Lateral habenula neurons signal errors in the prediction of reward information

Ethan S. Bromberg-Martin and **Okihide Hikosaka**

Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

Abstract

Humans and animals have a remarkable ability to predict future events, which they achieve by persistently searching their environment for sources of predictive information. Yet little is known about the neural systems that motivate this behavior. We hypothesized that information-seeking is assigned value by the same circuits that support reward-seeking, so that neural signals encoding conventional "reward prediction errors" include analogous "information prediction errors". To test this we recorded from neurons in the lateral habenula, a nucleus which encodes reward prediction errors, while monkeys chose between cues that provided different amounts of information about upcoming rewards. We found that a subpopulation of lateral habenula neurons transmitted signals resembling information prediction errors, responding when reward information was unexpectedly cued, delivered, or denied. Their signals evaluated information sources reliably even when the animal's decisions did not. These neurons could provide a common instructive signal for reward-seeking and information-seeking behavior.

How do we know what to seek and what to avoid? Conventional neural theories of motivated behavior propose that we assign value to each stimulus based on its predicted yield of future primary rewards (such as food and water)[1]. These values are then updated each time the stimulus produces an unpredicted outcome, an event called a "reward prediction error" (RPE)[1, 2]. If a stimulus leads to a larger reward than predicted (positive RPE), its value is increased; if it leads to a smaller reward than predicted (negative RPE), its value is decreased; and if it produces exactly the same reward as predicted (zero RPE), its value is maintained[1]. Much is known about the neural basis of this process. RPE signals were first discovered in midbrain dopamine neurons[1, 3] and have now been found in neurons in many structures including the medial prefrontal cortex[4, 5], striatum[6, 7], globus pallidus[8], and habenula[9, 10], as well as in ensemble electrical signals[11] and blood oxygenation-level dependent signals[12–14] in several of these structures.

Although such RPEs provide a good account of behavior directed toward gaining primary rewards, we often take actions to achieve other goals. A striking example is the preference for predictability. Humans[15–17] and animals[18–21] often choose to view sensory cues that have no effect on the delivery of primary rewards but which provide predictive information to help anticipate properties of the rewards in advance, such as their presence[18], size[22], and probability[23, 24]. This preference has been studied by economists as "temporal resolution of uncertainty"[25] and by psychologists as a form of "observing behavior"[26]. This preference cannot be learned from conventional RPEs; for example, if a predictive cue is equally likely to indicate a good outcome (positive RPE) or a bad outcome (negative RPE) then the net RPE would be zero, so the act of viewing the cue would not be reinforced[22] (Supplementary Figs. 1,2). Hence we initially expected that the preference to view reward-informative cues would be signaled by distinct neural circuits from RPEs. Yet our first investigation produced a very different result: the preference was signaled by midbrain dopamine neurons, the same neurons that transmit RPEs for primary rewards[22].

Based on this finding, we hypothesized that the preference to view reward-informative cues is generated by assigning them value in the same 'common currency' as primary rewards, and thus activating the same underlying RPE mechanism[22]. Then neural RPE signals would have two components: a conventional component that encodes errors in predicting the value of primary rewards, and a second component that encodes errors in predicting the reward value of viewing informative sensory cues. We will refer to the first component as a conventional reward prediction error (cRPE), and to the hypothesized second component as information-related reward prediction error (or "information prediction error", IPE). An action followed by the unexpected presentation of an informative cue would be reinforced (positive IPE); an action followed by the unexpected denial of an informative cue would be punished (negative IPE); and an action followed by a fully predictable presentation or denial of an informative cue would have no change in its value (zero IPE). Thus, whereas cRPEs would provide reinforcement for seeking primary rewards, IPEs would provide reinforcement for seeking reward-informative sensory cues.

In the present study we tested this hypothesis by designing an experiment to make a direct comparison between neural coding of cRPEs and IPEs. We recorded from neurons in the lateral habenula, a nucleus in the epithalamus involved in negative control of motivation[27]. Lateral habenula neurons transmit RPEs in an inverted manner[8–10] and can suppress dopamine neuron activity[27–29] and motivated behavior[30, 31]. We now show that a subset of lateral habenula neurons transmit IPEs in a similar manner, signaling when reward information is unexpectedly cued, delivered, or denied. These neurons can be highly reliable, providing accurate judgments of information quality even when the subject's decisions do not. Thus, these neurons are ideally positioned to provide a common instructive signal for reward- and information-seeking behavior.

RESULTS

Behavioral preference for reward information

We trained two monkeys to choose between different sources of reward information (Fig. 1a–c). On each trial two colored targets appeared on the screen and the monkey chose

between them with a saccadic eye movement. Then, after a delay of a few seconds, the monkey received either a big or a small water reward with equal probability. The monkey's choice had no effect on the water reward that was delivered but allowed the monkey to view different sets of visual cues (Fig. 1b). Choices of one target, the 100% info target, were always followed by *informative* reward cues whose shapes indicated the upcoming reward size. Choices of a second target, the 0% info target, were always followed by *random* reward cues whose shapes were randomized on each trial and conveyed no new information. Choices of a third target, the 50% info target, were followed by either informative or random cues with equal probability. There were two possible choice pairs: 100% info vs. 50% info, and 50% info vs. 0% info (Fig. 1c). After each of these *choice trials*, a later *forced trial* was scheduled in which the originally non-chosen target was presented alone and animals were forced to choose it[18, 21] (Fig. 1c). This ensured that animals had equal exposure to the different targets and cues.

Monkeys consistently chose the target that gave the best chance to view informative cues. They chose 100% info in preference to 50% info and chose 50% info in preference to 0% info (Fig. 1d). Their preferences were consistent across days, occurring during every recording session (Supplementary Fig. 3). This extends our previous findings[22] by showing that monkey preferences are sensitive to the degree of information probability.

In addition to measuring the animal's preference, our experiment was designed to allow two direct comparisons between neural coding of cRPEs and IPEs. The first comparison was between responses to predictive stimuli. Conventional RPEs could be measured in response to the reward cues, which indicated that the probability of a big reward increased, decreased, or remained unchanged. In the same way, IPEs could be measured in response to the target array, which indicated that the probability of viewing informative cues increased, decreased, or remained unchanged. The second comparison was between responses to outcome delivery. Conventional RPEs could be measured when rewards were bigger, smaller, or the same size as predicted. In the same way, IPEs could be measured when the cues were more, less, or equally informative as predicted.

Lateral habenula coding of reward prediction errors

While monkeys performed this task we recorded from 95 neurons in the lateral habenula (Methods). Consistent with previous studies[8–10], we found that many lateral habenula neurons encoded cRPEs and did so in an *inverted* manner - excited by negative cRPEs and inhibited by positive cRPEs (Fig. 2). The lateral habenula population was strongly responsive to reward cues: it was excited when they indicated a decrease in reward value (small reward cue, negative cRPE), inhibited when they indicated an increase in reward value (big reward cue, positive cRPE), and relatively non-responsive when they indicated no change in reward value (random cues, zero cRPE) (Fig. 2a, left). The population was also strongly responsive to rewards: it was excited when the reward was smaller than predicted (small reward after random cues, negative cRPE), inhibited when the reward was larger than predicted (large reward after random cues, positive cRPE), and had a much weaker response when the reward size was fully predictable (large or small reward after informative cues, zero cRPE) (Fig. 2a, right). Thus the population's response resembled an inverted version of

the RPE signal from formal models of reward prediction errors (Fig. 2a, bottom; see Supplementary Figs. 1,2 for a computational model fitted to the same response windows and magnitudes as the neural data).

This was the most common response pattern in single neurons. We quantified each neuron's cRPE coding using a "conventional reward prediction error index" (cRPE index; Methods). This index ranges between −1 and +1, with indexes < 0 for activity inversely related to cRPEs, indexes > 0 for activity directly related to cRPEs, and indexes = 0 for no coding of cRPEs. The index was calculated separately for responses to reward cues and reward delivery, and separately for positive and negative cRPEs. Most neurons had cRPE indexes < 0, indicating activity inversely related to cRPEs (Fig. 2b, neurons in lower left quadrant). We next calculated an overall measure of cRPE coding for each task event by taking the mean of the indexes for positive and negative cRPEs. Most neurons had mean indexes < 0 for both reward cues and reward delivery, indicating inverted coding of cRPEs for both task events (Fig. 2c, neurons in lower left quadrant).

