Title
Neurocognitive mechanisms of uncertainty reduction in value-based decision making

Author
Kobayashi, Kenji

Publication Date
2016

Peer reviewed|Thesis/dissertation

Neurocognitive Mechanisms of Uncertainty Reduction in Value-based Decision Making

By

Kenji Kobayashi

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Neuroscience

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ming Hsu, Chair
Professor Richard Ivry
Professor Jonathan Wallis
Professor Don A. Moore

Fall 2016

Abstract

Neurocognitive Mechanisms of Uncertainty Reduction in Value-based Decision Making

by

Kenji Kobayashi

Doctor of Philosophy in Neuroscience

University of California, Berkeley

Professor Ming Hsu, Chair

Adaptive decision making critically depends on agents' ability to reduce uncertainty. To reduce uncertainty, agents need to perceive relevant environmental signals, extract useful information, and incorporate it into their knowledge and behavior. This is not a trivial task for two reasons. First, the environment is full of sensory stimuli, but most of them are irrelevant. Agents need to evaluate relevance of signals so that they can prioritize perception and processing of the most relevant ones. Second, perceived signals can be analyzed in countless ways, but most of them do not provide useful information. Agents need to be selective in how to process perceived signals before trying to make use of them in behavior.

This thesis reports three attempts to understand neurocognitive mechanisms of uncertainty reduction in value-based decision making. It is relatively well understood how human brain produces decision-making behavior under uncertainty. It is also well understood how human brain perceives and processes sensory stimuli. However, little is known how these systems cooperate to enable adaptive uncertainty reduction. Three chapters in this thesis try to fill this gap by behavioral experiments, functional magnetic resonance imaging (fMRI) experiments, and computational modeling.

Chapter 1 examines the way human brain processes perceived signals under conditions of reducible and irreducible uncertainty. Because not all uncertainty can be reduced in the same way, agents should take into account the nature of uncertainty they are facing. This challenges a widespread idea that uncertainty reduction is driven by the extent to which signals violate prior expectancy. It is behaviorally shown that subjects are sensitive to reducibility of uncertainty, and it can be quantitative characterized by a Bayesian model where agents ignore expectancy violations that do not update beliefs or values. Furthermore, fMRI results reveal that neural processes underlying belief and value updating are separable from responses to expectancy violation, and that reducibility of uncertainty in value modulates connection from belief- to value-updating regions. These results illustrate how human brain uses knowledge on uncertainty to process signals adaptively.

Chapters 2 and 3 examine the way human brain evaluates incoming signals' relevance to economic choices. Human often seeks for information that is irrelevant to the task at hand. Although this

challenges a normative account in which agents assign positive value only to instrumental information, it has been unclear how general and important such preference is in decision-making settings. Chapter 2 reveals that, behaviorally, subjective value of information exceeds the normative prediction not only when information is non-instrumental but also when it is instrumental. Overvaluation depends on outcomes at stake, rejecting a popular theory of internal drive for entropy reduction. Observed valuation bias can be explained by introducing recursive utility, which penalizes choices under the lack of information, to the normative account. Using this novel model and fMRI, Chapter 3 reports that activation in striatum represents value of information, unifying information's relevance to the task and penalty for uninformed choice. These results propose a new neurocognitive account on how human brain acquires information with various degrees of uncertainty reduction.

TABLE OF CONTENTS

To my parents and my sister.

Chapter 1
Neural mechanisms of updating under reducible and irreducible uncertainty

Introduction

Adaptive decision-making in the real world depends critically on agents' ability to constantly make use of environmental signals to reduce uncertainty. Agents should take into account not only physical properties of signals but also the nature of uncertainty, as not all kinds of uncertainty can be reduced in the same way. In particular, two types of uncertainty, risk and ambiguity, have received widespread emphasis in the decision-making literature (Camerer & Weber, 1992; Ellsberg, 1961; Keynes, 1921; Knight, 1921). While ambiguity can be reduced by signals that carry new information and supplement agents' prior knowledge about the environment, risk cannot be reduced by any signals.

To illustrate the implication of these types of uncertainty for decision-making, consider a variation on the classic example of gambler's fallacy (Tversky & Kahneman, 1974). You are playing a gamble that depends on tosses of two coins, A and B. You know that coin A is fair (risk), but you do not know whether coin B is biased or fair (ambiguity). If you observe ten consecutive "head" tosses of B, it would suggest that B is biased, but observing the same sequence of A does not provide any new information. Although these two environmental signals are similar to each other and both surprising, appropriate decision-making requires agents to use them differently because of the different natures of uncertainty.

This, however, poses a challenge for theoretical frameworks such as reinforcement learning (RL), which do not incorporate explicit notions of uncertainty. Under RL, values of actions are updated to the extent that an observed outcome violates prior expectancy, such that agents cannot ignore merely expectancy-violating outcomes under risk (Sutton & Barto, 1998). One solution to this problem, which is possible under normative Bayesian and more recent model-based RL accounts, is to posit that agents construct internal models or beliefs about the environment, which may be sensitive to the nature of uncertainty (Behrens, Woolrich, Walton, & Rushworth, 2007; Itti & Baldi, 2009; Ma & Jazayeri, 2014; Nassar, Wilson, Heasly, & Gold, 2010; O'Reilly et al., 2013; Payzan-LeNestour & Bossaerts, 2011). This allows agents to update beliefs based on observed signals under ambiguity, but not under risk, and use these new beliefs in turn to update values.

Although there is growing evidence that the human brain constructs and makes use of beliefs to guide behavior, we know surprisingly little about how neural processes of updating take into account the prior knowledge about the reducibility of uncertainty. In particular, the presence of beliefs at the neural level (Behrens et al., 2007; Behrens, Hunt, Woolrich, & Rushworth, 2008; Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Gläscher, Daw, Dayan, & O'Doherty, 2010) is not sufficient for distinguishing between reducible and irreducible uncertainty, which previous studies have not empirically tested by manipulating reducibility of uncertainty.

To bridge this gap, we adapted the classic Ellsberg three-color urn problem, which manipulates reducibility of uncertainty. In this paradigm, if neural processes of belief updating and value updating are sensitive to the nature of uncertainty, they would be dissociated from neural

responses to expectancy violation (Itti & Baldi, 2009; O'Reilly et al., 2013). Our fMRI results revealed that neural correlates of belief updating and value updating are indeed distinct from expectancy violation. Furthermore, our connectivity analysis suggested that reducibility of uncertainty also modulates how neural processes of belief updating contribute to value updating.

In our paradigm, subjects are presented with a gamble involving a number of balls in an urn. They know the exact number of balls in one color (hereafter the risky color), but not in the other two colors (the ambiguous colors). For example, an urn contains four balls, two balls in yellow and two in either red or green; it could contain two red balls, one red and one green, or two green (Fig. 1a). Monetary outcome of a gamble is determined by a resolution draw from the urn; subjects win \$10 if the resolution draw matches a pre-determined color (the winning color), and nothing otherwise (Fig. 1b). We hereafter call a gamble ambiguous when its winning color is one of the ambiguous colors, and risky when it is the risky color.

We introduced environmental signals to this gamble in the form of an observed draw; prior to the resolution, a ball is drawn from the urn, reveals its color, and is returned back. We postulated that the observed draw first updates the prior belief about the urn content (belief updating, Fig. 1c) and then the value of the gamble (value updating, Fig. 1d).

This paradigm specifies and manipulates reducibility of uncertainty in beliefs and value as follows. First, because composition of ambiguous-color balls is unknown, belief should be updated by an ambiguous-color draw, but not by a risky-color draw. In our exemplar urn, a red draw updates belief because it demonstrates at least one red ball in the urn, increasing probability of a future draw in red ($\Delta P(R) > 0$) (Fig. 1c). On the other hand, a yellow draw does not carry any information, because it is already known that the urn contains two yellow balls ($\Delta P(Y) = 0$).

Second, value should be updated as a consequence of belief updating only in ambiguous gambles, but not in risky gambles. This is because the chance to win \$10 is not perfectly specified when the winning color is an ambiguous color. In our exemplar urn, a red draw increases the probability of a red resolution draw and decreases that of a green draw (Fig. 1d). As a consequence, the value of the gamble is increased if the winning color is red ($\$10 \cdot \Delta P(R) > 0$) and decreased if the winning color is green ($\$10 \cdot \Delta P(G) < 0$). On the other hand, if the winning color is yellow, a red draw does not update its value, because probability of a yellow draw is unaffected ($\$10 \cdot \Delta P(Y) = 0$).

Therefore, if subjects can rationally combine the prior knowledge about uncertainty with the color of the observed draw, they would update belief only after ambiguous-color draws, and update value only in ambiguous gambles (Fig. 1e). Such sensitivity would not be observed if updating is primarily driven by expectancy violation; since the draw's color is unpredictable, any draw in any gamble is associated with some level of expectancy violation, measured as $1 - P(\text{draw})$ (since $P(\text{draw}) < 1$, $1 - P(\text{draw}) > 0$ for any draw; Fig. 1c). To decouple updating from expectancy violation more clearly, we manipulated the urn composition across gambles (Table 1). For instance, increasing yellow balls in our exemplar urn would increase expectancy violation of a red draw, but decrease the magnitudes of belief updating and value updating. The manipulation of the urn composition therefore allowed us to look for neural correlates of belief updating and value updating while statistically controlling for expectancy violation, and vice versa, in fMRI analysis.

Results

Behavioral sensitivity to the nature of uncertainty. First, we tested the extent to which value updating is sensitive to the nature of uncertainty at the behavioral level. Subjective values of gambles were elicited via a BDM bidding procedure (see Methods), before and after the draw (Fig. 2a). If agents are sensitive to the nature of uncertainty as normatively predicted, they would update value only when they observe ambiguous-color draws in ambiguous gambles. On the other hand, if updating process is driven by expectancy violation, value would be also updated by risky-color draws or in risky gambles. We tested these predictions by classifying trials according to whether positive, negative, or zero value updating was normatively predicted given the draw color (Fig. 2b, Supplementary Table 1). Distribution of observed value updating indeed varied consistently with the normative prediction (chi-square test of independence, $\chi^2(4) = 1493.432$, $P < 10^{-5}$).

Quantitative modeling of updating. To quantitatively relate updating processes to BOLD responses in the following fMRI analysis, we next sought to provide a quantitative model of updating (Table 1). We postulated that agents first construct and update belief about probability of a future draw from the urn, and then use it to determine the expected value of the gamble ($10 × the probability of winning). We further assumed that, in belief, probability over all possible urn contents is considered and updated according to Bayes' rule (see Methods, Supplementary Text and Supplementary Fig. 1 for modeling details; note that this formalization instantiates so-called second-order probability, one of the most widely used approach to ambiguity in decision theory (Camerer & Weber, 1992; Ergin & Gul, 2009; Klibanoff, Marinacci, & Mukerji, 2005; Nau, 2006; Seo, 2009)). More specifically, prior probability over urn contents is assumed to follow a binomial distribution (we also used a uniform distribution to confirm robustness of our behavioral and fMRI findings; Supplementary Text and Supplementary Fig. 2).

We found that this model was able to explain subjective values of gambles well (Fig. 2c). Pre- and post-draw subjective values were consistent with model predictions in both ambiguous and risky gambles (log-likelihood ratio test, $P < 10^{-4}$, respectively; Fig. 2c, *left* and *middle*). More importantly, our model successfully predicted significant value updating after ambiguous-color draws in ambiguous gambles ($P < 10^{-4}$; Fig. 2c, *top right*), as well as negligible value updating in the other cases (+$0.068 after risky-color draws in ambiguous gambles, $P > .10$; −$0.228 after risky-color draws in risky gambles, $P > .10$; but −$0.143 after ambiguous-color draws in risky gambles, $P < .05$; Fig. 2c, *right*). We also noted that subjects exhibited ambiguity aversion (Ellsberg, 1961); while neither over- nor undervaluation was observed in risky gambles (−$0.063 before draws, −$0.234 after draws, $P > .10$ respectively), ambiguous gambles exhibited small undervaluation (−$0.252 before draws, $P < .10$; −$0.292 after ambiguous-color draws, $P < .05$; but −$0.184 after risky-color draws, $P > .10$). Overall, even though our model did not aim to account for ambiguity aversion, it is a successful first-order approximation of updating.

Neural correlates of belief updating. After establishing subjects' behavioral sensitivity to the reducibility of uncertainty, we examined how the brain processes environmental signals according to uncertainty in an fMRI experiment with a separate group of subjects. Specifically, we looked for neural correlates of belief updating and value updating, and tested whether they were dissociable from expectancy violation. During scanning, subjects observed a series of gambles (Fig. 3a); each trial started with presentation of the gamble's winning color, followed by the urn content and the observed draw. The gambles were resolved only after the scanning. This observational task was

adopted in order to isolate processes related to updating as opposed to choices. To ensure that subjects paid enough attention to the task, we elicited subjective assessment of the directionality of value updating and memory on the urn contents using auxiliary tasks (see Methods).

We first looked for brain regions where activation was correlated with belief updating. We quantified belief updating as the absolute difference between pre- and post-draw probability of ambiguous-color draws under our Bayesian model (Table 1). Even though our paradigm quantitatively dissociates belief updating from expectancy violation (measured as 1 – prior probability of the draw), their correlation is still not negligible in our parameter space ($r = 0.70$; Supplementary Fig. 3). Thus, we included both trial-wise belief updating and expectancy violation as parametric modulators in a single GLM and looked for regions where significant amount of variance could be explained uniquely by belief updating (i.e., adjusting for expectancy violation (Mumford, Poline, & Poldrack, 2015); see Methods). We found bilateral clusters in posterior middle frontal gyrus and superior frontal sulcus, bilateral clusters in intraparietal sulcus (IPS), and a cluster in precuneus (Fig. 3b, Supplementary Fig. 4; cluster-forming voxel-level threshold $P < .001$, uncorrected, and cluster-size threshold $k > 20$; see Table 2 for cluster-level P values corrected for whole-brain family-wise error). The clusters in frontal cortex may correspond to premotor region or frontal eye field (Vernet, Quentin, Chanes, Mitsumasu, & Valero-Cabre, 2014).

Neural correlates of value updating. We then examined neural correlates of value updating. While value should be updated only in ambiguous gambles, belief can be updated both in ambiguous and risky gambles. Given this difference, we expected that neural correlates of value updating are anatomically distinct from belief updating. We looked for brain regions which activation was correlated with value updating (quantified as the signed difference between pre- and post-draw expected values, see Table 1; this parametric modulator is orthogonal to belief updating and expectancy violation by design, see Supplementary Fig. 3). We found clusters in ventromedial prefrontal cortex, anterior and middle cingulate cortices, and left superior temporal gyrus (Fig. 3b, Supplementary Fig. 4, Table 2). These clusters did not overlap with the neural correlates of belief updating.

Neural correlates of expectancy violation. Next, we tested whether these updating regions responded to expectancy violation. Responses to expectancy violation, salience, or surprise has been long studied in cognitive neuroscience (Courchesne, Hillyard, & Galambos, 1975; Sokolov, 1963; N. K. Squires, Squires, & Hillyard, 1975). However, those studies have been motivated by the assumption that surprising signals tend to be relevant for agents, and dissociation between expectancy violation and uncertainty reduction has not been well studied. We found that activation in bilateral anterior insula (AI) was correlated with expectancy violation (adjusted for belief updating) (Fig. 3b, Supplementary Fig. 4, Table 2). Importantly, these clusters did not overlap with the neural correlates of either belief updating or value updating. This localization of expectancy violation is consistent with previous reports that AI responds to salient events in various domains (Corbetta, Patel, & Shulman, 2008; Menon & Uddin, 2010; Singer, Critchley, & Preuschoff, 2009).

ROI analysis of separable neural correlates. To further illustrate the dissociation among neural correlates of belief updating, value updating, and expectancy violation, we conducted region-of-interest (ROI) analysis, where ROIs were defined in a leave-one-subject-out fashion (see Methods).

We found that BOLD activation in belief-updating ROIs was correlated with belief updating as expected ($P < 10^{-4}$), but not with expectancy violation ($P > .10$) or value updating ($P > .05$; Fig 3c; see Supplementary Fig. 5 for ROI-wise results). Similarly, activation in value-updating ROIs was correlated with value updating ($P < 10^{-4}$), but not with belief updating or expectancy violation ($P > .10$). More critically, activation in expectancy-violation ROIs was correlated with expectancy violation ($P < 10^{-3}$), but not with belief updating or value updating ($P > .10$). These results show that neural processes of uncertainty reduction are anatomically dissociable from expectancy violation. We would have not found the dissociation between belief updating and expectancy violation if we had not adjusted for the correlation between their parametric regressors (Supplementary Fig. 6).

Interaction between updating regions. Lastly, we explored how these regions interact to drive appropriate value updating. Based on our results, we made two predictions about interregional interactions. First, we hypothesized that connections from belief-updating regions to value-updating regions would be modulated by the type of gamble. Since belief updating should contribute to value computation only in ambiguous gambles, interregional connections would be temporarily enhanced under ambiguous gambles (or temporarily weakened under risky gambles). Second, we hypothesized that connections from expectancy-violation regions to value-updating regions would not show such modulation as much, since expectancy violation does not drive value updating irrespective of the type of gamble, both theoretically and behaviorally. To test these predictions, we conducted dynamic causal modeling (DCM) analysis. It is appropriate for our purpose because it seeks to explain BOLD time-series from more than two ROIs simultaneously and can include directional connections modulated by experimental manipulations (Friston, Harrison, & Penny, 2003).

We compared three different scenarios: 1) belief-updating ROIs contribute to value updating, 2) expectancy-violation ROIs contribute to value updating, and 3) neither of them contributes (Fig. 4a; see Methods for details). To implement the first scenario, we constructed a family of models that instantiated every possible set of modulated connections from belief-updating ROIs to value-updating ROIs (Family I). Similarly, to implement the second scenario, another family of models instantiated every possible set of modulated connections from expectancy-violation ROIs to value-updating ROIs (Family II). The last scenario was implemented as a single model with no modulation in connections (Family III). We included four belief-updating ROIs from lateral frontal and parietal regions, two value-updating ROIs in medial prefrontal cortex, and two expectancy-violation ROIs from AI in DCM.

If neural processes of belief updating contribute to value updating in the way we hypothesized, our fMRI data should be best explained by Family I. This prediction was supported by the results of Bayesian model selection procedure (Penny et al., 2010; Stephan, Penny, Daunizeau, Moran, & Friston, 2009) (Fig. 4b). Probability that Family I outperformed both Families II and III ("exceedance probability") was over 80%, while that of Family II was around 10% and Family III below 5%. These results suggest that sensitivity to uncertainty reducibility is reflected in interaction from belief-updating regions, rather than expectancy-violation regions, to value-updating regions.

Discussion

In order to make adaptive decisions, we ubiquitously make use of environmental signals to reduce uncertainty. For appropriate uncertainty reduction, it is critical to understand whether current uncertainty is reducible, and by which signals it can be reduced. Specifically, we should not rely solely on signals' expectancy violation to determine whether and how much we can reduce uncertainty (Itti & Baldi, 2009; O'Reilly et al., 2013). Distinction between uncertainty reduction and expectancy violation is not clear in some traditionally prevalent frameworks such as reinforcement learning (RL) and Pearce-Hall (Pearce & Bouton, 2001; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Roesch, Esber, Li, Daw, & Schoenbaum, 2012; Sutton & Barto, 1998). In both, learning is driven by the degree to which prior expectancy about an event (e.g., the timing and amount of reward delivery) is violated. As a result, these theories do not explicitly state how agents can successfully ignore surprising, yet irrelevant, signals widespread in natural environments (Itti & Baldi, 2009).

In this study, we showed that the human brain is sensitive to the nature of uncertainty by demonstrating dissociation between uncertainty reduction and expectancy violation at behavioral and neural levels. This is related to two important lines of decision-making studies. First, model-based RL and Bayesian theories postulate that agents construct and update beliefs about the environment. Such agents may possess representation of the nature of uncertainty, by which uncertainty reduction could be decoupled from expectancy violation (Behrens et al., 2007; Itti & Baldi, 2009; Ma & Jazayeri, 2014; Nassar et al., 2010; O'Reilly et al., 2013; Payzan-LeNestour & Bossaerts, 2011). Unlike our study, however, most neuroimaging studies to date have examined the forms of learning that could be driven by expectancy violation, not manipulating reducibility of uncertainty (Daw et al., 2011; Gläscher et al., 2010).

Second, decision theory has long emphasized distinction between reducible and irreducible uncertainty, often referred to as ambiguity and risk respectively (Camerer & Weber, 1992; Ellsberg, 1961; Keynes, 1921; Knight, 1921). Past studies have mainly investigated influence of these types of uncertainty on choice and its neural basis (Bach, Hulme, Penny, & Dolan, 2011; Hsu, Bhatt, Adolphs, & Tranel, 2005; Huettel, Stowe, Gordon, Warner, & Platt, 2006), and these concepts have been applied to the problem of uncertainty reduction only recently (Chumbley et al., 2012). Our findings show that these types of uncertainty are not only modulators of value-based choices, but also determinants of neural processing and behavioral consequences of environmental signals.

Behaviorally, we found evidence that value updating is sensitive to the nature of uncertainty. Specifically, value was updated only under reducible uncertainty, irrespective of signals' expectancy violation (Fig. 2b), and could be quantitatively characterized by a Bayesian model (Fig. 2c). Similarly in the scanner, we found that subjects' responses to the auxiliary task on directionality of value updating conformed to the normative prediction, showing that surprising but irrelevant signals were successfully ignored (see Methods). Existence of representation of uncertainty in human has been shown in Bayesian sensory and sensorimotor literature (Ernst & Banks, 2002; Körding & Wolpert, 2004; Ma & Jazayeri, 2014; Pouget, Beck, Ma, & Latham, 2013), and our results extends it to value-based decision-making domains. At the same time, however, given the relatively simple nature of our task, it is also possible that our subjects did not carry out optimal full-Bayesian computation and instead used simpler heuristics. One possible heuristic approach is to consider a single "effective" urn composition, in which an ambiguous-color ball is treated as a pair of half balls

(Supplementary Fig. 1; see Supplementary Text for proof of its mathematical equivalence to full-Bayesian approach in our task). Thus, even if the actual computational processes of updating may not be fully Bayesian, subjects appear capable of incorporating relevant prior knowledge about uncertainty into updating.

At the neural level, we found that neural processes associated with belief updating and expectancy violation were anatomically separable, which have frequently been confounded with each other in past studies. Indeed, to our knowledge, there exist only two studies that explicitly decouple these two variables (O'Reilly et al., 2013; Schwartenbeck, FitzGerald, & Dolan, 2016). Interestingly, despite important differences in task design, these studies all found dissociation between belief updating and expectancy violation regions, and in particular representation of belief updating in lateral frontoparietal regions. This also accords well with results from Gläscher et al. (Gläscher et al., 2010), which used a Markov decision task to capture the degree to which a state prediction error (SPE) that updated the belief on state-action-state transition probabilities. Even though SPE was formally equivalent to expectancy violation in their particular task, our results make it unlikely that their results reflected expectancy violation alone.

However, a number of important differences exist between these findings in terms of the precise localization within frontoparietal regions, as well as recruitment of other regions. This may relate to more fine-grained task-level differences between studies. In O'Reilly et al.'s saccadic planning task, which did not include a reward-based component, the authors found both belief updating and expectancy violation responses in frontoparietal regions (O'Reilly et al., 2013); they reported belief-updating responses in pre-supplementary motor area and area 7a, and expectancy violation responses in IPS. Schwartenbeck & Dolan dissociated updating in belief about valid cue modality from expectancy violation by introducing trials where both cues predicted the same valence, such that even surprising outcomes could not discriminate their validity. While they mapped belief updating onto putative dopaminergic midbrain, inferior frontal gyrus, posterior parietal cortex, and ACC, and expectancy violation onto pre-supplementary motor area and dorsal ACC, it is possible that these mappings reflect reward-related processes to a certain extent.

Such task differences may well also explain differences in localization of expectancy violation across these studies. One particularly important nature of our task is the fact that it involves the presence of "mere" expectancy violation, which may not have consequences on subsequent behavior or valuation. Our localization of expectancy violation in AI is consistent with previous reports that AI responds to abrupt or rare stimuli in various domains (Menon & Uddin, 2010; Singer et al., 2009) and is involved in reorienting attention to surprising events (Corbetta et al., 2008; Sokolov, 1963). Our results raise the possibility that AI is primarily involved in detecting and rapidly broadcasting expectancy violation, which is not necessarily relevant for the ongoing tasks but may well still be important for agents' survival, through distinctly large bipolar neurons called Von Economo neurons (Allman et al., 2010; Evrard, Forro, & Logothetis, 2012).

In contrast, in O'Reilly et al's task, a saccadic target with high expectancy violation required subjects to reprogram their preprogrammed saccade, even if it did not update belief. Therefore, their localization of expectancy violation in IPS likely reflects saccade programming processes. Similarly, in the Schwartenbeck et al's task, expectancy violation was defined with respect to monetary outcome (specifically its valence), and thus their localization may reflect some reward processing,

including learning. Given the small number of studies that explicitly separate belief updating, expectancy violation, and value updating, future studies are necessary to more fully assess the functional mapping of belief updating and expectancy violation under different task demands.

We also dissociated neural correlates of value updating from expectancy violation; value updating was correlated with activation in medial prefrontal cortex (MPFC) and cingulate cortex, not AI (Fig. 3). It has been established that these regions are involved in value-related processing, such as valuation of choices or reward-based reinforcement (Bartra, McGuire, & Kable, 2013). Particularly, value updating in the current study is conceptually close to model-based reward prediction error (RPE), which has been associated with these regions (Behrens et al., 2008; Daw et al., 2011). Our results provide evidence that value-related computation in these regions may not be driven solely by expectancy violation. Interestingly, we did not find any evidence of value-updating representation in striatum, even at a liberal threshold, which has been also associated with model-based RPE in the past (Behrens et al., 2008; Daw et al., 2011; Gläscher et al., 2010). This discrepancy could be because, unlike past studies, we did not provide reward feedback to prevent learning over trials. Thus, while MPFC and cingulate cortex may be involved in reduction of value uncertainty irrespective of feedback existence (or independent of the task specifics in general), striatum might primarily respond to reward feedback. An alternative account is that striatum could be more involved in learning over time rather than one-shot updating, particularly through corticostriatal loops (Balleine, Delgado, & Hikosaka, 2007). These possibilities could be tested by providing trial-wise outcome (i.e., the resolution draw) in our task.

Although we found that belief- and value-updating regions were anatomically distinct, their computational processes should not be independent under a Bayesian framework of decision-making. However, in the extreme, activations in putative belief-updating regions (premotor/FEF and IPS) could be an epiphenomenal reflection of more general cognitive processes, such as working memory (Mohr, Goebel, & Linden, 2006; Reinhart et al., 2012), and irrelevant to updating processes. Using DCM, we found that processes in belief-updating ROIs in premotor/FEF and IPS contribute to MPFC regions involved in value computation via interregional interaction modulated by the nature of uncertainty (Fig. 4). Such modulation in connections might be biologically efficient; when value uncertainty is irreducible and valuation system can safely ignore incoming signals, energies to maintain synaptic transmission from belief-updating regions can be temporarily saved. Additionally, we did not find evidence for modulation in connections from expectancy-violation regions (AI) to value-updating regions. Even though interpretation of negative results requires caution, this suggests that reducibility-based modulation is not a general feature of connections across the cortex. It also further emphasizes the difference between functional roles of AI and lateral frontoparietal regions.

It is worth noting that, in general, the DCM results do not necessarily quantify or assume monosynaptic connections. In our case, although previous studies have found anatomical connections from premotor/FEF and IPS to cingulate motor area, connections to more anterior portion of MPFC, where we found value-updating representation, have not been reported (Beckmann, Johansen-Berg, & Rushworth, 2009; Eradath et al., 2015; Mars et al., 2011; Neubert, Mars, Sallet, & Rushworth, 2015; Tomassini et al., 2007). We speculate that modulated connections from premotor/FEF and IPS to MPFC could be mediated by anterior-posterior connections within

cingulate cortex (Margulies et al., 2007) or corticostriatal connections (Balleine et al., 2007; Di Martino et al., 2008).

Taken together, our findings suggest that the nature of uncertainty influences the way the human brain processes environmental signals. In order to understand the neural mechanisms of uncertainty reduction more comprehensively, future studies can expand this study's approach in a number of directions. For instance, while our paradigm delivered one environmental signal at a time, real world scenarios typically involve a multitude of signals. In these cases, prior knowledge about reducibility of uncertainty may in addition be crucial in determining how the brain allocates attentional resources, as it may not be possible for all signals to be attended to at all times (Gottlieb & Balan, 2010). Additionally, future studies are necessary to clarify the encoding scheme used by the brain to represent belief updating. In particular, due to the relatively small urn sizes used in the experiment, which constricts the range of belief updating values, we were unable to distinguish between different operationalizations of belief updating, e.g., between binomial or uniform priors over urn contents. Importantly, this includes the possibility that putative belief updating regions in fact categorically classify ambiguous or risky draw, which is akin to a median split of belief updating values in our model. Although our DCM results argue against the possibility that responses in these regions were driven solely by the draw type without contribution to updating processes, studies that use richer parameterization of urn sizes would be useful to more directly address this question, albeit at a cost of substantially increasing task complexity. Finally, unlike our paradigm, the prior knowledge about uncertainty reducibility in the real world is often far from complete and accurate. It is an open question how agents estimate uncertainty reducibility, particularly under non-stationary environments (Behrens et al., 2007; Payzan-LeNestour & Bossaerts, 2011; Yu & Dayan, 2005).

