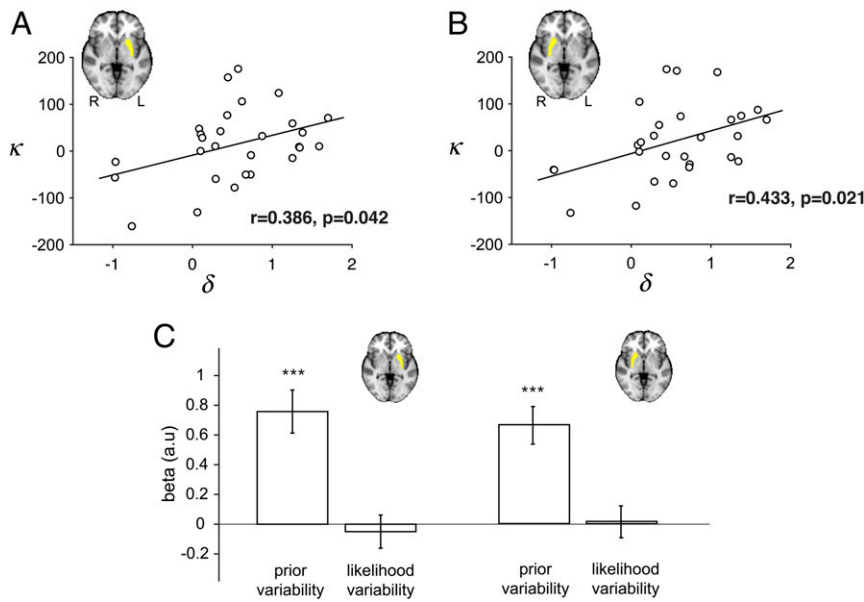
Fig. 7. Sensitivity to likelihood variability relative to prior variability in putamen activity (κ) represented individual subjects' overall tendency to underweight base rate (δ). (A) ROI results in the left putamen. Each data point represents a single subject. (B) ROI results in the right putamen. (C) Information-selective and variability-dependent functional connectivity profiles between putamen and mSFC (seed region). The beta value corresponding to the "prior variability" label and the "likelihood variability" label indicates the degree to which functional connectivity correlated with prior variability [left putamen, $t(27) = 5.232, P = 1.63\times10^{-5}$; right putamen, $t(27) = 5.253, P = 1.54\times10^{-5}$] and likelihood variability [left putamen, $t(27) = -0.449, P = 0.657$; right putamen, $t(27) = 0.148, P = 0.883$], respectively. Positive value indicates positive correlation. ***$P < 0.001$.

between putamen and mSFC that represented prior and likelihood variability also showed variability-dependent modulation selective to prior information (Fig. 7C). These findings indicate that a potential mechanism involving information-selective (prior but not likelihood) and variability-dependent (prior variability) functional connectivity between mSFC and putamen, a region that represented relative sensitivity to likelihood variability, might contribute to people's overall tendency to underweight base rate.

Decision Time Analysis. It is possible that our subjects already made a decision on which lottery—symbol or alternative lottery—to choose at the inference stage before the alternative lottery appeared at the choice stage. If this were the case, the fMRI results presented above, examining activity at the inference stage, might be driven by choice-related computations and thus did not purely reflect probabilistic inference. To address this issue, we analyzed the subjects' decision time as a function of decision difficulty and hypothesized that if the subjects already made a decision at the inference stage, decision time should vary little as a function of decision difficulty. Here we define decision difficulty, $\Delta_D = |\hat{\theta}_{sym} - \theta_{alt}|$, to be the absolute difference between reward probability of the alternative lottery (θ_{alt}) and subjective posterior reward probability of the symbol lottery ($\hat{\theta}_{sym}$) computed by $\hat{\theta}_{sym} = \varpi\mu_L + (1 - \varpi)\mu_\pi$, where ϖ is the subjective weight, μ_L is the likelihood of reward (proportion of red dots), and μ_π is the base rate of reward probability (0.5; the mean of the prior distributions). In principle, the closer Δ_D is to 0, the more difficult it is for subjects to make a decision. Hence, decision difficulty should increase as Δ_D approaches 0 and should decrease as Δ_D further deviates from 0. We found that decision time increased as a function of decision difficulty (red in Fig. 8A): subjects spent more time making decisions as Δ_D approached 0. This pattern persisted when we separately analyzed the early and late part of the fMRI session (green and blue, respectively, in Fig. 8A) and when we separately analyzed different conditions (Fig. 8B). To conclude, although we cannot completely rule out the possibility that subjects already made a decision at the inference stage, these results indicate that our design was effective in discouraging subjects from committing to a choice at the inference stage.