Coding of unpredicted changes in information probability

We next asked whether lateral habenula neurons encoded IPEs and whether they did so in a similar manner to cRPEs. The first information-related event in our task was the onset of the target array. The task was designed so that negative and positive IPEs could be measured with different target arrays. Negative IPEs could be measured on forced trials, by comparing forced 0% info trials (reduced information probability, negative IPE) to forced 50% info trials (no change in information probability, zero IPE) (Fig. 3a). Positive IPEs could be measured on choice trials, by comparing choice 100% info trials (increased information probability, positive IPE) to choice 50% info trials (no change in information probability, zero IPE) (Fig. 3b). According to our hypothesis lateral habenula neurons should encode IPEs in an inverted manner, with higher activity for negative IPEs and lower activity for positive IPEs.

Indeed, the lateral habenula population had activity inversely related to IPEs. For negative IPEs, activity was higher on forced 0% than on forced 50% info trials (Fig. 3a,c, $P < 10^{-3}$, signed-rank test). For positive IPEs, activity was lower on choice 100% trials than choice 50% info trials (Fig. 3b,c, $P = 0.032$). This was the most common response pattern in single neurons. We quantified each neuron's activity with an "information prediction error index", analogous to the cRPE index for conventional reward prediction errors, which was computed separately for negative IPEs and positive IPEs (Methods). Many lateral habenula neurons had indexes < 0, indicating inverted coding of IPEs (index for negative IPEs, mean = −0.16, $P < 10^{-3}$; index for positive IPEs, mean = −0.07, $P = 0.015$; signed-rank test; Fig. 3d,e).

This IPE-related activity was most prominent in a subpopulation of neurons. This could be seen by the fact that the IPE indexes were correlated, so that neurons with strong coding of negative IPEs also had strong coding of positive IPEs (rho = 0.35, $P < 10^{-3}$, permutation test, Fig. 3f). To examine them more closely, we selected the subpopulation of "information-predictive neurons" for which the mean of the two target IPE indexes was significantly below zero (n = 30; Fig. 3f, black circles). We then estimated their response to the target

arrays (Fig. 4a,b, using a cross-validation procedure to correct for selection bias; Methods). The subpopulation of information-predictive neurons had a clear differential response to the target array resembling a theoretical inverted IPE signal (Fig. 4a,b). This subpopulation was excited by negative IPEs (Fig. 4b, blue circle, $P < 10^{-3}$), inhibited by positive IPEs (Fig. 4b, red circle, $P = 0.017$), and had no significant response for zero IPEs (Fig. 4b, purple circles, all $P > 0.50$). The population of remaining neurons had weaker responses with no significant difference between the targets (n=65, Fig. 4c,d). There was a tendency for different response patterns on forced versus choice trials but no significant difference in overall response magnitudes (Fig. 4, Supplementary Figs. 4,5). Based on these findings we focused our further analysis on the subpopulation of information-predictive neurons, with the goal of testing whether they continued to encode IPEs throughout the behavioral task (for the remaining neurons, see Supplementary Figs. 6,7).

Coding of unpredicted denial of reward information

According to our hypothesis, lateral habenula neurons should encode negative IPEs when reward information is unexpectedly denied. This should occur in response to the randomized reward cues, which provided no new information. If the animal had chosen the 0% info target then the animal could fully predict that no information would be delivered so there would be little lateral habenula response ('predictable no-info', zero IPE; Fig. 5a, blue). But if the animal had chosen the 50% info target then the denial of information would be unpredicted and the animal would experience a negative information prediction error, causing lateral habenula neurons to be excited ('unpredicted no-info', negative IPE; Fig. 5a, purple). Note that the physical stimulus in both conditions, the random cue, was exactly the same. The only difference was the predicted probability that the cue would be informative or random.

Consistent with our hypothesis, the information-predictive neurons were excited when reward information was unexpectedly denied. They had little response on 0% info trials when the random cues were fully predictable (Fig. 5a, blue) but were excited on 50% info trials when the same random cues were presented unpredictably (Fig. 5a, purple). Thus, these neurons were excited by the first indication that reward information would be denied – either the 0% info target or unpredicted no-info delivery itself (Fig. 5a). We quantified each neuron's response to the random cues using an index for coding negative IPEs (Methods, Fig. 5b). Most information-predictive neurons had indexes < 0, indicating activity inversely related to information prediction errors (mean = −0.25, $P < 10^{-3}$; Fig. 5b). Furthermore, considering the lateral habenula population as a whole, the IPE indexes were correlated so that neurons with strong coding of negative IPEs for the random cues also had strong coding of both negative and positive IPEs for the targets (Fig. 5c).

Coding of unpredicted delivery of reward information

According to our hypothesis, lateral habenula neurons should encode positive IPEs when reward information is unexpectedly delivered. This should occur in response to the informative reward cues (Fig. 6a). If the animal had chosen the 100% info target then the animal could fully predict that an informative cue would appear so there would be no information prediction error ('predictable info', zero IPE; Fig. 6a, red). But if the animal had

chosen the 50% info target then the delivery of information would be unpredicted and the animal would experience a positive information prediction error, causing lateral habenula neurons to be inhibited ('unpredicted info', positive IPE; Fig. 6a, purple). Of course, the informative cues would simultaneously evoke *conventional* reward prediction errors encoding the cued reward size. Hence we hypothesized that lateral habenula neurons would respond with excitation or inhibition encoding cRPEs, but that 'unpredicted info' trials would have an additional inhibitory response encoding positive IPEs (Fig. 6a, right).

Consistent with our hypothesis, the information-predictive lateral habenula neurons were inhibited when reward information was unexpectedly delivered. They were inhibited by big-reward cues, and on 'unpredicted info' trials this inhibition was enhanced (Fig. 6a, −1.81 spikes/s, signed-rank test, $P = 0.001$). They were excited by small-reward cues, and on 'unpredicted info' trials this excitation was attenuated (Fig. 6a, −1.43 spikes/s, $P = 0.028$). These changes in firing rate were modest, perhaps due to competition with the strong cRPE signals that were present at the same time. Indeed, the IPE signal emerged with a similar latency to the cRPE signal but was strongest during the initial part of the response before the cRPE signal reached its peak (Fig. 6b). We quantified each neuron's response to the informative cues using an index for coding positive IPEs (Methods). The information-predictive neurons tended to have indexes < 0, indicating activity inversely related to information prediction errors (mean = −0.11; signed-rank test, $P = 0.001$; Fig. 6c). Furthermore, considering the lateral habenula population as a whole, these indexes were correlated with all three of the other IPE indexes for other task events (Fig. 6d). Hence the lateral habenula population encoded IPEs in a consistent manner throughout the task, including responses to information-predictive stimuli (Figs. 3,4), information denial (Fig. 5) and information delivery (Fig. 6).

Joint coding of reward and information prediction errors

We next tested whether IPEs and cRPEs were transmitted by the same neurons and whether they used a similar neural code. We measured each neuron's tendency for coding of inverted prediction errors by counting the total number of its IPE and cRPE indexes that were < 0. As described above, each neuron had a total of four IPE indexes and four cRPE indexes representing coding of positive and negative prediction errors evoked by predictive stimuli and outcome delivery.

The most common patterns in habenula neurons were to have all four IPE indexes < 0 and all four cRPE indexes < 0 (Fig. 7a; more common than expected by chance, binomial test, $P < 10^{-5}$). Considering IPEs and cRPEs together, most habenula neurons had six, seven, or all eight of their indexes < 0 (Fig. 7b; more common than chance, binomial test, $P = 0.006$, $P < 10^{-12}$, $P < 10^{-100}$). Most neurons with significant coding of IPEs also had significant coding of cRPEs, and the IPE and cRPE indexes were positively correlated (Supplementary Fig. 8). Thus many habenula neurons transmitted IPEs and cRPEs using a similar code.

