



The ventral striatum dissociates information expectation, reward anticipation, and reward receipt

Flavia Filimon^{a,1,2}, Jonathan D. Nelson^{a,b,1,2} , Terrence J. Sejnowski^{c,d,2} , Martin I. Sereno^e , and Garrison W. Cottrell^{d,f}

^aSchool of Psychology, University of Surrey, GU2 7XH Guildford, United Kingdom; ^bCenter for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, 14195 Berlin, Germany; ^cComputational Neurobiology Laboratory, Salk Institute, La Jolla, CA 92037; ^dTemporal Dynamics of Learning Center, University of California San Diego, La Jolla, CA 92093; ^eDepartment of Psychology, San Diego State University, San Diego, CA 92182; and ^fComputer Science & Engineering Department, University of California San Diego, La Jolla, CA 92093

Contributed by Terrence J. Sejnowski, March 28, 2020 (sent for review March 17, 2016; reviewed by Jacqueline Gottlieb and Read Montague)

Do dopaminergic reward structures represent the expected utility of information similarly to a reward? Optimal experimental design models from Bayesian decision theory and statistics have proposed a theoretical framework for quantifying the expected value of information that might result from a query. In particular, this formulation quantifies the value of information before the answer to that query is known, in situations where payoffs are unknown and the goal is purely epistemic: That is, to increase knowledge about the state of the world. Whether and how such a theoretical quantity is represented in the brain is unknown. Here we use an event-related functional MRI (fMRI) task design to disentangle information expectation, information revelation and categorization outcome anticipation, and response-contingent reward processing in a visual probabilistic categorization task. We identify a neural signature corresponding to the expectation of information, involving the left lateral ventral striatum. Moreover, we show a temporal dissociation in the activation of different reward-related regions, including the nucleus accumbens, medial prefrontal cortex, and orbitofrontal cortex, during information expectation versus reward-related processing.

ventral striatum | reward | expected value of information | fMRI | Bayesian optimal experimental design (OED)

Is information valuable only if it predicts reward? What is the role of dopaminergic structures in representing expected information about uncertain events? While searching for information may be intrinsically rewarding (1, 2), whether parts of the reward system represent expected information independent of external payoffs is unknown. Ventral striatum (VS), other subcortical (midbrain), and cortical (ventromedial prefrontal, orbitofrontal, anterior cingulate) regions form a reward circuit (3). Of particular interest is the VS, which includes nucleus accumbens (NAcc) and ventral putamen. The VS participates in many cognitive and affective processes; however, the extent to which the VS represents valence, salience, or other processes is debated (4–11). Whether the VS represents predictive information about the world's structure, valuing information on its own, is unknown.

Animal and human studies implicate VS in reward-based learning (3, 12–16). Dopaminergic cells in the substantia nigra (SN) and ventral tegmental area (VTA) that project to the VS respond to reward-predicting cues and unexpected rewards. Smaller- or larger-than-expected rewards suppress or increase firing rates (15–18). Reinforcement learning theories posit that reward prediction error (RPE) forms a teaching signal that updates the expected value of choices (13, 19, 20).

In human functional MRI (fMRI) studies, VS activations scale with RPE (3, 8, 21). VS blood-oxygen level-dependent (BOLD) responses and midbrain dopaminergic neuronal firing are correlated (22), supporting an RPE interpretation of the VS. However, RPE exemplifies “model-free” learning: That is, learning without an explicit model of the environment (14, 23). In contrast, “model-based” learning involves building a predictive model describing the

expected transitions between states given each possible action in those states (23, 24). Recent evidence shows both model-free and model-based responses in the VS (25, 26). NAcc lesions impair behavior driven by goals' motivational value, resulting in habitual, stimulus-driven behavior (12). This suggests that both reinforcement and goals drive the VS (10).

Dopamine is also linked to information. Macaque dopaminergic midbrain neurons signal both expectation of information about upcoming water rewards and expected reward amount (27–29).

Do reward structures such as the VS represent expected information about things other than reward, such as categorical structure in the world? Or is processing of information solely tied to reward expectation? Optimal experimental design (OED) theories from Bayesian decision theory provide a normative theoretical framework for quantifying the usefulness of expected information in situations without explicit external payoffs (30–32). OED models quantify the usefulness of information expected to result from an experiment, test, or query, before the

Significance

Whether the reward system represents the expected utility of information distinctly from reward remains unclear. Optimal experimental design models from statistics quantify the expected value of information that might result from a query before the answer is known, independent of payoffs, when the goal is purely to increase knowledge and reduce uncertainty. We disentangle information expectation, information revelation and categorization outcome anticipation, and reward processing in visual probabilistic categorization with event-related fMRI. We show dissociation between expectation of information, expectation of response-contingent reward, and reward receipt, involving different temporal and spatial signatures of reward regions, including lateral ventral striatum, nucleus accumbens, medial prefrontal cortex, and orbitofrontal cortex. These results suggest expected information and immediate reward are distinct in the brain.

Author contributions: F.F. and J.D.N. designed research; J.D.N. performed research; F.F. and J.D.N. analyzed data; and F.F., J.D.N., T.J.S., M.I.S., and G.W.C. wrote the paper.

Reviewers: J.G., Columbia University; and R.M., Fralin Biomedical Research Institute, Virginia Institute of Technology Carilion Research Institute.

The authors declare no competing interest.

Published under the PNAS license.

Data deposition: The data, stimulus, and analysis files are permanently archived and publicly available for download on the Open Science Framework, doi:10.17605/osf.io/aexv9 or <https://osf.io/aexv9/>.

¹F.F. and J.D.N. contributed equally to this work.

²To whom correspondence may be addressed. Email: flavia.filimon@gmail.com, jonathan.d.nelson@gmail.com, or terry@salk.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1911778117/-DCSupplemental>.

outcome of that query is known, when the goal is purely epistemic (i.e., reducing uncertainty about the state of the world). OED theories have successfully modeled people's eye movements and explicit queries (33–38). However, the neural substrates underlying the expected usefulness of information for reducing uncertainty about different hypotheses about the world are unknown.

We created an event-related fMRI design to disentangle the processing of information expectation, information revelation and response-contingent outcome anticipation, and feedback, in a probabilistic visual categorization task. We found that bilateral NAcc responds to advance information as a predictive cue for subsequent reward, with a prediction-error response during feedback. Left lateral VS, extending into the ventral putamen, selectively signals information expectation, without modulation by feedback. This suggests a dissociation between information and reward, and a VS role in information processing beyond reward-based reinforcement. Crucially, our results disentangle the value of information regarding upcoming rewards from expected information for predictions about the world, independent of subsequent reward reinforcement or feedback.

Results

Before the fMRI study, subjects learned to categorize stimuli where two features (claw or eye) were differentially predictive (85% or 60%) of the category (Fig. 1 *A* and *B*). Our event-related fMRI design (Fig. 1*C*) separated events into: 1) A blurred feature cue that predicted high (HI) or low (LO) information; 2) the revealing of the actual feature, categorization, and anticipation of response-contingent feedback; and 3) receipt of positive or negative feedback. This allowed us to differentiate the neural substrates underlying information expectation (InfoExpect), information revelation and categorization outcome anticipation (RevealAnticip), and rewarding (positive or negative) feedback.

Behavioral Results. All subjects achieved the stringent performance criterion during pre-fMRI behavioral training (*Materials and Methods*), choosing the optimal category in 98% of the last 200 training trials. All subjects correctly ranked the 85% feature as more informative than the 60% feature and correctly identified the more likely species, given each individual feature (*Materials and Methods*).

Subjects chose the more-probable category (the optimal response) within the allowed 1,500 ms in the vast majority (94.4%) of fMRI trials. In 4.0% of trials, responses were not on time. Only 1.6% of trials involved on-time selection of the less-probable category, given the presented stimulus.

fMRI Activations. During InfoExpect, subjects anticipated information of greater or lesser usefulness. Although it was already possible to anticipate the eventual level of classification accuracy, either 85% or 60%, the categories were equally likely during this stage and the most likely category could not be decided until information in the form of the specific feature would be revealed. Once the feature was revealed, subjects could categorize the stimulus as A or B, and started anticipating response-contingent feedback. The Feedback stage consisted of a smiley face emoticon for a correct classification, or a frowny face emoticon for an incorrect classification. fMRI results show only optimal response trials; suboptimal trials were modeled using a nuisance regressor.