At the choice stage, no brain region significantly represented the chosen value at the whole-brain level. Using anatomically

defined ROI in OFC (18), we found that the left medial OFC represented the subjective value of the chosen option (*SI Appendix, SI Methods: GLM-3*), consistent with previous studies (19, 20). Together with subjective-weight representations found in lOFC at the inference stage, these results indicate that OFC is involved in both probabilistic-inference (lOFC) and choice computations (mOFC) in decision under uncertainty (Fig. 8C).

Discussion

Humans often exhibit systematic biases in probabilistic inference, an essential computation for making decisions under uncertainty. In this study, we investigated the neurocomputational basis of base rate neglect, an important bias in probabilistic inference, in which people underweight base rate or prior information (21). At the behavioral level, we found that the degree to which humans underweight base rate was modulated by the variability of prior information: subjects significantly underweight the base rate of reward probability when they should trust the prior more, i.e., when prior variability was small, but the degree of underweighting was statistically indistinguishable from the ideal Bayesian when prior variability was large. This result suggests that it is the prior variability, not likelihood variability, that is the key statistical attribute contributing to base rate neglect. At the computational and neural implementation levels, we found that the OFC, mPFC, and putamen represented the relative subjective weight the participants assigned to the likelihood information that reflected base rate neglect, suggesting that base rate neglect arises from information-weighting computations in these brain regions.

Methodological Concerns for Base Rate Neglect. Since Kahneman and Tversky's landmark papers (1, 2), issues surrounding base rate neglect had been extensively discussed in the human judgment and decision making literature (22, 23). Grether (24) pointed out methodological concerns, highlighting the difficulty in controlling information presented as verbal descriptions or situations, the fact that subjects were not told the truth about the random process being examined, and that it was not clear that subjects had a positive incentive to give correct answers (incentive compatibility). In a task designed to address these concerns, Grether (24) found that subjects still showed robust underweighting base rate, although they did not completely neglect it as highlighted in the original findings. Our experimental design followed closely the approach of Grether
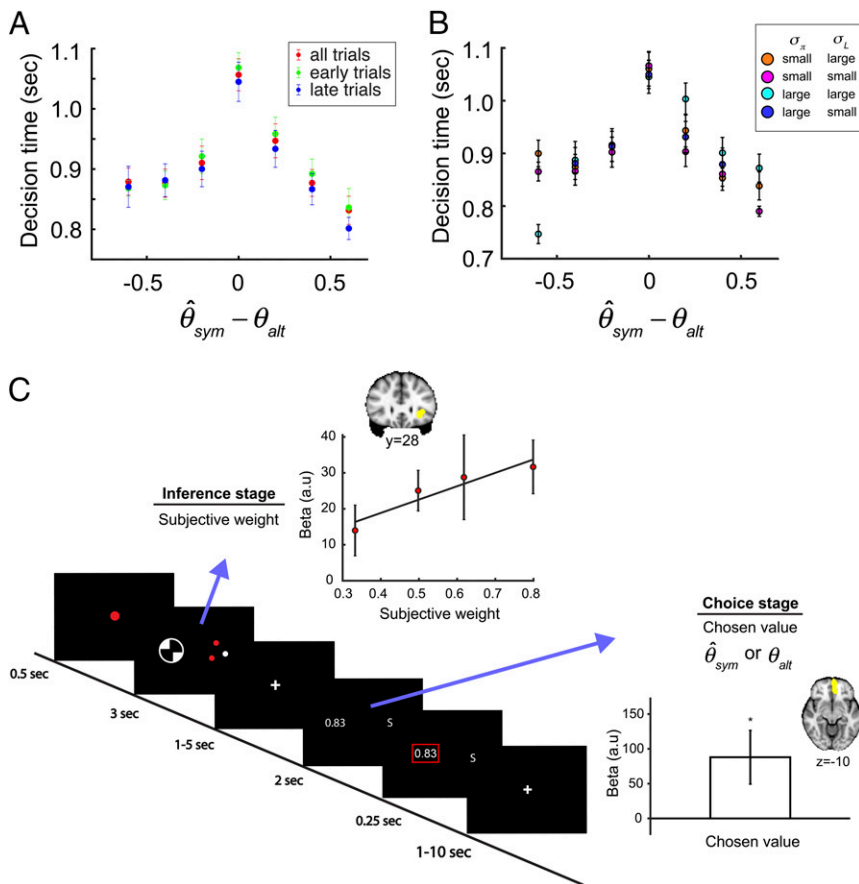
Fig. 8. Decision time analysis. (*A*) Average decision time (across subjects) plotted as a function of the difference between subjective posterior probability of reward associated with the symbol lottery ($\hat{\theta}_{sym}$) and the reward probability of the alternative lottery (θ_{alt}). Green indicates decision time computed based on data from the first half of the fMRI session, blue indicates decision time computed based on data from the second half of the fMRI session, and red indicates decision time computed based on data from the entire fMRI session. Error bars represent ±1 SEM. (*B*) Subjects' mean decision time plotted separately for each condition as a function of $\hat{\theta}_{sym} - \theta_{alt}$. (*C*) OFC represented both probabilistic-inference computations and choice computations. At the inference stage, lOFC represented the subjective weight (red line indicates the linear regression fit). At the choice stage, mOFC represented chosen value [$t(27) = 2.274, P = 0.0312$]. *$P < 0.05$.