Reliable prediction errors despite variable decisions

Our experiment is the first investigation of lateral habenula neurons during decision making, allowing us to test whether they are influenced by the decision process. In particular, we

found that animals made somewhat variable decisions: they usually chose the high-information probability target but occasionally chose the low-information probability target (Fig. 1d). One hypothesis is that animals chose the low-info target due to *noisy evaluation* – occasionally assigning it unusually high value. If so, then on trials when animals chose the low-info target, neurons should have treated it as unusually valuable (Supplementary Fig. 9). A second hypothesis is that animals chose the low-info target due to *noisy action selection* – that they always assigned it low value but chose it occasionally due to errors in motor execution or purposeful exploration[32, 33]. If so, then neurons should have always treated the low-info target as having low value.

To test these hypotheses, we analyzed the subset of information-predictive neurons that were recorded during choices of 0% > 50% info (n=25, Fig. 8a,d) and 50% > 100% info (n=17, Fig. 8b,e). This revealed that lateral habenula neurons had accurate prediction error signals even during low-info choices. Neurons had higher activity during low-info choices than high-info choices (Fig. 8a, $P < 10^{-3}$; Fig. 8b, $P = 0.006$; signed-rank test). Activity during choices of 0% info was similar to forced trials when the 0% info target was presented alone, and was higher than activity on all 50% or 100% info trials (Fig. 8d). Similarly, activity during low-info choices of 50% info was intermediate between trials with 0% or 100% information probability and was similar to (or slightly higher than) the other 50% info trial types (Fig. 8e).

Since the lateral habenula inhibits midbrain dopamine neurons[9, 28, 29], we wondered whether dopamine neurons had an opposite coding pattern. To test this, we repeated this analysis on our previous dataset[22], which was recorded in a simpler task where only the 100% and 0% info targets were available. Indeed, the activity of putative dopamine neurons was higher during high-info choices than low-info choices ($P = 0.002$) (Fig. 8c,f).

We next asked whether these neural signals emerged rapidly enough to contribute to the decision making process. Neural activity distinguishing between high-info and low-info choices emerged at a relatively long latency (Fig. 8a–c, difference between solid and dashed lines) after about 90% of the animal's choice saccades had already been initiated (Fig. 8a–c, colored horizontal lines above the x-axis indicate the central 90% of the saccadic reaction time distribution). This suggests a sequence in which (1) animals made rapid and partially noisy saccadic decisions, then (2) lateral habenula and dopamine neurons monitored the decision and signaled prediction errors based on the chosen option's value.

DISCUSSION

We found a subpopulation of lateral habenula neurons that signaled errors in the prediction of reward-informative sensory cues (IPEs) in addition to errors in the prediction of primary rewards themselves (conventional RPEs). These neurons signaled both types of prediction errors with a similar code: excitation for negative prediction errors and inhibition for positive prediction errors. And their signals were consistent across multiple task events including reward-and information-predictive stimuli and reward and information delivery themselves. These data support and extend current theories by showing that neural reward prediction error signals not only encode errors in predicting the value of primary rewards but

also errors in predicting the value of viewing reward-informative cues. This supports our hypothesis that viewing informative cues is assigned value in the same 'common currency' as primary rewards and could reinforce behavior through the same mechanism.

In addition to demonstrating the presence of IPE signals in the brain, our data also suggest a possible mechanism by which these signals could be translated into a behavioral preference. When animals learned that reward information would be denied, many lateral habenula neurons were activated. Activation of the lateral habenula inhibits dopamine neurons[27–29] and can act as a punishment to induce behavioral aversion[30, 31] (but see[34]). Conversely, when animals learned that reward information would be delivered, many lateral habenula neurons reduced their activity. Inactivation of the lateral habenula causes dopamine release in multiple brain regions including the nucleus accumbens[35] which can act as reinforcement to induce behavioral preferences[36]. Such a habenula-to-dopamine pathway is also consistent with our previous finding that dopamine neurons are activated by information-predictive stimuli[22], perhaps representing a component of the IPE signal. Our data also provide evidence that information-related neural signals are related to the motivational value that underlies behavioral preferences. Although we previously found information-related signals in dopamine neurons, recent studies suggested that some dopamine neurons encode general motivational salience rather than value[37–39]. Our present study, however, shows that IPE signals are present in the lateral habenula, where neurons predominantly transmit signals related to motivational value[10] and exert control over dopamine neurons that have properties consistent with value coding[37, 39]. Future studies may provide a conclusive test of whether IPE signals are first generated by lateral habenula neurons, dopamine neurons, or other brain structures, and how signals are elaborated and processed at each step in these neural pathways.

Our data raise several questions about how the brain assigns greater value to viewing informative cues than random cues. The brain might explicitly generate internal rewards based on a cue's 'informativeness' or reduction in uncertainty[15, 25, 40, 41]; alternately, there is evidence that the brain accomplishes this implicitly by having cues provide conditioned reinforcement that is disproportionate to their primary reward value[19–21, 42, 43]. Our neural data puts constraints on this process by excluding one recently proposed mechanism[43] (Supplementary Fig. 2), but many alternatives remain and deserve investigation. Our data also gives a rough estimate of the relative magnitude of the IPE and cRPE signals. IPE signals in information-predictive neurons were on average about 1/5th as large as cRPE signals, and could be accounted for by a model in which viewing an informative cue was as valuable as drinking ~0.17 ml of water reward (Supplementary Fig. 1). This neural relationship is likely to vary across species and tasks, and deserves further investigation and comparison with behavior. Finally, we found that some lateral habenula neurons did not significantly encode IPEs or encoded them in a different direction from cRPEs, and some neurons had additional idiosyncratic tonic and phasic forms of activity[44] (Supplementary Figs. 8,10,11). This might contribute to the heterogeneous response patterns in certain areas downstream of the lateral habenula, like the dorsal raphe nucleus[45, 46].

Our data show two new features of the relationship between lateral habenula neurons and decision making. First, we found that animals could make rapid saccadic decisions even

before prediction error signals emerged in lateral habenula and dopamine neurons. This supports the view that lateral habenula and dopamine prediction error signals do not contribute to rapid saccadic decision making[47], instead responding to the chosen option's value after the decision has been made. Second, we found that lateral habenula prediction error signals were accurate even when animals decided to view poor sources of information, a striking property also found in dopamine neurons[48, 49]. In this sense lateral habenula signals were more a reliable judge of information sources than the animal's outward behavior. This would allow the lateral habenula to be an effective 'critic' of saccadic decisions[1, 32, 50]: when the animal shifted its gaze to a poor source of reward information, lateral habenula neurons immediately sent a negative prediction error as if to criticize the decision. Thus the lateral habenula may transmit prediction error signals after each saccadic eye movement, providing immediate feedback to teach the eyes when and where primary rewards and reward information will appear.

METHODS

Subjects

Subjects were two male rhesus macaque monkeys (*Macaca mulatta*), monkey V (9.3 kg) and monkey Z (8.7 kg). All procedures for animal care and experimentation were approved by the Institute Animal Care and Use Committee and complied with the Public Health Service Policy on the humane care and use of laboratory animals. A plastic head holder, scleral search coils, and plastic recording chambers were implanted under general anesthesia and sterile surgical conditions.

Behavioral Tasks

Behavioral tasks were conducted using equipment described previously[22]. Monkeys sat in a primate chair facing a frontoparallel screen located 31 cm from the eyes in a sound-attenuated and electrically shielded room. Eye movements were monitored using a scleral search coil (monkey V) or SensoMotoric Instruments eye tracker (monkey Z). Stimuli were presented with a PJ550 projector (ViewSonic).