Methods

Behavioral experiment.

Subjects. 10 undergraduate students in University of California, Berkeley (6 women) participated. They provided written informed consents and all procedures were approved by UC Berkeley Committee for the Protection of Human Subjects. The experiment was conducted individually in a self-paced manner, in an isolated cubicle. The experiment program was written on Matlab (Mathworks) and Psychtoolbox (Brainard, 1997; Pelli, 1997), run on a laptop, and interacted with via its keyboard.

Task. Subjects were presented with gambles, each of which consisted of a winning color and an urn containing a number of balls in red, green, or yellow (Fig. 2a). One of the gambles was randomly selected and resolved at the end of the experiment; a ball was randomly drawn from the urn, and subjects received $10 only if it matched the winning color (in addition to the baseline payment for task completion). Subjective values of these gambles, both pre-draw and post-draw, were elicited as willingness to sell (WTS), i.e., the amount of money subjects were willing to give up the opportunity to play the gamble for, through a standard BDM bidding procedure (Becker & Brownson, 1964). In total, 18 gambles were presented in a randomized order (6 urn contents × 3 winning colors).

Each trial started with presentation of urn content and the winning color. Subjects were informed of the number of balls in one color (risky color) and the total number of balls in the other two colors (ambiguous colors), but not of the exact composition in the latter. Each ball in risky color was visually represented as a full circle, and each ball in ambiguous colors as a pair of half circles. The winning color was shown above the urn contents. After subjects indicated pre-draw WTS, the three possible colors of observed draw were presented in a randomized order, after each of which post-draw WTS was indicated. Thus, four WTSs in total were obtained in each gamble. Upon the resolution of the gamble, the experiment program randomly determined whether subjects observed the draw or not (50%), and which color the observed draw had (probability following the urn composition).

Data analysis. To examine value updating in a model-free manner, trial-wise difference between pre- and post-draw subjective values (WTSs) were calculated and categorized according to normatively prediction of directionality (Fig. 2b). To more quantitatively characterize subjective values, predictions from our quantitative Bayesian model (see below) were fitted to the observation in mixed-effect modeling implemented on R software and lmer package, with subjects as a random effect (Fig. 2c). To test one-sided deviation, the fixed-effect constant term (intercept) was compared against zero.

Bayesian modeling. Our quantitative Bayesian model consists of two stages, belief and valuation. The belief stage concerns probability distribution on a future draw's color. In order to generate unique point estimates of probability of ambiguous colors, all possible urn contents are considered, weighted according to their binomially distributed probability, and averaged (Supplementary Text, Supplementary Fig. 1). After the observed draw, probability over urn contents is updated according to Bayes' rule. In the valuation stage, pre- and post-draw values of gambles are calculated as expected outcome, i.e., $10 × probability of winning.

Note that this modeling is mathematically equivalent to a heuristic account, which only considers "effective" urn content (Supplementary Text). We also conducted behavioral and fMRI analyses using uniform prior distribution over urn contents (Supplementary Text, Supplementary Fig. 2).

fMRI experiment.

Subjects. 20 subjects (mean age = 21.7 years old, 11 women) participated after being screened for standard MRI contraindications. They provided written informed consents and all procedures were approved by UC Berkeley Committee for the Protection of Human Subjects. Among them, 1 subject declined the participation after the task instructions but before the scanning, and 2 subjects were discarded from analysis due to unsatisfactory performance in auxiliary tasks (see below), resulting in data from 17 subjects analyzed. During scanning, the experiment program was run on Matlab and Psychtoolbox, with which subjects interacted via an MRI-compatible button box.

Main task. During scanning, subjects observed gambles in a randomized order (Fig. 3a), one of which was randomly selected and resolved at the end of the experiment. 30 gambles were presented in each of the three EPI runs (90 in total), and the winning color was changed across runs (remained the same within each run). They covered all combination of the six urn content. Six urn contents (Table 1) was presented 15 times each, 5 times for risky gambles and 10 for ambiguous gambles ($6 \times 15 = 90$). Probability of the observed draw's color across gambles approximately followed the urn composition.

Each trial started with the fixation cross in the winning color (2s), followed by the urn content presentation (visual representation similar to the behavioral experiment). The urn content was presented for 5–12s. (After a variable delay of 4–6s, the urn opened its lid in a 0.5s animation, on which subjects were asked to press a button within 5s. Upon their button press, the balls moved into the urn in another 0.5s animation. This process was introduced to keep subjects' alertness, and is not shown in the Fig. 3a for simplicity.) After a variable interval (3–6s), a gray ball moved out of the urn in a 0.5s animation, and revealed its color after another variable interval (1–3s). After 3s, the drawn ball was returned back to the urn in a 0.5s animation, followed by a variable inter-trial interval (2.5–4.5s).

Auxiliary tasks. In order to verify subjects' engagement and understanding throughout the main observation task, we asked them to respond to three types of auxiliary tasks. Tasks was presented immediately after randomly selected 27 gambles (9 for each task type). In memory task, subjects were asked to choose the correct description of the previous gamble (the winning color, the urn content, and the observed draw) from two options. In value-updating judgment task, subjects were asked to indicate whether it is 1) less likely, 2) equally likely, or 3) more likely to win the gamble after the draw. In surprise rating task, subjects were asked to rate their surprise of the observed draw in 3-point scale. Since the surprise rating task was purely subjective, we used the memory task and the value-updating judgment task to test subjects' engagement and understanding; two subjects were excluded from the subsequent analyses because of their unsatisfactory performance (more than 2 wrong responses or more than 2 trials without responses within 10 seconds in either task). Subjects received up to $10 based on their performance in these two tasks.

fMRI data acquisition. MR images were acquired by a 3T Siemens Trio scanner and a 12-channel head coil. Functional images were obtained using T2*-weighted gradient-echo echo-planar imaging (EPI) pulse sequence (TR = 2000ms, TE = 30ms, voxel size = 3mm × 3mm × 3mm, inter-slice gap = 0.3mm, in-plane resolution = 64 × 64, 32 oblique axial slices). Slices were tilted by 30 degrees from AC-PC line to alleviate signal dropout from orbitofrontal cortex (Weiskopf, Hutton, Josephs, & Deichmann, 2006). T1-weighted structural images (1mm × 1mm × 1mm) were also obtained using magnetization-prepared rapid-acquisition gradient-echo (MPRAGE) pulse sequence.

Preprocessing. Motion correction, slice-time correction, normalization to MNI EPI template, and smoothing with a Gaussian kernel of 8-mm FWHM were applied to functional images using SPM8 (Wellcome Dept. of Cognitive Neurology, London, UK).

Whole-brain univariate analysis. Whole-brain analysis (Fig. 3b, Supplementary Fig. 4, Table 2) used general linear modeling (GLM) of BOLD time-series and group-level random-effect models on SPM8. To quantitatively relate BOLD signals to belief updating, value updating, and expectancy violation, these variables were parametrically defined under the Bayesian model: belief updating as the absolute difference between pre- and post-draw probability of ambiguous colors, value updating as signed difference between pre- and post-draw expected value of gambles, and expectancy violation as (1 – pre-draw probability of the observed draw). See Table 1 for their variations as functions of urn contents.

The three variables of interest, belief updating, value updating, and expectancy violation, were included as parametric modulators of a regressor modeling the events of the observed draws. In order to adjust for the correlation between belief updating and expectancy violation (Supplementary Fig. 3), we included both variables in a single GLM so that coefficient estimates captured only variance that were uniquely explained by each of them, removing shared variance (Mumford et al., 2015). Since SPM8 orthogonalizes the second parametric modulator against the first one by default, we actually implemented two GLMs: one in which the first parametric modulator was expectancy violation and the second was belief updating (GLM 1), and another one in which the order was reversed (GLM 2). GLM 1 was used to reveal neural correlates of belief updating adjusted for expectancy violation, and GLM 2 was used to reveal neural correlates of expectancy violation adjusted for belief updating. To illustrate how results are affected in cases where shared variance is not removed, we also reported coefficient estimates for expectancy violation in GLM 1 and belief updating in GLM 2 in Supplementary Fig. 6. GLM 3 included value updating as a sole parametric modulator.

All GLMs also included regressors that modeled events of gamble presentation, button press, question presentation, and question response, six movement parameters estimated in the motion correction procedure, 128s high-pass filtering, and AR(1) model of serial autocorrelation. All of the event-related regressors were convolved with the SPM's double-gamma canonical hemodynamic response function. Coefficient estimates of the parametric modulators of interest were then entered into group-level analysis. For clusters defined by voxel-level threshold $P < .001$, uncorrected, cluster-level P values with whole-brain correction for family-wise error (FWE) were calculated using nonparametric permutation in SnPM13 package (Hayasaka & Nichols, 2003; Nichols & Holmes, 2001; Woo, Krishnan, & Wager, 2014) (Table 2).

Regions-of-interest (ROI) analysis. Activations in ROIs were examined (Fig. 3c, Supplementary Fig. 5) using Marsbar package (Brett, Anton, Valabregue, & Poline, 2002) in the following steps. First, the ROIs were determined based on the group-level whole-brain analysis at the cluster-level threshold of $k > 20$ (Table 2). Second, subject-specific ROIs were defined from group-level activation maps based on the other 16 subjects' data (leave-one-subject-out) (Boorman, Rushworth, & Behrens, 2013; Hunt, Dolan, & Behrens, 2014). Clusters that survived uncorrected voxel-level $P < .001$, $k \geq 10$, with the local maxima located within 16 mm from all-subject group-level peaks were used (one cingulate value-updating ROI was discarded from ROI analysis because it could not be robustly identified in some iterations at this threshold). Third, mean BOLD time-series from each ROI was extracted from the hold-out subject's data, on which GLMs 1, 2, and 3 were fitted. Their coefficient estimates were normalized according to baseline of time-series, similarly to conventional calculation of percent signal change (although our coefficients are derived from parametric modulators and thus cannot be interpreted as percent signal change per se). Fourth, estimates were entered into mixed-effect modeling with subjects as a random effect, conducted on R and lmer package.

Please note that, although the circularity problem in ROI definition (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009) is slightly alleviated by the leave-one-subject-out procedure, it is not totally eliminated. This is because the sets of ROIs (Table 2) were determined based on data from all subjects, including the hold-out one. As a consequence, for belief-updating ROIs, coefficient estimates for belief updating (GLM 1) may be slightly positively biased, and estimates for expectancy violation (GLM 2) may be negatively biased; vice versa for expectancy-violation ROIs. (Since value updating is orthogonal to belief updating and expectancy violation by design, there is no bias in any coefficient estimates from value-updating ROIs, as well as coefficients estimates for value updating from any ROIs.)

Also note that statistical inference in ROI analysis did not compare coefficients of different parametric modulators (belief updating, value updating, and expectancy violation); it only concerned whether each coefficient was different from zero, not whether one coefficient was higher than others. The parametric modulators in GLMs are not in the same unit, and their coefficient estimates are not directly comparable.

Dynamic causal modeling (DCM). DCM analysis (Fig. 4) was conducted on SPM8. DCM is a generative model of BOLD time-series from multiple ROIs, and includes three types of factors: direct regional input, stationary interregional connections, and importantly for our purpose, temporal modulations in interregional connections (Friston et al., 2003). DCM is agnostic about whether connections are monosynaptic or not.

Since SPM8's DCM module can model up to 8 ROIs, we included four belief-updating ROIs (bilateral frontal and parietal cortex), two value-updating ROIs (MPFC and right VMPFC), and two expectancy-violation ROIs (bilateral AI) (Table 2) in our models. We hand-picked these ROIs primarily because we could identify these ROIs for every subject (see below), and secondly because these ROIs make the interpretation easier (we included belief-updating clusters in frontoparietal regions instead of precuneus, and value-updating clusters in MPFC instead of cingulate and STG). It is possible that the DCM results (Fig. 4) depend on this selection of ROIs; since conducting DCM with a huge model space such as ours takes long time even on multi-node multi-core computing

clusters, it is not feasible for us to conduct the same analysis with different sets of ROIs. However, as we did not select these ROIs based on DCM results, type-I error has not been affected by this ROI selection procedure.

The ROIs were defined in a subject-specific manner. From subject-wise activation maps (GLMs 1–3), ROIs that survived uncorrected voxel-wise threshold $P < .20$ with local maxima located within 16mm from all-subject group-level peaks were selected (Smith, Stephan, Rugg, & Dolan, 2006). 3 subjects were discarded from DCM analysis because some ROIs could not be identified (1 subject for left frontal belief-updating cluster, 1 for MPFC value-updating cluster, and 1 for left AI expectancy-violation cluster). Next, principle eigenvariate of BOLD time-series from 4mm-radius spheres centered on the local maxima were extracted.

To test modulation of interregional connections based on the type of gambles (ambiguous or risky), we constructed and compared three families of DCMs (Fig. 4a). We instantiated all possible sets of modulated connections from belief-updating ROIs to value-updating ROIs in Family 1 ($2^8 - 1 = 255$ models), and all possible sets of modulated connections from two expectancy-violation ROIs to two value-updating ROIs in Family 2 ($2^4 - 1 = 15$ models). Family 3 included only one model with no modulation. We adopted this family-wise approach because, while we were interested in testing existence of modulation in connections among the three sets of ROIs, we were not interested in discriminating contribution of specific ROIs. Particularly, we aimed to allow the possibility that ROIs that exhibit modulation in connections are heterogeneous across individuals (e.g., handedness might affect laterality). In all models, modulation in connections were implemented as differential strength of connections during ambiguous and risky gambles (boxcar functions from gambles' presentations to the trials' termination).

In addition to modulated connections, DCMs included regional inputs and stationary interregional connections, which were identical across all models. Regional inputs modeled the events of the observed draws. To explain away intraregional computational processes that were captured in univariate analysis, we included parametric modulator of value updating to inputs to value-updating ROIs, and parametric modulators of belief updating and expectancy violation to both belief-updating and expectancy-violation ROIs. Stationary interregional connections modeled bidirectionally influence among ROIs from the other category (belief-updating, value-updating, and expectancy-violation), but not within the same category.

Group-level, random-effect, family-level inference was conducted to compare the three families of models (Fig. 4b). Family comparison procedure takes into account both goodness of fit and model complexity, aggregates performance across all models in each family, and calculates exceedance probability, i.e., probability in which each family was better than the other families (Penny et al., 2010; Stephan et al., 2009).

Figures

Figure 1 Experimental paradigm. (**a**) An exemplar urn. This urn contains two balls in yellow (the risky color) and two balls in either red or green (the ambiguous colors). The composition within the two ambiguous-color balls is unknown. See Table 1 for manipulation over the urn contents in the actual experiments. (**b**) Gambles. Subjects receive \$10 if a resolution draw from the urn matches a pre-determined winning color. The gamble is called ambiguous when the winning color is an ambiguous color, and risky otherwise. (**c**) Belief updating and expectancy violation of observed draws. A ball is drawn from the urn, reveals its color, and is returned back. *Top*: A draw in red (ambiguous color) updates belief, because it demonstrates that at least one red ball is in the urn ($\Delta P(R) > 0$). The draw also violates prior expectancy to a certain extent ($1 - P(R) > 0$). *Bottom*: A draw in yellow (risky color) does not update belief, because the urn is already known to contain two yellow balls ($\Delta P(Y) = 0$). It is still associated with expectancy violation ($1 - P(Y) > 0$). (**d**) Value updating. *Left*: When the winning color is red, a red observed draw increases the expected value ($\$10 \cdot \Delta P(R) > 0$). *Middle*: When the winning color is green, a red draw decreases the expected value ($\$10 \cdot \Delta P(G) < 0$). *Right*: When the winning color is yellow, a red draw does not affect winning probability ($\$10 \cdot \Delta P(Y) = 0$). (**e**) Presence of belief updating, value updating, and expectancy violation. Expectancy violation does not guarantee belief updating or value updating.

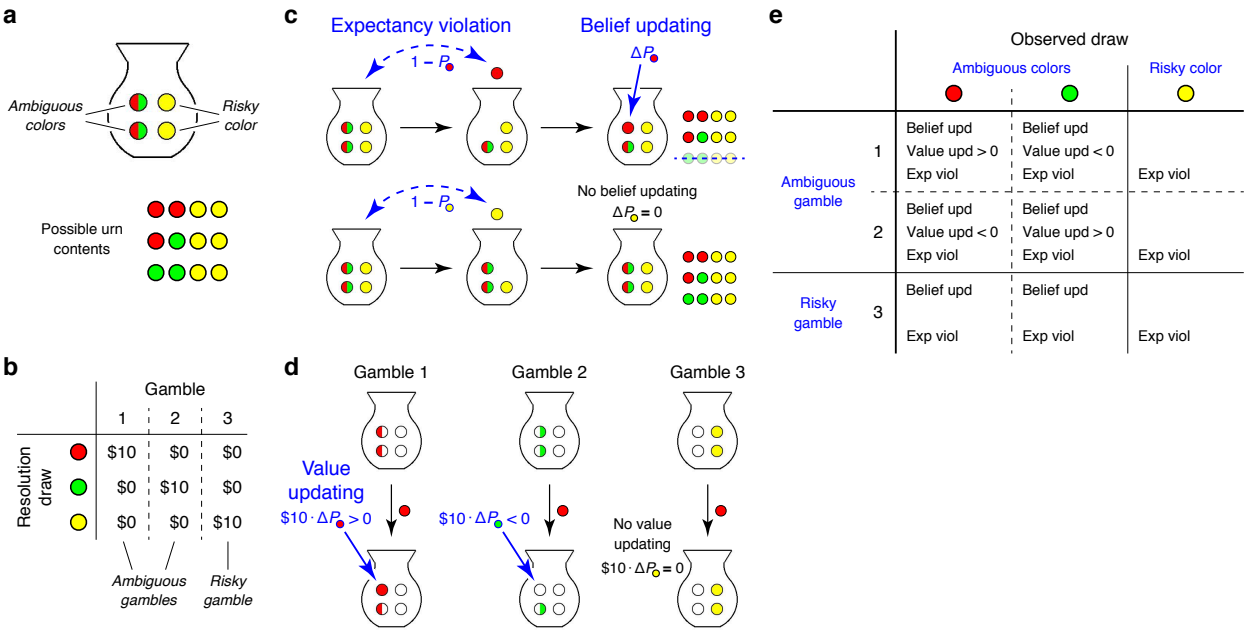


Figure 2 Behavioral results. (**a**) Experiment procedure. First, the urn content and the winning color were presented, and pre-draw subjective value of gambles were elicited. Next, the observed draw was presented, followed by elicitation of post-draw value. Post-draw valuation was repeated for all of the three colors. The urn content and the winning color were manipulated across gambles. (**b**) Histogram of value updating. Trials were classified based on the directionality of value updating predicted given the draw color (Fig. 1e). Proportion of observed direction agreed with the predictions (chi-square test of independence, $P < .05$). (**c**) Quantitative modeling of subjective values. Observation was compared against prediction of our Bayesian model. Error bars are standard errors, but not visible in some cases due to their small sizes. *Left*: Reported pre-draw values were successfully predicted in both ambiguous (*top*) and risky (*bottom*) gambles (log-likelihood ratio test, $P < .05$). *Middle*: Reported post-draw values were successfully predicted in both ambiguous (*top*) and risky (*bottom*) gambles, after both ambiguous- (blue open dots) and risky-color (red closed dots) draws ($P < .05$). *Right*: Non-zero updating was only predicted in ambiguous gambles after ambiguous-color draws (*top*, blue open dots), and linear prediction was successful ($P < .05$). Although ambiguous-color draws slightly decreased values of risky gambles (*bottom*, blue open dots; $P < .05$), its effect size was negligible compared to ambiguous gambles. Risky-color draws did not update values both in ambiguous and risky gambles (red closed dots, $P > .10$).

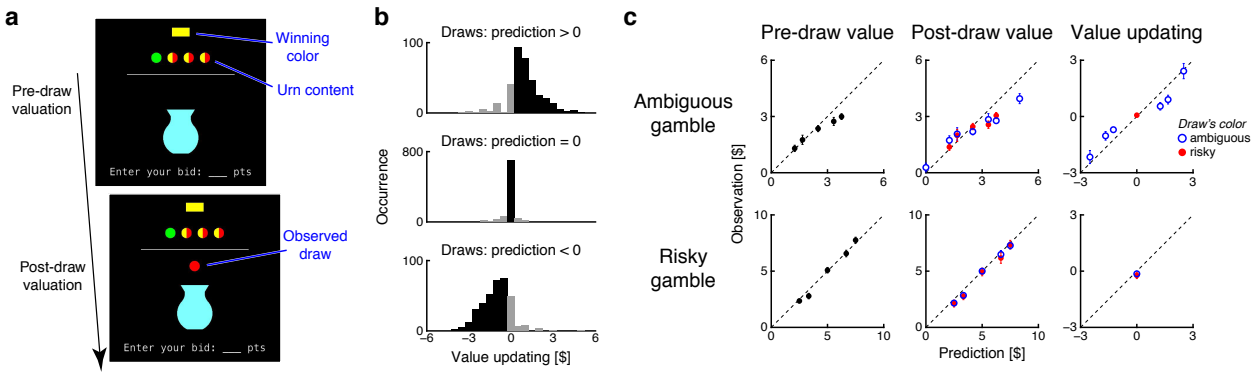


Figure 3 fMRI results. (**a**) Experiment procedure. The winning color was first presented in the fixation cross, followed by the urn content and then the observed draw. We analyzed BOLD signals time-locked to the observed draw. The urn content and the draw's color were manipulated across gambles, while the winning color was manipulated across the scan runs. (**b**) The activation maps. Shown are clusters that survived voxel-level threshold $P < .001$ (uncorrected) and cluster-size threshold $k > 20$. See Table 2 for cluster-level whole-brain FWE-corrected P values, and Supplementary Figure 4 for the whole-brain map. *Top:* Belief updating (adjusted for expectancy violation) was correlated with activation in posterior middle frontal gyrus / superior frontal sulcus cortex, intraparietal cortex, and precuneus. *Middle*: Value updating was correlated with activation in medial prefrontal cortex and cingulate cortex. *Bottom*: Expectancy violation (adjusted for belief updating) was correlated with activation in anterior insula. (**c**) ROI analysis. *Top:* Coefficients estimates for belief updating. Belief updating was not associated with value-updating ROIs or expectancy-violation ROIs ($P > .10$). *Middle:* Coefficients estimates for value updating. Value updating was not associated with belief-updating ROIs ($P > .05$) or expectancy-violation ROIs ($P > .10$). *Bottom:* Coefficients estimates for expectancy violation. Expectancy violation was not associated with belief-updating ROIs or value-updating ROIs ($P > .10$). Note that the coefficient estimates for belief updating (*top*), value updating (*middle*), and expectancy violation (*bottom*) are not directly comparable due to differences in scale of regressors. See Supplementary Figure 5 for ROI-wise analysis, and Supplementary Figure 6 for results when the correlation between belief updating and expectancy violation was not adjusted for. Error bars: SEMs, asterisks: $P < .05$.

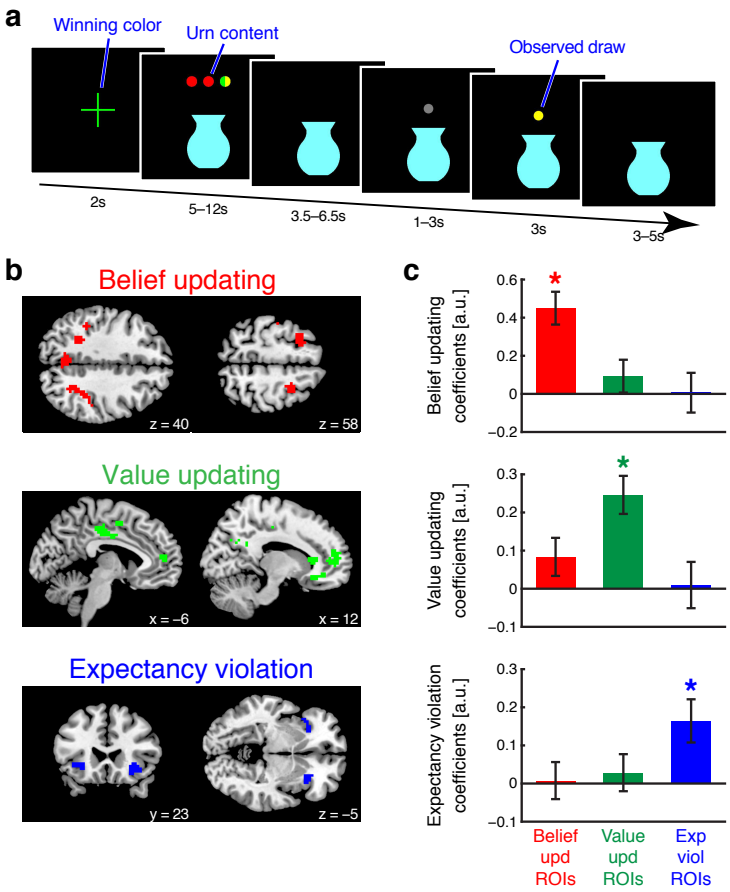
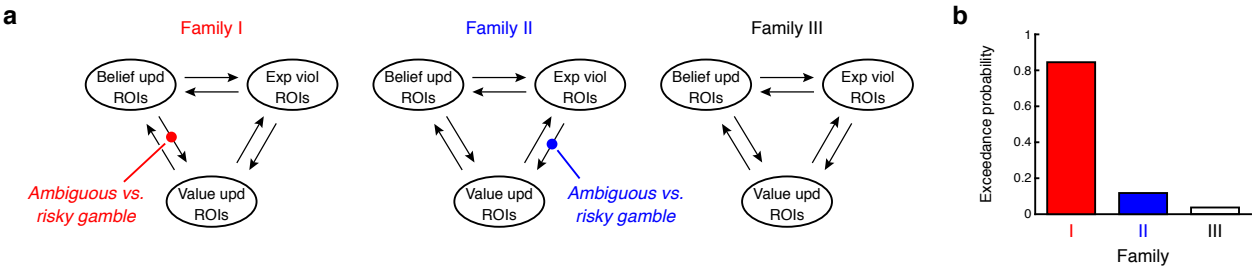


Figure 4 Dynamic causal modeling. (**a**) Families of models. Solid lines indicate intrinsic connections, and a dot on lines indicate their modulation according to the type of gambles. All modeled contained four belief-updating ROIs (bilateral frontal and parietal), two value-updating ROIs (MPFC and right VMPFC), and two expectancy-violation ROIs (bilateral insula). Families I and II instantiated all possible sets of modulation in connections from belief-updating ROIs or expectancy-violation ROIs to value-updating ROIs, respectively. Family III contained no modulation in connections. (**b**) Exceedance probability of the families, i.e., the probability of each family performing better than the other two families. Family 1 was supported by data.



Tables

Table 1. The list of the urn contents used in the behavioral and fMRI experiments, and the quantitative measurement of belief updating, value updating, and expectancy violation (derived under the Bayesian model with binomial prior probability distribution over urn contents; see Supplementary Text and Supplementary Fig. 1 for modeling details). See Supplementary Figure 3 for correlation among these variables.

Urn contents		Ambiguous-color draw				Risky-color draw			
Balls in ambiguous colors	Balls in risky color	Belief updating	Value updating		Expectancy violation	Belief updating	Value updating		Expectancy violation
			Ambiguous gamble	Risky gamble			Ambiguous gamble	Risky gamble	
1	1	0.250	±2.50	0	0.750	0	0	0	0.500
1	2	0.167	±1.67		0.833				0.333
2	1	0.167	±1.67		0.667				0.667
1	3	0.125	±1.25		0.875				0.250
2	2	0.125	±1.25		0.750				0.500
3	1	0.125	±1.25		0.625				0.750

Table 2 Clusters associated with belief updating, value updating, and expectancy violation. Clusters were formed at voxel-wise threshold $P < .001$, uncorrected. All clusters larger than 20 voxels are reported here. Shown are cluster sizes, cluster-level *P* values (based on permutation, corrected for whole-brain family-wise error), MNI coordinates of the peaks, voxel-level effect sizes, and voxel-level *P* values (uncorrected).