(24), and our results were consistent with what he found. In addition, our results indicate that the degree of underweighting the base rate is associated with the variability of prior information.

A potential limitation of the current study concerns the asymmetry in stimulus design between prior and likelihood information. Compared with prior information (each prior distribution was represented by a symbol icon), likelihood information was visually more capturing: colored dots with varying number to represent different levels of variability. For future studies, it is important to investigate the issue of information asymmetry and how it contributes to base rate neglect and the neural computations for subjective weight.

The Issue of Confidence in Probability Estimation. An alternative explanation for the subjective-weight findings in OFC, mPFC, and putamen is that these regions, instead of representing subjective weight, represented the level of confidence the subjects had in his or her probability estimates. Here the subjects might in general feel more confident if his or her estimates of reward probability (subjective posterior) were farther away from the base rate of reward probability (fixed at 0.5 throughout the experiment). In this case, a potential measure of confidence is the absolute deviation of subjective posterior from the base rate (|subjective posterior – 0.5|), which significantly correlates with subjective weight (*SI Appendix*, Fig. S3), and with prior and likelihood variability (*SI Appendix*, Fig. S4). Because of the significant correlations, it was challenging to tease apart confidence from subjective weight and from prior/likelihood variability.

We, however, believe that the confidence level is less likely to explain the findings described above for the following reasons. First, we found that after controlling for confidence, both the mPFC and lOFC still significantly represented individual differences in subjective weight (*SI Appendix*, Fig. S3 *B* and *C*). Second, subjects' confidence in probability estimates might be more complicated to capture than simply using the above definition. Consider the following example. It could be the case that the subjective posterior in a trial is close to 0.5 (e.g., when the symbol lottery comes from the small-variability prior distribution), and yet subjects are confident in his or her estimates because in that trial, the likelihood information was reliable due to the large sample size (e.g., 15 dots). Third, in contrast to subjective weight, it is unclear whether and how this definition of confidence contributes to subjects' choice behavior given that the alternative lottery was not fixed at 0.5 (randomly selected between 0.01 and 0.99). Finally, because we did not explicitly ask the subjects to report his or her estimation confidence during the experiment, it is hard to validate any potential measure of confidence, including the one described here.