Trials began with a central spot of light (1° diameter) which the monkey was required to fixate. After 800 ms of fixation, the spot disappeared and two colored targets appeared on the sides of the screen (2.5° diameter, 10–15° eccentricity; on forced trials, only one target appeared). The monkey had 710 ms to saccade to the chosen target, after which the non-chosen target immediately disappeared. After the 710 ms period the chosen target was replaced by a reward cue (14° diameter), which monkeys were not required to fixate. After 2250 ms the cue disappeared, a 200 ms tone began, and water reward delivery began using a gravity-based system (Crist Instruments). Small rewards lasted 50 ms (0.04 ml), large rewards lasted 700 ms (0.88 ml). The inter-trial interval was ~4.1–5.1 seconds beginning from reward onset. To minimize the effects of physical preparation, water was delivered directly into the mouth; licking the water spout was not required. We previously showed that animals received the same amount of water on informative and random cue trials and their preferences were not caused by the physical features of the targets and cues (e.g. shapes or colors)[22].

There were three targets, which had different probabilities of being followed by informative reward cues (100%, 50%, and 0%) and were distinguished by their colors (orange, purple, and green). The 100% and 0% info colors were counterbalanced across animals (monkey V: green 100% info, orange 0% info; monkey Z: the reverse). There were four cues. Two of the cues, the *informative cues*, had the same color as the 100% info target and their shapes indicated the size of the upcoming water reward. The other two cues, the *random cues*, had the same color as the 0% info target and their shapes were chosen randomly so that they were uncorrelated with the size of the upcoming water reward. For example, for monkey Z the orange "A" was the info-big cue indicating an upcoming big reward, the orange "Z" was the info-small cue indicating an upcoming small reward, and the green "ω" and "ς" shapes were followed by big and small rewards with equal probability.

There were two choice pairs: 100% vs. 50% info and 50% vs. 0% info. For each choice pair, there were $2^4 = 16$ types of choice trials, specified by four binary factors: reward size (big or small), cue type to show after a 50% info choice (informative or random), random cue shape to show if a random cue was presented (e.g. "ω" or "ς"), and location of the 50% info choice on the screen (left or right). The exposure to these 16 trial types was counterbalanced across each experimental session. Specifically, for each choice pair, choice trials were scheduled so that each successive 16 choice trials within a session included one trial of each type (presented in a randomized order).

To balance exposure to the targets, we used a compensation procedure in which each choice trial (e.g. choice of 100% over 50% info) was compensated by scheduling a later forced trial in which only the previously non-chosen target was available (e.g. forced 50% info)[18], as follows. Each session began with 8 consecutive choice trials (4 from each choice pair) presented in a random order. Each further set of 8 trials consisted of two trials from each of the 2×2 combinations of choice pair (100% vs. 50%, 50% vs. 0%) and choice type (choice, forced), presented in a random order. Starting at the beginning of the session, a list was maintained of 'uncompensated' completed choice trials. Each time a choice trial was completed it was added to the list. Each time a forced trial was presented its parameters were set equal to those of a randomly selected choice trial from the list (except that the previously chosen target was no longer available), after which that choice trial was considered 'compensated' and removed from the list.

In case of errors the trial was aborted: the screen went blank, a 200 ms error tone sounded, a 3000 ms delay occurred, and the inter-trial interval began. Errors occurred if the monkey (1) did not fixate within 5000 ms of fixation point onset, (2) broke fixation during the fixation period, (3) did not make a saccade, (4) made a saccade that did not land on a target, or (5) broke fixation on the target before the cue appeared. In a subset of sessions monkeys were not required to avoid error types 4–5 on forced trials and 5 on choice trials; they still continued to perform the task successfully. As in our previous study[22], monkeys were slightly reluctant to saccade toward and fixate the undesirable 0% info target and therefore made slightly more errors on forced 0% info trials; however, these errors were quite rare (< 0.3% of forced 0% info trials during neural recordings), so they probably did not have a meaningful influence on the results.

Neural Recording

Lateral habenula neurons were recorded as described previously[9, 10]. A recording chamber was placed over the midline of the parietal cortex, tilted posteriorly by 38 degrees (monkey V) or 40 degrees (monkey Z), and aimed at the habenula. A grid system allowed recording sites to be targeted with 1 mm spacing. We estimated the grid locations and depth of the habenula using magnetic resonance imaging (4.7 T, Bruker), its expected position 1.5–2.5 mm lateral to the midline, and its expected position relative to the visuomotor map of neural activity in the superior colliculus. The lateral habenula was identified based on neural recordings of that structure and nearby neural structures, as a small region that was located (1) directly beneath the corpus callosum and lateral ventricle, (2) medial to the adjacent mediodorsal thalamus (low rates, irregular bursty firing, and little or no task-related activity), (3) anterior to the presumed pretectum (~3 mm deeper with little or no phasic reward-related activity time-locked to task events), (4) symmetrically to the lateral habenula in the opposite hemisphere, and (5) had high frequency multiunit activity with stimulus-locked neural responses including negative reward signals (inhibitory response to fixation point and/or big rewards/big-reward cues, and/or excitatory response to small rewards/small-reward cues). Putative midbrain dopamine neurons was analyzed from a previous dataset recorded in and around the substantia nigra pars compacta and lateral ventral tegmental area[22]. Putative dopamine neurons were identified from their irregular tonic firing at 0.5–10 spikes/s, broad spike waveforms, and activity positively related to cue and/or outcome value[22].

Lateral habenula neurons were included in the database if they were recorded for at least 45 trials (monkey Z) or if they were recorded for at least 45 trials and were responsive to the task (monkey V). Neurons were recorded for 46–552 trials (mean and standard deviation: 170 ± 111 trials).

Data Analysis

We analyzed neural firing rates in three time windows representing responses to the target array (150–500 ms after target array onset), reward cue (100–350 ms after cue onset), and reward outcome (200–450 ms after reward onset). These time windows were chosen *a priori* based on our previous study of dopamine neurons[22] and were used here without modification. A neuron's baseline rate was defined as its firing rate during the 250 ms before target onset. Baseline rates ranged from 0.3 to 60.4 spikes/s (mean and standard deviation: 23.6 ± 14.0 spikes/s) and were similar for information-predictive neurons (25.2 ± 15.8 spikes/s) and the remaining neurons (22.9 ± 13.1 spikes/s). The analyses in Figs. 3,4 were performed on forced trials and high-info choice trials. Low-info choice trials were analyzed separately in Fig. 8. Correlations in scatterplots were calculated using all neurons.

We quantified neural prediction error signals using "cRPE indexes" for conventional reward prediction errors and "IPE indexes" for information prediction errors. Each index was calculated for each neuron based on the receiver operating characteristic area (ROC area) for discriminating between the single-trial neural firing rates from a pair of task conditions where the prediction errors of reward and/or reward information were theorized to be different. The ROC area was in the range between 0 and 1, representing the probability a

randomly chosen single-trial firing rate from the condition with more positive prediction errors was greater than a randomly chosen single-trial firing rate from the condition with more negative prediction errors. To convert the ROC area into a cRPE or IPE index, it was transformed to lie in the range [−1,+1] (using the equation: index = 2*(ROC area − 0.5)). The indexes were calculated using the following pairs of task conditions. Cue, index for negative cRPEs: info-small cue vs. random cues; index for positive cRPEs: random cues vs. info-big cue. Reward, index for negative cRPEs: random-small reward vs. informed-small reward; index for positive cRPEs: informed-big reward vs. random-big reward. Target, index for negative IPEs: forced 0% info vs. forced 50% info; index for positive IPEs: choice 50% > 0% info vs. choice 100% > 50% info. Cue, index for negative IPEs: random cue after 50% info target ('unpredicted no-info') vs. random cue after 0% info target ('predictable no-info'); index for positive IPEs: calculated using the mean of the ROC areas for the two informative cues (info-big cue after 100% info target ('predictable info, big') vs. info-big cue after 50% info target ('unpredicted info, big'), and info-small cue after 100% info target ('expected info, small') vs. info-small cue after 50% info target ('unpredicted info, small')). The single-neuron cRPE and IPE indexes were tested for significant differences from zero using Wilcoxon rank-sum tests (or, for indexes calculated using the mean of two ROC areas or the means of two indexes, permutation tests with 2000 permutations).