Fig. 2 shows group BOLD activations ($P < 0.005$, $k = 25$) from the whole-brain general linear modeling (GLM) analysis for each trial phase, contrasting 1) HI versus LO InfoExpect (InfoExpect_HI-LO); 2) revelation of the HI versus LO Info feature and response-contingent feedback anticipation (RevealAnticip_HI-LO); and 3) positive versus negative feedback (Feedback_

POS-NEG). During InfoExpect (Fig. 2*A*), the bilateral VS (including the NAcc) BOLD responses were significantly stronger to InfoExpect_HI than InfoExpect_LO. Additional regions revealed by the InfoExpect_HI-LO contrast included the left posterior putamen and regions not typically considered dopaminergic reward areas: The bilateral superior temporal sulcus (STS), anterior left superior frontal gyrus (SFG), left superior frontal sulcus (SFS)/SFG, left posterior parahippocampal gyrus, right cerebellum and cerebellar vermis, and right superior occipital gyrus (*SI Appendix, Table S1*).

In contrast, RevealAnticip_HI-LO activated a distinct set of brain regions from InfoExpect (Fig. 2*B*). RevealAnticip_HI activated the bilateral medial prefrontal cortex (MPFC)/rostral anterior cingulate cortex (rACC), bilateral posterior cingulate gyrus, left middle frontal gyrus/SFS, left posterior insula, left caudate, right cerebellar vermis, and right STS significantly more than RevealAnticip_LO (*SI Appendix, Table S2*). Importantly, the VS showed no modulation by RevealAnticip_HI-LO. Lowering the statistical threshold to a lenient $\alpha = 0.01$, uncorrected, revealed seven voxels in the left NAcc, extending laterally into the left putamen, which did not survive cluster correction. No modulation by RevealAnticip_HI-LO was observed in the right NAcc even at an uncorrected $\alpha = 0.05$ threshold.

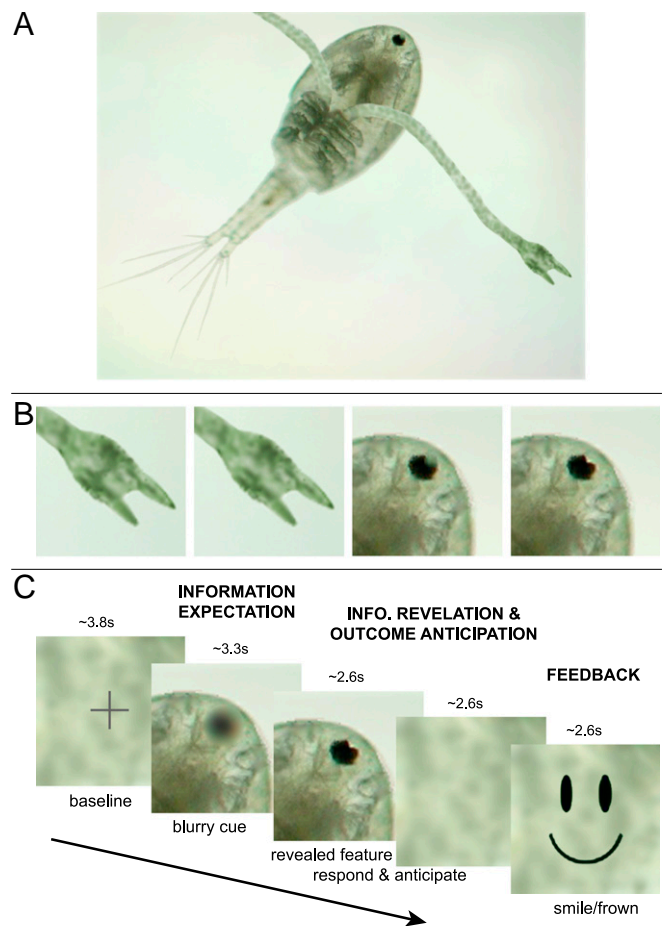


Fig. 1. Stimuli and trial design. (*A*) Example stimulus to be categorized as species A or B during behavioral training. (*B*) Two versions of each feature were used: claw1 = connected, claw2 = open; eye1 = dark, eye2 = dotted. Each image displayed one of the four possible claw-eye combinations, with different probabilities of category A or B. (*C*) An example event-related fMRI trial. Event durations were jittered between 1 and 7 TRs (1.5 to 10.5 s); average durations indicated in seconds.

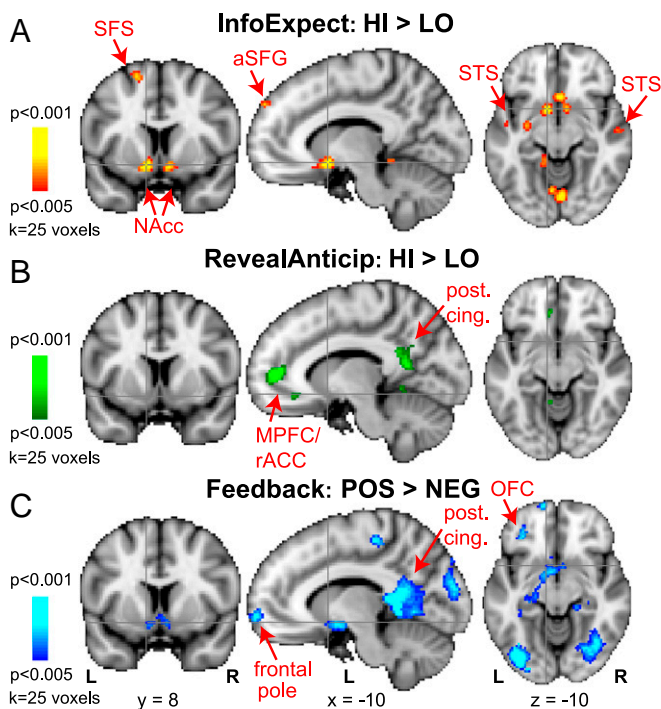


Fig. 2. Group BOLD activation contrasts during the three different stages of each trial: (A) InfoExpect, (B) RevealAnticip, (C) Feedback. (A) Red-to-yellow: Brain areas more active for HI than LO InfoExpect, when subjects view the blurry feature cue but cannot yet classify (species A versus B). (B) Green-to-light-green: Areas more active for Revelation of HI than LO Info features and concomitant response-contingent outcome anticipation when the species is categorized. (C) Blue-to-light-blue: Areas more active for positive than negative feedback (smile versus frown). Abbreviations: a, anterior; L, left; NEG, negative feedback (frown emoticon); POS, positive feedback (smile emoticon); post. cing., posterior cingulate; R, right. Cross-hairs are centered on left NAcc, to show lack of NAcc modulation during RevealAnticip. Colors represent different trial stages, not positive/negative BOLD signal.

Fig. 2C shows significantly higher activations following receipt of a smile versus frown (Feedback_POS-NEG). This contrast revealed the bilateral NAcc together with several regions previously implicated in reward processing, including the bilateral posterior cingulate gyrus, left orbitofrontal cortex (OFC), left frontal pole, right MPFC, right posterior insula, as well as visual and memory-related regions (right superior occipital gyrus, bilateral fusiform gyrus, bilateral hippocampus) (*SI Appendix, Table S3*).

The partially overlapping yet separate brain activations seen during InfoExpect, RevealAnticip, and Feedback suggest that expectation of information for categorization, and expectation and receipt of reward following categorization may be represented by distinct neural substrates. In particular, the NAcc appears to be involved in both expectation of useful information (InfoExpect_HI) and receipt of rewarding feedback (Feedback_POS), but not in expectation of response-contingent outcome/reward (RevealAnticip).

The MPFC, posterior cingulate, OFC, and other reward-related regions participate in reward anticipation and outcome processing, but not in information expectation, revealing distinct neural substrates for information expectation and reward processing. *SI Appendix, Fig. S4* shows additional percent signal change plots from these areas.

Importantly, BOLD activation differences between InfoExpect_HI-LO could be due either to differential deactivation relative to baseline, or to more positive (above-baseline) BOLD responses.

