nature neuroscience

Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task

Adam P Steiner[1] & A David Redish[2]

Disappointment entails the recognition that one did not get the value expected. In contrast, regret entails recognition that an alternative (counterfactual) action would have produced a more valued outcome. In humans, the orbitofrontal cortex is active during expressions of regret, and humans with damage to the orbitofrontal cortex do not express regret. In rats and nonhuman primates, both the orbitofrontal cortex and the ventral striatum have been implicated in reward computations. We recorded neural ensembles from orbitofrontal cortex and ventral striatum in rats encountering wait or skip choices for delayed delivery of different flavors using an economic framework. Economically, encountering a high-cost choice after skipping a low-cost choice should induce regret. In these situations, rats looked backwards toward the lost option, cells within orbitofrontal cortex and ventral striatum represented the missed action, rats were more likely to wait for the long delay, and rats rushed through eating the food after that delay.

Regret is a universal human experience[1–5]. The experience of regret modifies future actions[1,4,6]. However, regret in other mammals has never been identified; it is not known whether nonhuman mammals are capable of experiencing regret. Although nonhuman animals cannot verbally express regret, one can create regret-inducing situations and ask whether those regret-inducing situations influence neurophysiological representations or behavior: do nonhuman animals demonstrate the neural correlates of regret in potential regret-inducing situations?

When evaluating the experience of regret, it is important to differentiate regret from disappointment[4,7,8]. Disappointment is the realization that a realized outcome is worse than expected[7,8]; regret is the realization that the worse than expected outcome is due to one's own mistaken action[1–3,9]. Disappointment can be differentiated from regret through differences in the recognition of alternatives[2,6,8,10]. Regret can be defined as the recognition that the option taken resulted in a worse outcome than an alternative option or action would have. The revaluation of the previous choice in context of the current choice is the economic foundation of regret[4,6].

Humans with damage to the orbitofrontal cortex (OFC) do not express regret[2], and functional magnetic resonance imaging experiments reveal activity in the orbitofrontal cortex during regret[1,11]. In rats and nonhuman primates, the OFC has been implicated in decision-making, particularly in the role of expectations of future reward and the complex calculations of inferred reward[12–17]. Orbitofrontal cortical neurons represent the chosen value of an expected future reward[14,18,19], and earlier research suggested that an intact OFC is critical for reversal learning[20,21] (recent evidence suggests that OFC may have a more specialized role and is not necessary for reversal learning, at least in primates[22]). Orbitofrontal cortex has been hypothesized to be critical for learning and decision-making[10,15,23,24], particularly in the evaluation of expected outcomes[14,25].

The ventral striatum (vStr) has also been implicated in evaluation of outcomes[26–29], particularly in evaluation during the process of decision-making[23,29,30]. Neural recordings vStr and OFC in rats have found representations of reward, value and prediction of expected value in both structures[12,25,29,31–33]. In the rat, lesion studies suggest that orbitofrontal cortex is necessary for recognition of reward-related changes that require inference, such as flavor and kind, while vStr is necessary for recognition of any changes that affect value[15,23]. In rats deliberating at choice points, vStr reward representations are transiently active before and during the reorientation process[29], but reward representations in OFC are only active after the reorientation process is complete[25].

We developed a neuroeconomic spatial decision-making task for rats, Restaurant Row, in which the rat encounters a serial sequence of take or skip choices. The Restaurant Row task consisted of an inner loop approximately 1 m in diameter, with four spokes radiating from the inner loop (**Fig. 1a**). At the end of each spoke, a feeder supplied a different flavor of food: banana, cherry, chocolate or unflavored. Flavor locations remained constant throughout the experiment. Rats were trained to run around the loop, making stay or skip decisions as they passed each spoke.