By contrast, the definition of subjective weight is clear. It can be estimated from subjects' choice behavior, and it contributes to the subjects' decisions through the computation of subjective posterior. Nonetheless, it is important for future studies to investigate the neural representations of people's level of confidence in probability estimation and how it relates to the subjective-weight representations and information about prior and likelihood variability. This can be addressed through task design, for example, by asking the subjects to directly report confidence level in probability estimates and by manipulating both the mean and variance of the prior distributions.

Neural Computations for Base Rate Neglect. Our findings suggest that base rate neglect arises from information-weighting computations. Subjective weight, a behavioral metric that reflects the weight

individual participants assigned to the likelihood relative to prior information, allowed us to quantitatively define the degree to which subjects underweight or overweight base rate compared with the ideal Bayesian decision maker. We found that brain regions including OFC, mPFC, and putamen represented the subjective weight that characterized the individual differences in the degree of base rate neglect. This finding adds to a growing body of literature on the neural representations for prior, likelihood, and posterior information in probabilistic/statistical inference and decision tasks (7–11) by providing insights into the neural computations that give rise to the biases commonly observed in probabilistic inference.

Our findings also suggest that sensitivity of brain activity to information variability correlated with base rate neglect. We showed that putamen activity—its sensitivity in response to likelihood variability relative to prior variability—positively correlated with an individual subject's overall tendency to underweight base rate. This finding complements the subjective-weight findings in mPFC and OFC by highlighting the role of information-variability representations in contributing to base rate neglect. This finding also connects with a previous study by Vilares et al. (11) showing that putamen represents individual differences in weighting likelihood relative to prior information. Both studies suggest that sensitivity in putamen activity to information variability plays a crucial role in computing the subjective weight of likelihood relative to prior information.

Finally, our findings indicate that patterns of functional connectivity between variability-coding regions and subjective-weight regions modulated the degree of base rate neglect. In behavior, we found that the degree of base rate neglect depended primarily on the prior variability but not likelihood variability. Our fMRI results showed that mSFC (anterior portion of BA 8) represented both the prior and likelihood variability; mSFC is a region previously shown to be involved in task switching and selection of action sets, representing the degree of uncertainty in decision making (25–27), and is heavily connected with prefrontal and subcortical regions including OFC, mPFC, and putamen (28). We found that the strength in functional connectivity between mSFC and the subjective-weight regions was modulated by only the prior variability but not likelihood variability. This suggests that a mechanism involving information-selective (prior instead of likelihood) and variability-dependent (prior variability) functional connectivity between these regions plays a key role in affecting base rate neglect. This finding also connects with a previous study showing that uncertainty-dependent modulation of functional connectivity between ventromedial prefrontal cortex (vmPFC) and rostrolateral prefrontal cortex affected decision confidence (29). Together, these results indicate the possibility that the strength in functional connectivity can modulate how likely a piece of information is used by neural systems involved in combining different sources of information in probability estimation and decision making.

Implications to Value-Based Decision Making. Subjective-weight computation is essential when humans and animals face multiple sources of information and attempt to integrate them. The output of this computation reflects the degree to which a source of information or an attribute is weighted by the decision maker and therefore influences the summary statistic or the overall desirability of an option (subjective value). Hence, subjective-weight computation is critical not only to probabilistic inference, the focus of this study, but also to many decision problems that involve combining different sources of information or attributes. Although many studies had shown subjective-value representations in mPFC and OFC in value-based decision making (14, 30), few studies found subjective-weight representations. One notable exception is ref. 31, who found vmPFC to represent both subjective value and subjective weight associated with different food attributes (31). In addition to subjective-value representations, other studies also showed that OFC represents different statistics (32) or attributes of value

information (32, 33). Together, these findings highlight the rich representations for inference- and decision-related variables in mPFC/vmPFC and OFC and their involvement in combining multiple sources of information in probabilistic inference and decision making.

Materials and Methods

The data and analysis code are available at https://osf.io/ku97p/.