Region	# voxel	*Cluster-level P*	*Peaks*				
			x	*y*	*z*	*T*(16)	*Voxel-level P*
Belief updating							
R middle frontal gyrus / superior frontal sulcus	61	0.0568	27	2	55	5.77	0.000014
			27	14	46	4.30	0.000275
Precuneus	96	0.0280	6	-61	31	5.61	0.000020
			-6	-61	40	4.81	0.000096
			-6	-58	28	4.40	0.000224
R intraparietal sulcus	146	0.0128	30	-40	49	5.54	0.000022
			45	-31	46	5.06	0.000058
			27	-52	46	4.64	0.000136
L intraparietal sulcus	51	0.0724	-24	-43	46	5.18	0.000046
			-42	-37	40	4.66	0.000131
L middle frontal gyrus / superior frontal sulcus	39	0.1000	-30	11	58	4.67	0.000128
Value updating							
R ventromedial prefrontal cortex	39	0.0966	15	29	-14	6.36	0.000005
			15	38	-11	4.68	0.000125
			24	26	-11	4.13	0.000393
L superior temporal gyrus	43	0.0868	-48	-7	-2	5.75	0.000015
Medial prefrontal cortex	210	0.0118	-15	53	7	5.66	0.000018
			15	53	4	4.77	0.000104
			-12	41	4	4.66	0.000131
Cingulate cortex	107	0.0302	-9	-22	43	5.55	0.000022
			-18	-22	40	5.16	0.000047
			-6	-10	37	4.47	0.000193
Subgenual area	31	0.1194	9	26	7	5.53	0.000023
Cingulate cortex	29	0.1284	-6	-1	52	4.40	0.000224
			3	2	46	4.20	0.000339
Expectancy violation							
R insula	45	0.0660	30	20	-8	6.22	0.000003
L insula	27	0.1060	-30	26	-2	5.05	0.000059
			-42	20	-2	4.13	0.000393

Supplementary Materials

Supplementary Text: Bayesian modeling of updating and its alternatives.

In our Bayesian modeling, we postulated that point estimates of a future draw's color probability are calculated in belief. We specifically assumed that all possible urn contents are considered, weighted according to their "second-order" probability, and averaged. Second-order probability over urn contents is updated in a full Bayesian manner, and its prior is binomially distributed.

1. Heuristic account

Even though our behavioral (Fig. 2) and fMRI (Fig. 3) results were derived based on the Bayesian modeling, they are not dependent on specifics of this model. Specifically, predictions of this model are mathematically equivalent to another, more heuristic account, which considers only one "effective" urn content. In this account, each ambiguous ball in the urn (the ball which color could be one of the two ambiguous colors) is treated as a pair of half (0.5) balls in ambiguous colors. When an ambiguous-color draw is observed, one of such pairs is replaced with a full ball in the draw's color. Importantly, this heuristic account does not necessarily require that updating processes are full-Bayesian. Below is the proof of their mathematical equivalence.

Let γ, υ, and ρ be the numbers of green, yellow, and red balls in the urn respectively. Assume that the ambiguous colors are green and yellow, and the risky color is red. Thus, although subjects do not know γ and υ, they know $\gamma + \upsilon$ ($= \alpha$) and ρ (the total number of the balls in the urn is $\alpha + \rho$). Let g (G), y (Y), and r (R) be the event of the observed (resolution) draw in green, yellow, and red, respectively.

Prior. In the Bayesian model, prior second-order probability over urn contents $(\gamma, \upsilon, \rho) = (\gamma, \alpha - \gamma, \rho)$ follows a binomial distribution:

$$P(\gamma, \upsilon, \rho) = \frac{1}{2^{\alpha}}\binom{\alpha}{\gamma} \text{ for all } \gamma \text{ s.t. } 0 \le \gamma \le \alpha.$$

Prior probability of a future draw in green is, using binomial theorem,

$$P(g) = \sum_{\gamma} P(g|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho) = \sum_{\gamma} \frac{\gamma}{\alpha + \rho} \cdot \frac{1}{2^{\alpha}}\binom{\alpha}{\gamma} = \sum_{\gamma} \frac{\alpha}{\alpha + \rho} \cdot \frac{1}{2^{\alpha}}\binom{\alpha - 1}{\gamma - 1}$$
$$= \frac{\alpha}{2(\alpha + \rho)} \sum_{\gamma} \frac{1}{2^{\alpha - 1}}\binom{\alpha - 1}{\gamma - 1} = \frac{\alpha}{2(\alpha + \rho)}.$$

It is obvious that

$$P(G) = P(y) = P(Y) = P(g) = \frac{\alpha}{2(\alpha + \rho)}.$$

Since $P(r \mid \gamma, \upsilon, \rho) = \rho/(\alpha + \rho)$ does not depend on γ (and υ), it is also obvious that

$$P(r) = P(R) = \sum_{\gamma} P(r|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho) = \frac{\rho}{\alpha + \rho} \sum_{\gamma} P(\gamma, \upsilon, \rho) = \frac{\rho}{\alpha + \rho}.$$

One can easily see that this prior probability distribution over a future draw derived from the Bayesian model is equivalent to the heuristic model, in which prior "effective" urn content is $(\gamma, \upsilon, \rho) = (\alpha/2\,, \alpha/2\,, \rho)$.

Posterior. In the Bayesian model, second-order probability is updated following Bayes' rule. After an observed draw in green g,

$$\begin{aligned} P(\gamma, \upsilon, \rho|g) &= \frac{P(g|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho)}{\sum_{\gamma} P(g|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho)} = \frac{P(g|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho)}{P(g)} \\ &= \frac{\gamma}{\alpha + \rho} \cdot \frac{1}{2^{\alpha}} \binom{\alpha}{\gamma} \cdot \frac{2(\alpha + \rho)}{\alpha} = \frac{1}{2^{\alpha-1}} \binom{\alpha - 1}{\gamma - 1}. \end{aligned}$$

Thus, posterior probability of a future draw in green after g is updated as

$$\begin{aligned} P(G|g) &= \sum_{\gamma} P(G|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho|g) = \sum_{\gamma} \frac{\gamma}{\alpha + \rho} \cdot \frac{1}{2^{\alpha-1}} \binom{\alpha - 1}{\gamma - 1} \\ &= \sum_{\gamma} \frac{\gamma - 1}{\alpha + \rho} \cdot \frac{1}{2^{\alpha-1}} \binom{\alpha - 1}{\gamma - 1} + \sum_{\gamma} \frac{1}{\alpha + \rho} \cdot \frac{1}{2^{\alpha-1}} \binom{\alpha - 1}{\gamma - 1} \\ &= \frac{\alpha - 1}{2(\alpha + \rho)} \sum_{\gamma} \frac{1}{2^{\alpha-2}} \binom{\alpha - 2}{\gamma - 2} + \frac{1}{\alpha + \rho} \sum_{\gamma} \frac{1}{2^{\alpha-1}} \binom{\alpha - 1}{\gamma - 1} = \frac{\alpha - 1}{2(\alpha + \rho)} + \frac{1}{\alpha + \rho} \\ &= \frac{\alpha + 1}{2(\alpha + \rho)} \end{aligned}$$

(binomial theorem was used twice). Similarly,

$$\begin{aligned} P(Y|g) &= \sum_{\gamma} P(Y|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho|g) = \sum_{\gamma} \frac{\alpha - \gamma}{\alpha + \rho} \cdot \frac{1}{2^{\alpha-1}} \binom{\alpha - 1}{\gamma - 1} \\ &= \frac{\alpha - 1}{2(\alpha + \rho)} \sum_{\gamma} \frac{1}{2^{\alpha-2}} \binom{\alpha - 2}{\gamma - 1} = \frac{\alpha - 1}{2(\alpha + \rho)}. \end{aligned}$$

$$P(R|g) = \sum_{\gamma} P(R|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho|g) = \frac{\rho}{\alpha + \rho} \sum_{\gamma} P(\gamma, \upsilon, \rho|g) = \frac{\rho}{\alpha + \rho}.$$

It is obvious that the posterior probability is equivalent to the heuristic account, in which posterior "effective" urn content after g is $(\gamma, \upsilon, \rho|g) = (\alpha/2 + 1/2\,, \alpha/2 - 1/2\,, \rho)$. The same applies to the posterior to y.

Lastly, a red (risky) draw r does not update the belief. In the Bayesian model, since $P(r \,|\, \gamma, \upsilon, \rho) = \rho/(\alpha + \rho)$ is constant,

$$P(\gamma, \upsilon, \rho|r) = \frac{P(r|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho)}{\sum_\gamma P(r|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho)} = \frac{P(\gamma, \upsilon, \rho)}{\sum_\gamma P(\gamma, \upsilon, \rho)} = P(\gamma, \upsilon, \rho)$$

while in the heuristic account, $(\gamma, \upsilon, \rho|r) = (\alpha/2\,, \alpha/2\,, \rho) = (\gamma, \upsilon, \rho)$. Thus, $P(G|r) = P(G), P(Y|r) = P(Y), P(R|r) = P(R)$ in both models. ■

2. Uniform second-order prior

We also noted that similar behavioral and fMRI results could be yielded when we used uniform prior, another natural choice of second-order probability. In this case,

$$P(\gamma, \upsilon, \rho) = \frac{1}{\alpha + 1} \text{ for all } \gamma \text{ s.t. } 0 \leq \gamma \leq \alpha.$$

Prior probability in this model is actually identical to the model with binomial prior:

$$P(g) = P(G) = P(y) = P(Y) = \sum_\gamma P(g|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho) = \sum_\gamma \frac{\gamma}{\alpha + \rho} \cdot \frac{1}{\alpha + 1} = \frac{\alpha}{2(\alpha + \rho)}.$$

$$P(r) = P(R) = \sum_\gamma P(r|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho) = \frac{\rho}{\alpha + \rho} \sum_\gamma P(\gamma, \upsilon, \rho) = \frac{\rho}{\alpha + \rho}.$$

It is obvious that a risky draw r does not update the belief irrespective of the prior. Thus, the only difference between predictions of the binomial and uniform prior lies in posterior probability after ambiguous-color draws:

$$P(\gamma, \upsilon, \rho|g) = \frac{P(g|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho)}{\sum_\gamma P(g|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho)} = \frac{P(g|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho)}{P(g)} = \frac{\gamma}{\alpha + \rho} \cdot \frac{1}{\alpha + 1} \cdot \frac{2(\alpha + \rho)}{\alpha}$$
$$= \frac{2\gamma}{\alpha(\alpha + 1)}$$

$$P(G|g) = \sum_\gamma P(G|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho|g) = \sum_\gamma \frac{\gamma}{\alpha + \rho} \cdot \frac{2\gamma}{\alpha(\alpha + 1)} = \frac{2\alpha + 1}{3(\alpha + \rho)}$$

$$P(Y|g) = \sum_\gamma P(Y|\gamma, \upsilon, \rho) \cdot P(\gamma, \upsilon, \rho|g) = \sum_\gamma \frac{\alpha - \gamma}{\alpha + \rho} \cdot \frac{2\gamma}{\alpha(\alpha + 1)} = \frac{\alpha - 1}{3(\alpha + \rho)}$$

while $P(R|g)$ remains the same as the prior $P(R)$.

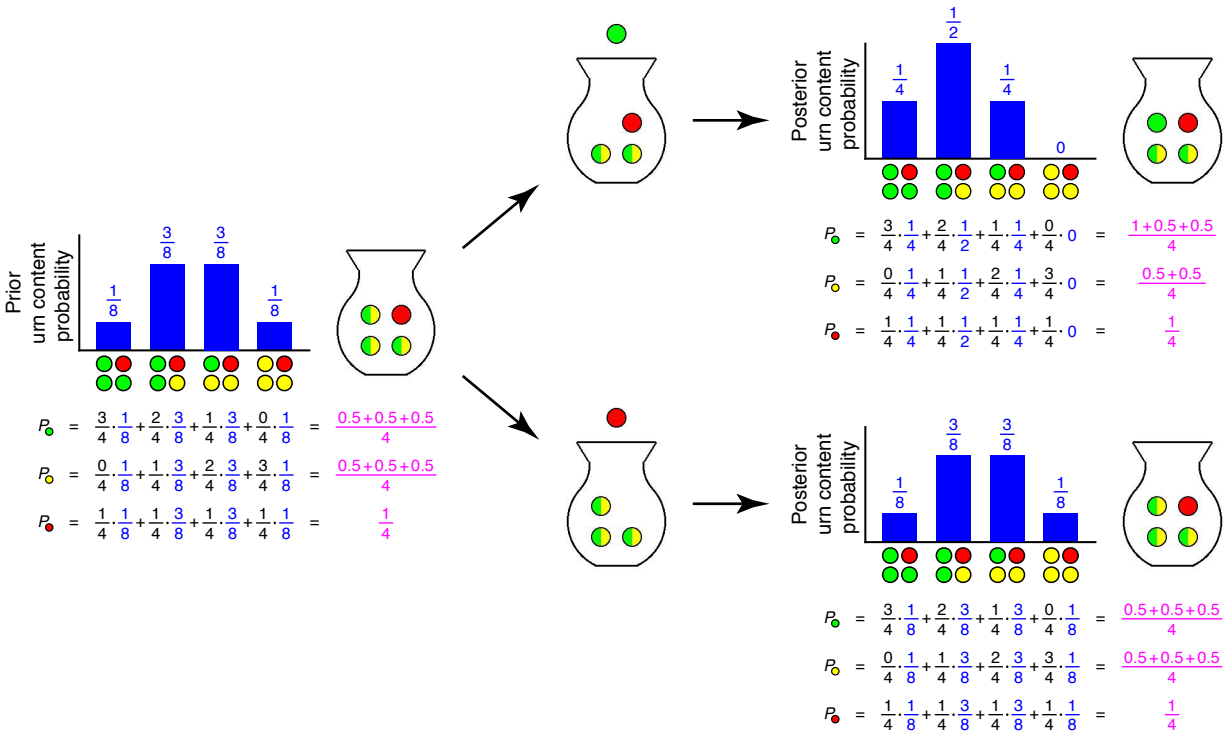
Note that equivalence to the heuristic account does not hold when the Bayesian model adopts uniform prior (or any non-binomial prior).

Behaviorally, this uniform-prior model successfully predicted values of ambiguous gambles after ambiguous-color draws ($P < .05$; Supplementary Fig. 2a), but we noted that it was outperformed by the original binomial (residual sum of squares, 109727 vs. 158925). Prediction of value updating based using the uniform model was also successful ($P < .05$), but was outperformed by binomial prior (residual sum of squares, 97861 vs. 147059). In fMRI analyses, activation maps under the uniform model were overall quite similar to the original results under the binomial model (Supplementary Fig. 2b).

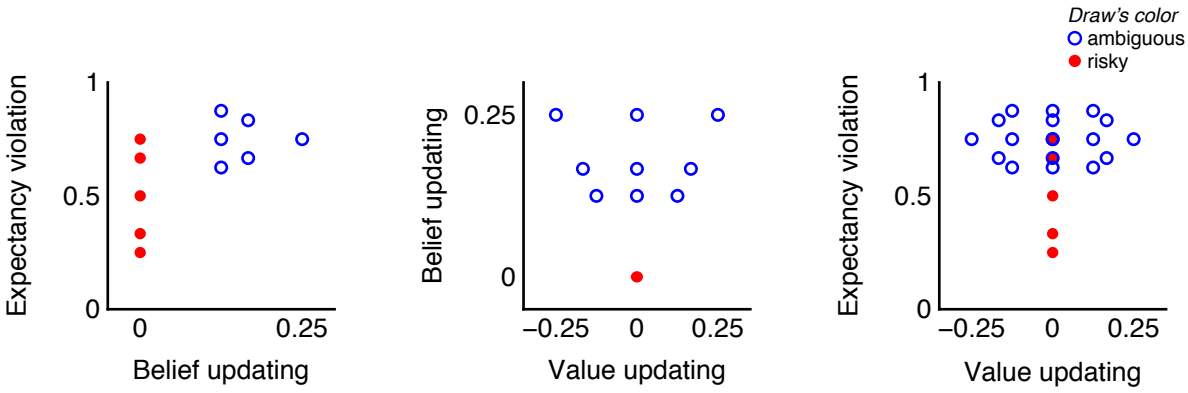
Together, these show that our conclusions are overall robust with respect to the specifics of our original Bayesian modeling.

Supplementary Figure

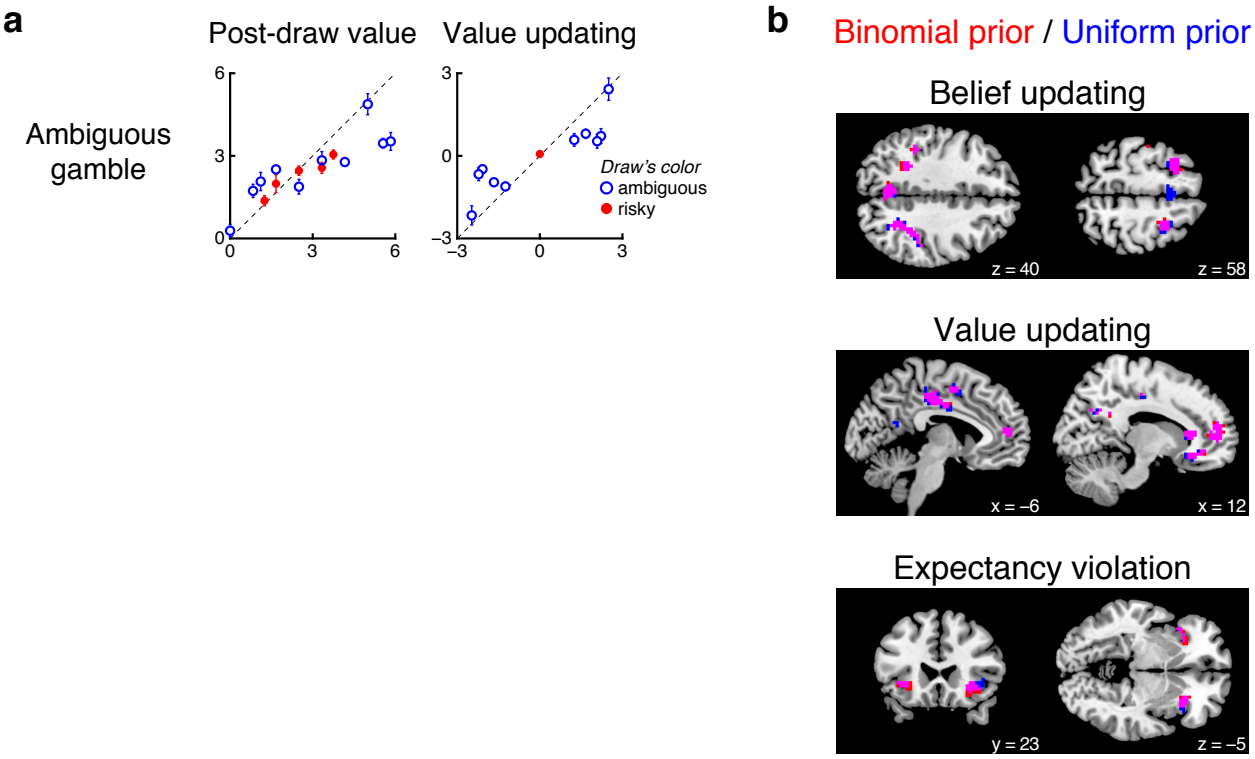
Supplementary Figure 1 Bayesian model of belief updating. Point estimates of probability of a future draw from the urn are generated in belief. To do so, second-order probability over all possible urn contents is constructed and updated according to Bayes' rule (*blue*). Shown is an exemplar urn with four possible contents (*left*). An ambiguous-color draw updates the second-order probability, eliminating the possibility of one possible urn content (*top right*). A risky-color draw does not update the prior (*bottom right*). This Bayesian account is actually mathematically equivalent to a heuristic account, which only considers "effective" urn content by treating each ambiguous ball as a pair of half (0.5) balls (*magenta*). See Supplementary Text for formal descriptions of models.



Supplementary Figure 2 Bayesian modeling with uniform prior. While results shown in Fig. 2c and Fig. 3b were obtained using the Bayesian model with second-order binomial prior, we also conducted the same analyses with uniform prior to confirm their robustness. See *Supplementary Text* for modeling details. (**a**) Behavioral results. Predictions from the uniform model differ from the binomial (original) model only in post-draw value and value updating in ambiguous gambles with ambiguous-color draws. Prediction of the uniform model was successful ($P < .05$, respectively), but was outperformed by the binomial model (Supplementary Text). Error bars: standard errors. (**b**) fMRI results. Activation maps of belief updating, value updating, and expectancy violation, obtained using the binomial model (*red*, the same as Fig. 3b) were quite similar to those obtained using the uniform model (*blue*, overlaps in *magenta*). Shown are clusters that survived voxel-wise threshold $P < .001$ (uncorrected) and cluster-size threshold $k > 10$.

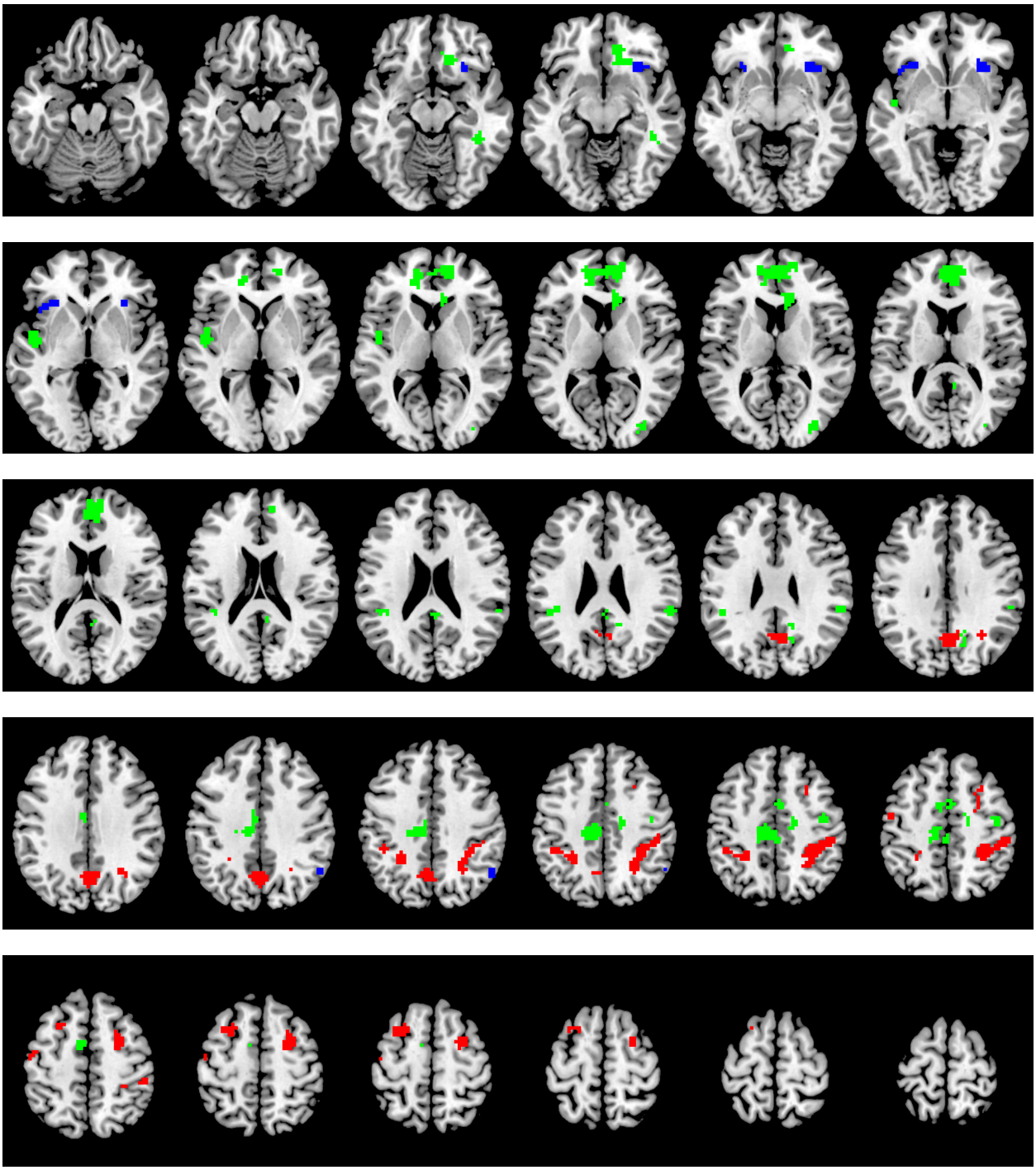


Supplementary Figure 3 Correlation among belief updating, value updating, and expectancy violation. The variables are derived under the Bayesian model with second-order binomial prior (Supplementary Fig. 1; Supplementary Text). *Left:* Even though belief updating does not coincide with expectancy violation, their correlation is not negligible in our parameter space ($r = 0.70$). *Middle and right:* As value updating could be positive or negative, it is orthogonal to both belief updating and expectancy violation by design. See also Supplementary Table 1.

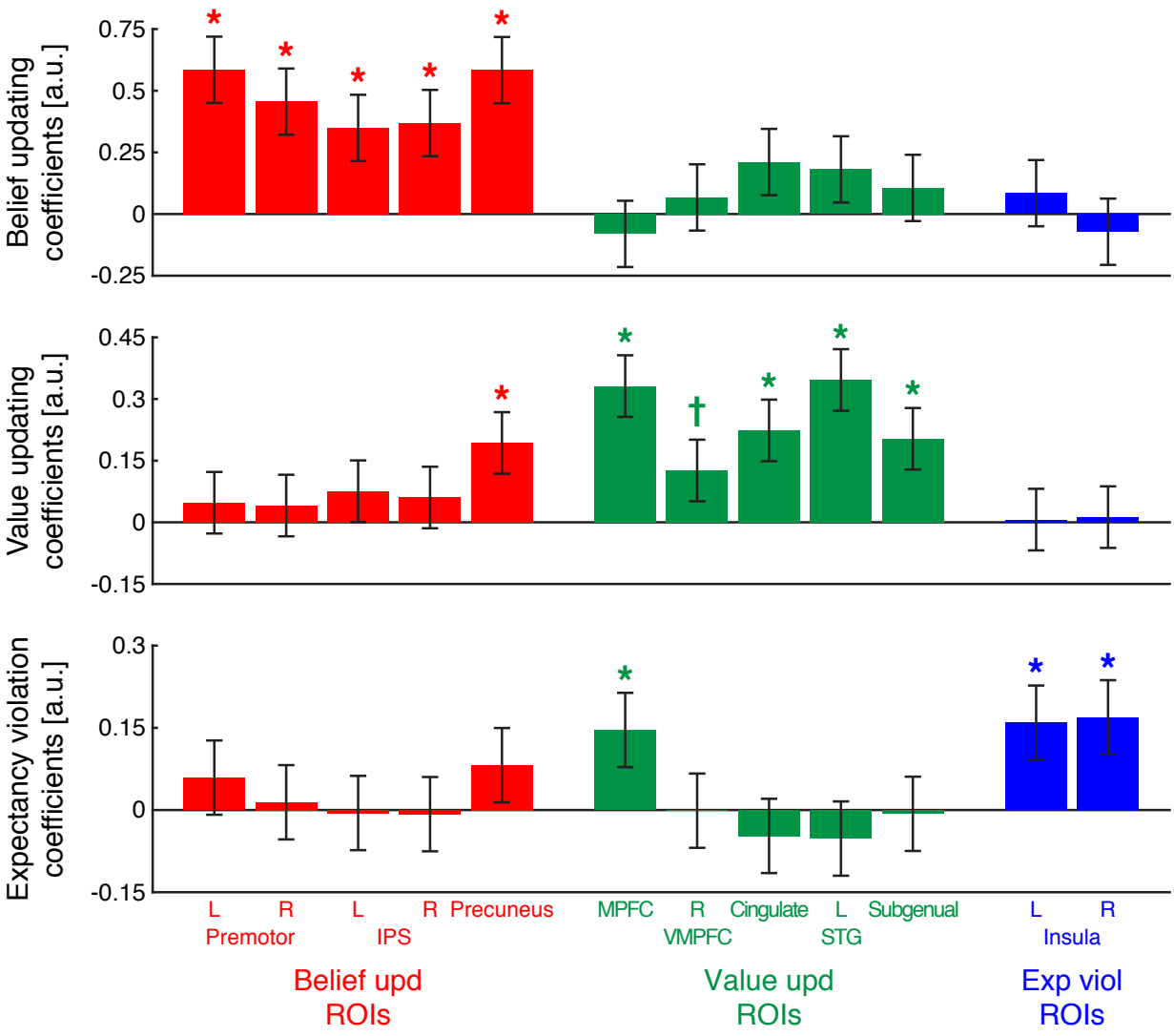


Supplementary Figure 4 Whole-brain activation maps of belief updating, value updating, and expectancy violation. The same results as Fig. 3b. Shown are clusters that survived voxel-wise threshold $P < .001$ (uncorrected) and cluster-size threshold $k > 10$.