Zones were defined to differentiate each spoke. Upon entering each zone, rats encountered different offers of delays. Zone entries were defined entirely by the detected position of the rat's head and were not explicitly marked on the track. On entry into a zone, a tone sounded; the pitch of the tone indicated the delay the rat had to wait to receive a reward (higher pitch representing longer delay). As long as the rat stayed within the zone, the delay counted down, with each subsequent second indicated by a lower pitch tone. If the rat left the zone, the offer was rescinded: the countdown stopped, no sound was played and the rat's only option was to proceed on to the next spoke and the next zone.

The delays were independently selected pseudorandomly from a uniform distribution ranging from 1 to 30 s (for two of the rats) or 1 to 45 s (for the other two). The delay offered at each zone encounter was independent of that offered at other zones for that lap. When making

[1]Graduate Program in Neuroscience, University of Minnesota, Minneapolis, Minnesota, USA. [2]Department of Neuroscience, University of Minnesota, Minneapolis, Minnesota, USA. Correspondence should be addressed to A.D.R. (redish@umn.edu).

Received 3 March; accepted 14 May; published online 8 June 2014; doi:10.1038/nn.3740

a decision to stay or skip at a given zone (when offered a given delay), the only information the rat had was the flavor of the food offered (flavor locations remained constant throughout the experiment), the delay it would have to wait (delay signaled by pitch of the auditory cue) and the probability distribution of any future offers (offers were drawn from a uniform distribution of 1–30 s or 1–45 s).

Rats ran one 60-min session per day. This time limit meant that rats had a time budget of 60 min to spend foraging for food. Because the session was time-limited, the decision to stay or skip a zone was not independent of the other zones: waiting at one zone was time that could have been spent at another zone. An economically maximizing rat should distribute its time among the offers, waiting for valuable offers but skipping expensive offers. Assuming that an animal likes some flavors more than others, the economic value of an offer should depend on the delay offered and the animal's preferences.

RESULTS

Revealed preferences

We trained four rats on the Restaurant Row task (**Fig. 1**). Thresholds and preferences were determined by using an economic framework. All four rats showed similar behaviors in that they were likely to wait through the delay for delays less than a threshold, but unlikely to wait through the delay for delays greater than a threshold. When rats skipped an option, they left within the first ~5 s, independent of delay (**Supplementary Fig. 1**). The threshold between waiting and skipping tended to be different for the different flavors for a given rat (**Fig. 1b** and **Supplementary Fig. 1**). The fact that rats either stayed through the entire delay or left after a very stable 3 s implies that rats were not waiting for a specific delay cue but were making economic decisions based on the delay offered (**Supplementary Fig. 1**). These thresholds were consistent within a rat but differed among rats (**Fig. 1c–e**), indicating an underlying revealed, economic preference for each flavor of food that did not change across a session (**Supplementary Fig. 2**). There were no differences in reward handling between delays; rats generally waited 20–25 s after reward delivery before leaving for the next zone (**Supplementary Fig. 3**).

To directly test whether the rats were making economic decisions (comparing value and cost), we ran two of the rats, after completing the primary Restaurant Row experiment, on a variant of the task in which one reward site provided three times as much food as the other three sites. In this control task, rats were run in four 20-min blocks, so that each site could be the large reward site for one block. (The order of which reward site provided excess reward was varied pseudorandomly. Rats were removed to a nearby resting location for 1 min between blocks.) Rats were consistently willing to wait longer for more food (**Supplementary Fig. 4**). All results reported here except for those in **Supplementary Figure 4** are from the primary Restaurant Row experiment.

Reward responses

We recorded 951 neurons from orbitofrontal cortex (OFC) and 633 neurons from ventral striatum (vStr) (see **Supplementary Fig. 5** for recording locations). Neurons were identified as reward-responsive if their activity during the 3 s following reward delivery was significantly different ($P < 0.05$, Wilcoxon) than a bootstrapped ($n = 500$) sample of activity during 3-s windows taken randomly across the entire session[25,29]. 81% of OFC neurons responded to reward; 86% of vStr neurons responded to reward. Responses in both OFC and vStr often differentiated among the four reward sites (**Supplementary Figs. 6** and **7**).