We performed one fMRI experiment (experiment 1) and one behavioral control experiment (experiment 2). The design of the behavioral control experiment was identical to the fMRI experiment except that the gap between session 1 and session 2 was shortened to 1 h and that session 2 was performed in a behavioral testing room.

Subjects. For the fMRI experiment (experiment 1; $n = 28$ subjects; 14 males; mean age, 24.6 y; age range, 21 to 30 y), subjects were paid 620 New Taiwan dollar (NTD; 1 US dollar = 30 NTD) for their participation (NTD 500 for the behavioral session and 120 for the fMRI session) and additional monetary bonus (average, 383 NTD) based on their performance in the experiment. For the behavioral control experiment (experiment 2; $n = 28$ subjects; 14 males; mean age, 23 y; age range, 21 to 31 y), subjects were paid 240 NTD for their participation (for two behavioral sessions) and an additional monetary bonus (average: 406 NTD) based on their performance in the experiment. All participants had no psychiatric or neurological disorders and gave written informed consent prior to participation; all study procedures were approved by the Taipei Veterans General Hospital Institutional Review Board (experiment 1) and by the National Yang-Ming University Institutional Review Board (experiment 2).

Procedure. There were two sessions in both experiments. Session 1 was conducted in a behavioral testing room. Session 2 was conducted in the MRI scanner (experiment 1) and in a behavioral testing room (experiment 2). The tasks were programmed using the Psychophysics Toolbox in MATLAB (34, 35).

Session 1: Learning Prior Distributions. The goal of the session was to establish knowledge about the probability of reward associated with different visual stimuli. There were two visual stimuli, each representing a unique probability density function on probability of reward. Both were beta distributions with two parameters α and β. Critically, we manipulated the variability of prior knowledge by varying the variance of the density functions while keeping the mean fixed. The SDs of the two prior distributions (σ_π) were 0.1, ($\alpha = 12, \beta = 12$), and 0.2236, ($\alpha = 2, \beta = 2$). The means of both distributions were 0.5, indicating that both stimuli had an average of 50% chance to receive a fixed monetary reward. Prior to the experiment, the subjects did not know about the probability distributions associated with the two stimuli and had to acquire this knowledge through experience. There were 10 blocks of trials, each consisting of 30 trials. In each block, only one of the two visual stimuli was presented. The small-σ_π stimulus was presented in five blocks, and the large-σ_π stimulus was presented in the other five blocks. The ordering of the blocks was semirandomized so that subjects encountered no more than two successive blocks with the same distribution.

In each trial, the subjects were instructed to estimate the probability of reward associated with the presented visual stimulus with key presses (an integer from 0 to 100 where 100 represents 100% reward), which was sampled from its corresponding probability density function. They were rewarded based on how close his or her estimate was to the true reward probability—reward probability sampled from the density function—in an attempt to motivate them to learn the probability distributions (see *SI Appendix, SI Methods,* for more details).

Session 2: Integrating Prior and Likelihood Information. The goal of the session was to investigate how subjects combined prior knowledge, established through session 1, with likelihood information about probability of reward. In each trial, the subjects were asked to choose between two lotteries that differed only in the probability of receiving a small monetary reward. The magnitude of reward associated with both options was the same and fixed throughout the experiment, so that subjects should make their decisions based on which option carried a larger probability of reward. For one of the lottery options, referred to as the symbol lottery, its probability of reward was not explicitly stated in numeric or graphical forms. The subjects therefore needed to infer the probability based on two pieces of information: prior and likelihood. In each trial, the prior information was represented by one of the two visual stimuli (a symbol icon) that the subjects encountered in session 1. Each stimulus represented a probability density function on probability of reward. Subjects were told that the probability of reward associated with the

symbol lottery was sampled from the density function. Meanwhile, the likelihood information was represented by a set of colored dots (red or white). Given the probability of reward associated with the symbol lottery in the current trial, the dots summarized the sample drawn from it. Each red dot represented a reward outcome, and each white dot represented a no-reward outcome. Hence, the proportion of red dots indicates the likelihood of reward. Note that if the sample size were infinitely large, the proportion of red dots would be equal to the reward probability of the symbol lottery in that trial.