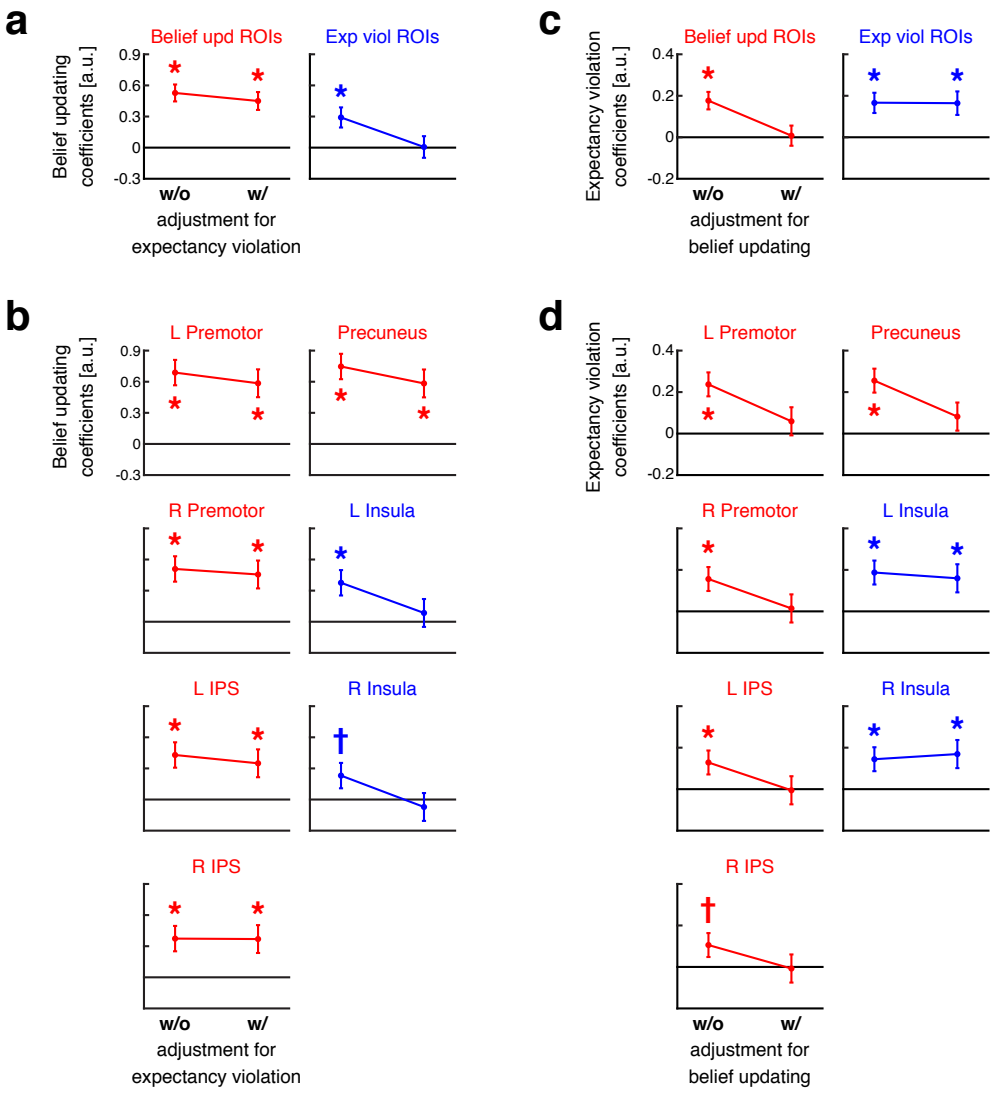
Belief updating / Value updating / Expectancy violation



Supplementary Figure 5 ROI-wise analysis results. While analysis in Fig. 3c aggregated activations from all ROIs in the same categories (belief updating, value updating, or expectancy violation), results in this figure were obtained on ROI-wise basis. *Top:* Belief updating was correlated with activation in all of the belief-updating ROIs ($P < .05$), but none of value-updating ROIs or expectancy-violation ROIs ($P > .10$). *Middle*: Value updating was correlated with activation in all of the value-updating ROIs ($P < .05$; $P < .10$ in right VMPFC) and a belief-updating ROI in precuneus ($P < .05$), but not in the rest of the belief-updating ROIs or any of the expectancy-violation ROIs ($P > .10$). Note that the precuneus cluster was close to a value-updating ROI in cingulate cortex (Table 2, Supplementary Fig. 4). *Bottom*: Expectancy violation was correlated with activation in both of the expectancy-violation ROIs ($P < .05$) as well as a value-updating ROI in MPFC ($P < .05$), but not in the rest of the value-updating ROIs or any of the belief-updating ROIs ($P > .10$). Error bars: SEMs, asterisks: $P < .05$, daggers: $P < .10$.



Supplementary Figure 6 The effect of adjustment for correlation between belief updating and expectancy violation. Belief updating and expectancy violation are correlated in our parameter space (Supplementary Fig. 3), and if we had not controlled for this correlation, we would have been unable to dissociate their neural correlates. (**a**) Coefficients estimates for belief updating in belief-updating ROIs (*left*) and expectancy-violation ROIs (*right*). The left plots in each panel show estimates without adjustment for expectancy violation (from GLM 2; see Methods), and the right plots show estimates with adjustment (from GLM 1). Even expectancy-violation ROIs showed significant association with belief updating when adjustment was not applied. (**b**) The same as (**a**), for every ROI. (**c**) Coefficients estimates for expectancy violation in belief-updating ROIs (*left*) and expectancy-violation ROIs (*right*). The left plots show estimates without adjustment for belief updating (from GLM 1), and the right plots show estimates with adjustment (from GLM 2). Even belief-updating ROIs showed significant association with expectancy violation when adjustment was not applied. (**d**) The same as (**c**), for every ROI. Error bars: SEMs, asterisks: $P < .05$, daggers: $P < .10$.



Supplementary Table 1. Distribution of observed value updating in the behavioral experiment. Shaded cells conform to normative prediction (Fig. 1e). See Fig. 2b for histograms.

		> +¢50	(0, +¢50]	0	[-¢50,0)	< -¢50
ambiguous gambles	Ambiguous-color draws: positive	56.11%	26.94%	9.17%	0.28%	7.50%
	Ambiguous-color draws: negative	3.89%	1.94%	11.67%	22.78%	59.72%
	Risky-color draws	6.39%	6.67%	71.39%	10.83%	4.72%
risky gambles	Ambiguous-color draws	2.50%	2.78%	79.17%	7.50%	8.06%
	Risky-color draws	2.22%	3.33%	79.44%	3.89%	11.11%

Chapter 2
Unified account on preference for instrumental and non-instrumental information

Introduction

Human ubiquitously attempts to acquire information relevant to their decisions and goals. By acquiring information about the state of the environment and incorporating it into their choices, agents can increase their chance to obtain rewards and avoid punishments. However, information is often associated with various forms of costs, such as time, effort, or money, and agents need to consider if a piece of information is worth its cost. Moreover, there are often multiple sources of information in the environment, and agents need to pick up one that would be the most useful. Therefore, while adaptive acquisition of information is a part of value-based decision-making, it constitutes a class of challenging decision-making problems by itself.

Information acquisition has been studied in various forms, in various fields, and in various conceptual schemes (Fig. 1). Most notably, economists have postulated a normative theory on how rational agents should acquire information to maximize expected utility (EU) (Bandyopadhyay, 1977; Edwards, 1965; Howard, 1966; 1967; Marschak, 1971; 1973; Medlin, 1979; Schepanski & Uecker, 1984; Wendt, 1969). The theory prescribes that agents compute value of information (VoI) and acquires the information only if its value exceeds its cost (or values of alternative information pieces). It further states that information is valued only to the extent to which it improves agents' choices and their EU.

An important tenet of this normative theory is that information should be positively valued only if it is instrumental, i.e., only if it would affect agents' future choices. For instance, if an agent considers what to do this weekend, the weather forecast may directly affect its decision (e.g., it may go hiking if it is predicted to be sunny, but it may instead go to a theater if it is expected to be rainy). The weather forecast is valuable in this case because it will influence the agent's choice and improve overall expected utility. On the other hand, if the agent has already decided to go to a theater and purchased a ticket, the weather forecast will not affect the agent's choice anymore and thus is of little value. Such non-instrumental information should not be acquired when associated with non-zero costs.

This normative theory on instrumental VoI provides precise predictions on information acquisition behavior, and it has achieved a certain degree of descriptive validity (Medlin, 1979; Schepanski & Uecker, 1984; Shanteau & Anderson, 1972; Wendt, 1969). Particularly, because the theory assumes EU-maximizing agents, it explains the influence of agents' risk attitudes on their information acquisition (Abbas, Bakir, Klutke, & Sun, 2013; Bakır & Klutke, 2011; Berg & Eisenberger, 1996; Blair & Romano, 1988; Mehrez, 1985; Nadiminti, Mukhopadhyay, & Kriebel, 1996). However, the theory has three major drawbacks that have limited its influence. First, precise calculation of instrumental VoI requires agents to simulate their choices and utilities under all possible information states, which is a computationally demanding process and particularly problematic in the artificial intelligence (AI) field. Second, the theory does not provide any predictions when the utility function is not available. The idea of EU maximization is not widely

used by psychologists studying curiosity, i.e., people's tendency to acquire information for its own sake. Third, studies in economics and animal psychology have reported that that human and animal exhibit consistent preference for information that resolves uncertainty earlier than later, even if it is not instrumental (Ahlbrecht & Weber, 1996; Beierholm & Dayan, 2010; Bennett, Bode, Brydevall, Warren, & Murawski, 2016; Blanchard, Hayden, & Bromberg-Martin, 2015; Bromberg-Martin & Hikosaka, 2009; 2011; Chew & Ho, 1994; Eliaz & Schotter, 2007; Kreps & Porteus, 1978; McDevitt, Dunn, Spetch, & Ludvig, 2016; Wu, 1999). Such preference has been observed even when the information is costly, violating the tenet of the normative theory.

Instead of the normative theory, various fields have proposed their own theoretical accounts for information acquisition. Some of them attribute information acquisition to a non-EU-maximizing drive for information that human is endowed with. For instance, AI literature has proposed a bonus utility term added to actions that might lead to information acquisition (Dayan & Sejnowski, 1996; Ishii, Yoshida, & Yoshimoto, 2002; Moore & Atkeson, 1993; Sutton & Barto, 1998; Yoshida & Ishii, 2005). In its most basic form, the bonus is constant across all actions or those that have not been taken in the recent past. Moreover, studies on curiosity postulate a drive to reduce the gap between what is already known and what could be potentially known, called information gap (Loewenstein, 1994). Although quantitative modeling of curiosity has not been established, information gap is often assumed to be the difference in entropy. Agents with the constant information bonus or the entropy reduction drive would obtain information more often than rational EU-maximizing agents, while not necessarily taking into account specific future choices and outcomes.

Economists, on the other hand, have proposed penalty for states or choices in which agents could have obtained, but have not obtained, some information(Kreps & Porteus, 1978; Wu, 1999). It was originally proposed by studies on temporal lotteries, in which majority of subjects preferred lotteries resolved earlier to those resolved later, even if they were otherwise identical (i.e., timing of resolution was non-instrumental), and even if earlier resolution was costly (Ahlbrecht & Weber, 1996; Bennett et al., 2016). Although such preference is inconsistent with the normative theory, it could be explained by introducing penalty for being uninformed; agents prefer earlier resolution because they are averse to being uninformed for a longer time. Importantly, it has been observed that preference on timing of uncertainty resolution is dependent on outcomes at stake; most subjects prefer earlier resolution when a large positive outcome is at stake (e.g., big monetary prizes), but some prefer later resolution when a large negative outcome is at stake (e.g., cancer diagnosis) (Ahlbrecht & Weber, 1996; Chew & Ho, 1994; Eliaz & Schotter, 2007). Such sensitivity to outcomes cannot be explained by the constant information bonus or the drive for entropy reduction.

Therefore, a number of conceptual schemes on information acquisition have been proposed in related fields. So far, however, these concepts have been studied largely separately. For instance, because studies on temporal lotteries have focused on non-instrumental information, it remains unclear whether and how much preference for earlier resolution affects behavior in cases under the presence of instrumental VoI (Beierholm & Dayan, 2010). Similarly, it has not been well investigated how curiosity is sensitive to decision-related variables, such as possible future outcomes, and to what extent it is prioritized over EU-maximizing information acquisition. Literature on the instrumental VoI theory, on the other hand, has largely ignored preference for non-

instrumental information, leaving the theory normative rather than descriptive. More problematically, most existing accounts are either qualitative or axiomatic and thus difficult to falsify.

Intuitively, human information acquisition may be driven by more than one factors, and instrumental VoI (or some form of its approximation) may be one of them. To understand information acquisition, therefore, we need to examine what non-normative factors additionally contribute to behavior and to what extent. From the theoretical perspective, this requires falsifiable models that unify instrumental VoI and one or more non-normative accounts. From the empirical perspective, we need tasks that include instances of instrumental information and rich parameterization of decision-related variables, such that the contribution of a non-normative factor could be quantified and compared against normative and other non-normative factors. These approaches have rarely been taken in the past.

This current study aims to address these issues by adopting a new experimental paradigm on information purchase and postulating quantitative, yet descriptive, modeling of subjective VoI. The normative model of instrumental VoI is used as a standard (baseline) model, to which the non-normative accounts from AI, psychology, and economics are added and tested against the empirical data.

Results

This study adopted a new task paradigm to quantify information acquisition behavior (Fig. 2a). Subjects' primary decision in the task was to choose whether they would accept or reject a monetary gamble with two possible outcomes (x_1, x_2). The gamble was visually presented as a roulette wheel, and when it was played, a dot appeared on the perimeter and its side determined the outcome. Because the dot's location followed uniform distribution across the perimeter, the two outcomes were equally likely at this point: $P_{\text{uninfo}}(x_1) = P_{\text{uninfo}}(x_2) = 0.5$.

After making the choice on the gamble with P_{uninfo} (hereafter an uninformed choice), subjects were presented with a piece of information and indicated whether, or how much, they were willing to purchase it. The information, if purchased, would update the outcome's probability distribution and improve its predictability. Visually, the information was presented as a partition line on the wheel, and if subjects paid the information cost, the side of the partition line in which the dot would appear was revealed. This means that the information revealed the true outcome probability among two possible distributions: $P_{\text{info}}(x_1) = 1 - P_{\text{info}}(x_2) = p$ or $1 - p$, where $p\ (\leq 0.5)$ corresponds to the angle of the partition line. The lower p was, the more diagnostic the information was; the diagnosticity was the highest when $p = 0$ (vertical partition line) and the lowest when $p = 0.5$ (horizontal partition line). Being informed of the true probability distribution P_{info}, subjects made the choice on the gamble again (hereafter an informed choice), overriding their prior uninformed choice. This opportunity for subjects to change their choices is exactly the source of instrumental VoI.

Two independent dataset were analyzed, one during fMRI scanning and one from a purely behavioral experiment. Paradigms adopted in these datasets were almost identical. They differed in the ways subjects elicited their preference on information purchase; fMRI subjects made a binary choice on whether to purchase information at a given monetary cost (variable across trials), while behavioral subjects reported their willingness-to-pay for information in BDM bidding procedure

(Becker & Brownson, 1964). They also differed in the gamble outcomes and partition line angles (see Methods and Fig. 3c, d). Findings on neural representation of VoI from the fMRI experiment are reported in Chapter 3.

Estimation of utility function.
First, the descriptive validity of the normative theory on instrumental VoI was evaluated. This model yields prediction of VoI, hereafter w, based on the agent's utility function u. To obtain w, group-level utility function was first estimated based on subjects' uninformed choices on gambles. Subjects exhibited risk aversion in both behavioral and fMRI experiments, which was successfully captured by asymmetric power utility functions (Supplementary Fig. 1).

Standard model of instrumental VoI.
VoI prediction w under the normative, standard model can be derived using a decision tree (Fig. 2b). The first branches in the decision tree concern the informed choice (Fig. 2b, *right*) and the uninformed choice (*left*), or in other words, whether the information should be purchased at the cost c or not. When the cost is equal to the normative prediction of VoI ($c = w$), agents should be indifferent to information purchase, i.e., EU of the uninformed choice EU_{uninfo} should be equal to the informed choice EU_{info}.

Assuming that EU-maximizing policy is fully deterministic (i.e., agents always choose the action with the highest EU), EU_{uninfo} is given as

$$EU_{\text{uninfo}}(x_1, x_2) = \max(0.5 \cdot u(x_1) + 0.5 \cdot u(x_2), 0) \quad (1)$$

where utility of rejecting the gamble is 0 as agents do not gain or lose anything.

When the information is purchased, its content would be stochastic; $P_{\text{info}}(x_1)$ could turn out to be either p or $1 - p$ ($p \leq 0.5$) at the equal likelihood. Hereafter the former information state is called s_1, and the latter s_2. Overall EU of the informed choice EU_{info} is the average of EU-maximizing choices under the two information states:

$$EU_{\text{info}}(x_1, x_2, p, c) = 0.5 \cdot EU_{\text{info}}(x_1, x_2, p, c|s_1) + 0.5 \cdot EU_{\text{info}}(x_1, x_2, p, c|s_2) \quad (2)$$

where EU under each information state is

$$EU_{\text{info}}(x_1, x_2, p, c|s_1) = \max\big(p \cdot u(x_1 - c) + (1 - p) \cdot u(x_2 - c), u(-c)\big) \quad (3)$$
$$EU_{\text{info}}(x_1, x_2, p, c|s_2) = \max\big((1 - p) \cdot u(x_1 - c) + p \cdot u(x_2 - c), u(-c)\big) \quad (4)$$

because agents need to pay the sunk cost of information purchase c even if they end up rejecting the gamble.

EU-maximizing agents are supposed to purchase the information when EU_{info} is higher than EU_{uninfo}, and forgo otherwise. This prediction was supported by fMRI experiment dataset, in which subjects indicated their preference for information purchase in binary choices. They purchased information more frequently when the difference between EU_{info} and EU_{uninfo} was larger, which was successfully modeled by the logit (softmax) function of the EU difference (negative log-likelihood

= 2889.7; $p < .0005$, permutation; Fig. 3a). This means that the amount to which subjects chose to pay for the information was overall sensitive to x_1, x_2, and p, as predicted by Eqs. 1-4 (Fig. 3c).

When $c = w$, EU-maximizing agents are supposed to be indifferent to information purchase. Therefore, the normative prediction of VoI w is the solution of

$$EU_{\text{uninfo}}(x_1, x_2) = EU_{\text{info}}(x_1, x_2, p, w) \quad (5)$$

The solution w, indeed, successfully explained willingness-to-pay reported by subjects in the behavioral experiment's (mean squared error = 2.2116, $p < .0005$, permutation; Fig. 3b). Again, this indicates that willingness-to-pay exhibited dependence on x_1, x_2, and p that was approximately consistent with Eq. 5 (Fig. 3d).

Another important feature of the standard model is its sensitivity to agents' risk attitude. To test this, it was examined whether the information purchase behavior could be modeled under the assumption of expected-value (EV) maximization (i.e., linear u) equally well. Even though EV-based predictions also performed decently well, they were outperformed by EU-based predictions in both fMRI and behavior experiments (Fig. 3e, f). Note that this was not due to overfitting; the utility functions were obtained from uninformed choices on gambles, not from the information purchase behavior *per se*. This illustrates the importance of taking risk attitude into account, which is rarely done outside behavioral economics and decision-making studies.

Violation of the standard model's prediction.

Analysis so far has shown that the standard model provides a decent approximation of information acquisition. To test the standard model more critically, we next examined its most important prediction; when information is non-instrumental, i.e., when the informed choices under s_1 or s_2 are the same to the uninformed choice, its value w should be zero (see Appendix for proof). This is an important tenet of the normative theory, which sees information acquisition as nothing more than a part of EU maximization. No other existing accounts on information acquisition make the same claim that value of non-instrumental information is always zero.

To test this key prediction that non-instrumental information's value is zero, let us consider an exemplar gamble in which one of the two outcomes was 0 (such gambles were included only in the behavioral experiment). It is obvious that any information in such a gamble is non-instrumental, because the gamble should be always accepted if the other outcome is positive, and rejected if the other outcome is negative. Contrary to this prediction, observed willingness-to-play was significantly higher than zero in the ($10, $0) gamble ($t(95) = 7.2955$, $p < 10^{-10}$; Fig. 4a). This cannot be explained by subjects' misunderstanding of the gamble, because most subjects accepted to play it as expected (Fig. 4b). This evidences that non-standard preference for non-instrumental information does indeed exist. More interestingly, however, while willingness-to-play in the ($0, −$10) gamble was also larger than zero ($t(95) = 3.1238$, $p < .01$; Fig. 4a), it was significantly smaller than ($10, $0) ($t(95) = 6.8185$, $p < 10^{-9}$). This illustrates that the preference for non-instrumental information is sensitive to outcomes at stake.

Preference for non-instrumental information itself has been already documented by previous studies. However, it has not been well established how important such preference is under the

presence of the instrumental VoI. Most past studies adopted tasks in which subjects never acquire instrumental information, and as pointed out in (Beierholm & Dayan, 2010), it would be possible that the behavioral effect of such preference disappears when information could be instrumental; subjects may value information no more than the normative predictions under such circumstances. If this is the case, subjects in the current task would not indicate higher willingness-to-pay than the normative prediction w when the latter is non-zero. On the other hand, if some non-standard factors indeed contribute to the behavior even when information is instrumental, subjects would exhibit consistent overvaluation.

To test these possibilities, the difference between willingness-to-pay and the standard predictions w (or overvaluation) was examined across all gambles (Fig. 4c, d). Indeed, subjects consistently exhibited overvaluation ($t(95) = 10.7569$, $p < 10^{-10}$), evidencing the general behavioral effect of non-standard preference. Furthermore, overvaluation was larger in the gambles with positive EU than the ones with negative EU ($t(95) = 6.4501$, $p < 10^{-8}$), and was an inverted-U function of p (Fig. 4c). Sensitivity to EU and p was consistently observed over the majority of trials (Fig. 4d). These results show that the non-standard factor's contribution is dependent on outcomes at stake and information diagnosticity.

These conclusions were also supported by the independent dataset from fMRI experiment. Subjects purchased information more often than normatively predicted ($t(36) = 5.2996$, $p < 10^{-5}$; Fig. 4e, f), and such overpurchase was even higher among the gambles with positive EU than negative EU ($t(36) = 7.1152$, $p < 10^{-7}$). Even though the fMRI experiment had only three levels of p, non-monotonic relationship between the overpurchase and p was observed (Fig. 4e), more or less consistently over gambles (Fig. 4f).

Taken together, these results show that the standard model is not enough to account for information purchase behavior. Its descriptive validity could be improved by combining it with some non-standard accounts, which would allow overvaluation and its sensitive to outcomes and information diagnosticity.

Non-standard models of VoI.

To account for the observed violation of the standard prediction, we next examined alternative accounts on information acquisition. All of the examined accounts allow preference for non-instrumental information and overvaluation of instrumental information. The current study does not aim to compare these accounts against the standard model alone; rather, it aims to explore how to improve the standard model by combining it with the alternative accounts. Therefore, each account was implemented as a modification of the standard model; the size of its influence was controlled by a free parameter.

To our knowledge, little to no attempt has been made to combine the non-standard accounts with the normative theory on instrumental VoI to develop quantitative, concrete, and falsifiable models. Therefore, before describing the results of formal modeling analysis, the way each account could be implemented into the standard model is first described below. The examined accounts were: 1) constant information bonus, which has no systematic sensitivity to decision-related variables (x_1, x_2, p), 2) drive for entropy reduction, which is sensitive to p but not to outcomes at stake (x_1, x_2),

and 3) a class of economic models on penalty for uninformativeness, which are sensitive to outcomes at stake (x_1, x_2), and some of which are also sensitive to p (Fig. 4).

1. Constant information bonus

AI literature has long studied how to design agents that acquire information efficiently. A problem of particular interest is the conflict between exploration and exploitation; even when agents already know that some actions lead to rewards, they should not keep taking those actions (exploitation), and instead examine other actions' outcomes from time to time (exploration), because some of them might be better the known associations. One common solution to this conflict, which is not necessarily optimal but excels at computational efficiency, is "exploration bonus," which is a constant term added to utility of non-exploiting actions (Dayan & Sejnowski, 1996; Ishii et al., 2002; Moore & Atkeson, 1993; Sutton & Barto, 1998; Yoshida & Ishii, 2005). This idea has also been used by neuroscience studies due to its simplicity, but its descriptive validity has been questioned (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006).

In our paradigm, this account assumes that a constant term is added the utility of informed choice (Eq. 2):

$$EU^*_{\text{info}}(x_1, x_2, p, c) = 0.5 \cdot EU_{\text{info}}(x_1, x_2, p, c|s_1) + 0.5 \cdot EU_{\text{info}}(x_1, x_2, p, c|s_2) + C_{\text{bonus}}$$

where C_{bonus} is a free parameter. VoI prediction under this model w^* satisfies $EU_{\text{uninfo}}(x_1, x_2) = EU^*_{\text{info}}(x_1, x_2, p, w^*)$.

A key prediction of this model is that agents are more willing to purchase information than the standard model (i.e., $w^* \geq w$). However, it also predicts that the overvaluation ($w^* - w$) is not distinguishable between gambles with positive and negative EUs, and is not sensitive to p (Fig. 5). Thus, this model fails to characterize the behavioral overvaluation, which was larger in gambles with positive EU and ones with negative EU.

2. Entropy reduction

In psychology, information acquisition is often seen as a manifestation of curiosity, or an internal drive to obtain information for its own sake. One widespread idea is that curious agents attempt to reduce the gap between the amount of information they currently possess and the amount of information they could potentially possess, and that such information gap could be measureable as the difference of entropy (Berlyne, 1957; 1966; Loewenstein, 1994).

In the current task, the information updates outcome probability from $P_{\text{uninfo}}(x_1) = 0.5$ to $P_{\text{info}}(x_1) = p$ or $1 - p$. Thus, entropy reduction is zero when $p = 0.5$, and it grows as p decreases. To implement this account, a bonus term of entropy reduction can be added to the utility of informed choice (Eq. 2):

$$EU^*_{\text{info}}(x_1, x_2, p, c) = 0.5 \cdot EU_{\text{info}}(x_1, x_2, p, c|s_1) + 0.5 \cdot EU_{\text{info}}(x_1, x_2, p, c|s_2) + C_{\text{ent}}(-\Delta\text{entropy})$$

where C_{ent} is a free parameter.

As the constant bonus model, this model also predicts overvaluation that is not systematically dependent on outcomes (Fig. 5). One critical difference from the constant bonus model is that this model predicts overvaluation's sensitivity to p. However, since the entropy reduction is a decreasing function of p, so is the predicted overvaluation, which is inconsistent with the observed inverted-U relationship.

3. Economic models: penalty for uninformed choices

Analyses so far highlighted the need for models that allow sensitivity to outcomes at stake. This has been addressed by theoretical accounts proposed from economic studies on temporal lotteries, namely subjects' preference for earlier resolution of uncertainty. Evidence has shown that such preference is prominent when the lottery involves a big positive outcome, but not necessarily when it involves a big negative outcome (Ahlbrecht & Weber, 1996; Berns et al., 2006; Chew & Ho, 1994).

Theoretical accounts for such preference have been proposed by Kreps & Porteus (1978) (Kreps & Porteus, 1978) and Wu (Wu, 1999). Both theories introduce penalties for making uninformed choice; utility is reduced when agents could have obtained, but did not obtain, a piece of information. Kreps & Porteus reduce EU of uninformed choices or states via recursive utility, while Wu does so by introducing non-linear probability weighting function. These accounts were originally proposed only axiomatically, and subsequent studies that empirically tested them remained largely qualitative. Moreover, to our knowledge, they have not been applied to settings where information could be instrumental. Therefore, the current study marks the first attempt to quantify penalty for uninformed choices in the presence of instrumental information.

3-1. Recursive utility

In Kreps & Porteus, informed and uninformed choices are evaluated based not on utility u but on recursive utility $v(u)$. Furthermore, while first-order utility is averaged over outcomes or states in informed choice ($v(\overline{u})$), recursive utility for each outcome or state is first obtained and then averaged in uninformed choice ($\overline{v(u)}$). Therefore, concavity of v explains preference for informed choice, even when the information is non-instrumental. The original theory does not specify whether recursive utility should be averaged over outcomes (x_1, x_2) or information states (s_1, s_2); this study considers both formulations due to the lack of empirical evidence for either one.