Because responses differentiated among rewards, a decoding algorithm applied to these neural ensembles should be able to distinguish among the reward sites. We used a Bayesian decoding algorithm with a training set defined by the neuronal firing rate in the 3 s following delivery of reward (which we used to calculate p(spikes | reward)) or a training set defined by the neuronal firing rate in the 3 s following entry into a zone (which we used to calculate p(spikes | zone)). To provide a control for unrelated activity, we also included a fifth condition in our calculation, the average neuronal firing rate during times the rat was not in any countdown zone. Thus, the training set consisted of five expected firing rates: firing rate after reward receipt or zone entry (i) at banana, (ii) at cherry, (iii) at chocolate, (iv) at unflavored and (v) on the rest of the maze. From this training set, Bayesian decoding uses the population firing rate at a given time to derive the posterior probability of the representation p(reward | spikes) or p(zone | spikes). For simplicity, we will refer to these two measures as p(reward) and p(zone).

To pool data from all four sites, we categorized and rotated each reward site on the basis of the current position of the rat. This gave us four sites that progressed in a serial manner: the previous site, the

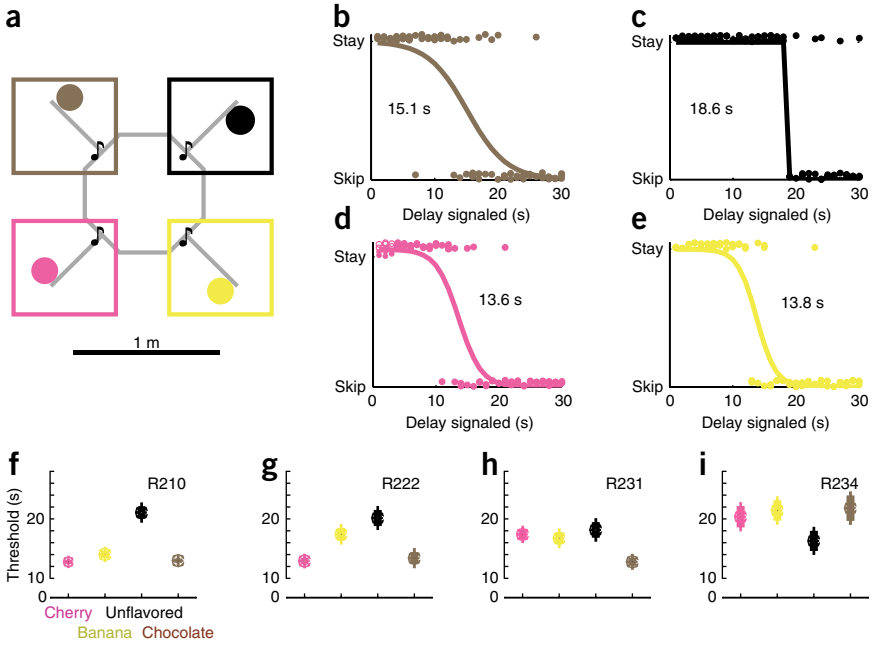


Figure 1 Restaurant Row and revealed preferences in rats. (**a**) The Restaurant Row task consisted of a central ring with four connected spokes leading to individual food flavors. Rats ran counterclockwise around the ring, encountering the four invisible zones (square boxes) sequentially. Color reflects flavor: pink, cherry; yellow, banana; black, unflavored (plain); brown, chocolate. (**b–e**) Rats typically waited through short delays but skipped long delays. Each panel shows the stay or go decisions for all encounters of a single rat running a single session (R210-2011-02-02). A small vertical jitter has been added for display purposes. Thresholds were fit as described in the Online Methods. (**f–i**) Rats R210 (**f**), R222 (**g**), R231 (**h**) and R234 (**i**) each demonstrated a different revealed preference that was consistent within a rat across all sessions but differed among rats. Thresholds were fit for each flavor for each session. Each panel shows the mean fit threshold for a given rat, with s.e.m. over sessions. An important consideration is to control for the possibility that rats were waiting for a specific cue before leaving the zone.