We manipulated the variability of the likelihood information by varying the sample size (number of dots, 3 or 15) used to draw from the probability of reward. Together with the manipulation of the variance of prior distribution, we achieved a 2 (prior variability, σ_π, small and large) × 2 (likelihood variability, σ_L, small and large) factorial design. We use σ_π and σ_L to denote the SD of the prior distribution and the likelihood function, respectively. We refer to each combination of prior and likelihood variability as a condition. The average SD of the likelihood function for the smaller sample size (3 dots) and larger sample size (15 dots) was 0.2722 and 0.1205, respectively. These values are close to the SD of the prior information (0.2236 and 0.1). As a result, the difference in the range of variance between the prior and likelihood information was controlled.

In each trial, the prior and likelihood information were presented on the left and right side of the screen, with the locations randomized across trials. Following the presentation of the symbol lottery, there was a fixation period (1 to 5 s, discrete uniform distribution in steps of 1 s). This was followed by the presentation of the second lottery, also referred to as the alternative lottery. Information about its probability of reward was explicitly revealed in numeric form and determined randomly (*SI Appendix, SI Methods*, for more details).

When the alternative lottery was presented, subjects were instructed to choose between the symbol lottery (indicated by S) and the alternative lottery (a number) within 2 s. The location of the two lotteries (left or right) was randomized and balanced across trials. Once the subjects indicated his or her decision with a button press, the chosen option was revealed on the screen for 250 ms. No feedback on the reward outcome of the chosen option was revealed. This was to prevent the subjects from updating knowledge about the prior distribution associated with the visual stimulus through feedback and from learning how to integrate prior and likelihood information through feedback.

There were six blocks in the session. Each block had 40 trials. Each combination of the prior symbol icon (high prior variability, low prior variability) and sample size (3 dots, 15 dots) for the likelihood information had 10 trials in each block. The order of the trials was randomized.

fMRI Analysis. The GLMs of BOLD response can be seen in *SI Appendix*. All whole-brain fMRI results were based on cluster-level inference and family-wise error corrected for multiple testing at $P < 0.05$. Two procedures were implemented. First, for cluster-level inference using Gaussian random field theory, we used $z > 3.1$ ($P < 0.001$) as the cluster-forming threshold (36). Second, for cluster-level inference based on nonparametric permutation test, we used the threshold-free-cluster-enhancement procedure (37). Tables showing significant clusters of activation can be seen in *SI Appendix*.