Under the outcome-recursive-utility model, Eqs. 1 and 2 should be rewritten as

$$EU^*_{\text{uninfo}}(x_1, x_2) = \max\big(0.5 \cdot v\big(u(x_1)\big) + 0.5 \cdot v\big(u(x_2)\big), 0\big)$$
$$EU^*_{\text{info}}(x_1, x_2, p, c) = 0.5 \cdot v\big(EU_{\text{info}}(x_1, x_2, p, c|s_1)\big) + 0.5 \cdot v\big(EU_{\text{info}}(x_1, x_2, p, c|s_2)\big)$$

and w^* satisfies $EU^*_{\text{uninfo}}(x_1, x_2) = EU^*_{\text{info}}(x_1, x_2, p, w^*)$. Concavity of v is captured by a free parameter C_{KP} in a power function $v = u^{(1-C_{\text{KP}})}$. (Note that the power function is adopted only because of its simplicity; there has been no attempt to empirically specify the form of underlying v to our knowledge.)

To construct the state-recursive-utility model, we first need to rewrite Eq. 1 to separate EU of the uninformed acceptance under the information states (s_1 and s_2):

$$0.5 \cdot u(x_1) + 0.5 \cdot u(x_2) = 0.5\big(p \cdot u(x_1) + (1-p) \cdot u(x_2)\big) + 0.5\big((1-p) \cdot u(x_1) + p \cdot u(x_2)\big)$$

and then apply recursive utility:

$$EU^*_{\text{uninfo}} = \begin{cases} 0.5 \cdot v\big(p \cdot u(x_1) + (1-p) \cdot u(x_2)\big) + 0.5 \cdot v\big((1-p) \cdot u(x_1) + p \cdot u(x_2)\big) & \text{if } EU_{\text{uninfo}} > 0 \\ 0 & \text{if } EU_{\text{uninfo}} \leq 0 \end{cases}$$

$EU^*_{\text{info}}(x_1, x_2, p, c)$ is the same as the outcome-recursive-utility model above.

Under these models with concave v (i.e., positive C_{KP}), overvaluation would occur ($w^* \geq w$) only when $EU_{\text{uninfo}} > 0$ (Fig. 5). This is qualitatively consistent with the behaviorally observed overvaluation. An important difference between the outcome model and the state model lies in predicted overvaluation's sensitivity to p; while the outcome model predicts constant overvaluation over p, the state model predicts that overvaluation decreases with p. However, even though subjects indeed exhibited overvaluation that was sensitive to p, it was not a monotonic decreasing function as predicted by the state model; it showed inverted-U relationship. The fact that the state model allows sensitivity to p, therefore, does not necessarily make it attractive compared to the outcome model.

3-2 Probability weighting

According to Wu's theory, EU of uninformed choices is calculated based on probability weighting function. Probability weighting could be applied to outcomes or states. In general, predictions of this model critically depend on the form or weighting function, which has not been empirically specified. However, since the two outcomes (x_1, x_2) or two states (s_1, s_2) are equiprobable in the current paradigm, we only need one free parameter that encodes weight of probability 0.5 (C_{w} below).

In the outcome-probability-weighting model, predicted VoI w^* satisfies $EU^*_{\text{uninfo}}(x_1, x_2) = EU_{\text{uninfo}}(x_1, x_2, p, w^*)$, where

$$EU^*_{\text{uninfo}}(x_1, x_2) = \max\big((0.5 - C_{\text{w}}) \cdot u(x_1) + (0.5 + C_{\text{w}}) \cdot u(x_2), 0\big)$$

assuming $x_1 > x_2$. EU^*_{uninfo} under the state-probability-weighting model is derived similarly to the state-recursive-utility model:

$$EU^*_{\text{uninfo}} = \begin{cases} (0.5 + C_{\text{w}}) \cdot \big(p \cdot u(x_1) + (1-p) \cdot u(x_2)\big) + (0.5 - C_{\text{w}}) \cdot \big((1-p) \cdot u(x_1) + p \cdot u(x_2)\big) & \text{if } EU_{\text{uninfo}} > 0 \\ 0 & \text{if } EU_{\text{uninfo}} \leq 0 \end{cases}$$

These models' predictions are quantitatively similar to the recursive utility models (Fig. 5), even though they are conceptually different. The main difference between them lies in quantitative sensitivity of overvaluation to outcomes, but it also critically depends on the form of underlying recursive utility function or weighting function, rigorous specification of which is beyond the scope of the current study. If goodness-of-fit drastically differs among these economic models, it would suggest that empirical specification of the underlying functions, rather than axiomatic derivation, is important in the future modeling effort.

Non-standard model evaluation – fMRI experiment dataset.
Model-free analysis suggested that the standard model could be improved by allowing overvaluation, or contribution of non-standard factors. Accounts from economics seem particularly promising because they predict that overvaluation is sensitive to outcomes at stake. To formally test these observations, the modified models were formally evaluated based on the information purchase behavior from fMRI experiment.

Subjects' behavior supported the non-standard accounts. All modified models achieved lower AICs than the standard model (Fig. 6a). Estimates of the free parameters (C_{bonus}, C_{ent}, C_{KP}, and C_{w}) were all significantly higher from zero, which is the standard model's assumption ($p < .001$, bootstrap, Fig. 6b), confirming their contributions to the model fit. These echo the earlier observation that the standard model can be improved by allowing overvaluation.

To statistically evaluate the difference in model fit while addressing the issue of overfitting, 10-fold cross-validation (200 iterations) was conducted. As a result, all of the four economic models outperformed the standard model ($p < .05$, Bonferroni corrected). To the contrary, the constant bonus or the entropy reduction did not improve the standard model. These confirm that overvaluation is sensitive to outcomes at stake, which cannot be explained by constant bonus or entropy reduction alone.

The economic models' advantage over the constant bonus and entropy reduction models was, however, not clear in the direct model comparison. AICs of the former were overall lower than the latter (Fig. 6a). However, statistical testing using 10-fold cross-validation revealed that, while three of the four economics models (outcome-recursive-utility, state-recursive-utility, and outcome-probability-weighting) outperformed the entropy reduction model (Table 1; $p < .05$, Bonferroni corrected), none of them outperformed the constant bonus model ($p > .05$). The difference among the economics models' goodness-of-fit was also minor ($p > .05$, uncorrected). Experiment designs with higher power would be needed to specify the form of functions underlying the economics models and to discriminate them against alternatives rigorously.

Non-standard model evaluation – behavioral experiment dataset.
The modified models were also evaluated and compared based on willingness-to-pay in the behavioral experiment. In this analysis, goodness-of-fit was measured as mean squared error (MSE), and statistical significance of its difference was calculated by 10-fold cross-validation procedure (200 iterations).

The behavioral experiment dataset also supported some non-standard accounts. All modified models, except for entropy reduction and state-probability-weighting, yielded statistically better goodness-of-fit than the standard model ($p < .05$, Bonferroni corrected; Fig. 6c). Estimates of the free parameters in all models except for entropy reduction were significantly higher than zero ($p < .001$, bootstrap, Fig. 6d). Even though those estimates quantitatively differed from the fMRI experiment (Fig. 6b), their valence was consistent. These results reconfirm the earlier observation that the information purchase behavior can be explained better by combining the standard instrumental VoI and non-standard factors.

Unlike the fMRI experiment dataset, however, the behavioral experiment dataset strongly preferred the constant bonus model to the economic models. The constant bonus model fitted the behavior significantly better than probability weighting models ($p < .05$, Bonferroni corrected; Fig. 4d, Table 2) and indistinguishable from recursive utility models ($p > .05$, Bonferroni corrected). This indicates that behavioral subjects showed overvaluation irrespective of outcomes and information diagnosticity. The difference between the two datasets might be attributable to the elicitation methods; sensitivity of BDM bidding procedure to underlying subjective values might be imperfect, possibly yielding overbidding when the actual subjective value is low (Medlin, 1979; Wendt, 1969). Alternatively, the task design in the fMRI experiment might not be statistically powerful enough to detect the contribution of the constant bonus.

Even though the constant bonus model was well supported by behavioral experiment data, it is still possible that its model fit could be improved even further by incorporating the uninformativeness penalty. While variance explained by the constant bonus may be larger than variance explained by recursive utility or probability weighting, the latter may still be innegligible. To test this, the free parameters from the economics models (C_{KP} or C_{w}) were added to the modified model that already included the constant bonus (C_{bonus}), resulting in models with two free parameters. If the constant bonus is enough to explain willingness-to-pay, goodness-of-fit of these two-free-parameter models would not be better than constant-bonus-alone model.

The results of model fitting revealed that state-recursive-utility indeed significantly improved the goodness-of-fit of the constant-bonus-only model ($p < .05$, Bonferroni corrected; Fig. 6e). Outcome-recursive-utility also yielded mild improvement ($p < .10$, Bonferroni corrected). In both models, parameter estimates were higher than zero ($p < .05$, bootstrap; Fig. 6f), confirming the contribution of uninformativeness penalty above and beyond the constant bonus. Therefore, to account for the observed overvaluation as much as we could in the current scheme, we need a model that unifies three factors; instrumental VoI, constant bonus, and recursive utility.

On the other hand, probability weighting did not significantly improve goodness-of-fit of the constant-bonus-only model ($p > .05$, uncorrected; Fig. 6e). This result is noteworthy given that predictions of the recursive utility models and the probability weighting models are quantitatively similar (Fig. 5). This highlights the importance of specifying the economics models and underlying utility or weighting functions more accurately and rigorously based on the empirical data.

Together, the formal model evaluation and comparison have revealed that 1) the non-standard portion of subjective VoI exists on top of instrumental VoI, and 2) it is sensitive to the outcomes at stake, which can be explained by penalty for uninformed choice due to concavity of recursive utility.

Discussion

Value-based decision-making is primarily defined by choices and their outcomes. Most studies in the field have focused on choices which outcomes are immediate delivery of reward or punishment. However, the modern world is full of choices about information that is not rewarding or punishing by itself. The biggest challenge in studying information acquisition lies in quantification of utility or value of information (VoI); it is much more difficult to measure or manipulate than value of reward or punishment. Even though the classic economics theory provides the normative predictions of VoI, it is fairly complex, makes a number of assumptions, and is applicable only to the settings where

decision-related variables are clearly well defined. Because of these limitations, the normative model has not been influential, and instead, various conceptual accounts have been proposed in respective fields. However, there has been little attempt to unify or relate these accounts with each other or with the normative model, either theoretically or empirically, hindering establishment of a general and falsifiable modeling scheme.

The current study aimed to propose descriptive, yet quantitative, modeling of information acquisition. The normative model of instrumental VoI was used as a standard model, to which the alternative accounts were implemented as modifications. This approach allowed us to provide concrete, parametric, and falsifiable predictions on how information acquisition depends on decision-related variables. To test them empirically, the current study adopted a new task paradigm, which was able to capture the behavioral effect of both instrumental VoI and preference for non-instrumental information simultaneously. It was richly parameterized so that we could statistically compare descriptive validity of the standard and modified models.

The normative theory indeed explained subjects' behavior to a certain extent, which showed that it was an acceptable standard or baseline model. At the same time, however, subjects' behavior systematically deviated from its prediction; subjects exhibited overvaluation of information, both instrumental and non-instrumental, which indicated the contribution of non-standard factors. Critically, although preference for non-instrumental information has been long documented in various literature, it has not yet been established to what extent such preference contributes to behavior above and beyond instrumental VoI. Our results show that such preference indeed exists and affects behavior even under the presence of instrumental information, illustrating the limitation of the normative theory.

Observed overvaluation had two prominent features; first, it was sensitive to outcomes at stake (higher when EU of uninformed choice was positive), and second, it was an inverted-U function of information's diagnosticity (p). Sensitivity to outcomes was consistent with existing accounts proposed in economics, which penalizes choices or states in which agents could have obtained, but did not obtain, an additional piece of information. Such penalty for uninformativeness was theoretically introduced by recursive utility or probability weighting, rather than uniform discounting, which explained the sensitivity to outcomes. These accounts have been used almost exclusively by "temporal lotteries" studies, which have largely focused on situations where timing of uncertainty resolution is non-instrumental; subjects typically cannot choose whether to accept or reject a given lottery, but just whether to see uncertainty resolved earlier or later. Such a task design may be well-suited for detecting (possibly subtle) preference for earlier uncertainty resolution, but it has left the possibility that such preference would not be generalizable to situations where information may be instrumental (Beierholm & Dayan, 2010). Our results reject this possibility and show that these accounts are applicable to a wider range of information acquisition problems.

On the other hand, overvaluation's non-monotonic sensitivity to p cannot be explained by any of the examined accounts. Although it might seem superficially similar to a classic notion that curiosity is an inverted-U function of information gap (Berlyne, 1966), a careful examination reveals a subtle but important conceptual distinction. In the information gap theory, the inverted-U function is usually interpreted as the lack of curiosity when agents feel that they currently possess too little, or too much, knowledge. This idea is not directly applicable to the current paradigm where subjects

had the same amount of prior knowledge (two equiprobable outcomes), irrespective of p. More importantly, even if information gap theory could be modified to explain the inverted-U function, it is unclear how it could also explain sensitivity to outcomes, unless it takes the principle of EU maximization into account seriously.

While the reason behind the non-monotonic sensitivity to p is unclear, one straightforward interpretation is that subjects' representation of p was biased; while they were accurate in estimating diagnosticity when it is the highest ($p = 0.5$) or the lowest ($p = 0$), they overestimated the intermediate levels of diagnosticity. This could be explained by probability weighting function as presumed in prospect theory (Kahneman & Tversky, 1979). (It is not to be confused with probability weighting in Wu (1999); the inverted-U function could be ascribed to weighting within $EU_{\text{info}}(s_1)$ and $EU_{\text{info}}(s_2)$, while Wu introduced weighting to averaging over them.) Another possibility is that subjects' perception of the partition line's angle itself was biased, known as oblique effect (Appelle, 1972). These hypotheses could be tested by future studies that manipulate diagnosticity more finely and present it to subjects differently. If neither of them is supported empirically, we need to consider how existing accounts, particularly recursive utility or probability weighting, should be modified.

Importantly, the sensitivity of overvaluation to outcomes could not be explained by the non-economics accounts. Formal modeling analysis confirmed this, preferring the economic models to non-economic ones. This finding particularly poses a challenge to curiosity studies, most of which examine situations where decision-related variables (especially outcomes at stake) and agents' utility are not well defined, and thus the economic models are not directly applicable. Although adoption of simplistic accounts such as entropy reduction may be justifiable in such settings, it might be more promising to make reasonable assumptions about underlying utilities or to infer them from external evidence (e.g., from subjects' own choices, from independent subjects' choices, or from some other measures such as decoding of eye movement or neural activity). Such approaches are important in two ways; first, it would allow us to explore how to approximate complicated situations in a tractable way, and second, it would shed light on what kind of strategy or heuristics human brain might use when it is too demanding to compute EU under different actions and information states.

The current study analyzed two independent datasets, one fMRI and one behavioral. Although they painted largely similar pictures, formal modeling analysis revealed one difference; the behavioral study dataset, but not fMRI, preferred the constant bonus model to the alternative accounts. This difference may be attributable to empirical paradigms. For instance, fMRI and behavioral experiments used different sets of decision-related variables, and fMRI experiment design might lack statistical power to detect the contribution of the constant bonus. Another, perhaps more important, difference lies in the elicitation methods; the behavioral experiment adopted BDM bidding, and while it is a de facto standard in decision-making literature, it might have imperfect sensitivity to variation of underlying subjective values (Medlin, 1979; Wendt, 1969). In particular, it may be susceptible to overbidding with low subjective values, which could explain a positive constant bonus. This issue could be clarified by future studies that use alternative elicitation methods, including choices between multiple pieces of information (rather than between information and monetary costs). Nonetheless, the economic notion of uninformativeness penalty, particularly in the form of recursive utility, improved the model fit above and beyond the constant bonus. The support for the economics models is thus by and large consistent between the two independent datasets.

This study examined two non-standard economic accounts, recursive utility (Ahlbrecht & Weber, 1996; Kreps & Porteus, 1978) and probability weighting (Wu, 1999). Although a small difference was found between their goodness-of-fit, it is worth emphasizing that their performance may critically depend on underlying recursive utility or probability weighting function's form. Little to no attempt has been made to specify them based on empirical data, and our dataset may not be large and thorough enough to do so either. Empirical studies with richer parametric designs are definitely needed to characterize the underlying function forms rigorously. That said, the fact that our results supported these economic models, even though they are only prematurely constructed, would suggest that these approaches are promising.

There are several ways to position these economic models within larger conceptual schemes. First, attempts to explain non-canonical preference using curvature of utility function or non-linear probability weighting function are akin to traditional approaches to risk preference; non-neutral risk attitude has been explained by either concavity of utility function (Pratt, 1964) or probability weighting (Kahneman & Tversky, 1979). Second, the idea of recursive utility has also been adopted by studies on decision-making under ambiguity, raising a possibility that ambiguity attitudes and information acquisition behavior might be linked (Berg & Eisenberger, 1996; Camerer & Weber, 1992; Snow, 2010). Third, as a possible psychological basis of non-linear probability weighting, Wu proposed that people tend to pay more attention to the worst outcomes or states than others (Wu, 1999; Yechiam & Hochman, 2013). Relationships among these concepts could be empirically explored by characterizing individual differences using the current task and comparing them against other measures.

The task paradigm in this study was designed such that decision-related variables, particularly prior and posterior outcome probability distributions, were graphically and intuitively presented. Although this was necessary to reliably elicit subjective VoI, the task design had a few limitations because of such consideration, and future studies may need to address them. First, the outcomes were always equally likely in uninformed choices. Manipulation of prior probability would be critical in order to test not only instrumental VoI theory but also the modified models with recursive utility or probability weighting (Wendt, 1969). Second, each trial contained only two outcomes, two information states, and two possible choices. Computation of instrumental VoI becomes complex quite easily once any of these constraints is relaxed, and it is not known how subjective VoI keeps up with such complexity. Third, subjects were endowed with the ability to choose whether to accept or reject gambles in all trials. In the real world, however, we are often faced with a mixture of choices to be made and choices that are already made, and it has not been well studied how people prioritize information about the former (could be instrumental) to the latter (never instrumental).

In conclusion, we presented a new way to conceptualize multiple factors behind information acquisition behavior in a unified, quantitative, and falsifiable, modeling scheme. Empirically we found that, although the standard model of instrumental VoI could serve as a good baseline model, subjects exhibited overvaluation that was systematically sensitive to outcomes at stake. We further found that such overvaluation could be explained by economic accounts on penalty for uninformative choices. This is the first evidence that such penalty affects behavior in tandem with instrumental VoI. More generally, our findings would have implications on how to acquire and

distribute information adaptively in the digital age, and particularly, how to help people focus on important information while ignoring unimportant, but psychologically attractive, information.

Methods

fMRI experiment

Subjects. 47 healthy, naïve subjects participated in the experiment. 10 were removed from the analysis due to excessive motion during scanning, resulting in 37 subjects used in the data analysis.

Procedure. The choice phase of the task (Fig. 2a) was conducted during the scanning, followed by the outcome phase outside the scanner. Subjects received instructions and practiced the task prior to the scanning. Subjects interacted with the experiment program via a MR-compatible button box in the scanner and a keyboard outside. The program was run by Matlab and Psychtoolbox (Brainard, 1997; Pelli, 1997) on a Windows PC. See Chapter 3 for acquisition and analysis of MR signal.

Task. Each subject's scanning consisted of five sessions, each of which included 30 trials in a randomized order (150 trials in total). In each trial, one gamble with two outcomes was visually presented as a roulette wheel partitioned by a vertical while line at the middle. Ten combination of outcomes were used: (+$12, −$9), (+$9, −$12), (+$9, −$9), (+$12, −$6), (+$6, −$12), (+$9, −$6), (+$6, −$9), (+$6, −$6), (+$12, −$3), and (+$3, −$12). Subjects chose whether to play the gamble or not (two-alternative-forced choice; "uninformed choice"). In the outcome phase after the scanning, five trials were randomly selected by the experiment program and implemented into the actual monetary payment. If they had chosen to play a selected gamble, a green dot appeared on the perimeter of the roulette and its side on the wheel determined the outcome. Without the information delivery, the green dot's location followed uniform distribution over the perimeter, making the two possible outcomes equally likely.

The choice on gamble during scanning was followed by information presentation. Information was visually represented as a magenta partition line on the roulette wheel running through its center. The information display was followed by presentation of monetary cost, and subjects chose whether to purchase the information or not (two-alternative-forced choice). The information line was either vertical, slanted by 30°, or slanted by 60°. Each angle was presented once per each pair of outcomes in each session (3 angle × 10 outcomes = 30 trials). The monetary cost was variable across trials and determined independently from the outcomes and partition angle. Five costs were used: ¢5 (3 trials per run), $1 (8 trials), $2 (8 trials), $3 (8 trials), and $9 (3 trials). Subjects were instructed to make uninformed choice on gamble and information purchase choice within 2 seconds. The trials in which they failed to respond were discarded from the analysis.

In the outcome phase, subjects received the actual content of information in the trials selected for payment, but only if subjects had chosen to purchase it. The delivered information revealed the side of the magenta partition the green dot would appear by brightening it. This changed the outcome probability distribution from (0.5, 0.5) to either (1, 0) or (0, 1) (vertical line), (5/6, 1/6) or (1/6, 5/6) (30°-slanted line), or (2/3, 1/3) or (1/3,2/3) (60°-slanted line). The brighter side was chosen randomly. Subjects made the decision on whether to play the gamble or not based on the delivered information ("informed choice"), disregarding their own uninformed choice during the scanning. If subjects had chosen not to purchase the information, the brighter side was not presented, and their uninformed choice was implemented.

Analysis. Group-level utility function was estimated based on subjects' uninformed choices on gambles (Supplementary Fig. 1). Utility function estimation had four free parameters: the powers on

the positive and negative domains, a multiplicative term on the negative domain (capturing loss aversion), and the temperature parameter in the logit (softmax) function that maps expected utility onto binary choices. Parameter estimates were obtained by maximum likelihood estimation procedure (MLE) on Matlab.

To analyze information purchase behavior under the standard and modified models, subjects' binary choices were fitted to the logit function of difference between informed choice's EU and uninformed choice's EU using MLE (Eqs. 1 and 2, and their variants). Analysis under the standard model (Fig. 3) contained only one free parameter (logit temperature). Goodness-of-fit of the standard model, measured by negative log-likelihood, was statistically evaluated by permutation, shuffling the outcome labels within each subject dataset (1999 iterations). Fitting of the modified models (Fig. 6) involved two free parameters, the logit temperature and a parameter calibrating non-instrumental VoI (C_{bonus}, C_{ent}, C_{KP}, or C_{w}). 95% CI of parameter estimates was obtained by bootstrap, sampling subjects with replacement (1999 iterations).

To statistically test the difference in goodness-of-fit while addressing the issue of overfitting (Fig. 6, Table 1), 10-fold cross-validation was conducted. First, choices from each subject were split into ten datasets, and nine of them was collated over subjects and used as the training dataset to estimate the parameter (C_{bonus}, C_{ent}, C_{KP}, or C_{w}). Second, goodness-of-fit (negative log-likelihood) in the hold-out dataset was obtained by logit function with the non-instrumental VoI parameter fixed at the training estimate (one free parameter remaining: logit temperature), as well as with the instrumental-VoI-only standard model. This was repeated for every hold-out dataset (10 times). Third, the whole procedure was repeated 200 times, resulting in 2000 goodness-of-fit measurements for each model. Fourth, those measures were compared between models in a pairwise manner to compute statistical significance.

The behavior was compared with the standard model's predictions in a model-free manner as well (Fig. 3c, Fig. 4e, f). The rational EU-maximizing choices on information purchase were derived on a trial-by-trial basis as a function of monetary costs, and then aggregated over trials to obtain the total amount of money each subject should have agreed to pay throughout the entire sessions (Fig. 3c). The degree to which subjects "overpurchased" information was measured as the difference between these predictions and the total amount each subject actually agreed to pay (Fig. 4e, f). Note that, since only five gambles were selected for actual monetary payment, subjects did not actually pay the total costs of all information they agreed to purchase.

Behavioral experiment

Subjects. 119 healthy, naïve subjects participated in the experiment. 23 were removed from the analysis because they showed misunderstanding or were unable to complete the task within allotted time, resulting in 96 subjects in the data analysis.

Procedure. A group of subjects (five to eighteen) participated simultaneously. Subjects received instructions and practiced the task prior to the experiment. Subjects interacted with the experiment program running on Matlab and Psychtoolbox via keyboards of Windows laptops on partitioned desks. An online survey of personality measures was conducted prior to the day of the experiment (data not analyzed in this study).

Task. The task was almost identical to the fMRI experiment. The choice phase consisted of 100 trials in a randomized order. 20 outcomes were used, 16 of which were combinations of one gain (+$12, +$10, +$8, or +$6) and one loss (−$12, −$10, −$8, or −$6), and the remaining four were (+$10, 0), (+$6, 0), (−$10, 0), and (−$6, 0). The information line was either slanted by 0° (vertical), 22.5°, 45°, 67.5°, or 90° (horizontal). Each angle was presented once per each pair of outcomes (5 angle × 20 outcomes = 100 trials). In each trial, after the information presentation, subjects indicated their willingness-to-pay for the information in BDM bidding procedure. Whether they purchased the information or not in the trials selected for payment (i.e., the outcome of bidding) was determined in the outcome phase.

Analysis. Group-level utility function was estimated in the same way as fMRI experiment (Supplementary Fig. 1). To analyze subjects' information purchase behavior under the standard and modified models, VoI predictions were obtained (Eq. 5 and their variants) and compared against willingness-to-pay. Since Eq. 5 does not have an analytical solution due to maximum operations and non-linear utility function, VoI predictions were numerically obtained by minimizing the sum of squared difference between RHS and LHS on Matlab (fminsearch).

Analysis under the standard model (Figs. 3, 4) contained no free parameter. Goodness-of-fit of the standard model, measured by mean squared error (MSE), was statistically evaluated by permutation (1999 iterations). Overvaluation (Fig. 4a-d) was obtained as the difference between willingness-to-pay and the normative VoI predictions, averaged over subjects. Free parameters in the modified models (C_{bonus}, C_{ent}, C_{KP}, or C_{w}; Fig. 6) were estimated using grid search; at each level of the free parameters, VoI predictions were obtained by fminsearch and its MSE against willingness-to-pay were calculated. The parameter value that minimized MSE was then obtained. 95% CI of parameter estimates was similarly obtained by bootstrap, sampling subjects with replacement (1999 iterations). MSEs were compared between models by cross-validation (2000 iterations) as described in the fMRI experiment's analysis section above.

Appendix

Derivation of the standard model's prediction that non-instrumental information is valueless.
The standard, normative theory prescribes that value of non-instrumental information should be zero. This can be easily derived from the Eqs. 1-5. Consider a gamble which EU is positive irrespective of the information states:

$$EU_{\text{uninfo}}(x_1, x_2) = 0.5 \cdot u(x_1) + 0.5 \cdot u(x_2)$$
$$EU_{\text{info}}(x_1, x_2, p, c|s_1) = p \cdot u(x_1 - c) + (1-p) \cdot u(x_2 - c)$$
$$EU_{\text{info}}(x_1, x_2, p, c|s_2) = (1-p) \cdot u(x_1 - c) + p \cdot u(x_2 - c)$$

According to Eq. 5, w in this gamble satisfies:

$$0.5 \cdot u(x_1) + 0.5 \cdot u(x_2) = 0.5 \cdot u(x_1 - w) + 0.5 \cdot u(x_2 - w)$$

Since u is a strictly increasing function, the only solution to this equation is $w = 0$.
Similarly, in a gamble which EU is negative irrespective of the information states,

$$EU_{uninfo}(x_1, x_2) = 0$$
$$EU_{\text{info}}(x_1, x_2, p, c \mid s_1) = EU_{\text{info}}(x_1, x_2, p, c \mid s_2) = u(-c)$$

Again, using Eq. 5 and the fact that u is a strictly increasing function, the only solution in this case is $w = 0$.

Figures

Figure 1. Theoretical accounts on information acquisition. (*Left*) The normative theory of value of information (VoI) prescribes how rational expected utility (EU) maximizer acquires information. In this theory, information should be regarded as valuable only to the extent to which it is instrumental, i.e., affects EU-maximizing choices. The theory incorporates the influence of risk attitude on information acquisition through utility function. (*Right*) Alternative accounts from AI, psychology, and economics allow that agents acquire information even when if non-instrumental. They make different predictions about dependence on decision-related variables such as outcomes at stake or diagnosticity of information (see also Fig. 5).