We used a region-of-interest (ROI) analysis (*Materials and Methods*) in the left and right VS to calculate percent signal change during each trial phase. Voxels that differed significantly from baseline during either InfoExpect_HI or InfoExpect_LO were selected; this revealed only positive BOLD activations in the left lateral VS (Fig. 3, red), specifically for InfoExpect_HI, and only negative BOLD activations in the right VS/NAcc (Fig. 3, yellow, extending bilaterally into both NAcc nuclei), specifically for InfoExpect_LO. Fig. 3B, *Left* shows that the left lateral VS was more active during InfoExpect_HI both versus InfoExpect_LO and versus baseline; that is, the difference between InfoExpect_HI-LO is due to positive BOLD activations for InfoExpect_HI [paired two-tailed t test InfoExpect_HI versus InfoExpect_LO: $t(9) = 4.14$, $P = 0.003$]. (Note that ROI selection was orthogonal to the HI versus LO t test by selecting voxels versus baseline rather than only voxels significant for HI versus LO, thereby avoiding biasing P values with multiple testing.) Positive BOLD activations during InfoExpect_HI extended laterally from the left NAcc to ventral putamen.

In contrast, the right VS/NAcc (Fig. 3B, *Right*) was inhibited or deactivated versus baseline, with significantly more negative BOLD responses to InfoExpect_LO than to InfoExpect_HI [paired two-tailed t test, $t(9) = 2.53$, $P = 0.03$]. This deactivation to InfoExpect_LO was centered on the right NAcc but extended medially toward the contralateral (left) NAcc (i.e., bilaterally). In other words, the left lateral VS is positively activated by expectation of a more informative stimulus, whereas the right VS/NAcc is modulated during the InfoExpect stage via greater deactivation versus baseline for InfoExpect_LO than for InfoExpect_HI (see also *SI Appendix, Figs. S1–S3*).

Fig. 3B, *Right* also shows that the right VS/NAcc displayed a similar suppression pattern during feedback (reward receipt). Right VS/NAcc activation was not significantly different from baseline during Feedback_POS [paired two-tailed t test $t(9) = -0.28$, not significant], but was significantly suppressed or deactivated during Feedback_NEG, similar to when the less informative cue (InfoExpect_LO) was received [paired two-tailed t test Feedback_POS versus Feedback_NEG, $t(9) = -6.98$, $P < 0.0001$]. This is consistent with neuronal firing results in primates, where negative prediction error (reward expected, but not received) leads to suppressed neuronal firing (15, 16).

Other than the left lateral VS, the left posterior putamen, cerebellar vermis, and bilateral STS also showed both greater positive (above-baseline) BOLD responses to InfoExpect_HI and were significantly more active for InfoExpect_HI than for InfoExpect_LO (*SI Appendix, Fig. S4* and *Table S1*). *SI Appendix, Fig. S5* shows reverse contrasts (LO-HI) for InfoExpect, RevealAnticip, and Feedback_NEG-POS. No brain region was significantly more active for InfoExpect_LO-HI.

The above results show that VS subregions respond differentially to information expectation versus reward receipt. The right VS/NAcc appears to be treating expectation of information as a predictor for trial-by-trial reward. In contrast, the left lateral VS is significantly modulated only during the InfoExpect stage, and does not show an RPE-type modulation in the Feedback stage.

To further explore the activations in the left lateral VS and right VS/NAcc, we calculated theoretical response profiles across the trial stages, for each experimental condition, for hypothetical expected information gain (IG) and RPE models (Fig. 3C and D and *SI Appendix, Supplementary Results*). The stages of the trial were split according to information level or type of feedback and included InfoExpect_HI, InfoExpect_LO, RevealAnticip_HI, RevealAnticip_LO, Feedback_POS, and Feedback_NEG.

We correlated group-average BOLD percent (BOLD%) signal change in each trial stage, in the left lateral VS and right VS/NAcc (Fig. 3B), with predicted responses of both the theoretical IG and RPE models (Fig. 3C, *II*, Fig. 3D, and *SI Appendix, Supplementary Results*). We used both the group-averaged BOLD% signal change

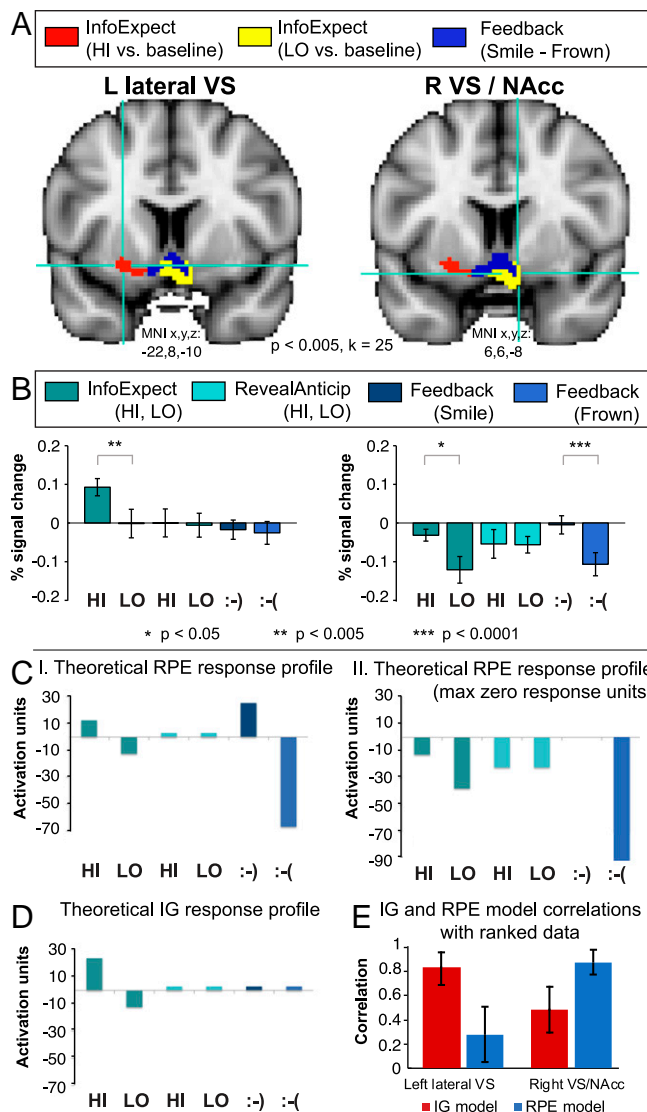


Fig. 3. fMRI BOLD results and percent signal change from the left and right VS and model predictions for hypothetical IG versus RPE responses. (A) Group BOLD activations ($P < 0.005$, whole-brain cluster threshold = 25 voxels) in left and right VS. Red: Left VS ROI selected based on InfoExpect_HI versus baseline. This contrast revealed only positive BOLD activations, all in the left lateral VS, extending laterally from the left NAcc to the left ventral putamen (left crosshairs, centered on MNI $x = -22, y = 8, z = -10$). Yellow: right VS ROI selected based on InfoExpect_LO versus baseline. This contrast revealed only negative BOLD activations (yellow) in the right VS, centered on the right NAcc, extending medially toward the left NAcc (right crosshairs, centered on MNI $x = 6, y = 6, z = -8$). Blue: Significant activations for Feedback_POS > NEG (smile > frown), shown superimposed. These overlapped the bilateral NAcc, but not the left lateral VS. (B) Percent signal change plots for the left and right VS ROIs (Left and Right, respectively) during each trial stage (see legend). Error bars represent SEM. [“:-”) represents “smile” and “;-)” represents “frown”]. (C, I) shows the response profile of a hypothetical brain area that responds exactly in accord with the RPE model. For calculations used to derive this response profile see *SI Appendix*. (C, II) depicts the response profile of a hypothetical RPE brain area, shifted so that the maximum level of activation is zero. (D) The response profile of a hypothetical brain area that responds exactly in accord with the IG model. For calculations used to derive this response profile, see *SI Appendix*. (E) The correlation between the hypothetical RPE and IG models and the left lateral VS and right VS/NAcc ROIs, respectively. Error bars represent SEM, obtained via bootstrap sampling.

data, and within-subject-ranked data (i.e., by ranking the magnitude of the percent signal change for each subject across the six trial stages) in these analyses. In the group data, the left lateral VS's Pearson correlation with the IG model was $r = 0.845$, and with the RPE model was $r = 0.386$. In the right VS/NAcc, the data's correlation with the RPE model was $r = 0.780$, and with the IG model was $r = 0.557$. In the ranked data, the left lateral VS's Pearson correlation with the IG model was $r = 0.830$, and with the RPE model was $r = 0.280$. In the right VS/NAcc, the average correlation with the RPE model was $r = 0.877$, and with the IG model was $r = 0.484$ (Fig. 3E). Thus, both the group-average data correlations and the individual-subject ranked data correlations suggest that the IG model better explains the left lateral VS, and the RPE model better explains the right VS/NAcc.