1. D. Kahneman, A. Tversky, On the psychology of prediction. *Psychol. Rev.* **80**, 237 (1973).
2. A. Tversky, D. Kahneman, Judgment under uncertainty: Heuristics and biases. *Science* **185**, 1124–1131 (1974).
3. W. Edwards, H. Lindman, L. J. Savage, Bayesian statistical inference for psychological research. *Psychol. Rev.* **70**, 193–242 (1963).
4. B. U. Forstmann, S. Brown, G. Dutilh, J. Neumann, E.-J. Wagenmakers, The neural substrate of prior information in perceptual decision making: A model-based analysis. *Front. Hum. Neurosci.* **4**, 40 (2010).
5. M. J. Mulder, E.-J. Wagenmakers, R. Ratcliff, W. Boekel, B. U. Forstmann, Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *J. Neurosci.* **32**, 2335–2343 (2012).
6. E. Fouragnan *et al.*, Reputational priors magnify striatal responses to violations of trust. *J. Neurosci.* **33**, 3602–3611 (2013).
7. M. d'Acremont, W. Schultz, P. Bossaerts, The human brain encodes event frequencies while forming subjective beliefs. *J. Neurosci.* **33**, 10887–10897 (2013).
8. C.-C. Ting, C.-C. Yu, L. T. Maloney, S.-W. Wu, Neural mechanisms for integrating prior knowledge and likelihood in value-based probabilistic inference. *J. Neurosci.* **35**, 1792–1805 (2015).
9. S. C. Y. Chan, Y. Niv, K. A. Norman, A probability distribution over latent causes, in the orbitofrontal cortex. *J. Neurosci.* **36**, 7817–7828 (2016).
10. V. Chambon *et al.*, Neural coding of prior expectations in hierarchical intention inference. *Sci. Rep.* **7**, 1278 (2017).
11. I. Vilares, J. D. Howard, H. L. Fernandes, J. A. Gottfried, K. P. Kording, Differential representations of prior and likelihood uncertainty in the human brain. *Curr. Biol.* **22**, 1641–1648 (2012).
12. O. Bartra, J. T. McGuire, J. W. Kable, The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* **76**, 412–427 (2013).
13. J. A. Clithero, A. Rangel, Informatic parcellation of the network involved in the computation of subjective value. *Soc. Cogn. Affect. Neurosci.* **9**, 1289–1302 (2014).
14. J. W. Kable, P. W. Glimcher, The neurobiology of decision: Consensus and controversy. *Neuron* **63**, 733–745 (2009).
15. C. Padoa-Schioppa, K. E. Conen, Orbitofrontal cortex: A neural circuit for economic decisions. *Neuron* **96**, 736–754 (2017).
16. D. Griffin, A. Tversky, The weighing of evidence and the determinants of confidence. *Cognit. Psychol.* **24**, 411–435 (1992).
17. K. J. Friston *et al.*, Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* **6**, 218–229 (1997).
18. N. Tzourio-Mazoyer *et al.*, Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273–289 (2002).
19. T. Larsen, J. P. O'Doherty, Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: A combined fMRI-EEG study. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130473 (2014).
20. K. Wunderlich, A. Rangel, J. P. O'Doherty, Economic choices can be made using only stimulus values. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15005–15010 (2010).
21. D. Kahneman, P. Slovic, A. Tversky, *Judgment under Uncertainty: Heuristics and Biases* (Cambridge University Press, Cambridge, 1982).
22. G. Gigerenzer, W. Hell, H. Blank, Presentation and content: The use of base rates as a continuous variable. *J. Exp. Psychol. Hum. Percept. Perform.* **14**, 513–525 (1988).
23. J. J. Koehler, The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behav. Brain Sci.* **19**, 1–17 (1996).
24. D. M. Grether, Bayes rule as a descriptive model: The representativeness heuristic. *Q. J. Econ.* **95**, 537–557 (1980).
25. K. R. Ridderinkhof, M. Ullsperger, E. A. Crone, S. Nieuwenhuis, The role of the medial frontal cortex in cognitive control. *Science* **306**, 443–447 (2004).
26. K. G. Volz, R. I. Schubotz, D. Y. von Cramon, Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. *Neuroimage* **21**, 848–857 (2004).
27. K. G. Volz, R. I. Schubotz, D. Y. von Cramon, Variants of uncertainty in decision-making and their neural correlates. *Brain Res. Bull.* **67**, 403–412 (2005).
28. S. Zhang, J. S. Ide, C. S. Li, Resting-state functional connectivity of the medial superior frontal cortex. *Cereb. Cortex* **22**, 99–111 (2012).
29. B. De Martino, S. M. Fleming, N. Garrett, R. J. Dolan, Confidence in value-based choice. *Nat. Neurosci.* **16**, 105–110 (2013).
30. C. Padoa-Schioppa, Neurobiology of economic choice: A good-based model. *Annu. Rev. Neurosci.* **34**, 333–359 (2011).
31. T. A. Hare, C. F. Camerer, A. Rangel, Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* **324**, 646–648 (2009).
32. M. O'Neill, W. Schultz, Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* **68**, 789–800 (2010).
33. T. C. Blanchard, B. Y. Hayden, E. S. Bromberg-Martin, Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* **85**, 602–614 (2015).
34. D. H. Brainard, The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
35. D. G. Pelli, The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat. Vis.* **10**, 437–442 (1997).
36. A. Eklund, T. E. Nichols, H. Knutsson, Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7900–7905 (2016).
37. S. M. Smith, T. E. Nichols, Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* **44**, 83–98 (2009).