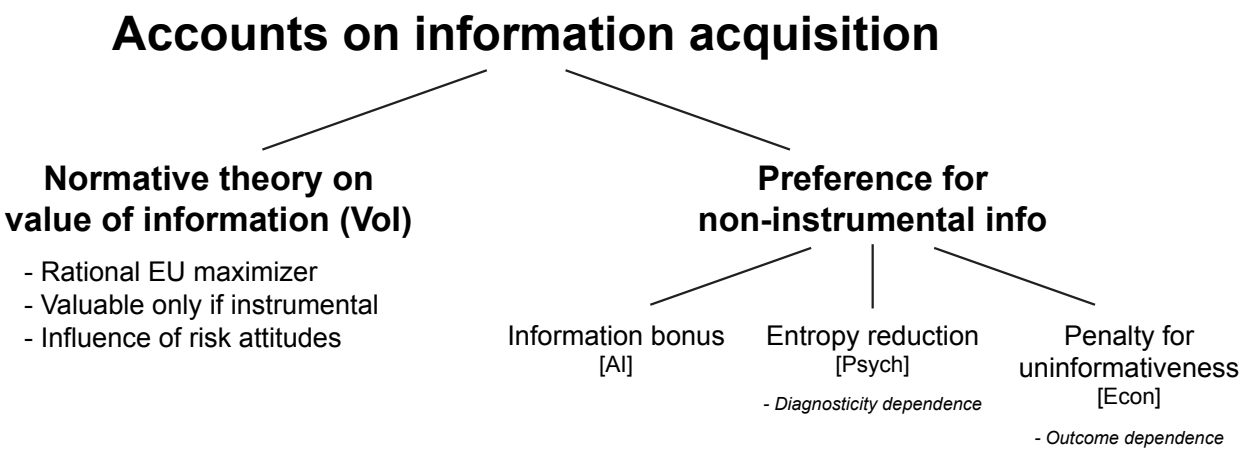
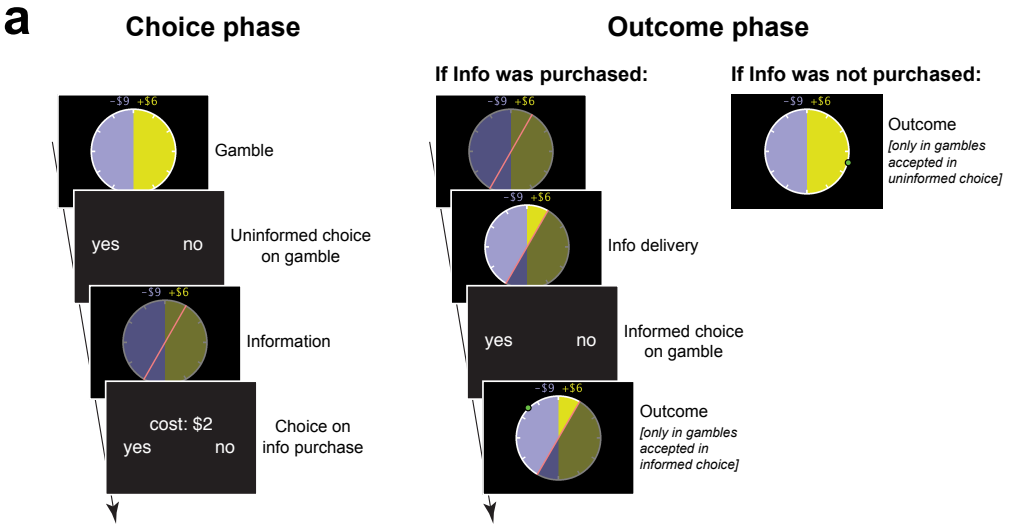


Figure 2. Experimental paradigm. **(a)** *Left*: During the choice phase, gambles with two monetary outcomes (x_1 and x_2) were presented in the form of a roulette wheel. Subjects' primarily decision was to choose whether they would accept or reject each gamble. They first made the decision without an additional information (uninformed choice). Information was then presented as a magenta partition line on the wheel, followed by information purchase choice. In fMRI experiment, the information cost was presented, and subjects chose whether they would purchase the information or forgo it. In behavioral experiment, they indicated willingness-to-pay for the information (not shown). *Right*: In the outcome phase, the content of the information was revealed only if it had been purchased during the choice phase. It brightened one side of the partition, revealing the outcome probability distribution. Subjects were allowed to change their choice on the gamble based on the information (informed choice). If the gamble had been accepted, a green dot appeared at a random location on the perimeter of the wheel (or within in brighter side after information delivery), and its side determined the monetary outcome. **(b)** Decision tree. EU-maximizing agents purchase information if they prefer informed choice to uninformed choice ($EU_{\text{info}} > EU_{\text{uninfo}}$). EU_{info} is the average over the two information states, s_1 and s_2.



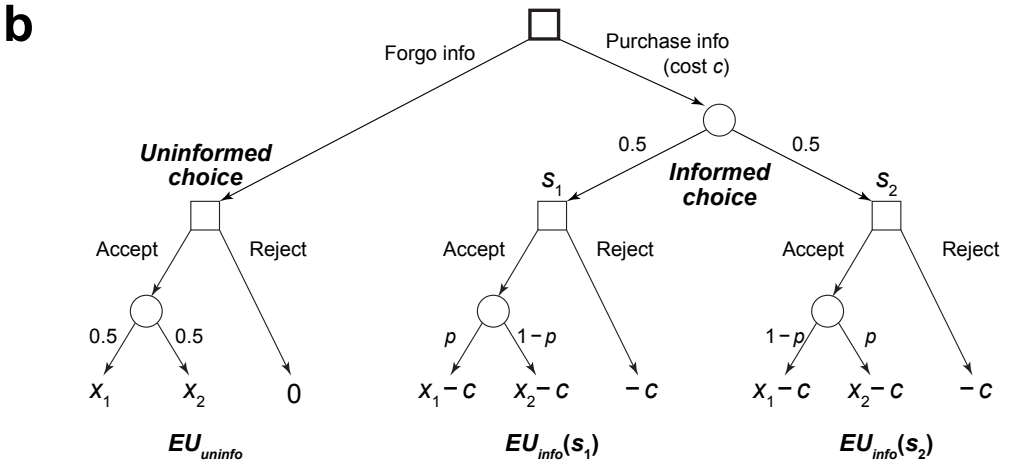


Figure 3. Consistency between behavior and the normative model prediction. **(a)** Model prediction was fitted to fMRI experiment dataset. As predicted, subjects purchased the information more frequently with the larger difference between EU of informed and uninformed choices ($p < .0005$). Each dot represents a unique combination of x_1, x_2, p, and c. **(b)** Model prediction was compared against behavioral experiment dataset. Willingness-to-pay was well predicted by the model ($p < .0005$). Each dot represents a unique combination of x_1, x_2, and p. **(c, d)** The observed and predicted amounts of information purchase in fMRI (c) and behavioral (d) dataset are compared. Rows are ordered by EU (high to low). Subjects were sensitive to outcomes at stake and information diagnosticity, overall as predicted. **(e, f)** Goodness-of-fit was better under the assumption of EU maximization than EV in fMRI (e) and behavioral (f) dataset.

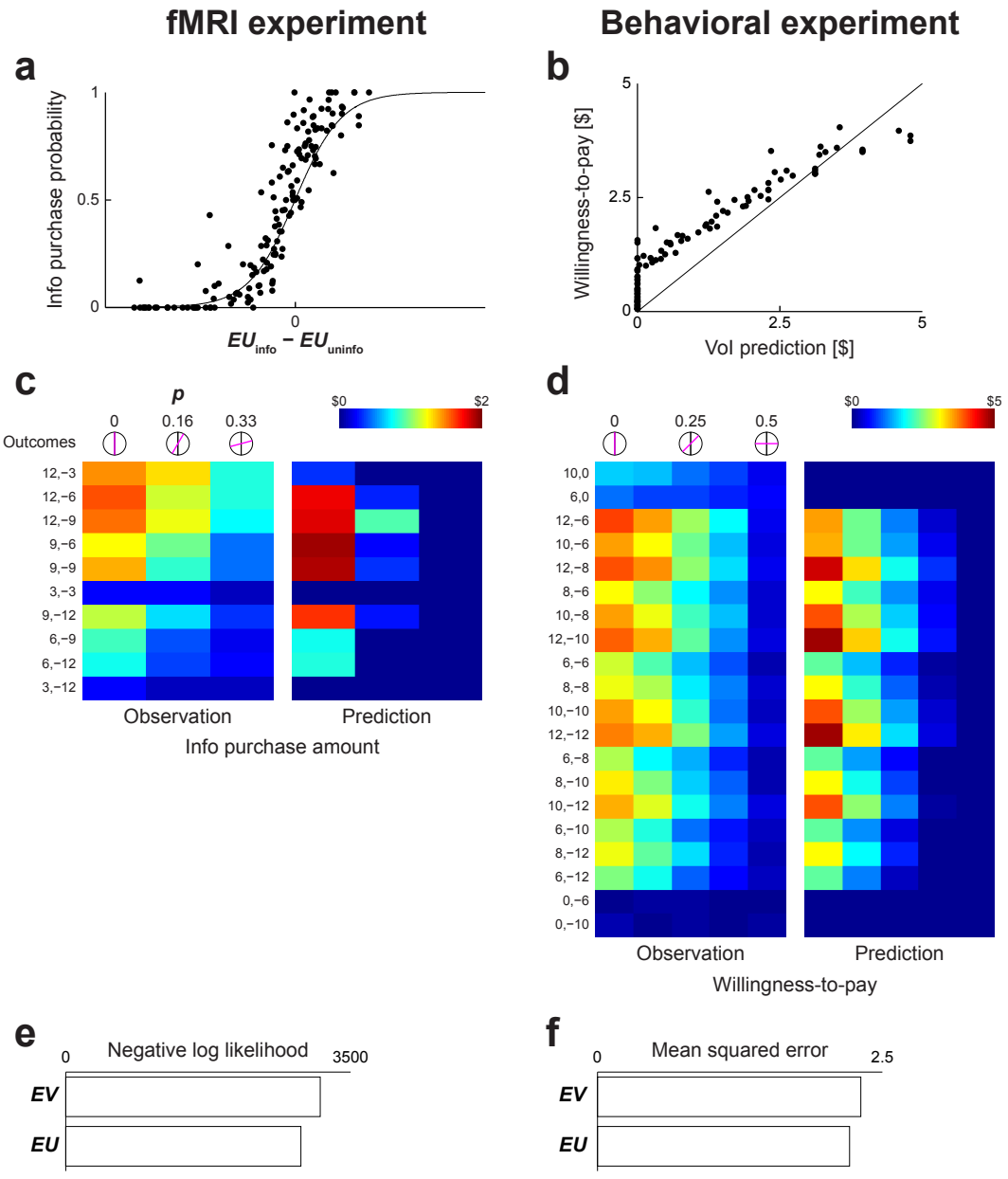


Figure 4. Violation of the normative prediction. **(a)** Willingness-to-pay was higher than zero in the gambles (+$10, $0) (*blue*) and ($0, −$10) (*red*). This violates the normative prediction that VoI is zero in these gambles. **(b)** Histogram of willingness-to-pay (*top*) and uninformed choice (*bottom*), collapsed across p. **(c)** Overvaluation, or the difference between willingness-to-pay and the normative prediction, was higher than zero. It was higher in the gambles with positive EU (*blue*) than negative EU (*red*). Overvaluation also showed non-monotonic dependence on information diagnosticity. **(d)** Overvaluation across all trials are shown. This corresponds to the difference between observation and prediction in Fig. 3d. **(e)** The amount of overpurchase, or the difference between the total amount subjects chose to pay and the normative prediction, was higher than zero. It was higher in the gambles with positive EU (*blue*) than negative EU (*red*). Non-monotonic dependence on information diagnosticity was also observed. **(f)** Overpurchase across all trials are shown. This corresponds to the difference between observation and prediction in Fig. 3c. Error bars: SEM.

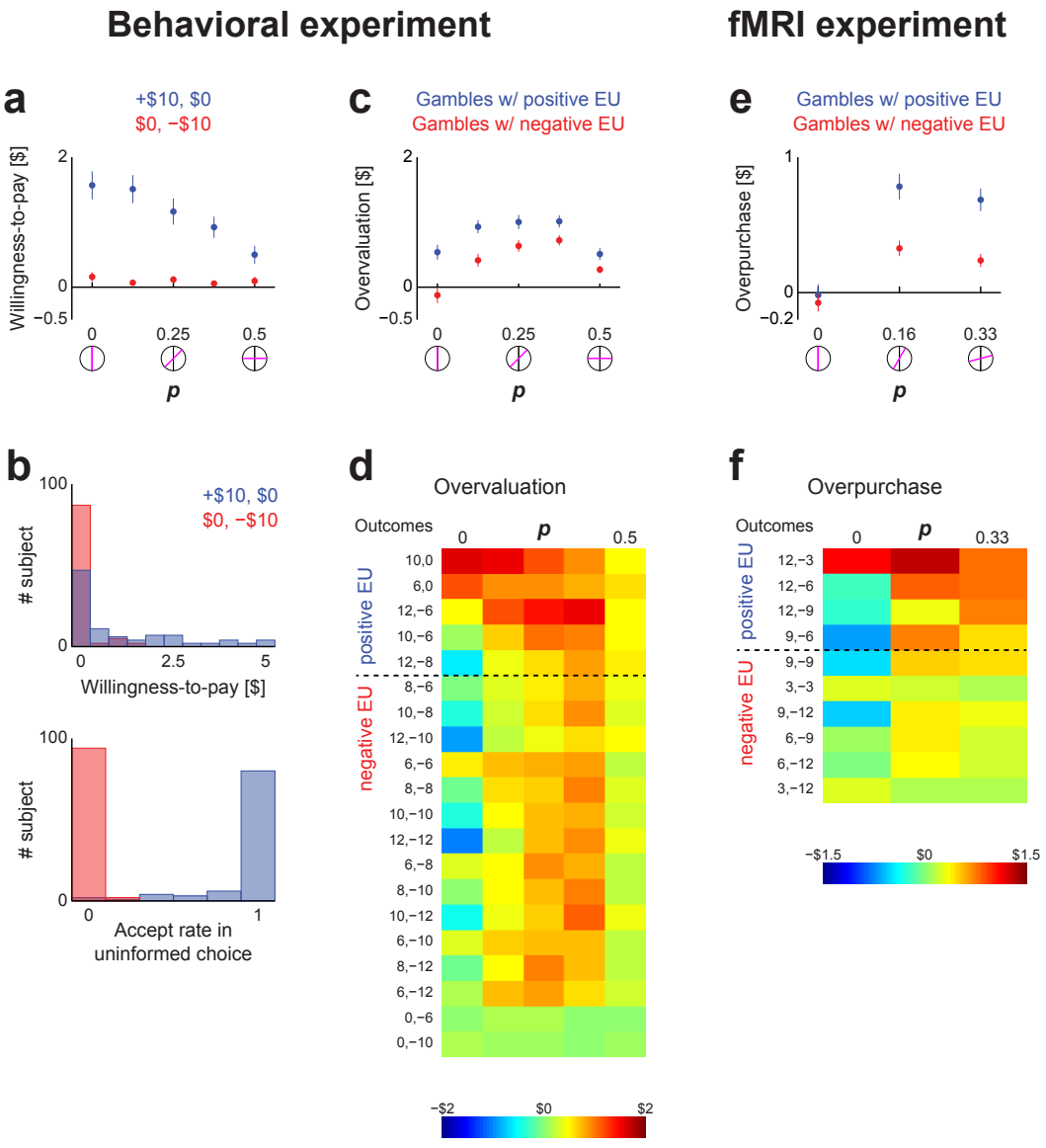


Figure 5. Schematic illustration of predicted overvaluation, i.e., the modified models' prediction w^* minus the standard prediction w.

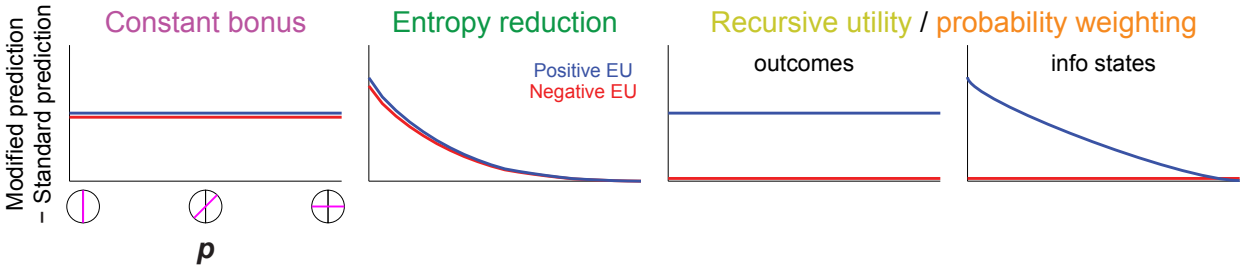
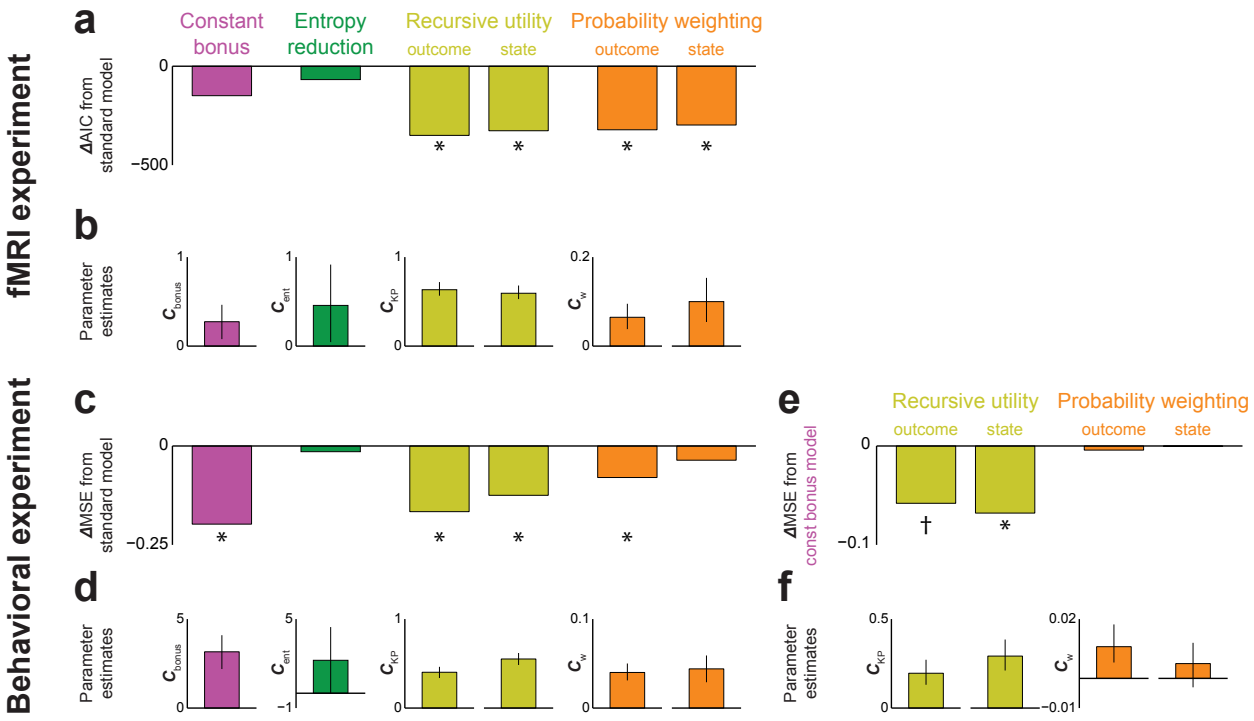


Figure 6. Formal modeling analysis. **(a)** The modified models' AICs were lower than the standard model. The difference was significant with economic models (recursive utility and probability weighting). See Table 1 for comparison among the modified models. **(b)** Parameter estimates of the models in (a) were significantly different from zero. **(c)** The modified models' mean squared errors (MSE) were lower than the standard model. Constant bonus model exhibited the largest improvement, followed by economic models. See Table 2 for comparison among the modified models. **(d)** Parameter estimates of the models in (c) were significantly different from zero except for entropy reduction. **(e)** When recursive utility was added to the constant-bonus-only model (*leftmost* in c), MSE was reduced even further. **(f)** Estimated values of recursive utility models in (e) were larger than zero. *: $p < .05$, †: $p < .10$, Bonferroni corrected. Error bars: 95% CI.



Tables

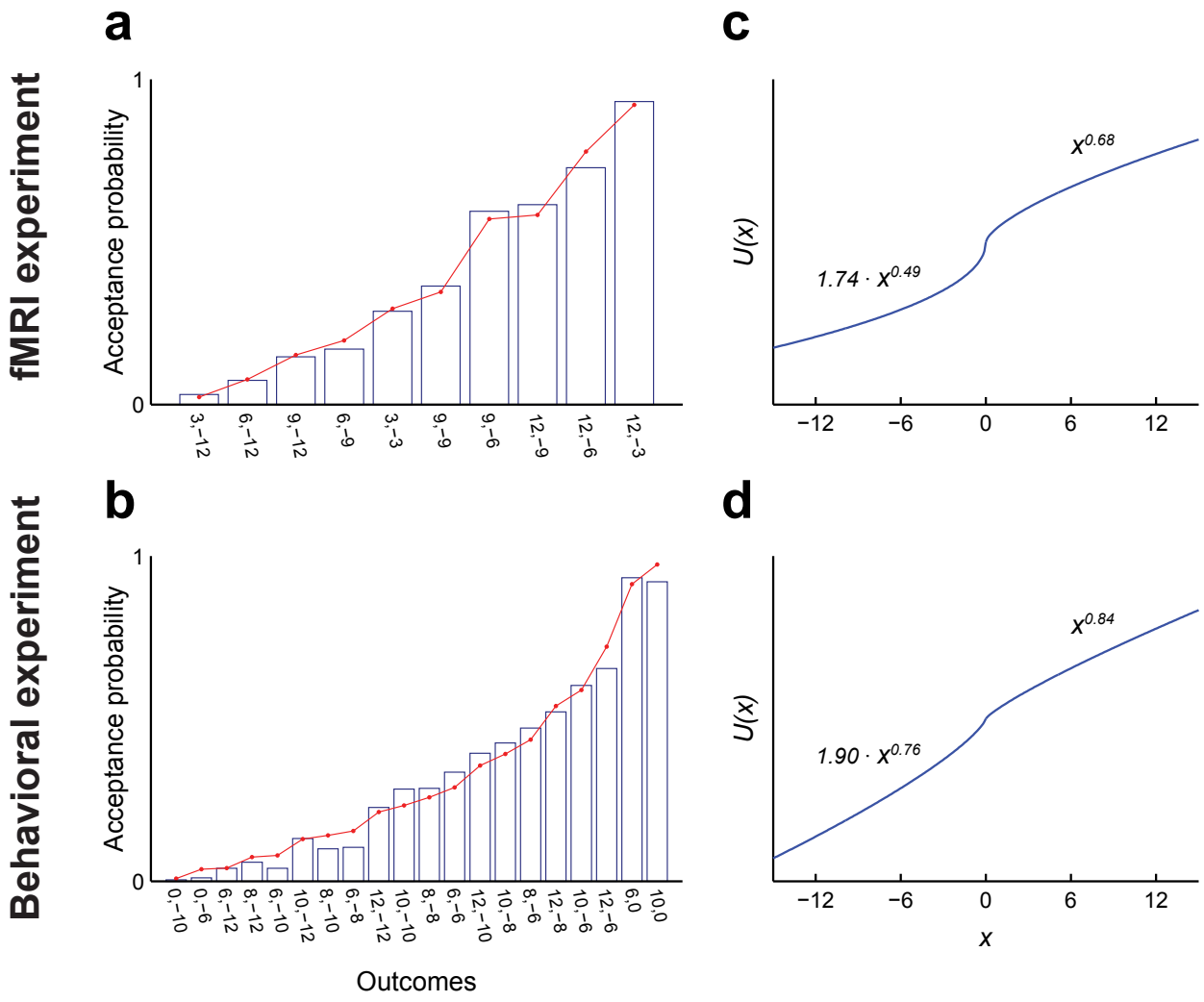
Table 1. Comparison between economic and non-economic models in fMRI experiment. Show are uncorrected *p*-values evaluated by 10-fold cross-validation (2000 iterations total). See Fig. 6a for effect sizes. Bold: $p < .05$ after Bonferroni correction.

	Recursive utility		Probability weighting	
	Outcome	State	Outcome	State
Information bonus	.0245	.0340	.0160	.0430
Entropy reduction	**.0015**	**.0025**	**.0030**	.0095

Table 2. Comparison between economic and non-economic models in behavioral experiment. Show are uncorrected *p*-values evaluated by 10-fold cross-validation (2000 iterations total). See Fig. 6c for effect sizes. Bold: $p < .05$ after Bonferroni correction.

	Recursive utility		Probability weighting	
	Outcome	State	Outcome	State
Information bonus	.2305	.0510	**<.0005**	**<.0005**
Entropy reduction	**.0030**	**<.0005**	**.0050**	.1210

Supplementary Figure 1. Estimation of utility function. **(a, b)** Probability of uninformed gamble acceptance in fMRI (a) and behavioral (b) datasets. *Blue bars*: observed, *red lines*: logit model fit. **(c, d)** Estimated utility function in fMRI (c) and behavioral (d) datasets.



Chapter 3
Neural representation of value of economic information

Introduction

To make adaptive decisions under uncertainty, we often seek for information about the state of the world. Because it is rare that all information relevant to a decision is available at once, acquiring information from the environment over time is a critical part of decision-making. For instance, scholars would read a scientific paper's abstract and a couple of first paragraphs to evaluate whether it is relevant enough for themselves to keep reading. In the modern world, where a tremendous amount of information is available at our fingertips, adaptive evaluation and acquisition of information is more ubiquitous and important than ever.

Adaptive acquisition of information is, at the same time, a challenging decision-making problem by itself. Information is usually associated with costs, such as money, time, or effort, and agents need to judge if information's utility is worth its cost. For instance, people with health problems would consider consulting with doctors based not only on how serious their issues are but also on how much they need to pay. Furthermore, agents often need to choose an information source and abandon the other sources; people usually consult with only one or two doctors, even if other doctors might be more trustworthy. How do people solve these problems and acquire information adaptively?

One well-established normative solution from economics is value of information (VoI) (Edwards, 1965; Howard, 1966; Marschak, 1971; Wendt, 1969). The normative theory of VoI sees information acquisition as a form of expected-utility-maximizing behavior. Agents first compute value of a piece of information, and acquire it only if its value exceeds the cost. Similarly, to select an information source, agents compute VoIs of all sources available and pick the best one.

An important assumption in this theory is that information is not intrinsically valuable. Instead, information is valuable only to the extent to which it helps agents obtain future rewards and avoid punishments, or in other words, the extent to which it is instrumental. For instance, the value of a weather forecast could be measured by how much it helps the agent make a plan for a weekend; the agent can go hiking when it is expected to be sunny, and go to a theater when a rainfall is predicted. According to this theory, VoI is non-zero only if information is instrumental, and non-instrumental information should not be acquired when associated with costs.

This normative theory not only provides specific predictions of information acquisition behavior but also proposes underlying neurocognitive processes. Most notably, it raises the possibility that VoI is represented in human brain and guides behavior, which is a concrete and falsifiable hypothesis on neural basis of information acquisition. However, the normative theory has rarely been used in neuroscience literature; instead, most neuroscience studies, particularly human fMRI, have focused on specific forms of information acquisition such as exploration-exploitation dilemma (Badre, Doll, Long, & Frank, 2012; Boorman, Behrens, Woolrich, & Rushworth, 2009; Daw et al., 2006; Kolling, Behrens, Mars, & Rushworth, 2012; Quilodran, Rothé, & Procyk, 2008; Shenhav, Straccia, Cohen, & Botvinick, 2014). While these studies have

provided important insights on neural mechanisms behind specific actions, it is not yet clear how their findings are generalizable to other information acquisition problems.

This chapter reports a new fMRI study that takes an alternative approach; it aims to find neural representation of VoI. Since VoI itself is a versatile concept and applicable to various situations, finding VoI representation would form a basis for a more general neuroscientific theory of information acquisition. Moreover, it would be directly relatable to other types of value-based decision-making. In particular, although there are abundant studies on brain regions involved in valuation of objects or services, such as an apple or a dinner at a restaurant (Gross et al., 2014; Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Lau & Glimcher, 2008; McNamee, Rangel, & O'Doherty, 2013; Padoa-Schioppa & Assad, 2006; 2007; Plassmann, O'Doherty, & Rangel, 2007; Rangel, Camerer, & Montague, 2008; Rudebeck et al., 2008; M. F. S. Rushworth & Behrens, 2008; M. Rushworth, Noonan, & Boorman, 2011; Tremblay & Schultz, 1999), it is not clear if the same regions are involved in VoI. It is possible that VoI recruits totally different brain regions because of the unique underlying processes; the normative theory predicts that, upon calculating VoI, agents need to simulate their own choices with and without the information and compare their expected utilities. Furthermore, since the content of information is not perfectly predictable, agents need to simulate their choices under all possible information contents (e.g., what to do when a sunny, windy, rainy, or snowy weekend is expected). Due to such computational processes, valuation of information may recruit different brain regions from other types of valuation.