For the analysis of target responses in information-predictive neurons (Fig. 4a,b), we used a cross-validation procedure to remove the effects of selection bias. We split the data into two sets containing the even-numbered and odd-numbered trials, separately for each neuron in the lateral habenula population and for each target condition (forced 0%, forced 50%, choice 50%, choice 100%). Then, each neuron's 'even data' was used to select whether its 'odd data' would be included in the analysis, and vice versa. This ensured that the decision about whether to include a trial in the analysis was always made without looking at that trial's data. Specifically, we calculated each neuron's mean target IPE index separately for the even and odd datasets. If neither index was significantly < 0, the neuron was excluded; if the 'even data' index was significantly < 0, then its 'odd data' was included; if the 'odd data' index was significantly < 0, then its 'even data' was included; if both indexes were significantly < 0, then both even and odd data were included. This removed the effects of selection bias in simulated datasets. For the remaining neurons (Fig. 4c,d), we used the same procedure except the selection criterion was that the mean target IPE index was *not* significantly < 0.

To analyze low-info choices, for each choice pair we selected the neurons that were recorded during at least one low-info choice and that were lateral habenula information-predictive neurons or that were putative dopamine neurons with positive information signals (defined as having higher activity on 100% info trials than forced 0% info trials, $P < 0.05$, rank-sum test).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997; 275:1593–9. [PubMed: 9054347]

2. Rescorla, RA.; Wagner, AR. Classical Conditioning II: Current Research and Theory. Black, AH.; Prokasy, WF., editors. Appleton Century Crofts; New York, New York: 1972. p. 64-99.

3. Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron. 2005; 47:129–41. [PubMed: 15996553]

4. Seo H, Lee D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. J Neurosci. 2007; 27:8366–77. [PubMed: 17670983]

5. Matsumoto M, Matsumoto K, Abe H, Tanaka K. Medial prefrontal cell activity signaling prediction errors of action values. Nat Neurosci. 2007; 10:647–56. [PubMed: 17450137]

6. Kim H, Sul JH, Huh N, Lee D, Jung MW. Role of striatum in updating values of chosen actions. J Neurosci. 2009; 29:14701–12. [PubMed: 19940165]

7. Oyama K, Hernadi I, Iijima T, Tsutsui K. Reward prediction error coding in dorsal striatal neurons. J Neurosci. 2010; 30:11447–57. [PubMed: 20739566]

8. Hong S, Hikosaka O. The globus pallidus sends reward-related signals to the lateral habenula. Neuron. 2008; 60:720–9. [PubMed: 19038227]

9. Matsumoto M, Hikosaka O. Lateral habenula as a source of negative reward signals in dopamine neurons. Nature. 2007; 447:1111–5. [PubMed: 17522629]

10. Matsumoto M, Hikosaka O. Representation of negative motivational value in the primate lateral habenula. Nat Neurosci. 2009; 12:77–84. [PubMed: 19043410]

11. Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev. 2002; 109:679–709. [PubMed: 12374324]

12. O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. Neuron. 2003; 38:329–37. [PubMed: 12718865]

13. McClure SM, Berns GS, Montague PR. Temporal prediction errors in a passive learning task activate human striatum. Neuron. 2003; 38:339–46. [PubMed: 12718866]

14. Salas R, Baldwin P, de Biasi M, Montague PR. BOLD Responses to Negative Reward Prediction Errors in Human Habenula. Front Hum Neurosci. 2010; 4:36. [PubMed: 20485575]

15. Chew SH, Ho JL. Hope: an empirical study of attitude toward the timing of uncertainty resolution. Journal of Risk and Uncertainty. 1994; 8:267–288.

16. Eliaz K, Schotter A. Experimental testing of intrinsic preferences for noninstrumental information. American Economic Review (Papers and Proceedings). 2007; 97:166–169.

17. Luhmann CC, Chun MM, Yi DJ, Lee D, Wang XJ. Neural dissociation of delay and uncertainty in inter-temporal choice. Journal of Neuroscience. 2008; 28:14459–14466. [PubMed: 19118180]

18. Prokasy WF Jr. The acquisition of observing responses in the absence of differential external reinforcement. J Comp Physiol Psychol. 1956; 49:131–4. [PubMed: 13319523]

19. Fantino, E. Handbook of operant behavior. Honig, WK.; Staddon, JER., editors. Prentice Hall; Englewood Cliffs, NJ: 1977.

20. Dinsmoor JA. Observing and conditioned reinforcement. The Behavioral and Brain Sciences. 1983; 6:693–728.

21. Daly, HB. Learning and Memory: The Behavioral and Biological Substrates. Gormezano, I.; Wasserman, EA., editors. L.E. Associates; 1992. p. 81-104.

22. Bromberg-Martin ES, Hikosaka O. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. Neuron. 2009; 63:119–126. [PubMed: 19607797]

23. Ward E. Acquisition and extinction of the observing response as a function of stimulus predictive validity. Psychon Sci. 1971; 24:139–141.

24. Hayden BY, Heilbronner SR, Platt ML. Ambiguity aversion in rhesus macaques. Front Neurosci. 2010; 4

25. Kreps DM, Porteus EL. Temporal resolution of uncertainty and dynamic choice theory. Econometrica. 1978; 46:185–200.

26. Wyckoff LB Jr. The role of observing responses in discrimination learning. Psychol Rev. 1952; 59:431–42. [PubMed: 13004146]

27. Hikosaka O. The habenula: from stress evasion to value-based decision-making. Nat Rev Neurosci. 2010; 11:503–13. [PubMed: 20559337]

28. Christoph GR, Leonzio RJ, Wilcox KS. Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. J Neurosci. 1986; 6:613–9. [PubMed: 3958786]

29. Ji H, Shepard PD. Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA(A) receptor-mediated mechanism. J Neurosci. 2007; 27:6923–30. [PubMed: 17596440]

30. Shumake J, Ilango A, Scheich H, Wetzel W, Ohl FW. Differential neuromodulation of acquisition and retrieval of avoidance learning by the lateral habenula and ventral tegmental area. J Neurosci. 2010; 30:5876–83. [PubMed: 20427648]

31. Friedman A, et al. Electrical stimulation of the lateral habenula produces an inhibitory effect on sucrose self-administration. Neuropharmacology. 2010

32. Doya K. Metalearning and neuromodulation. Neural Networks. 2002; 15:495–506. [PubMed: 12371507]

33. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. Nature. 2006; 441:876–9. [PubMed: 16778890]

34. Sutherland RJ, Nakajima S. Self-stimulation of the habenular complex in the rat. J Comp Physiol Psychol. 1981; 95:781–91. [PubMed: 6975784]

35. Lecourtier L, Defrancesco A, Moghaddam B. Differential tonic influence of lateral habenula on prefrontal cortex and nucleus accumbens dopamine release. Eur J Neurosci. 2008; 27:1755–62. [PubMed: 18380670]

36. Ikemoto S. Brain reward circuitry beyond the mesolimbic dopamine system: A neurobiological theory. Neurosci Biobehav Rev. 2010

37. Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature. 2009; 459:837–841. [PubMed: 19448610]

38. Joshua M, Adler A, Bergman H. The dynamics of dopamine in control of motor behavior. Curr Opin Neurobiol. 2009; 19:615–20. [PubMed: 19896833]

39. Bromberg-Martin ES, Matsumoto M, Hikosaka O. Dopamine in motivational control: rewarding, aversive, and alerting. Neuron. 2010; 68:815–834. [PubMed: 21144997]

40. Caplin A, Leahy J. Psychological expected utility theory and anticipatory feelings. The Quarterly Journal of Economics. 2001

41. Butko NJ, Movellan JR. Infomax control of eye movements. IEEE Transactions on Autonomous Mental Development. 2010; 2:91–107.