Given variability in BOLD magnitude and correlation strength in individual subjects, we also conducted further analyses of the single-subject ranked data in relation to the theoretical IG and RPE models (*SI Appendix, Supplementary Results*). To check which theoretical model better explained an individual subject in a particular ROI, we took that subject's data's correlation to the IG model minus the correlation to the RPE model. In the left lateral VS, 7 of 10 individual subjects showed higher correlation with the IG model than with the RPE model (95% CI for correlation difference, -0.029 to 0.533). (The reported 95% CI was obtained by bootstrap sampling, with 1 million samples with simple bootstrap, taking individual subjects' IG-RPE difference scores as inputs.) In the right VS/NAcc, 8 of 10 individual subjects showed higher correlation with the RPE model than with the IG model (95% CI for correlation difference -0.350 to -0.057 , by bootstrap sampling). We also used bootstrap sampling of the whole dataset to further explore which theoretical model better accounts for each area, again measuring the difference between the data's correlation with the IG and RPE models (*SI Appendix, Supplementary Results*). In the left lateral VS, in 95.83% of 1 million bootstrap samples of the dataset, the IG model better explained the data than the RPE model (*SI Appendix, Fig. S6, Left*). In the right VS/NAcc, in 99.83% of the 1 million bootstrap samples, the RPE model explained the data better than the IG model (*SI Appendix, Fig. S6, Center*).

To address whether there could be differential sensitivity to information versus reward between the two ROIs, irrespective of whether a particular ROI correlates more highly with the IG or RPE model, we conducted further analyses of the ranked data (*SI Appendix, Supplementary Results*). We computed the difference (as described above) in each subject's data's correlation to the IG model, minus the correlation to the RPE model, in each ROI. We then took the left lateral VS difference score and subtracted the right VS/NAcc difference score from it, thus giving a differential information sensitivity score. In 9 of 10 individual subjects, this differential information sensitivity score was positive [95% CI for differential information sensitivity score 0.198 to 0.685 , by bootstrap sampling; two-tailed paired t test $t(9) = 2.90, P = 0.0176$]. We also conducted this analysis in bootstrap sampling of the whole dataset (*SI Appendix, Supplementary Results*). In 99.81% of 1 million bootstrap samples, the differential information sensitivity score was positive (*SI Appendix, Fig. S6, Right*).

In summary, our results suggest that different neural mechanisms process expectation of information, expectation of reward (outcome), and receipt of reward. The left lateral VS signals the informativeness of a cue but does not participate in reward (feedback) processing, whereas the NAcc signals both the information value of a cue and RPEs.

Discussion

Using a probabilistic visual categorization task and whole-brain event-related fMRI, we identified distinct temporal and spatial processing signatures of regions subserving information expectation,

information revelation and response-contingent outcome anticipation, and feedback processing.

The InfoExpect stage predicts HI or LO information and expected classification accuracy, but provides no information about the stimulus's category. Categorization can only occur upon revelation of the specific feature. Following categorization, response-contingent positive or negative feedback can be anticipated (RevealAnticip stage). The Feedback stage then allows for prediction error calculation. Statistical OED models state that the value of expected information can be calculated before the specific feature is revealed. Importantly, the left lateral VS was modulated by the value of expected information but not by reward receipt, consistent with key predictions of the OED models.

Average classification accuracy, and hence the probability of positive or negative feedback, can be anticipated, and do not change, during the InfoExpect (blurry feature) and RevealAnticip (feature revelation and categorization) stages. From the perspective of OED models, however, the former stage represents curiosity about a category (or query, or experiment, or test), whereas the latter stage represents curiosity satisfied, since the feature that predicts which category will be most likely has now been revealed. Since the expected performance accuracy did not change between the InfoExpect and RevealAnticip stages, whereas brain activations did, this suggests that it is not simply the anticipated response accuracy, but curiosity about the stimulus's category modulating brain activations to these two stages. This would be akin to two different medical tests being used to diagnose a disease as A or B, with one test having a higher accuracy than the other. It is not just anticipation of the particular accuracy prior to executing the medical test, but also curiosity about the actual diagnosis, that matters. The differential responses in the left lateral VS and right VS/NAcc support this distinction between the value of expectation of information and outcome of a query, as predicted by OED models.

The right VS/NAcc (extending bilaterally) indicated sensitivity to different levels of expected information by deactivating versus baseline for LO compared to HI Info expectation. Moreover, the NAcc responded to subsequent feedback with a prediction error-like response, showing deactivation for negative feedback, and no modulation versus baseline for positive feedback. This is consistent with subjects expecting mostly positive feedback across trials, since HI and LO features led to ~85% and 60% correct categorizations, respectively. Although the relationship between negative BOLD responses and neuronal inhibition is not fully established, BOLD deactivations correlate with decreases in multiunit activity, whereas positive BOLD responses correlate with an increase in neuronal firing and postsynaptic potentials (39). The similarity of negative BOLD during negative feedback to macaque neuronal firing suppression for negative prediction errors supports an inhibition interpretation of the right VS/NAcc negative BOLD response. In contrast, the left lateral VS showed greater positive BOLD responses during HI versus LO InfoExpect and showed no prediction-error response (no modulation) during feedback, selectively signaling higher information expectation with increased BOLD activity versus baseline. This suggests a functional difference between the left lateral VS and NAcc. Specifically, the left lateral VS is involved in information expectation, distinct from an RPE mechanism.

Subject payment was independent of classification performance. Subjects consistently chose the optimal category, for both LO and HI Info features. Thus, anticipation during the InfoExpect stage is not related to expectation of greater monetary gain. However, the blurry cue did already predict higher or lower accuracy. Could left lateral VS activation during InfoExpect be due to anticipation of classification accuracy or positive feedback? The left lateral VS was only modulated during InfoExpect, but not during RevealAnticip or Feedback. If anticipatory activity in the left lateral VS during InfoExpect were

simply predictive of more correct performance, positive or negative feedback should modulate its activity, as in the right VS/NAcc. No such prediction error-type response was observed in the left lateral VS. This suggests that the left lateral VS is excited by, or "values," information for its own sake. The left lateral VS ignored noisy probabilistic trial-by-trial feedback while maintaining optimal categorization performance, consistent with model-based learning (10) and with macaque basal ganglia representations of long-term object-value independent of immediate reward outcome (40). Ultimately, valuing information could be adaptive for overall survival fitness, regardless of noisy probabilistic feedback (1). Crucially, our results suggest that distinct neural substrates process expected information value versus reinforcement feedback, and that information expectation does not fully overlap with general reward expectation.

Nonhuman Neurophysiology and Connectivity. Our results are also consistent with macaque neurophysiology in related brain areas. Monkeys strongly prefer advance knowledge in a task where cues predict different amounts of information about upcoming water rewards, despite advance information not affecting the reward amount (28). However, recordings in lateral habenula neurons show decreased firing to cues that predicted more information, similar to their response to large versus small rewards (28). Habenula neurons inhibit dopamine neurons, responding to negative events while being inhibited by positive events (41).

Macaque dopaminergic midbrain neurons show enhanced firing to advance information about food and drink rewards, and RPE-type responses upon reward receipt (27), similar to our right VS/NAcc results. Ventral striatal BOLD activations correlate with excitatory dopaminergic input from midbrain nuclei (22). Our VS activations are consistent with dopaminergic inputs modulated by expectation of advance information. Importantly, the left lateral VS modulation by InfoExpect but not by Feedback in our study suggests a different functional role compared to the midbrain (27), lateral habenula (28), and orbitofrontal neurons (29), which respond both to cues signaling advance information about food or water rewards, and to unexpected rewards or the absence of rewards. Our study identifies an information-specific response independent of trial-by-trial reward feedback.