One important reason why the normative theory has not gained popularity in neuroscience is its limited descriptive validity. It is widely accepted that human ubiquitously acquires non-instrumental information, even if it is costly, violating the most important tenet of the normative theory. This observation has led to a common conception that information is actually intrinsically valuable to human; human may have non-economic, psychological internal drive to acquire information for its own sake (Berlyne, 1966; Loewenstein, 1994). However, this does not necessarily mean that the normative theory needs to be entirely abandoned. Empirically, it has not been established how important the preference for non-instrumental information is when agents could also obtain instrumental information; empirical studies on preference for non-instrumental information have mostly focused on settings where information is never instrumental. Theoretically, little to no attempt has been made to modify the normative theory to improve its descriptive power. It is entirely possible that the preference for non-instrumental information could be accounted for by modifying the normative theory, rather than replacing it with alternative psychological theories such as curiosity.

The behavioral study reported in Chapter 2, indeed, has proposed a specific way in which the normative model could be modified and improved. The study found that subjects deviated from the normative model's prediction by showing consistent overvaluation, suggesting that the non-normative preference for information operates even when information could be instrumental. Moreover, we found that such overvaluation was sensitive to outcomes at stake, which was consistent with an existing economic account on penalty for the lack of information (hereafter uninformativeness penalty) via recursive utility (Ahlbrecht & Weber, 1996; Eliaz & Schotter, 2007; Kreps & Porteus, 1978). When recursive utility was introduced to the standard model, it significantly improved the goodness of fit.

Critically, this newly proposed model (hereafter "modified model") provides a quantitative yet descriptive prediction of VoI. Moreover, it explains preference for both instrumental and non-instrumental information in a unified way under the principle of expected utility maximization. This enables us to look for neural representation of VoI without relying on the assumption of rationality. Furthermore, since the modified model still uses the normative model as a baseline, it is possible to divide the predicted VoI into two components; the standard component, which is identical to the normative model's prediction, and the non-standard component, which is introduced by uninformativeness penalty. This raises two hypotheses on how VoI could be computed; first, one mechanism calculates the standard VoI, which output is combined with uninformativeness penalty calculated by another mechanism; and second, the modified VoI is computed as a whole by a single mechanism without clear segregation.

In this chapter, the new model of VoI proposed in Chapter 2 and human fMRI were used to uncover neural representation of VoI. Specifically, two possible scenarios were tested; first, the standard and non-standard components are represented in distinct brain regions, and second, the whole modified VoI is represented without clear segregation. The results supported the second scenario; while no convincing evidence for representation of the standard or non-standard component alone was discovered, it was found that striatum represented the modified VoI as a whole. As striatum has been long associated with valuation processes in literature, our results suggest that VoI recruits neural machinery that is at least partly overlapped with other types of valuation, in spite of the stark conceptual difference.

Results

This study adopted a newly designed task of value-based information acquisition (Fig. 1a). During scanning, 37 subjects were presented with monetary gambles and asked to decide whether to play them or not. Each gamble had two possible monetary outcomes, one gain and one loss (e.g., $12 gain and $6 loss, variable across trials). Gambles were visually presented as a roulette wheel. When subjects accepted to play the gamble, a dot appeared at a random location on the perimeter, and its side (left or right) determined the monetary outcome. Thus, the two outcomes were equally likely at this point.

After subjects made their initial decisions on gambles ("uninformed choice"), an additional piece of information was presented, and subjects chose whether to buy the information at a given monetary cost (variable across trials) or not. This information was visually represented as a magenta partition line running through the wheel's center (vertical, slanted by 30 degrees, or slanted by 60 degrees; angle was variable across trials). If purchased, the information revealed which side with respect to the partition line the dot would appear. This disclosed the true probability distribution out of the two symmetrical ones (e.g., if the line was slanted by 60 degrees, probability of gain could be either 1/3 or 2/3). It is important to note that subjects knew the two possible distributions beforehand based on the partition line's angle. Subjects who purchased the information could change their uninformed choices after seeing the delivered information ("informed choice"; this took place after the scanning).

Subjects' information purchase behavior was compared against the normative model's prediction (see Chapter 2 for details). Under the normative model, information should be purchased only in trials where expected utility of the choice would be improved when information

was purchased, even after taking the information cost into account. Consistently with this prediction, probability of information purchase increased with the difference between expected utility of informed and uninformed choices, the former of which involved the sunk cost of information purchase (Fig. 1b).

Although the normative model made a decent success in explaining the observed behavior, subjects violated its prediction systematically by overvaluing the information. This was successfully explained by a modification of the normative model, in which expected utility of uninformed choice was reduced (compared to informed choice) via concave recursive utility function (Fig. 1b; $p < .00005$, 10-fold cross-validation iterated for 2000 times). This finding was confirmed by an independent dataset, in which 96 subjects reported their subjective VoI as willingness-to-pay (Chapter 2, Fig. 6c).

Neural representation of the standard component of VoI

Neural representation of VoI was examined based on the modified model obtained in the behavioral analysis. BOLD responses during presentation of the information were analyzed. The modified model provides a numeric prediction of VoI as a function of the gamble's outcomes and the information diagnosticity (i.e., the partition line's angle). One notable difference between the predictions of the modified and normative models is that, under the modified model, non-instrumental information's value could be non-zero due to uninformativeness penalty. Importantly, the modified VoI can be devided into two parts, the standard component (the same as the normative model's predictions) and the non-standard component (the modification due to uninformativeness penalty).

Based on this feature of the modified model, two possible scenarios about neural representation of VoI were examined. The first scenario is that VoI is computed according to the normative model in some region(s), while uninformativeness penalty is introduced somewhere else. This scenario is conceived based on a widespread idea that information acquisition is driven by multiple distinct factors, one of which is the improvement in economic utility and the others are some intrinsic, psychological drive for information. The second scenario is that some region(s) incorporates both the standard and non-standard components of VoI in a unified manner, or in other words, represents VoI as predicted by the modified model as a whole. This scenario would suggest that the neural processes of information valuation are not necessarily separable into multiple factors, at least at the spatio-temporal scale of fMRI.

To test the first scenario, representation of the standard component of VoI was examined. Representational similarity analysis (RSA) and searchlight approach was used for this purpose (Kriegeskorte & Kievit, 2013; Kriegeskorte, Mur, & Bandettini, 2008). RSA examines association between local patterns of BOLD signals and the variables of interest (VoI in this case) by comparing their distances among instances, or more formally, representational dissimilarity matrix (RDM). An important advantage of RSA compared to the conventional fMRI analysis using general linear model (GLM) is its sensitivity. GLM-based analysis makes an assumption that VoI is encoded by neighboring voxels in a consistent manner, but such an assumption has not yet been empirically verified. It is possible that voxels that encode positively are spatially intermingled with voxels that encode VoI negatively, and even though these voxels certainly contain information about VoI, it cannot be reliably detected by GLM-based analysis due to spatial smoothing (Kahnt,

Heinzle, Park, & Haynes, 2011). RSA, on the other hand, may be sensitive even when such a representational scheme was used in the brain.

In RSA, RDM was constructed based on the standard component of VoI (Fig. 2a, left) and compared against RDM of BOLD signals within a spherical searchlight (8mm radius, moved across the whole brain). Their similarity was measured as Spearman's ρ, which was then z-transformed and averaged over 37 subjects, yielding a group-level, whole-brain correlation map (Fib. 2b). This analysis revealed significant association between the standard VoI and BOLD signals from bilateral superior frontal sulcus (SFS), bilateral postcentral sulcus (PCS), right superior occipital gyrus (SOG), and bilateral lingual gyrus (LG) (Fig. 2b, Table 1; voxel-wise threshold $p < .05$, corrected for whole-brain family wise error evaluated by nonparametric permutation, and cluster size threshold $k > 10$).

However, this analysis suffers from a serious coufound. By design, the standard VoI is highly correlated with information's diagnosticity, which was visually represented as the partition line's angle (Spearman's $\rho = 0.84$, Fig. 3a, top). For instance, the vertical line provided perfect prediction of the outcome, and hence higher VoI, while the 60°-slanted line provided lower predictability and hence lower VoI. Since the partition line was a very salient visual feature during scanning, it is critical to account for its effect on neural responses, particularly in parietal and occipital cortices.

Another RSA was conducted to control for this confound. In this analysis, the similarity between the standard VoI RDM and BOLD RDM were evaluated by partial correlation, removing the contribution of the partition line's angle. This analysis, however, did not provide any evidence for association between the standard VoI and any clusters identified in the initial analysis (Wilcoxon signed rank test, $p > .05$; Fig. 3b, top). Therefore, we cannot reject the possibility that the initial RSA's results reflect visually evoked responses rather than valuation processes.

Neural representation of the non-standard component of VoI

In addition to the standard component of VoI, representation of the non-standard component was also predicted in the first scenario. Searchlight RSA with RDM based on the non-standard component of VoI was conducted to test this (Fig. 2a, middle). However, there was no region in which activation pattern represented the non-standard VoI (voxel-wise $p > .001$, uncorrected). Together with the lack of compelling evidence for the standard VoI representation (Fig. 3b, top), the results so far did not support the first scenario, which predicted that the standard and non-standard components of VoI were mapped onto distinct brain regions.

Neural representation of the modified VoI

Next, the second scenario was examined, which predicted that the modified VoI was represented as a whole, i.e., without clear segregation between the standard and non-standard components. To look for such representation, searchlight RSA was conducted with RDM based on the modified VoI (Fig. 2a right). This analysis revealed significant association between the modified VoI and BOLD signals from right SFS, left PCS, and right striatum (Fig. 2c, Table 2; $p < .05$, $k > 10$). While clusters in SFS and PCS were overlapped with those obtained in the RSA with the standard VoI (Fig. 2b), the striatum cluster was unique to this analysis.

However, this analysis suffers from the same confound as the previous RSA with the standard VoI; the modified VoI was correlated with the partition angle, a salient visual feature (Spearman's $\rho = 0.71$; Fig. 3a, bottom). To address this issue, partial correlation between the modified VoI RDMs and BOLD RDMs in each cluster was evaluated, removing the contribution of the partition angle. SFS and PCS did not exhibit significant partial correlation ($p > .05$; Fig. 3b, bottom), consistent with earlier results. Importantly, on the other hand, association between striatum and the modified RDM was still statistically significant ($p < .05$). Therefore, visually evoked responses could explain activation in SFS and PCS, but not activation in striatum.

Model comparison: the standard versus modified VoI

Results so far suggest that striatum computes VoI as predicted by the modified model as a whole, rather than its standard or non-standard component alone. In other words, striatum may evaluate information by taking into account both the degree to which the information is instrumental and the degree to which the uninformed choice is disliked.

To provide further evidence that the modified VoI is computed in the brain as a whole rather than the standard component, regions-of-interest (ROI) analysis was conducted, which examined whether BOLD RDM was more similar to the modified VoI's RDM or to the standard VoI's RDM. To define ROIs, all voxels that reached significance ($p < .05$, FWE corrected) in either of the two RSAs were included (i.e., the union of voxels in Fig. 2b and c). This procedure was adopted to alleviate the issue of circularity in statistical inference (Kriegeskorte et al., 2009); under the null hypothesis, ROIs defined by the union would favor either RDM at the equal chance. Eight ROIs were obtained in total: two SFS, two PCS, three occipital, and one striatum.

As a result of this direct model comparison, striatum, but not any other ROIs, exhibited stronger association with the modified VoI than the standard VoI (Wilcoxon signed rank test, $p > .05$; Fig. 4). This provides further evidence that valuation processes in striatum calculate VoI as a whole, as predicted by the modified model.

Discussions

Information acquisition is a critical part of decision-making. Almost all decisions we make in daily life involve uncertainty and thus could be improved based on additional relevant information. However, it is challenging to acquire information adaptively, because information can be rarely obtained without cost. Although information acquisition is ubiquitous in the modern age, little is known about its neural basis, particularly the underlying valuation processes.

The current fMRI study aimed to understand value of information (VoI) is computed in human brain. It found that VoI was represented by activation in striatum. More specifically, activation in striatum was associated with VoI predicted by a newly developed model, which combines the traditional notion of instrumental VoI (Edwards, 1965; Howard, 1966; Marschak, 1971; Wendt, 1969) and recursive utility that penalizes uninformed choices or states (Ahlbrecht & Weber, 1996; Kreps & Porteus, 1978). While the former has been long regarded as a gold standard in economics, it has never been widely used in neuroscience due to its unsatisfactory descriptive validity. The current study addresses this issue by introducing uninformativeness penalty, which improved the modified model's ability to predict behavior. This modified model allowed us to look for neural representation of VoI that was descriptively, rather than normatively, estimated.

VoI in the modified model contains a non-standard component, namely penalty for uninformativeness introduced by recursive utility. This feature is relevant to a widespread notion that information could be acquired by reasons other than expected utility maximization, particularly curiosity (Berlyne, 1966; Loewenstein, 1994). Interestingly, the current study did not find any evidence for representation of the non-standard component alone. Although we need to be careful in interpreting negative results, this suggest that, at least at the spatiotemporal scale of fMRI, dichotomy between economic and non-economic drives, or rational and irrational mechanisms, might not be useful to understand neural basis of information acquisition.

Striatum is one of the regions that have been the most robustly associated with valuation (Lau & Glimcher, 2008; Levy, Snell, Nelson, Rustichini, & Glimcher, 2010; Strait, Sleezer, & Hayden, 2015; Yamada et al., 2013). Valuation of information is conceptually quite different from other types of valuation; to compute VoI, agents need to stimulate choices under possible information states, evaluate their utilities, and combine them to determine how much information improves expected utility. Such processes are not necessary for other types of valuation. It is worth pointing out here that, in the real world, human may well learn value of information over time through trial and error. However, because the current task did not provide any feedback about usefulness of information to subjects, the findings in the current study cannot be explained by such learning. Our results therefore suggest that striatum may play a critical role in valuation even when it involves complex computational processes.

A series of studies particularly relevant to the current findings has been conducted by Bromberg-Martin and colleagues (Blanchard et al., 2015; Bromberg-Martin & Hikosaka, 2009; 2011). They found that monkeys prefer to know the amount of juice they are going to receive earlier than later, even though that information does not affect juice amount itself, and such preference is reflected by activity of dopaminergic neurons in midbrain. Even though their studies did not involve instrumental information, an economic account on the preference for earlier uncertainty resolution, i.e., penalty for uninformativeness due to recursive utility, also explains the non-standard component of VoI in the current task. Since striatum receives direct dopaminergic inputs from midbrain (Morris, Arkadir, Nevet, Vaadia, & Bergman, 2004; Schultz, 1998), our results are consistent with their findings. Our results suggest that their results may not be explained by some biases in reinforcement learning alone (Beierholm & Dayan, 2010; Iigaya, Story, Kurth-Nelson, Dolan, & Dayan, 2016), and raise the possibility that dopaminergic neurons also encode the standard (i.e., instrumental) component of VoI.

Interestingly, the current study did not find representation of VoI in cortical regions. In particular, even though anterior cingulate cortex (ACC), ventromedial prefrontal cortex (vmPFC), and orbitofrontal cortex (OFC) have been robustly associated with valuation along with striatum (Hare et al., 2008; Padoa-Schioppa & Assad, 2006; 2007; Plassmann et al., 2007; Rangel et al., 2008; Rudebeck et al., 2008; M. Rushworth et al., 2011; M. F. S. Rushworth & Behrens, 2008; Tremblay & Schultz, 1999), none of them was found to be associated with VoI, at least at the stringent statistical criteria. These results, however, would not necessarily imply that these regions are not involved in information valuation; it is quite likely that these regions play critical roles in valuation, but in such a way that the current study's approach (i.e., mapping of VoI) cannot detect it. One possibility is that striatum communicates with ACC, vmPFC, and/or OFC, and aggregates

these regions' outputs in order to compute VoI (Balleine et al., 2007; Di Martino et al., 2008; Friedman et al., 2015). These cortical regions, on the other hand, may serve specific roles within the process, such as comparison of possible outcomes (Boorman et al., 2009; 2013), simulation of hypothetical actions (Fermin et al., 2016; Hayden, Pearson, & Platt, 2009), evaluation of status quo (Shenhav, Straccia, Botvinick, & Cohen, 2016), or action selection based on VoI calculated in striatum (Hayden, Pearson, & Platt, 2011; Quilodran et al., 2008).

The task design adopted in this study has a few limitations that need to be addressed by future studies. First, interpretability of RSA results was hindered by the confound of visually evoked responses; diagnosticity of information was presented as the angle of the partition line, which was a quite salient visual feature. Even though the effect of the partition angle was controlled for by RSAs based on partial correlation, this issue could be more directly addressed by improved task designs, e.g., the roulette wheel being rotated between trials. Second, all gambles in the current task had two outcomes and two information states that were equally likely. To provide more compelling evidence that striatum computes VoI, it is critical to examine how striatum responds to probability distributions over outcomes and states; both the normative theory (Medlin, 1979; Wendt, 1969) and the recursive utility theory (Ahlbrecht & Weber, 1996; Eliaz & Schotter, 2007) provide specific predictions on how VoI should be sensitive to probability distributions.

In conclusion, the current study revealed, for the first time, that activation in striatum represents VoI. To represent VoI, striatum is suggested to combine two factors: the degree to which information is instrumental, and the degree to which the uninformative choice is penalized through recursive utility. This complements previous studies on specific information acquisition actions (e.g., exploration-exploitation dilemma) and sheds light on a general functional role of striatum in valuation process. Conceptually, the current study illustrates the advantage of combining multiple theoretical accounts into a unified model of information acquisition.

Methods

Subjects. 47 subjects participated in the study after providing informed consent and being screened for standard MRI contraindications. 10 were removed from the analysis due to excessive motion during scanning, resulting in 37 subjects used in the data analysis (mean: 24.5 years old, SD: 5.2, 21 female).

Task. Inside the scanner, subjects interacted with the experiment program on Matlab and Psychtoolbox (Brainard, 1997; Pelli, 1997) via an MRI-compatible button box. Each subject's scanning consisted of five EPI runs, each of which included 30 trials in a randomized order (150 in total).

Each trial started with presentation a gamble with two outcomes. The gamble was visually presented as a roulette wheel partitioned by a vertical while line at the middle (3 seconds). Ten outcome combinations were used: (+$12, −$9), (+$9, −$12), (+$9, −$9), (+$12, −$6), (+$6, −$12), (+$9, −$6), (+$6, −$9), (+$6, −$6), (+$12, −$3), and (+$3, −$12). After fixation screen (1–4 seconds), subjects chose whether to play the gamble or not (two-alternative-forced choice; "uninformed choice", within 2 seconds), followed by another presentation of the same gamble (1–4 seconds). The information was then presented as a magenta partition line running through the wheel's center (2 seconds). The information line was either slanted by 0° (vertical), 30°, or 60°. Each angle was presented once per each pair of outcomes in each EPI session (3 angle × 10 outcomes = 30 trials). The information display was followed by presentation of monetary cost, and subjects decided whether to purchase the information or not (two-alternative-forced choice, within 2 seconds). The monetary cost was variable across trials and determined independently from gambles and partition angle; it was either ¢5 (3 trials per session), $1 (8 trials), $2 (8 trials), $3 (8 trials), or $9 (3 trials). The trials in which subjects failed to respond within 2 seconds were discarded from the analysis.

After the scanning, five trials were randomly selected by the experiment program and implemented into the actual monetary payments. If they had chosen to play a selected gamble during scanning without purchasing the information, a green dot appeared on the perimeter of the roulette, which side on the wheel determined the outcome. Without the information delivery, the green dot's location followed uniform distribution over the perimeter, making the two possible outcomes equally likely. If subjects had chosen to purchase the information during the scanning, subjects received its actual content of information; the side of the magenta partition the green dot would appear was revealed by being brightened. This changed the outcome probability distribution from (0.5, 0.5) to either (1, 0) or (0, 1) (vertical line), (5/6, 1/6) or (1/6, 5/6) (30°-slanted line), or (2/3, 1/3) or (1/3,2/3) (60°-slanted line). The brighter side was chosen randomly. Subjects then made the decision on whether to play the gamble based on the delivered information ("informed choice"), disregarding their own uninformed choice during the scanning. If they chose to buy the information, the green dot appeared randomly within the brightened area, which side determined the outcome.

Behavioral analysis. All analyses below were conducted on Matlab, using maximum likelihood estimation procedure. Subjects' uninformed choices on gambles during scanning were first analyzed to obtain the group-level utility function. Binary choices were predicted by softmax function of the expected utility (i.e., average of two outcomes' utility), which temperature

parameter was estimated with the utility function simultaneously. The obtained utility function was $U(x) = x^{0.68}\ (x > 0), 1.74 \cdot x^{0.49}(x < 0)$, indicating strong risk aversion and loss aversion.

To model information purchase behavior by the normative theory, binary choices were predicted by another softmax function based on the difference between expected utility of informed and uninformed choices. Expected utilities were obtained on a trial-by-trial basis under the assumptions that 1) agents adopted a fully deterministic expected-utility-maximizing policy, and 2) the information cost was sunk in the informed choice:

$$\begin{aligned} EU_{\text{uninfo}} &= \max(0.5 \cdot u(x_1) + 0.5 \cdot u(x_2), 0) \\ EU_{\text{info}} &= 0.5 \cdot EU_{\text{info1}} + 0.5 \cdot EU_{\text{info2}} \\ EU_{\text{info1}} &= \max\big(p \cdot u(x_1 - c) + (1-p) \cdot u(x_2 - c), u(-c)\big) \\ EU_{\text{info2}} &= \max\big((1-p) \cdot u(x_1 - c) + p \cdot u(x_2 - c), u(-c)\big) \end{aligned}$$

where x_1, x_2 are the gamble's outcomes, p is the posterior probability of an outcome after the information delivery, and c is the information cost. The goodness of fit of the softmax function, measured by negative log likelihood, was statistically evaluated by permutation over trials (1999 iterations).

The modified model included an additional free parameter, the concavity of recursive utility function $v(u)$, which controls the size of penalty for uninformed choice. Expected utilities under the modified model are:

$$\begin{aligned} EU_{\text{uninfo}} &= \max\big(0.5 \cdot v\big(u(x_1)\big) + 0.5 \cdot v\big(u(x_2)\big), 0\big) \\ EU_{\text{info}} &= 0.5 \cdot v(EU_{\text{info1}}) + 0.5 \cdot v(EU_{\text{info2}}) \end{aligned}$$

where EU_{info1}, EU_{info2} are the same as above. A symmetric power function was used for v, and its concavity was estimated simultaneously with the softmax function's temperature parameter. The estimated recursive utility was $v(u) = u^{0.47}$.

To statistically compare the standard and the modified models, 10-fold cross-validation analysis was conducted. First, trials were split into ten datasets (10 trials per subject), and nine of them was collated as the training dataset. Second, v was estimated from the training dataset. Third, negative log-likelihood in the hold-out dataset was obtained with the standard model and the modified models respectively, the latter of which used v estimated in the previous step. This was repeated for every hold-out dataset (10 times), and then repeated 200 times, i.e., 2000 iterations in total. This procedure circumvented the problem of overfitting because the goodness of fit was compared only in the hold-out sample.

fMRI data acquisition. MR images were acquired by a 3T Siemens Trio scanner and a 12-channel head coil. Functional images were obtained using T2*-weighted gradient-echo echo-planar imaging (EPI) pulse sequence (TR = 2000ms, TE = 30ms, voxel size = 3mm × 3mm × 3mm, inter-slice gap = 0.3mm, in-plane resolution = 64 × 64, 32 oblique axial slices). Slices were tilted by 30 degrees from AC-PC line to alleviate signal dropout from orbitofrontal cortex (Weiskopf et al.,

2006). T1-weighted structural images (1mm × 1mm × 1mm) were also obtained using magnetization-prepared rapid-acquisition gradient-echo (MPRAGE) pulse sequence.

Preprocessing. Motion correction and slice-time correction were applied to EPI images on SPM12 (Wellcome Dept. of Cognitive Neurology, London, UK) prior to further analyses described below.

General linear modeling (GLM). In order to obtain representational dissimilarity matrix (RDM) for searchlight RSA, voxel-wise signals for each type of information (30 in total) were first estimated by GLM on SPM12. All EPI runs from each subject were concatenated and analyzed by a single GLM. The GLM included 30 regressors modeling information presentation (2-second boxcar, up to 5 times, once per each EPI run), 30 regressors modeling information presentation (2-second boxcar, up to 5 times), regressors modeling gamble presentation and information presentation in all trials in which subjects did not respond, and regressors modeling button press events (right-hand and left-hand responses separately). All of these event-related regressors were convolved with the SPM's double-gamma canonical hemodynamic response function. The GLM also included dummy regressors modeling distinct EPI runs, movement parameters estimated in the motion-correction procedure, 128-sec high-pass filtering, and AR(1) model of serial autocorrelation. Coefficient estimates of the regressors modeling information presentation were used as the activation pattern in RSA.

Representational similarity analysis (RSA). Association between the variables of interest (the standard VoI, the modified VoI, or their difference) and the BOLD activation pattern was examined by searchlight RSA implemented by a custom-made Matlab script. Each entry in VoI RDMs (Fig. 2a) was the absolute difference between predicted VoI among 30 instances. Each entry in BOLD RDM was Euclid distance of activation patterns within a spherical searchlight (8mm radius) among 30 instances. Spearman's correlation between model-based RDM and searchlight BOLD RDM was calculated, z-transformed, and assigned to the voxel at the center of the searchlight. The searchlight was moved across the entire brain covered by EPI, resulting in subject-wise whole-brain correlation maps.

These subject-wise correlation maps were normalized to MNI template using SPM12's DARTEL procedure, which consists of two steps: non-linear normalization to the average brain among 37 subjects and affine transformation to MNI template. Normalized maps were then smoothed with a Gaussian kernel (8-mm FWHM) and averaged over subjects. Statistical significance in these group-level correlation maps (Fig. 2b, Tables 1 and 2) was evaluated at the voxel-level threshold $p < .05$, corrected for whole-brain family-wise error (FWE) based on non-parametric permutation, estimated by SnPM13 package (Nichols & Holmes, 2001).

To account for the confound of the partition line's angle (Fig. 3), RSA was repeated in which partial correlation between VoI RDMs and searchlight BOLD RDMs was computed while controlling for RDM based on the line's angle. To evaluate statistical significance within each cluster (ROI), z-transformed correlation coefficients were averaged over voxels that survived $p < .05$ in the initial analysis, and then compared against zero at the group level (Wilcoxon signed rank test).

Figures

Figure 1. (a) Experimental paradigm. *Top*: During scanning, gambles with two monetary outcomes were presented. Each gamble took the form of a roulette wheel. Subjects chose whether to accept or reject the gamble without the additional information (uninformed choice). After a delay, information was presented as a magenta partition on the wheel, followed by the information cost. Subjects indicated whether they to purchase it or forgo it. *Bottom*: If subjects purchased the information, its content was delivered after the scanning. One side of the partition was brightened, revealing the true outcome probability distribution. Subjects were then allowed to change their prior choice on the gamble based on the delivered information (informed choice). **(b)** Behavioral results. *Left*: The standard, normative model explained information purchase probability decently well. As predicted, subjects purchased the information more frequently with the larger difference between expected utilities of informed and uninformed choices ($p < .0005$). Right*:* The modified model achieved the better model fit than the standard model ($p < .05$). In the modified model, uninformed choice was penalized via recursive utility. EU_{info}: expected utility of informed choice, EU_{uninfo}: expected utility of uninformed choice (see Methods for details).