42. Wyckoff LB Jr. Toward a quantitative theory of secondary reinforcement. Psychol Rev. 1959; 66:68–78. [PubMed: 13645851]

43. Beierholm UR, Dayan P. Pavlovian-instrumental interaction in 'observing behavior'. PLoS Comput Biol. 2010; 6

44. Bromberg-Martin ES, Matsumoto M, Nakahara H, Hikosaka O. Multiple timescales of memory in lateral habenula and dopamine neurons. Neuron. 2010; 67:499–510. [PubMed: 20696385]

45. Nakamura K, Matsumoto M, Hikosaka O. Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. J Neurosci. 2008; 28:5331–43. [PubMed: 18480289]

46. Ranade SP, Mainen ZF. Transient firing of dorsal raphe neurons encodes diverse and specific sensory, motor, and reward events. J Neurophysiol. 2009; 102:3026–37. [PubMed: 19710375]

47. Redgrave P, Gurney K. The short-latency dopamine signal: a role in discovering novel actions? Nat Rev Neurosci. 2006; 7:967–975. [PubMed: 17115078]

48. Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H. Midbrain dopamine neurons encode decisions for future action. Nat Neurosci. 2006; 9:1057–63. [PubMed: 16862149]

49. Roesch MR, Calu DJ, Schoenbaum G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat Neurosci. 2007; 10:1615–1624. [PubMed: 18026098]

50. Houk, JC.; Adams, JL.; Barto, AG. Models of Information Processing in the Basal Ganglia. Houk, JC.; Davis, JL.; Beiser, DG., editors. MIT Press; Cambridge, MA: 1995. p. 249-274.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript
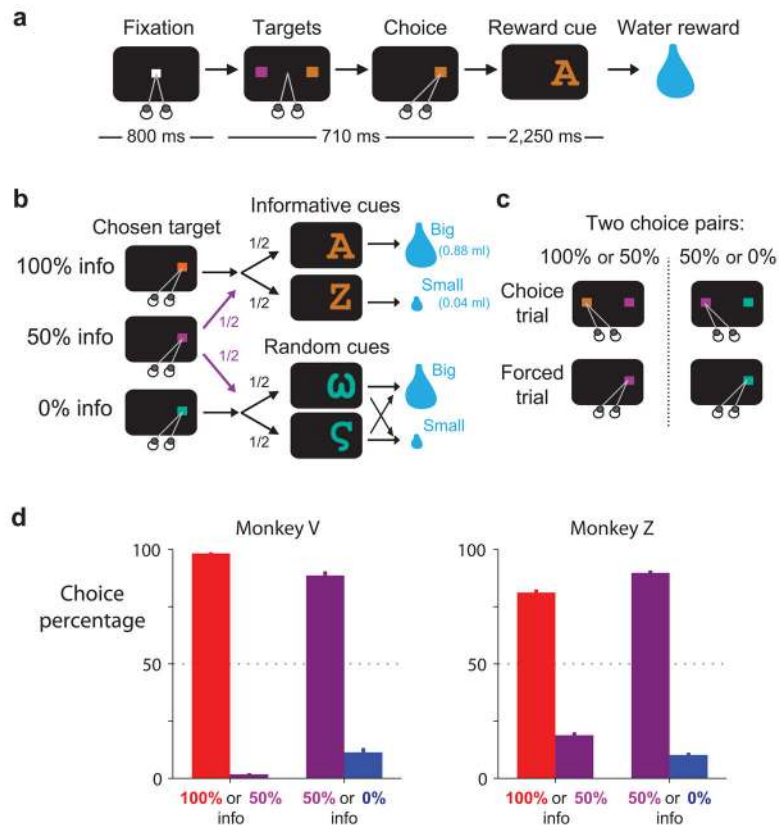
Figure 1.
Behavioral preference to view informative reward cues. (**a**) On each trial animals viewed a fixation point, used a saccadic eye movement to choose a colored visual target, viewed a visual cue, and received a big or small water reward. (**b**) The three potential targets led to informative reward cues with 100%, 50%, or 0% probability. (**c**) Targets were presented in two choice pairs, 100% versus 50% info and 50% versus 0% info. Choice trials were followed by forced trials to equalize exposure to the non-chosen option. (**d**) Animals expressed a strong preference to choose the target that led to a higher probability of viewing informative cues. Bars are mean choice percentages ±1 SE.
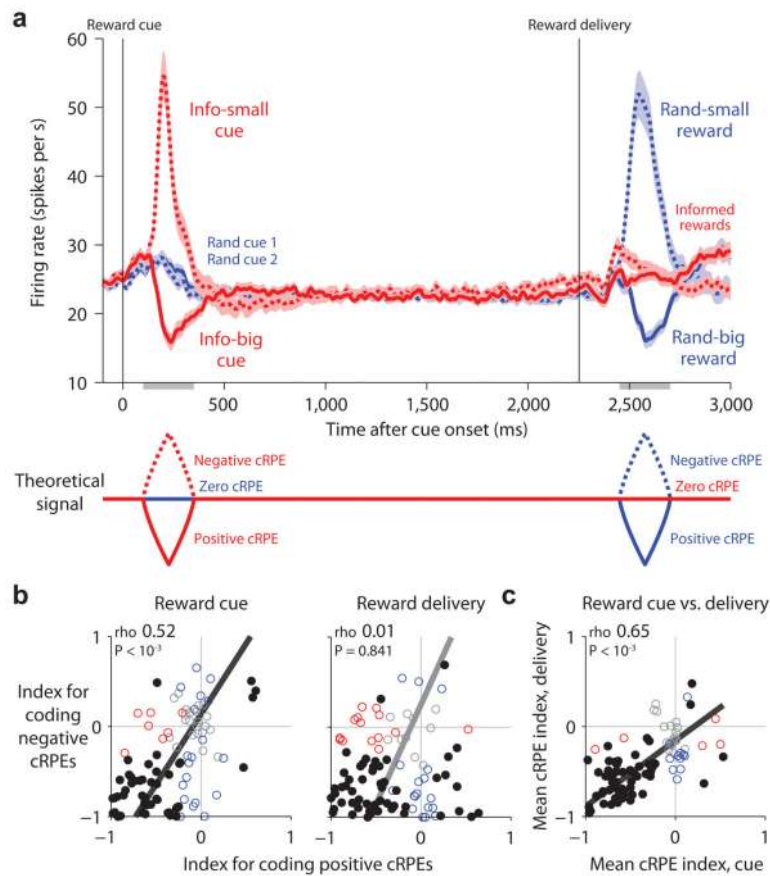
Figure 2.