Anatomically, the VS is heterogeneous, with differences between subregions such as the core and shell (12), and different activation patterns depending on which limbic and cortical inputs are integrated in different contexts (5). Our medial-lateral subdivision of the VS is in line with recent macaque *ex vivo* high-resolution diffusion tensor MRI findings (42), which demonstrated that the VS contains medial, lateral, ventral, and dorsal subdivisions. Whereas the medial tractography-based subdivision corresponds well to the histological NAcc, the lateral tractography-based VS region corresponds to the histological "neurochemically unique domains of the accumbens and putamen (NUDAPs)," which show a unique pattern of μ - and κ -opioid as well as D1-like dopamine receptors (42). This tractography-based medial-lateral distinction in the macaque matches our fMRI activations, suggesting a lateral VS–NAcc subdivision in humans. Moreover, the finding that the lateral "NUDAPs-like" VS region contains a unique signature of opioid and dopamine D1-like receptors, which in our study was activated by the anticipation of information but not by simple rewards, such as positive trial-by-trial reinforcement, suggests that information may indeed function as a different type of reward that is processed by a specific combination of dopamine and opioid receptors different from those found in other parts of the reward system.

The tractography-based subdivisions between lateral and medial parts of the VS could be due to different topographical connections from the VTA (43). In the mouse midbrain this area has been found to include functionally diverse clusters of spatially organized dopaminergic neurons representing sensory,

motor, and cognitive variables, as well as RPEs based on multiple factors, such as trial difficulty and previous trial outcomes. Our NAcc results may thus reflect a mixed RPE signal, while the left lateral VS may represent a different mixture of perhaps more cognitive and sensory signals.

Probabilistic Categorization: Accuracy and Internal Feedback Signals. NAcc involvement in our task is consistent with neuropsychological and neuroimaging findings that similar neuronal mechanisms are involved in reward prediction and in incremental, feedback-based learning of probabilistic categories (44).

Although basal ganglia are typically involved in habitual stimulus–response associations or reward-driven habits, the VS shows both model-free and model-based responses (10, 14, 20, 26, 44). The VS is more activated by correct than incorrect performance in n-back working memory tasks, even without external feedback or reward, and even in trials without motor responses (45). This is consistent with our InfoExpect activations, since HI Info features are more likely to lead to correct categorization.

A VS response independent of external feedback was also observed in an observational learning task where subjects learned to categorize stimuli into two categories by observation alone, with the category label presented before each stimulus (8). The VS showed greater activity for correct than incorrect categorization trials, even without external feedback. Moreover, the VS and putamen were more activated for error trials in which subjects indicated greater confidence than usual, suggesting sensitivity to internally generated confidence signals. Our event-related fMRI study disentangles probabilistic information expectation from categorization and feedback. Our results extend these findings and dissociate between prediction error signals in the NAcc, and information expectation independent of feedback in the left lateral VS.

Instrumental Versus Noninstrumental Information Seeking. Is expected information in our study a type of instrumental information (i.e., information that shapes learning by guiding future actions)? Whereas instrumental information is the reduction of uncertainty regarding which action to take next (e.g., ref. 46), noninstrumental information seeking (e.g., refs. 47–50) involves wanting to find out about an outcome (e.g., whether a reward will be obtained or not), even if there are no more actions to take that could change the outcome. Curiosity-based information-seeking (9) is also a form of noninstrumental information seeking. Trivia questions about which participants are more curious lead to greater activity in the VS, SN, VTA, and cerebellum (9). This increase in activity occurs during anticipation of interesting information (i.e., during anticipation of answers), but not during presentation of answers, at which point curiosity is satisfied. This is consistent with our stronger VS and cerebellar activation for more informative cues during the InfoExpect than during the RevealAnticip stage.

OED models attempt to quantify the value of expected information independent of outcomes (30, 31, 34, 37), which would seem noninstrumental. However, OED models quantify the expected usefulness of a query or test (e.g., a medical test) before that test is carried out, with the implication being that one could then choose between different tests by calculating each test's a priori expected informational value. This assessment of the expected usefulness of possible future tests, and the quantification of the expectation of such information could also be conceived as instrumental, because it would lead to the choice of one test over another, or in our case one category over the other. Given that the outcome (feedback) did not change our subjects' behavior on the next trial (since our subjects followed optimal choice strategies and ignored trial-by-trial feedback, as did the left lateral VS), one could argue that the valuation of information in our experiment is independent of outcomes and as such is noninstrumental. We would argue that gathering

information is ultimately adaptive and as such rewarding, and that this is a case of long-term instrumental information seeking that is not necessarily reflected in immediately following actions. Future work could address such subtle distinctions between these different types of information seeking.

Saliency Versus Valence. Our results extend the “incentive saliency” view of the VS (4, 6). The VS is equally activated by anticipation of uncertain gains, uncertain losses, and certain gains, with lower activations during certain losses; thus, neither valence nor saliency alone explain the VS (6).

Dopaminergic midbrain (VTA/SN) regions have been found to track valenced (positive or negative) information prediction errors (IPEs), for example, when probabilistic information about a possible monetary win or loss is promised but not delivered (51). In contrast, in that study NAcc only tracked traditional RPEs, consistent with our NAcc results. (Note that our study did not involve IPEs: If, for example, a blurry 85% cue was presented in the initial trial stage, then an 85% information cue always followed.) Interestingly, no neural representation of IPEs independent of valence were identified (51), suggesting that perhaps the brain tracks simpler information-related representations rather than an IPE per se. The information expectation signal we obtained in the left lateral VS is consistent with this interpretation.

Parietofrontal Networks. Our BOLD activation contrast InfoExpect_HI-LO did not show posterior parietal activations typical of attention networks (52–55) or motor preparation (56–58). This is consistent with all stimuli being presented foveally, with no need for spatial target selection. Attention has also been found to bias striatal RPEs during learning, and to correlate with a frontoparietal network during attentional switches between different valuable stimuli (59). However, in our task, learning occurred well prior to fMRI scanning, with subject behavior at ceiling, and only one stimulus was presented at a time, rather than requiring choosing between multiple stimuli (59). Moreover, greater sensory uncertainty stimuli (more difficult, equivalent to the LO Info feature) activate parieto-frontal networks more strongly than easier stimuli (the equivalent of the HI Info feature) (58). In contrast, we found no attention network or posterior parietal modulation by either LO or HI InfoExpect. Moreover, there was no VS modulation during RevealAnticip, when attentional demands should increase, given the need to categorize the stimulus at that point. Attention is thus unlikely to explain our results. There are many different kinds of saliency, potentially represented by different striatum subregions (6), consistent with the functional heterogeneity we report. Moreover, there is evidence that attention for value is distinct from attention for action and attention for learning (53).

Two recent macaque neurophysiology studies investigated the lateral intraparietal (LIP) area's role in representing expectation of instrumental information (information that guides subsequent actions) versus reward (46, 54). In the first study (54), monkeys made two saccades on each trial: The first saccade was to a cue located either outside or inside the neuronal receptive field, upon which a field of dots would start moving either up or down, indicating whether the upper or lower saccade target was correct; the second saccade was to the target indicated by the cue's dot motion. A colored border around the cue indicated its validity, namely a probability of 100%, 80%, or 55% that the cue's up or down motion correctly indicated the correct (rewarded) direction of the final saccade. The second study (46) used a similar task, except that the cue to which the first saccade was made was either informative, correctly indicating the saccade target, or uninformative, with the correct saccade target being known in advance. In addition, reward size was manipulated to be either large or small.

Both studies (46, 54) found that prior to the first saccade, LIP neurons encode the expected instrumental information (information gain) that the saccade to the first cue is expected to bring for the following action: That is, how useful the first saccade is expected to be in reducing uncertainty about which next action (saccade up or down) would be rewarded. Importantly, neuronal information gain sensitivity was unrelated to neuronal reward sensitivity; moreover, many LIP cells showed enhanced firing for smaller rather than larger rewards. These results show that in situations demanding active sampling via eye movements, area LIP is involved in representing the expectation of instrumental information, and that this expected information gain can indeed be separated from reward per se. This is consistent with our study, in which information-gain sensitivity was also dissociable from reward sensitivity, in our case in the VS. In our study, we did not use an oculomotor decision-making task, and there was no active sampling; rather, all targets and informational cues were presented foveally. Thus, while the LIP may be part of a priority map implementing the required cognitive effort in an active information-sampling context (46), our lack of active sampling may explain the absence of LIP and parietal activations, and suggests that perhaps for situations where information is presented foveally and simply needs “digested” passively (rather than actively sought out via sequences of actions), different neural substrates may represent the expectation of useful information.