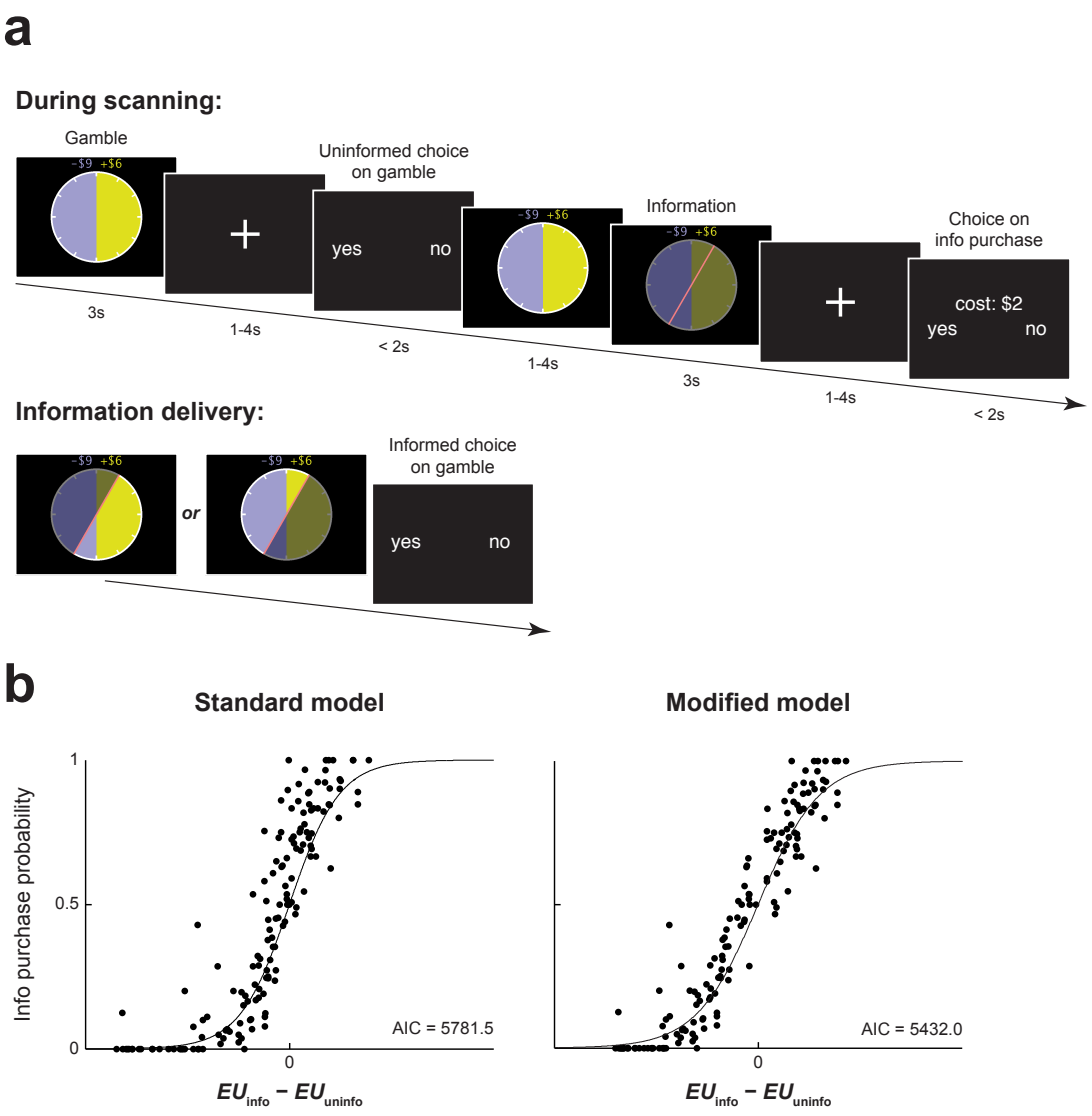


Figure 2. Representational similarity analysis (RSA). **(a)** Representational dissimilarity matrix (RDM) was constructed based on the predicted VoI under the standard model (*left*), predicted VoI under the modified model (*right*), and their difference ("non-standard component", *middle*). Each matrix has the size of 30 × 30, as there were 30 unique instances of information (10 gambles × 3 partition angle). Each entry was the absolute value difference. Each RDM was compared with BOLD RDMs in the whole-brain searchlight RSAs. **(b, c)** Searchlight RSA results with the RDMs of the standard (b) and modified (c) VoI. Shown are clusters that survived voxel-wise $p < .05$, whole-brain corrected for family wise error (FWE), and cluster size $k > 10$. No cluster was found with the non-standard VoI's RDM. SFS: superior frontal sulcus, PCS: postcentral sulcus, SOG: superior occipital gyrus, LG: lingual gyrus. See Tables 1 and 2 for peak coordinates and effect sizes.

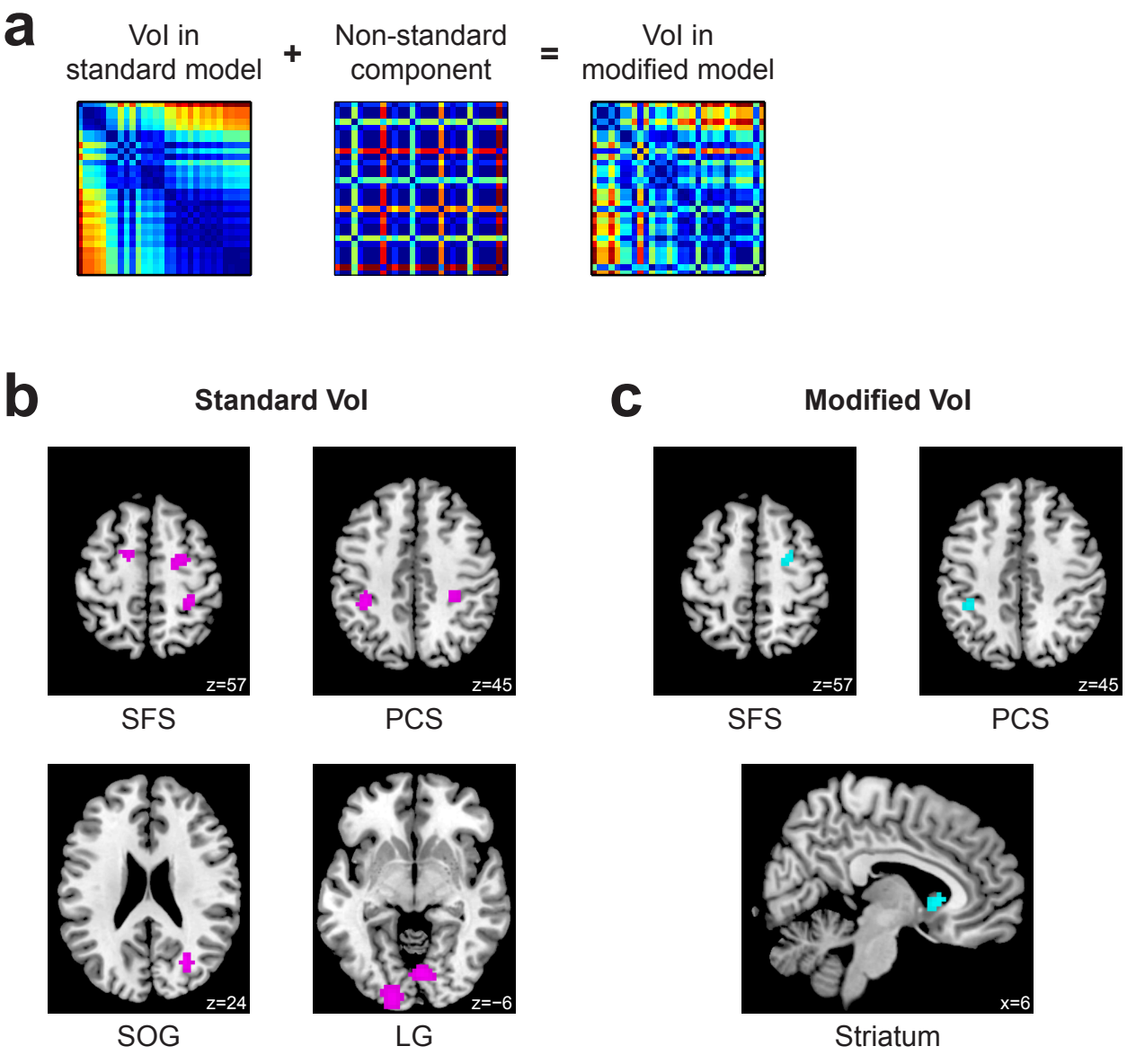


Figure 3. Controlling for visual features, in particular effect of the partition angle on RSA. **(a)** The partition angle was correlated with the standard VoI (*top*, Spearman's ρ = 0.84) and the modified VoI (*bottom*, 0.71). **(b)** ROI analysis. Averaged correlation coefficients were shown for each cluster identified in Fig. 2b, c. Blue: non-partial correlation (the same as Fig. 2b, c), red: partial correlation controlling for the partition angle RDM. *: $p < .05$, Wilcoxon signed rank test. Error bars: SEM.

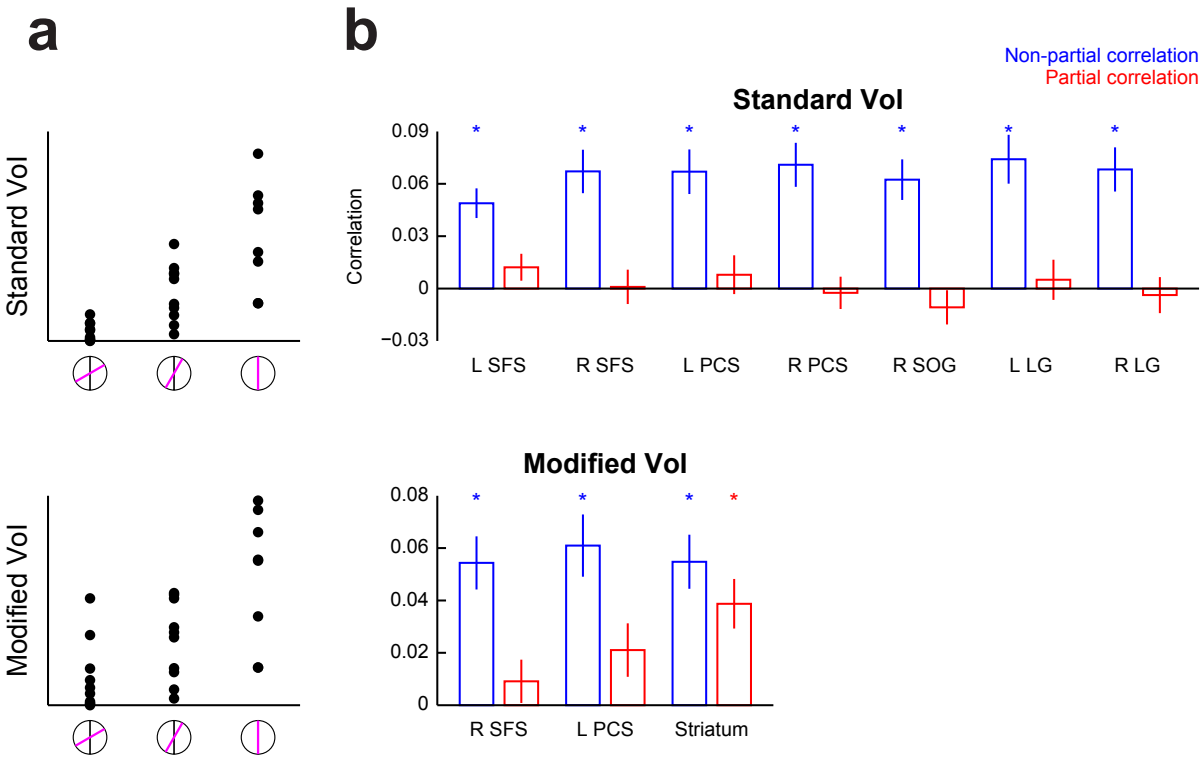
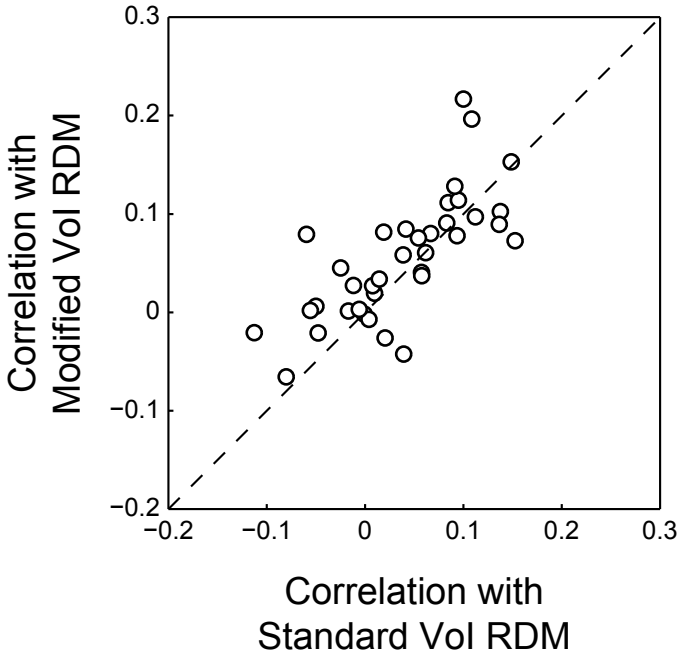


Figure 4. Comparison between the standard and modified VoIs in striatum. Each dot corresponds to each subject. RDM based on the modified VoI exhibited higher correlation with striatum activation than RDM based on the standard VoI ($p < .05$, Wilcoxon signed rank test).



Tables

Table 1. Results of RSA with the standard VoI. Shown are the clusters that survived voxel-wise $p < .05$, whole-brain corrected for family wise error (FWE), and cluster size $k > 10$. Peak coordinates are in MNI space. SFS: superior frontal sulcus, PCS: postcentral sulcus, SOG: superior occipital gyrus, LG: lingual gyrus, L: left, R: right. See also Fig. 2b.

Region	# voxel	*Peaks*				
		x	*y*	*z*	*T*(36)	*Corrected P*
L SFS	35	-18	0	60	5.62	0.0060
R SFS	65	24	-6	54	5.20	0.0205
L PCS	63	-36	-39	45	5.29	0.0140
R PCS	94	30	-36	54	5.63	0.0060
R SOG	42	27	-66	24	5.39	0.0105
L LG	89	-15	-99	-9	5.37	0.0110
R LG	81	9	-78	-3	5.49	0.0085

Table 2. Results of RSA with the modified VoI. Shown are the clusters that survived voxel-wise $p < .05$, whole-brain corrected for family wise error (FWE), and cluster size $k > 10$. Peak coordinates are in MNI space. SFS: superior frontal sulcus, PCS: postcentral sulcus, L: left, R: right. See also Fig. 2c.

Region	# voxel	*Peaks*				
		x	*y*	*z*	*T*(36)	*Corrected P*
R SFS	34	24	-6	54	5.17	0.0240
L PCS	33	-36	-39	42	5.15	0.0255
Striatum	34	3	18	3	5.46	0.0130

References

Abbas, A. E., Bakir, N. O., Klutke, G.-A., & Sun, Z. (2013). Effects of risk aversion on the value of information in two-action decision problems. *Decision Analysis*, *10*(3), 257–275.

Ahlbrecht, M., & Weber, M. (1996). The Resolution of uncertainty: An experimental study. *Journal of Institutional and Theoretical Economics*, *152*(4), 593–607.

Allman, J. M., Tetreault, N. A., Hakeem, A. Y., Manaye, K. F., Semendeferi, K., Erwin, J. M., et al. (2010). The von Economo neurons in frontoinsular and anterior cingulate cortex in great apes and humans. *Brain Structure and Function*, *214*(5-6), 495–517.

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique effect" in man and animals. *Psychological Bulletin*, *78*(4), 266–278.

Bach, D. R., Hulme, O., Penny, W. D., & Dolan, R. J. (2011). The known unknowns: Neural representation of second-order uncertainty, and ambiguity. *The Journal of Neuroscience*, *31*(13), 4811–4820.

Badre, D., Doll, B. B., Long, N. M., & Frank, M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron*, *73*(3), 595–607.

Bakır, N. O., & Klutke, G.-A. (2011). Information and preference reversals in lotteries. *European Journal of Operational Research*, *210*(3), 752–756.

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience*, *27*(31), 8161–8165.

Bandyopadhyay, R. (1977). Information for organizational decisionmaking—A literature review. *IEEE Transactions on Systems, Man, and Cybernetics*, *7*(1), 1–15.

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

Becker, S. W., & Brownson, F. O. (1964). What price ambiguity? Or the role of ambiguity in decision-making. *The Journal of Political Economy*, *72*, 62–73.

Beckmann, M., Johansen-Berg, H., & Rushworth, M. F. S. (2009). Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *Journal of Neuroscience*, *29*(4), 1175–1190.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–249.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221.

Beierholm, U. R., & Dayan, P. (2010). Pavlovian-Instrumental Interaction in "Observing Behavior." *PLoS Computational Biology*, *6*(9), e1000903.

Bennett, D., Bode, S., Brydevall, M., Warren, H., & Murawski, C. (2016). Intrinsic valuation of information in decision making under uncertainty. *PLoS Computational Biology*, *12*(7), e1005020.

Berg, M., & Eisenberger, R. (1996). An experiment on the evaluation of information under risk and ambiguity. *OR Spektrum*, *18*(3), 179–186.

Berlyne, D. E. (1957). Uncertainty and conflict: A point of contact between information-theory and behavior-theory concepts. *Psychological Review*, *64*(6), 329–339.

Berlyne, D. E. (1966). Curiosity and exploration. *Science*, *153*(3731), 25–33.

Berns, G. S., Chappelow, J., Cekic, M., Zink, C. F., Pagnoni, G., & Martin-Skurski, M. E. (2006). Neurobiological Substrates of Dread. *Science*, *312*(5774), 754–758.

Blair, R. D., & Romano, R. E. (1988). The influence of attitudes toward risk on the value of forecasting. *The Quarterly Journal of Economics*, *103*(2), 387–396.

Blanchard, T. C., Hayden, B. Y., & Bromberg-Martin, E. S. (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron*, *85*(3), 602–614.

Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., & Rushworth, M. F. S. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron*, *62*(5), 733–743.

Boorman, E. D., Rushworth, M. F., & Behrens, T. E. (2013). Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *The Journal of Neuroscience*, *33*(6), 2242–2253.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J. B. (2002). Region of interest analysis using an SPM toolbox. *NeuroImage*, *16*(2).

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, *63*(1), 119–126.

Bromberg-Martin, E. S., & Hikosaka, O. (2011). Lateral habenula neurons signal errors in the prediction of reward information. *Nature Neuroscience*, *14*(9), 1209–1216.

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, *5*(4), 325–370.

Chew, S. H., & Ho, J. L. (1994). Hope: An empirical study of attitude toward the timing of uncertainty resolution. *Journal of Risk and Uncertainty*, *8*(3), 267–288.

Chumbley, J. R., Flandin, G., Bach, D. R., Daunizeau, J., Fehr, E., Dolan, R. J., & Friston, K. J. (2012). Learning and generalization under ambiguity: an fMRI study. *PLoS Computational Biology*, *8*(1), e1002346.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, *58*(3), 306–324.

Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology*, *39*(2), 131–143.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879.

Dayan, P., & Sejnowski, T. J. (1996). Exploration bonuses and dual control. *Machine Learning*, *25*(1), 5–22.

Di Martino, A., Scheres, A., Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Shehzad, Z., et al. (2008). Functional connectivity of human striatum: A resting state fmri study. *Cerebral Cortex*, *18*(12), 2735–2747.

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

Eliaz, K., & Schotter, A. (2007). Experimental testing of intrinsic preferences for noninstrumental information. *American Economic Review*, *97*(2), 166–169.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, *75*(4), 643–669.

Eradath, M. K., Abe, H., Matsumoto, M., Matsumoto, K., Tanaka, K., & Ichinohe, N. (2015). Anatomical inputs to sulcal portions of areas 9m and 8Bm in the macaque monkey. *Frontiers in Neuroanatomy*, *9*(123), 30.

Ergin, H., & Gul, F. (2009). A theory of subjective compound lotteries. *Journal of Economic Theory*, *144*(3), 899–929.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.

Evrard, H. C., Forro, T., & Logothetis, N. K. (2012). Von Economo neurons in the anterior insula of the macaque monkey. *Neuron*, *74*(3), 482–489.

Fermin, A. S. R., Yoshida, T., Yoshimoto, J., Ito, M., Tanaka, S. C., & Doya, K. (2016). Model-based action planning involves cortico-cerebellar and basal ganglia networks. *Scientific Reports*, *6*, 31378.

Friedman, A., Homma, D., Gibb, L. G., Amemori, K., Rubin, S. J., Hood, A. S., et al. (2015). A corticostriatal path targeting striosomes controls decision-making under conflict. *Cell*, *161*, 1320–1333.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302.

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595.

Gottlieb, J., & Balan, P. (2010). Attention as a decision in information space. *Trends in Cognitive Sciences*, *14*(6), 240–248.

Gross, J., Woelbert, E., Zimmermann, J., Okamoto-Barth, S., Riedl, A., & Goebel, R. (2014). Value signals in the prefrontal cortex predict individual preferences across reward categories. *The Journal of Neuroscience*, *34*(22), 7580–7586.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of Neuroscience*, *28*(22), 5623–5630.

Hayasaka, S., & Nichols, T. E. (2003). Validating cluster size inference: random field and permutation methods. *NeuroImage*, *20*(4), 2343–2356.

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive reward signals in the anterior cingulate cortex. *Science*, *324*(5929), 948–950.

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, *14*(7), 933–939.

Howard, R. A. (1966). Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, *2*(1), 22–26.

Howard, R. A. (1967). Value of information lotteries. *IEEE Transactions on Systems Science and Cybernetics*, *3*(1), 54–60.

Hsu, M., Bhatt, M., Adolphs, R., & Tranel, D. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, *310*(5754), 1680–1683.

Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., & Platt, M. L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron*, *49*(5), 765–775.

Hunt, L. T., Dolan, R. J., & Behrens, T. E. J. (2014). Hierarchical competitions subserving multi-attribute choice. *Nature Neuroscience*, *17*(11), 1613–1622.

Iigaya, K., Story, G. W., Kurth-Nelson, Z., Dolan, R. J., & Dayan, P. (2016). The modulation of savouring by prediction error and its effects on choice. *eLife*, *5*, e13747.

Ishii, S., Yoshida, W., & Yoshimoto, J. (2002). Control of exploitation–exploration meta-parameter in reinforcement learning. *Neural Networks*, *15*(4-6), 665–687.
Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.
Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J.-D. (2011). Decoding the formation of reward predictions across learning. *Journal of Neuroscience*, *31*(41), 14624–14630.
Keynes, J. M. (1921). A treatise on money. London: McMillan.
Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, *73*(6), 1849–1892.
Knight, F. H. (1921). Risk, uncertainty and profit. Boston, MA: Houghton Mifflin.
Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural mechanisms of foraging. *Science*, *336*(6077), 95–98.
Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*(6971), 244–247.
Kreps, D. M., & Porteus, E. L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, *46*(1), 185–200.
Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401–412.
Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis–connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4.
Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, *12*(5), 535–540.
Lau, B., & Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron, 58*(3), 451–463.
Levy, I., Snell, J., Nelson, A. J., Rustichini, A., & Glimcher, P. W. (2010). Neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology*, *103*, 1036–1047.
Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98.
Ma, W. J., & Jazayeri, M. (2014). Neural coding of uncertainty and probability. *The Annual Review of Neuroscience*, *37*, 205–220.
Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2007). Mapping the functional connectivity of anterior cingulate cortex. *NeuroImage*, *37*(2), 579–588.
Mars, R. B., Jbabdi, S., Sallet, J., O'Reilly, J. X., Croxson, P. L., Olivier, E., et al. (2011). Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *The Journal of Neuroscience*, *31*(11), 4087–4100.
Marschak, J. (1971). Economics of information systems. *Journal of the American Statistical Association*, *66*(333), 192–219.
Marschak, J. (1973). Limited role of entropy in information economics. *Theory and Decision*, *5*(1), 1–7.
McDevitt, M. A., Dunn, R. M., Spetch, M. L., & Ludvig, E. A. (2016). When good news leads to bad choices. *Journal of the Experimental Analysis of Behavior*, *105*(1), 23–40.

McNamee, D., Rangel, A., & O'Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience*, *16*(4), 479–485.
Medlin, S. M. (1979). The value of information in a sealed bidding experiment. *Organizational Behavior and Human Performance*, *24*(3), 269–286.
Mehrez, A. (1985). The effect of risk aversion on the expected value of perfect information. *Operations Research*, *33*(2), 455–458.
Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, *214*(5-6), 655–667.
Mohr, H. M., Goebel, R., & Linden, D. E. (2006). Content-and task-specific dissociations of frontal activity during maintenance and manipulation in visual working memory. *Journal of Neuroscience*, *26*(17), 4465–4471.
Moore, A. W., & Atkeson, C. G. (1993). Prioritized sweeping: Reinforcement learning with less data and less time. *Machine Learning*, *13*(1), 103–130.
Morris, G., Arkadir, D., Nevet, A., Vaadia, E., & Bergman, H. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron*, *43*(1), 133–143.
Mumford, J. A., Poline, J.-B., & Poldrack, R. A. (2015). Orthogonalization of regressors in fMRI models. *PLoS ONE*, *10*(4), e0126255.
Nadiminti, R., Mukhopadhyay, T., & Kriebel, C. H. (1996). Risk aversion and the value of information. *Decision Support Systems*, *16*(3), 241–254.
Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *The Journal of Neuroscience*, *30*(37), 12366–12378.
Nau, R. F. (2006). Uncertainty aversion with second-order utilities and probabilities. *Management Science*, *52*(1), 136–145.
Neubert, F.-X., Mars, R. B., Sallet, J., & Rushworth, M. F. S. (2015). Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(20), E2695–E2704.
Nichols, T. E., & Holmes, A. P. (2001). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*(1), 1–25.
O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, *110*(38), E3660–E3669.
Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226.
Padoa-Schioppa, C., & Assad, J. A. (2007). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, *11*(1), 95–102.
Payzan-LeNestour, E., & Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Computational Biology*, *7*(1), e1001048.
Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, *52*, 111–139.
Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552.
Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers

into movies. *Spatial Vision*, *10*(4), 437–442.
Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M., & Leff, A. P. (2010). Comparing families of dynamic causal models. *PLoS Computational Biology*, *6*(3), e1000709.
Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience*, *27*(37), 9984–9988.
Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nature Neuroscience*, *16*(9), 1170–1178.
Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, *32*(1/2), 122–136.
Quilodran, R., Rothé, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, *57*(2), 314–325.
Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556.
Reinhart, R. M. G., Heitz, R. P., Purcell, B. A., Weigand, P. K., Schall, J. D., & Woodman, G. F. (2012). Homologous mechanisms of visuospatial working memory maintenance in macaque and human: Properties and sources. *The Journal of Neuroscience*, *32*(22), 7711–7722.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In B. AH & P. WF (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64–99). New York: Appleton Century Crofts.
Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., & Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain. *European Journal of Neuroscience*, *35*, 1190–1200.
Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. *The Journal of Neuroscience*, *28*(51), 13775–13785.
Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389–397.
Rushworth, M., Noonan, M., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron, 70*(6)*,* 1054–1069.
Schepanski, A., & Uecker, W. C. (1984). The value of information in decision making. *Journal of Economic Psychology*, *5*(2), 177–194.
Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*, 1–27.
Schwartenbeck, P., FitzGerald, T. H. B., & Dolan, R. (2016). Neural signals encoding shifts in beliefs. *NeuroImage*, *125*, 578–586.
Seo, K. (2009). Ambiguity and second-order belief. *Econometrica*, *77*(5), 1575–1605.
Shanteau, J., & Anderson, N. H. (1972). Integration theory applied to judgments of the value of information. *Journal of Experimental Psychology*, *92*(2), 266–275.
Shenhav, A., Straccia, M. A., Botvinick, M. M., & Cohen, J. D. (2016). Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(6), 1127–1139.
Shenhav, A., Straccia, M. A., Cohen, J. D., & Botvinick, M. M. (2014). Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nature Neuroscience*, *17*(9), 1249–1254.
Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings,

empathy and uncertainty. *Trends in Cognitive Sciences*, *13*(8), 334–340.
Smith, A. P. R., Stephan, K. E., Rugg, M. D., & Dolan, R. J. (2006). Task and content modulate amygdala-hippocampal connectivity in emotional retrieval. *Neuron*, *49*(4), 631–638.
Snow, A. (2010). Ambiguity and the value of information. *Journal of Risk and Uncertainty*, *40*(2), 133–145.
Sokolov, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology*, *25*, 545–580.
Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387–401.
Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, *46*(4), 1004–1017.
Strait, C. E., Sleezer, B. J., & Hayden, B. Y. (2015). Signatures of value comparison in ventral striatum neurons. *PLoS Biology*, *13*(6), e1002173.
Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: an introduction. Cambridge, MA: The MIT Press.
Tomassini, V., Jbabdi, S., Klein, J. C., Behrens, T. E. J., Pozzilli, C., Matthews, P. M., et al. (2007). Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *Journal of Neuroscience*, *27*(38), 10259–10269.
Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature*, *398*(6729), 704–708.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.
Vernet, M., Quentin, R., Chanes, L., Mitsumasu, A., & Valero-Cabre, A. (2014). Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations. *Frontiers in Integrative Neuroscience*, *8*, 66.
Weiskopf, N., Hutton, C., Josephs, O., & Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-brain analysis at 3 T and 1.5 T. *NeuroImage*, *33*(2), 493–504.
Wendt, D. (1969). Value of information for decisions. *Journal of Mathematical Psychology*, *6*(3), 430–443.
Woo, C. W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*, *91*, 412–419.
Wu, G. (1999). Anxiety and decision making with delayed resolution of uncertainty. *Theory and Decision*, *46*(2), 159–199.
Yamada, H., Inokawa, H., Matsumoto, N., Ueda, Y., Enomoto, K., & Kimura, M. (2013). Coding of the long-term value of multiple future rewards in the primate striatum. *Journal of Neurophysiology*, *109*(4), 1140–1151.
Yechiam, E., & Hochman, G. (2013). Loss-aversion or loss-attention: The impact of losses on cognitive performance. *Cognitive Psychology*, *66*(2), 212–231.
Yoshida, W., & Ishii, S. (2005). Model-based reinforcement learning: a computational model and an fMRI study. *Neurocomputing*, *63*, 253–269.
Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692.