Lateral habenula neurons transmit an inverted reward prediction error signal. (**a**) Average firing rate of lateral habenula neurons in response to the reward cues (left) and reward outcomes (right) resembled a theoretical inverted cRPE signal (bottom; generated from the model in Supplementary Fig. 1c using the same response windows as for the neural data). Left: activity aligned at cue onset for the informative cues (red lines; big-reward cue, solid; small-reward cue, dashed), and the two random cues (blue solid and dashed lines). Right: activity aligned at reward onset for informed rewards (red lines; big reward, solid; small reward, dashed) and randomized rewards (blue; big reward, solid; small reward, dashed). Activity was smoothed with a Gaussian kernel ($\sigma = 10$ ms). Shaded regions represent the mean firing rate ± 1 SE of the baseline-subtracted neural firing rate. Gray bars below the x-axis indicate the analysis windows. (**b**) Single-neuron cRPE indexes for responses to the cues (left) and reward delivery (right), calculated separately for positive cRPEs (x-axis) and negative cRPEs (y-axis). Colored dots indicate neurons with significant indexes along the x-axis (red), y-axis (blue), or both (black) ($P < 0.05$, rank-sum test). Text indicates rank correlation (rho) and its significance (permutation test); solid line indicates the best-fitting linear relationship using type 2 regression. (**c**) Single-neuron mean cRPE indexes for responses to the cues (x-axis) and reward delivery (y-axis). Same format as (b). Most neurons had consistent coding of cRPEs for both the cues and reward delivery.
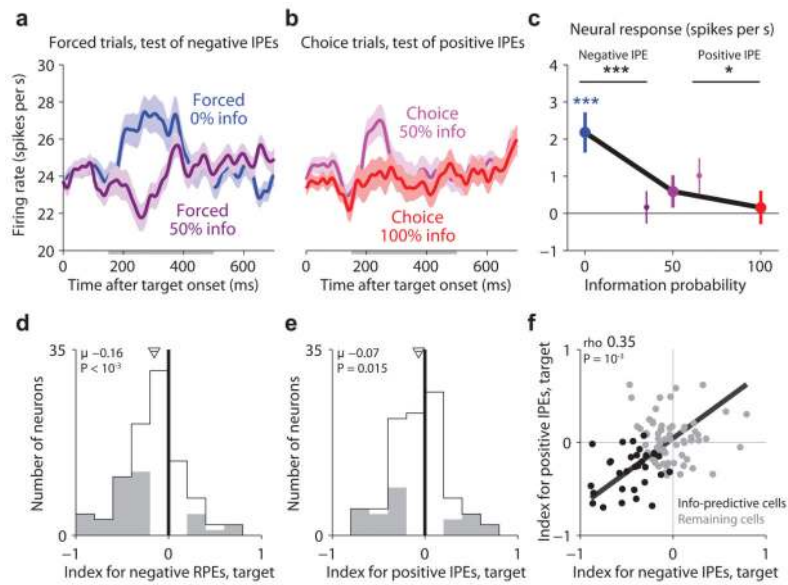
Figure 3.

Lateral habenula activity related to information prediction errors evoked by the information-predictive targets. (**a**) Average lateral habenula activity was higher on forced 0% info trials (blue) than forced 50% info trials (dark purple). Same format for smoothing and error regions as Fig. 2a. Gray bar below the x-axis indicates the analysis window. (**b**) Average lateral habenula activity was higher on choice trials when the animal chose 50% info in preference to 0% info (light purple) than when the animal chose 100% info in preference to 50% info (red). Same format as (a). (**c**) Mean baseline-subtracted activity in response to the target array on trials with 0%, 50%, and 100% info probability (blue, purple, red dots). Error bars are ±1 SE. Small data points next to the 50% info data point represent forced 50% info trials (dark purple, left) and choice 50% > 0% info trials (light purple, right). Colored asterisks are responses significantly different from baseline (*/**/*** for $P <$ 0.05/0.01/0.001, signed-rank test). Black asterisks indicate significant effects of information probability. (**d,e**) Single neuron indexes for coding negative IPEs (d, forced trials) and positive IPEs (e, choice trials). Gray indicates indexes significantly different from 0 ($P <$ 0.05, rank-sum test). Arrow and horizontal line indicate mean±SE, text indicates mean and significance (signed-rank test). (**f**) Correlation between indexes for negative and positive IPEs. Text indicates rank correlation and its significance (permutation test). Black dots are the "information-predictive neurons" (mean IPE index < 0, $P <$ 0.05, permutation test; n=30).
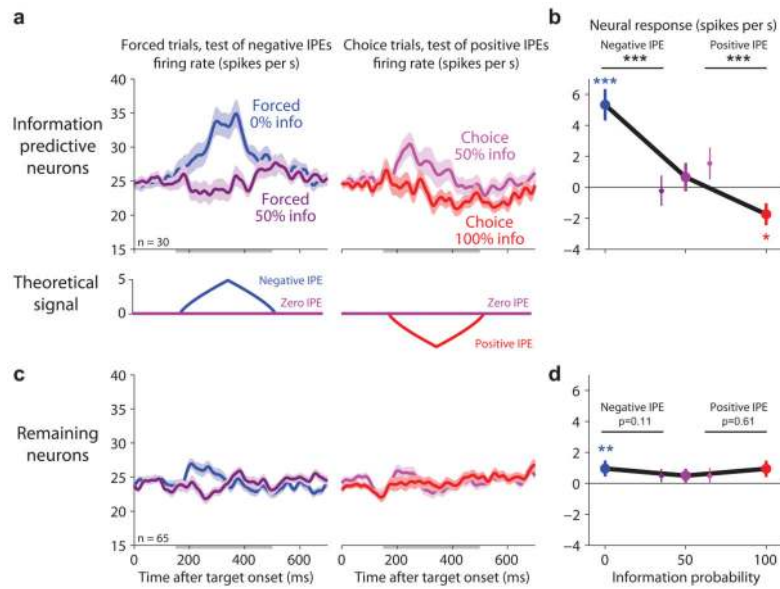
Figure 4.
Information-related signals are strongest in a subpopulation of neurons. (**a,b**) Average activity of the subpopulation of information-predictive neurons (top) resembles the theoretical inverted IPE signal (bottom, generated from the model in Supplementary Fig. 1b using the same response windows as for the neural data and plotting model rate on same scale as neural rate). Same format as Fig. 3a–c, but only showing activity from the information-predictive cells. These activity measurements were cross-validated to remove selection bias (Methods). (**c,d**) Average activity of the remaining neurons shows little or no sensitivity to IPEs. Same format as (a,b).
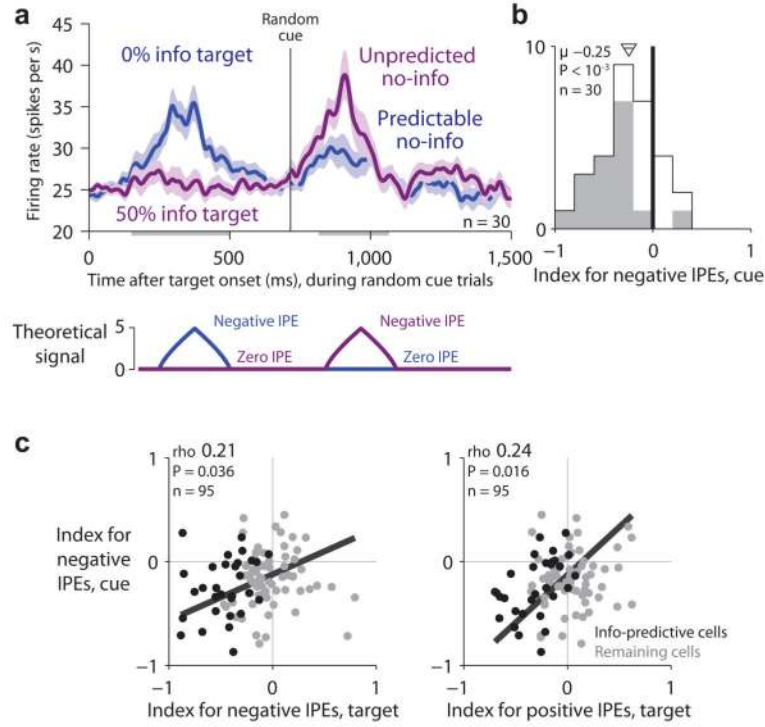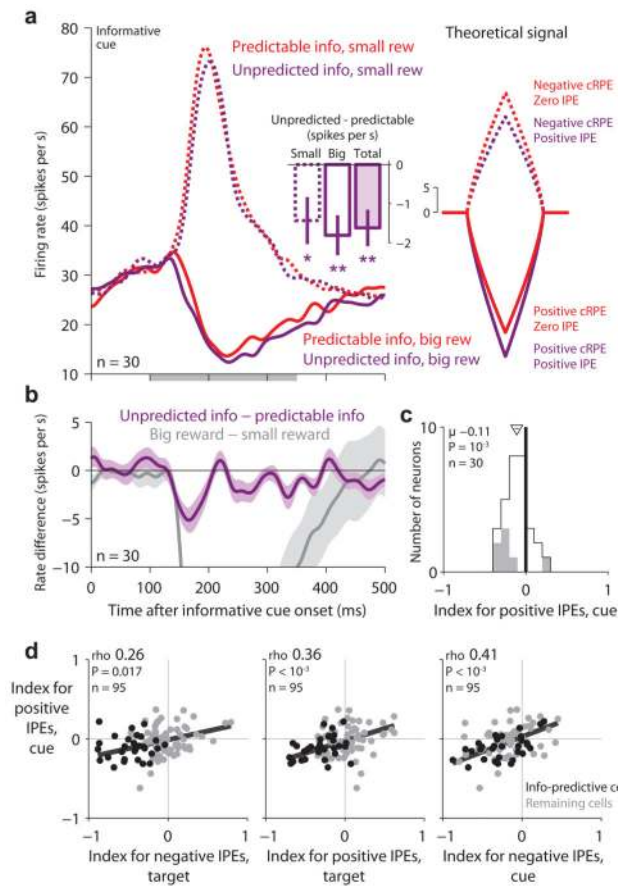
Figure 5.