Other Reward-Related Regions. In addition to the VS, other reward circuit regions (3, 11) were activated across different trial stages, including the MPFC, OFC, and posterior cingulate cortex (PCC) (Fig. 2 and *SI Appendix, Fig. S4*). The MPFC and PCC were modulated by feature informativeness during RevealAnticip and by Feedback, but not during InfoExpect, in contrast to the left lateral VS and right VS/NAcc. The OFC showed modulation by Feedback POS > NEG, similar to the right VS/NAcc, but was not involved in InfoExpect or RevealAnticip (*SI Appendix, Fig. S4*). This temporal dissociation of different reward circuits is consistent with different processes being active during information expectation than during response-contingent outcome anticipation, suggesting that the left lateral VS is not simply anticipating positive feedback during the InfoExpect stage. The ACC/MPFC, which in our study were activated during the RevealAnticip stage but not during the InfoExpect stage, have previously been correlated with reward anticipation (60). This supports our interpretation of (response-contingent) reward anticipation occurring during the RevealAnticip stage but not during the InfoExpect stage, even though the general accuracy could be anticipated during both of those stages.

Alternative Explanations. A study of facial attractiveness found nonlinearities in the responses of the NAcc, lateral OFC, VTA, and other parts of the reward circuit (61). Could our results be explained by nonlinearities in reward processing? We used an event-related task design to separate different stages of processing; due to the long trials that this required, we used two levels of information expectation (HI and LO) and two levels of reward (smile or frown). Nonlinearities in the reward or information expectation response could thus be hard to uncover in our dataset. Future experiments could employ a range of information and reward levels to examine whether there are nonlinearities in information expectation or reward coding.

Could temporal discounting explain our results? Although we did not compensate subjects according to their performance, subjects could have valued overall experimental performance more greatly than trial-by-trial feedback. However, the temporal discounting reported in animals and humans usually involves devaluing (discounting) larger but temporally more distant rewards, relative to smaller immediate rewards (62). If temporal

discounting were to apply in our data, then immediate smile or frown feedback on a trial should have led to greater activation (or deactivation) than the possibility of better overall performance throughout the whole experiment. This is not what we found in the left lateral VS, where the presentation of the probabilistic blurry information expectation cue that informed optimal overall performance throughout the experiment led to higher activation than trial-by-trial smile or frown feedback. Thus, a generalized valuation of information is a more plausible explanation of our left lateral VS results than temporal discounting.

Summary. Using an event-related fMRI task we dissociated information expectation, information receipt, and response-contingent outcome anticipation, and feedback processing, during probabilistic categorization. Lateral aspects of the left VS signaled information expectation independent of reward feedback, whereas the NAcc was modulated during both InfoExpect and Feedback stages. This suggests dissociation between information and immediate reward in different VS subregions, and that probabilistic information for classification is different from rewards such as food, money, or positive feedback.

Materials and Methods

Participants. Ten right-handed participants with normal or corrected-to-normal vision (6 males; ages 22 to 33 y; mean 25.7 y) gave informed consent and were paid for their time. The study protocol was approved by the University of California, San Diego Institutional Review Board.

Behavioral Prescanning Task and Stimuli. Subjects implicitly learned the relative usefulness of two different features via an experience-based multiple cue probabilistic classification task (35, 63). The initial behavioral training session took place 1 to 4 d before the fMRI experiment in 9 of 10 subjects, with refresh training immediately before scanning, to ensure effective learning of the probabilistic task (*SI Appendix, Supplementary Methods*).

During behavioral training subjects classified images of plankton-like stimuli (Fig. 1A) as either “species A” or “species B” by pressing one of two buttons with their right hand. The category (A or B) of each stimulus depended in a probabilistic way on two binary features (eye and claw), which could each take one of two forms (eye1, eye2, and claw1, claw2) (Fig. 1B). In each trial, a specimen was chosen at random, with 50/50 probability for each category. Immediate feedback on the true species was then provided via a smile or frown emoticon. Note that feedback depended on whether the correct species was chosen in a given trial, not on whether the optimal classification choice was made. Due to intrinsic uncertainty in the probabilistic task environment, the maximal possible classification accuracy was 85%. Thus, choosing the optimal (most probable) species, given the stimulus provided, led to negative feedback in 15% of trials on average. Learning was self-paced, with no time constraints.

Subjects could achieve 85% classification accuracy using the more informative feature and 60% using the less informative feature alone, for either A or B (*SI Appendix, Supplementary Methods*). Which feature (eye or claw) was most informative, and which feature version (e.g., eye1 or eye2) tended to predict species A, was counterbalanced across subjects.

The number of training trials was adapted automatically by the software according to each subject’s performance. Training continued until the subject achieved a stringent learning criterion, namely 1) choosing the most-probable category given the presented stimulus in 98% of the last 200 trials, and 2) choosing the most-probable category in all five of the most recent presentations of each of the four stimulus configurations. This learning criterion (35, 63) was chosen to ensure that participants had good implicit understanding of the probabilistic task (*SI Appendix, Supplementary Methods*). In addition, participants’ explicit knowledge of the features was tested. Participants were asked to indicate, for each version of each feature (version 1 or 2 of the claw and the eye), whether that feature (if viewed alone) would tend to indicate A or B. Participants were also asked which feature, if viewed alone, would be more useful on average. All subjects correctly ranked the 85% feature as more informative than the 60% feature and correctly identified the more likely species, given each individual feature value. This is consistent with previous behavioral results (63), suggesting that our experience-based training method can meaningfully convey environmental statistics on probabilistic classification tasks.

fMRI Task and Stimuli. Following behavioral refresh training, subjects participated in an event-related fMRI task. Subjects were asked to categorize stimuli based on one foveally-presented feature in each trial (eye or claw). Each trial (Fig. 1C) started with a variable-duration fixation cross (baseline), followed by presentation of a blurred eye or claw image. The blurry feature image revealed no information whatsoever about the true feature version that followed (i.e., the exact same blurry eye image was used, irrespective of whether eye1 or eye2 would subsequently be presented). This InfoExpect stage thus allowed subjects to expect receiving either more useful (HI Info) or less useful information (LO Info) for categorization in the next phase. The best classification choice (species A or B) was maximally uncertain before the specific feature was revealed. Furthermore, species A and B were each the best choice for classification in half of the trials. This ensured that prior to seeing the revealed feature subjects could not anticipate a species A or B response, preventing specific motor preparation. Although the specific classification decision could not be anticipated in this stage, the anticipated level of classification accuracy could be anticipated (i.e., 85% or 60%, in the case of the HI or LO information-predicting blurry cue, respectively).

The specific feature version was then revealed, allowing subjects to choose category A or B, and to then immediately start anticipating positive or negative choice-contingent feedback depending on the probability of the most likely category. Since revelation of a specific feature immediately triggers response planning and response-contingent outcome anticipation, this stage was termed the RevealAnticip stage. Subjects then received feedback, as described above (Feedback stage). Subjects were required to respond within 1,500 ms of the specific feature version being revealed, using a button box (right index or middle finger, counterbalanced across subjects). To ensure subjects were able to respond within the allotted time, subjects practiced the fMRI task for at least one fMRI-length run before scanning. Subjects preferentially chose the more-likely species for each feature in the fMRI practice runs, for HI Info and LO Info features alike.

The informativeness of each blurry feature cue (which predicts HI or LO Info) can be quantified using OED (34, 38) or heuristic (64) models. Expected probability gain (classification error minimization) emerged as the most psychologically plausible model in a related behavioral task (35). Expected probability gain of the HI Info cue is 0.35 and of the LO Info cue is 0.10 (SI Appendix, Supplemental Methods). We designed the probabilities in our experiment so that other models (38) would also agree with the ranking of the HI and LO cue (e.g., expected information gain of HI cue = 0.39 bits, LO cue = 0.03 bits; expected impact of HI cue = 0.70, LO cue = 0.20).

The fMRI experiment was programmed in PsychToolbox 3.0.8, and run on a Windows XP Pro SP3 Dell laptop. Stimuli subtended a square $\sim 3 \times 3^\circ$ of visual angle and were projected on a screen attached to the scanner bore, viewed via a mirror attached to the headcoil.

Each functional run lasted 480 s (320 repetition times [TRs]; 1 TR = 1,500 ms) and contained 32 trials. Each event's duration was jittered according to a geometric distribution, (minimum = 1 TR, truncated to maximum = 7 TRs per event). The geometric distribution provides for maximal uncertainty about event duration and allows for separation of the hemodynamic response to each event (58). To prevent visual anticipation of when the revealed feature would disappear, the presentation duration of the revealed feature was also jittered (Fig. 1). The mean event durations were: baseline, 3.8 s; InfoExpect, 3.3 s; RevealAnticip, 2.6 s each (these two events were combined into one period, as explained above); Feedback, 2.6 s. Each run's sequence of stimuli was unique for each subject.

fMRI Data Acquisition. Imaging was performed at the Center for fMRI, University of California, San Diego (3T GE scanner; 8-channel head coil; functional imaging parameters: echoplanar T2* gradient echo pulse sequence, 23 contiguous axial slices, interleaved, bottom-up order, $3.44 \times 3.44 \times 5$ -mm voxel size, 64×64 matrix, echo time [TE] = 31.1 ms, flip angle [FA] = 70° , TR = 1,500 ms, 320 volumes/functional run, bandwidth = 62.5 kHz; six dummy volumes; structural imaging parameters: T1-weighted MPRAGE, $1 \times 1 \times 1$ -mm voxel size, 256×256 matrix). Most subjects (7 of 10) completed five functional runs; two subjects completed four functional runs; due to technical difficulties, one subject completed only two functional runs.

fMRI Data Analysis. fMRI analyses were carried out using the FMRIB Software Library (FSL 4.1; <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/index.html>). Standard

fMRI preprocessing was performed (brain extraction, motion correction, grand mean intensity scaling, prewhitening, high-pass filtering >100 s, slice timing correction, and spatial smoothing [8 mm full-width half-maximum Gaussian kernel]). In each subject, all functional images were first registered to the middle image of the middle functional run (or the run following the midpoint, for subjects with an even number of runs), followed by registration to the subject's high-resolution anatomical scan via a six-parameter rigid body transformation, using FMRIB's Linear Image Registration Tool. Registration to standard space (Montreal Neurological Institute, MNI) was carried out with FMRIB's Nonlinear Image Registration Tool. Statistical analyses were performed using FMRIB's GLM analysis tool FEAT.

A whole-brain first-level GLM analysis modeled the following regressors using boxcar functions (length = each event's duration): InfoExpect_HI, InfoExpect_LO, RevealAnticip_HI, RevealAnticip_LO, Response, Feedback_POS_HI, Feedback_POS_LO, Feedback_NEG. Events where a frown emoticon followed either the 60% or 85% probabilistic stimulus were modeled as a single Feedback_NEG regressor, due to very few trials in which a frown emoticon followed 85% stimuli. The response regressor was orthogonalized with respect to RevealAnticip_HI and RevealAnticip_LO. Temporal derivatives were added to the model for each regressor. Premature, late, or sub-optimal responses (i.e., choice of the less-probable category given the stimulus) were captured by three additional regressors of no interest (InfoExpect_incorrect, RevealAnticip_incorrect, Feedback_incorrect). The model additionally included six motion-correction parameters. Regressors were convolved with a preset double- γ hemodynamic response function. Contrasts of interest were computed using linear combinations of regressors, including: InfoExpect_HI-LO, RevealAnticip_HI-LO, Feedback_POS_HI-LO, Feedback_POS (HI+LO), Feedback_POS-NEG. Functional runs were combined in a fixed-effects second-level analysis per subject. Data were averaged across subjects using FMRIB's FLAME1+2 mixed-effects third-level analysis.

ROI Analysis. Group-level activations were thresholded at $z = 2.6$ ($\alpha < 0.005$), whole-brain cluster-corrected with $k = 25$ contiguous voxels. One set of ROIs was identified using the contrasts InfoExpect_HI-LO, RevealAnticip_HI-LO, and Feedback_POS-NEG, in order to illustrate whether brain regions activated during one stage of the trial (e.g., InfoExpect) were also activated during other stages of the trial (e.g., RevealAnticip or Feedback). For the VS, to avoid biasing voxel selection toward one condition (e.g., InfoExpect_HI) and to avoid biasing significance tests when contrasting percent signal change estimates between two conditions (65), an additional set of ROIs was selected based on functional and anatomical criteria. For the functional criterion, in each condition (InfoExpect_HI and InfoExpect_LO), we searched for voxel clusters ($P < 0.005$, $k = 25$) that were significantly different from baseline (either above or below). For the anatomical criterion, voxels were restricted to the ventral parts of the striatum (i.e., we ensured no selected voxels were located in the caudate nucleus dorsal to the internal capsule, in the globus pallidus, or in adjacent anatomical regions, such as the insula, which is adjacent to the putamen). This revealed only above-baseline voxels for InfoExpect_HI and only below-baseline voxels for InfoExpect_LO. FMRIB's "featquery" function was used to calculate percent signal change for each (standard-space) ROI by scaling each regressor's parameter estimate by the peak to peak height of the regressor, multiplied by 100, and dividing by the average (across runs, per subject) of the mean of the filtered timeseries.

Data Availability. The data, stimulus and analysis files are permanently archived and publicly available for download on the Open Science Framework, doi:10.17605/osf.io/aexv9 or <https://osf.io/aexv9/> (66).

ACKNOWLEDGMENTS. We thank Taru Flagan and the University of California, San Diego fMRI Center staff for support; and Jorge Rey and Sheila O'Connell (University of Florida Medical Entomology Laboratory) for allowing us to adapt their copepod images. This study was supported by NIH Grant NIH T32 MH020002-04 (to T.J.S.), NIH MH57075-08 (to G.W.C.); NSF SBE-0542013 (Temporal Dynamics of Learning Center; to G.W.C.); Deutsche Forschungsgemeinschaft NE 1713/1 and NE 1713/2, part of the SPP 1516 "New Frameworks of Rationality" priority program (to J.D.N.); and University of California, San Diego Academic Senate Grant RH094G-COTTRELL (to G.W.C.).

1. J. Gottlieb, P. Y. Oudeyer, M. Lopes, A. Baranes, Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends Cognit. Sci.* **17**, 585–593 (2013).
2. J. Gottlieb, P. Y. Oudeyer, Towards a neuroscience of active sampling and curiosity. *Nat. Rev. Neurosci.* **19**, 758–770 (2018).

3. S. N. Haber, B. Knutson, The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology* **35**, 4–26 (2010).
4. K. C. Berridge, The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology (Berl.)* **191**, 391–431 (2007).