Lateral habenula activity related to negative information prediction errors evoked by denial of reward information. (**a**) Average activity of information-predictive neurons on trials when random cues were presented (top) resembled the theoretical inverted IPE signal (bottom; model uses the same conventions as in Fig. 4). Activity is shown in response to the target array (left) and the onset of the random cues (right) for trials when the information probability was 0% (blue, 'predictable no-info') or 50% (purple, 'unpredicted no-info'). Same format for smoothing and error regions as Fig. 2a. This population was excited by unpredicted no-info. (**b**) Single-neuron indexes for coding negative IPEs in response to the random cues. Same format as Fig. 3d,e, but showing the subpopulation of information-predictive neurons. (**c**) Neurons with strong negative IPE signals in response to the random cues tended to have strong negative IPE signals (left) and positive IPE signals (right) in response to the targets. Same format as Fig. 3f.

Figure 6.
Lateral habenula activity related to positive information prediction errors evoked by delivery of reward information. (**a**) Average activity of information-predictive neurons on trials when informative cues were presented (left) resembled the theoretical combined inverted IPE +cRPE signal (right; model uses the same conventions as in Fig. 4). Activity is shown in response to the onset of the informative cues (right) for trials when the information probability was 100% (red, 'predictable info') or 50% (purple, 'unpredicted info') and when the informative cue indicated a big reward (solid lines) or small reward (dashed lines). This population had strong excitation or inhibition encoding cRPEs, and also had a lower firing rate on 'unpredicted info' than 'predictable info' trials (inset: difference in firing rate between 'unpredicted info' and 'predictable info', calculated separately for small-reward, big-reward, and all trials; error bars are ± 1 SE; */** for $P < 0.05/0.01$, signed-rank test). (**b**) Activity difference related to information probability (purple, unpredicted info – predictable info) and cued reward value (gray, big reward – small reward). Shaded area indicates ± 1 SE. (**c**) Single-neuron indexes for coding positive IPEs in response to the informative cues. Same format as Fig. 3d,e, but showing the subpopulation of information-predictive neurons. (**d**) Neurons with strong coding of positive IPEs evoked by the informative cues (y-axis) also tended to have strong coding of positive IPEs evoked by the targets (middle) and negative IPEs evoked by the random cues (left) and targets (right). Same format as Fig. 3f.
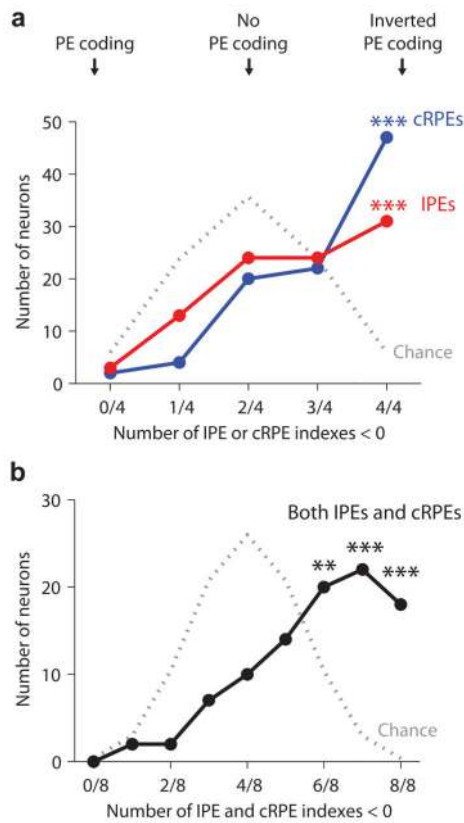
Figure 7.
Joint coding of IPEs and conventional RPEs in single neurons. (**a**) Histogram of all lateral habenula neurons sorted by their total number of IPE indexes (red) or cRPE indexes (blue) that were below zero, indicating coding of inverted prediction errors (counting all four cRPE indexes from Fig. 2b and all four IPE indexes from Figs. 3f,5b,6c). Gray dotted line indicates the null hypothesis that the indexes were randomly distributed above or below zero by chance. Asterisks indicate response patterns that occurred in more neurons than expected by chance (*/**/*** for $P < 0.05/0.01/0.001$, binomial test). (**b**) Same as (a) for the combined count of cRPE and IPE indexes that were below zero (considering all eight indexes). The most common patterns were to have six, seven, or eight indexes below zero, indicating inverted coding cRPEs and IPEs.
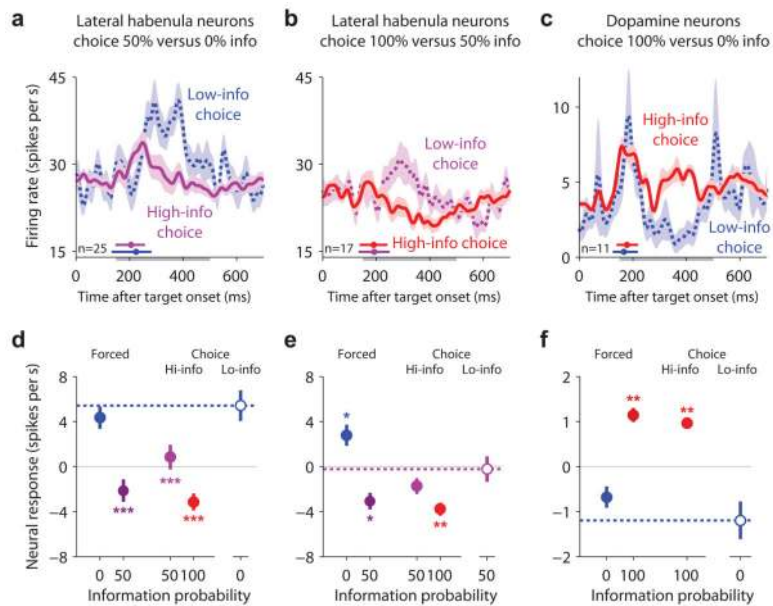
Figure 8.
Lateral habenula and dopamine neurons signal information probability reliably despite variable decisions. (**a**) Information-predictive lateral habenula neurons had higher activity when animals made low-info choices of 0% info > 50% info (dashed blue line) than when they made high-info choices of 50% info > 0% info (purple). Activity is shown for all neurons recorded during at least one choice of 0% info > 50% info. Same format for smoothing and error regions as Fig. 2a. The small colored circles above the x-axis indicate the median saccadic reaction time for each condition; the horizontal colored lines indicate the central 90% of the reaction time distribution. (**b**) Same as (a), for choices between 50% info (purple) vs. 100% info (red). (**c**) Same as (a,b), for putative dopamine neurons recorded during choices between 0% info (blue) vs. 100% info (red). (**d**) Mean baseline-subtracted activity during forced trials (left), high-info choice trials (middle), and low-info choice trials (right, open circle and dashed line), for information-predictive lateral habenula neurons recorded during at least one low-info choice of 0% info > 50% info. Error bars are ± 1 SE. Asterisks indicate responses that were significantly different from those on low-info choice trials (*/**/*** for $P < 0.05/0.01/0.001$, signed-rank test). (**e,f**) Same as (d), for the neurons shown in (b,c).