5. Y. Goto, A. A. Grace, Limbic and cortical information processing in the nucleus accumbens. *Trends Neurosci.* **31**, 552–558 (2008).
6. J. C. Cooper, B. Knutson, Valence and salience contribute to nucleus accumbens activation. *Neuroimage* **39**, 538–547 (2008).
7. T. Zaehele et al., Nucleus accumbens activity dissociates different forms of salience: Evidence from human intracranial recordings. *J. Neurosci.* **33**, 8764–8771 (2013).
8. R. Daniel, S. Pollmann, Striatal activations signal prediction errors on confidence in the absence of external feedback. *Neuroimage* **59**, 3457–3467 (2012).
9. M. J. Gruber, B. D. Gelman, C. Ranganath, States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* **84**, 486–496 (2014).
10. P. Dayan, K. C. Berridge, Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cogn. Affect. Behav. Neurosci.* **14**, 473–492 (2014).
11. J. P. O'Doherty, The problem with value. *Neurosci. Biobehav. Rev.* **43**, 259–268 (2014).
12. A. C. Burton, K. Nakamura, M. R. Roesch, From ventral-medial to dorsal-lateral striatum: Neural correlates of reward-guided decision-making. *Neurobiol. Learn. Mem.* **117**, 51–59 (2015).
13. R. Daniel, S. Pollmann, A universal role of the ventral striatum in reward-based learning: Evidence from human studies. *Neurobiol. Learn. Mem.* **114**, 90–100 (2014).
14. D. Shohamy, Learning and motivation in the human striatum. *Curr. Opin. Neurobiol.* **21**, 408–414 (2011).
15. W. Schultz, Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87–115 (2006).
16. W. Schultz, Behavioral dopamine signals. *Trends Neurosci.* **30**, 203–210 (2007).
17. B. Knutson, J. C. Cooper, Functional magnetic resonance imaging of reward prediction. *Curr. Opin. Neurol.* **18**, 411–417 (2005).
18. P. R. Montague, P. Dayan, T. J. Sejnowski, A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).
19. Y. Niv, R. Montague, "Theoretical and empirical studies of learning" in *Neuroeconomics: Decision Making and the Brain*, P. W. Glimcher, C. Camerer, R. A. Poldrack, E. Fehr, Eds. (Academic Press, London, 2008), pp. 329–350.
20. M. A. A. van der Meer, A. D. Redish, Ventral striatum: A critical look at models of learning and evaluation. *Curr. Opin. Neurobiol.* **21**, 387–392 (2011).
21. B. Abler, H. Walter, S. Erk, H. Kammerer, M. Spitzer, Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage* **31**, 790–795 (2006).
22. B. Knutson, S. E. B. Gibbs, Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology (Berl.)* **191**, 813–822 (2007).
23. R. Sutton, A. Barto, *Reinforcement Learning: An Introduction*, (MIT Press, Cambridge, MA, 1998).
24. J. Gläscher, N. Daw, P. Dayan, J. P. O'Doherty, States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* **66**, 585–595 (2010).
25. A. M. Bornstein, N. D. Daw, Multiplicity of control in the basal ganglia: Computational roles of striatal subregions. *Curr. Opin. Neurobiol.* **21**, 374–380 (2011).
26. N. D. Daw, S. J. Gershman, B. Seymour, P. Dayan, R. J. Dolan, Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
27. E. S. Bromberg-Martin, O. Hikosaka, Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* **63**, 119–126 (2009).
28. E. S. Bromberg-Martin, O. Hikosaka, Lateral habenula neurons signal errors in the prediction of reward information. *Nat. Neurosci.* **14**, 1209–1216 (2011).
29. T. C. Blanchard, B. Y. Hayden, E. S. Bromberg-Martin, Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* **85**, 602–614 (2015).
30. I. J. Good, *Probability and the Weighing of Evidence*, (Griffin, New York, 1950).
31. D. V. Lindley, On a measure of the information provided by an experiment. *Ann. Math. Stat.* **27**, 986–1005 (1956).
32. W. Kim, M. A. Pitt, Z. L. Lu, M. Steyvers, J. I. Myung, A hierarchical adaptive approach to optimal experimental design. *Neural Comput.* **26**, 2465–2492 (2014).
33. J. Najemnik, W. S. Geisler, Optimal eye movement strategies in visual search. *Nature* **434**, 387–391 (2005).
34. J. D. Nelson, Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychol. Rev.* **112**, 979–999 (2005).
35. J. D. Nelson, C. R. M. McKenzie, G. W. Cottrell, T. J. Sejnowski, Experience matters: Information acquisition optimizes probability gain. *Psychol. Sci.* **21**, 960–969 (2010).
36. N. R. Bramley, D. A. Lagnado, M. Speekenbrink, Conservative forgetful scholars: How people learn causal structure through sequences of interventions. *J. Exp. Psychol. Learn. Mem. Cogn.* **41**, 708–731 (2015).
37. A. Coenen, J. D. Nelson, T. M. Gureckis, Asking the right questions about the psychology of human inquiry: Nine open challenges. *Psychon. Bull. Rev.* **26**, 1548–1587 (2019).
38. V. Crupi, J. D. Nelson, B. Meder, G. Cevolani, K. Tentori, Generalized information theory meets human cognition: Introducing a unified framework to model uncertainty and information search. *Cogn. Sci.* **42**, 1410–1456 (2018).
39. D. J. Hayes, A. G. Huxtable, Interpreting deactivations in neuroimaging. *Front. Psychol.* **3**, 27 (2012).
40. J. Gottlieb, M. Hayhoe, O. Hikosaka, A. Rangel, Attention, reward, and information seeking. *J. Neurosci.* **34**, 15497–15504 (2014).
41. Y. Niv, S. Chan, On the value of information and other rewards. *Nat. Neurosci.* **14**, 1095–1097 (2011).
42. X. Xia et al., Fine-grained parcellation of the macaque nucleus accumbens by high-resolution diffusion tensor tractography. *Front. Neurosci.* **13**, 709 (2019).
43. B. Engelhard et al., Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* **570**, 509–513 (2019).
44. D. Shohamy, C. E. Myers, J. Kalanithi, M. A. Gluck, Basal ganglia and dopamine contributions to probabilistic category learning. *Neurosci. Biobehav. Rev.* **32**, 219–236 (2008).
45. T. D. Satterthwaite et al., Being right is its own reward: Load and performance related ventral striatum activation to correct responses during a working memory task in youth. *Neuroimage* **61**, 723–729 (2012).
46. M. Horan, N. Daddaoua, J. Gottlieb, Parietal neurons encode information sampling based on decision uncertainty. *Nat. Neurosci.* **22**, 1327–1335 (2019).
47. J. A. M. Rodriguez Cabrero, J.-Q. Zhu, E. A. Ludvig, Costly curiosity: People pay a price to resolve an uncertain gamble early. *Behav. Processes* **160**, 20–25 (2019).
48. K. Iigaya, G. W. Story, Z. Kurth-Nelson, R. J. Dolan, P. Dayan, The modulation of savouring by prediction error and its effects on choice. *eLife* **5**, e13747 (2016).
49. M. Brydevall, D. Bennett, C. Murawski, S. Bode, The neural encoding of information prediction errors during non-instrumental information seeking. *Sci. Rep.* **8**, 6134 (2018).
50. D. Bennett, S. Bode, M. Brydevall, H. Warren, C. Murawski, Intrinsic valuation of information in decision making under uncertainty. *PLoS Comput. Biol.* **12**, e1005020 (2016).
51. C. J. Charpentier, E. S. Bromberg-Martin, T. Sharot, Valuation of knowledge and ignorance in mesolimbic reward circuitry. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E7255–E7264 (2018).
52. M. Corbetta, G. L. Shulman, Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
53. J. Gottlieb, Attention, learning, and the value of information. *Neuron* **76**, 281–295 (2012).
54. N. C. Foley, S. P. Kelly, H. Mhatre, M. Lopes, J. Gottlieb, Parietal neurons encode expected gains in instrumental information. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E3315–E3323 (2017).
55. J. Gottlieb, Understanding active sampling strategies: Empirical approaches and implications for attention and decision research. *Cortex* **102**, 150–160 (2018).
56. F. Filimon, J. D. Nelson, R. S. Huang, M. I. Sereno, Multiple parietal reach regions in humans: Cortical representations for visual and proprioceptive feedback during online reaching. *J. Neurosci.* **29**, 2961–2971 (2009).
57. F. Filimon, Human cortical control of hand movements: Parietofrontal networks for reaching, grasping, and pointing. *Neuroscientist* **16**, 388–407 (2010).
58. F. Filimon, M. G. Philiastides, J. D. Nelson, N. A. Kloosterman, H. R. Heekeren, How embodied is perceptual decision making? Evidence for separate processing of perceptual and motor decisions. *J. Neurosci.* **33**, 2121–2136 (2013).
59. Y. C. Leong, A. Radulescu, R. Daniel, V. DeWoskin, Y. Niv, Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* **93**, 451–463 (2017).
60. S. M. Gorka, K. L. Phan, S. A. Shankman, Convergence of EEG and fMRI measures of reward anticipation. *Biol. Psychol.* **112**, 12–19 (2015).
61. X. Liang, L. A. Zebrowitz, Y. Zhang, Neural activation in the "reward circuit" shows a nonlinear response to facial attractiveness. *Soc. Neurosci.* **5**, 320–334 (2010).
62. A. R. Hariri et al., Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. *J. Neurosci.* **26**, 13213–13217 (2006).
63. B. Meder, J. D. Nelson, Information search with situation-specific reward functions. *Judgm. Decis. Mak.* **7**, 119 (2012).
64. L. Martignon, K. V. Katsikopoulos, J. K. Woike, Categorization with limited resources: A family of simple heuristics. *J. Math. Psychol.* **52**, 352–361 (2008).
65. N. Kriegeskorte, W. K. Simmons, P. S. F. Bellgowan, C. I. Baker, Circular analysis in systems neuroscience: The dangers of double dipping. *Nat. Neurosci.* **12**, 535–540 (2009).
66. F. Filimon, J. D. Nelson, T. J. Sejnowski, M. I. Sereno, G. W. Cottrell, The Ventral Striatum dissociates information expectation, reward anticipation, and reward receipt. Open Science Framework. <https://doi.org/10.17605/OSF.IO/AEXV9>. Deposited 26 March 2020.