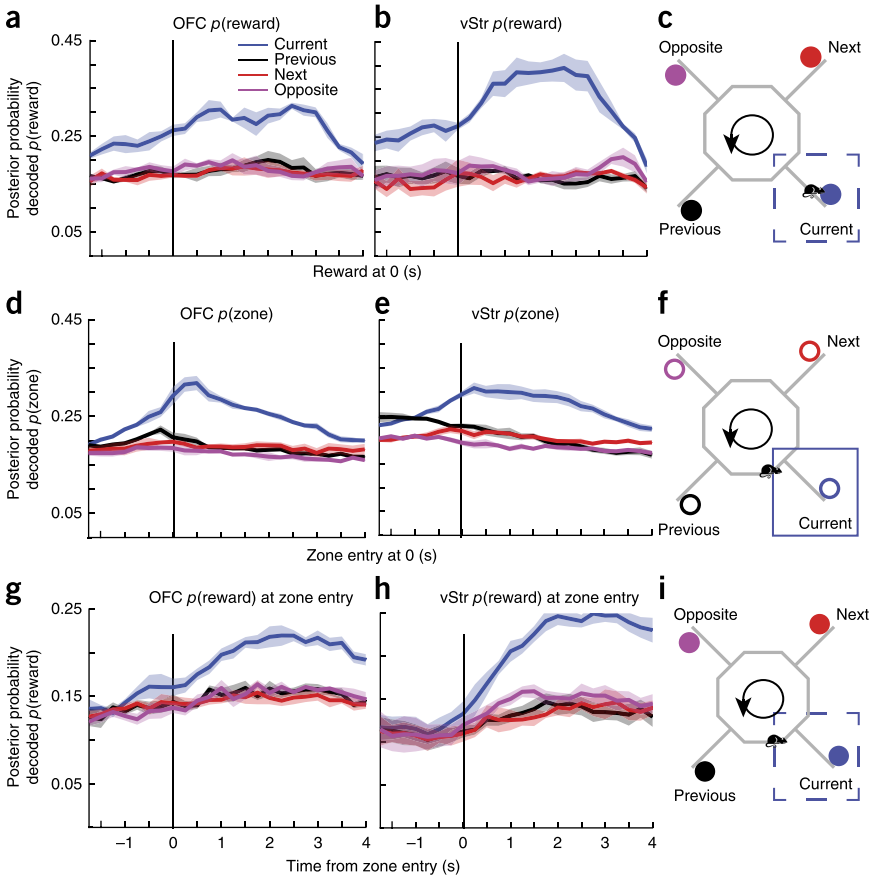


Figure 2 Ensembles in OFC and vStr represent the current reward and the current zone. (**a**,**b**) *p*(reward) at each reward for OFC (**a**) and vStr (**b**), defining the training set for decoding as activity at reward delivery and the test set as activity at each moment surrounding reward delivery (shaded area, s.e.m.). The neural ensemble decoded the current reward reliably (distribution of current reward was determined to be significantly different, empirical cumulative distribution function, significant at $\alpha = 0.05$). *p*(reward) is the posterior probability indicating the likelihood of representing a given reward flavor as calculated by the Bayesian decoding. (**c**) For **a**,**b**, the training set is the reward types and the test set is activity when the rat receives reward. Rat icon indicates that decoding aligned to reward delivery (when the rat is already at feeders). Filled-circle feeder locations indicate that the training set for the decoder is based on responses to reward delivery. Dashed lines indicate zone location. (**d**,**e**) *p*(zone) at each zone for OFC (**d**) and vStr (**e**), defining the training set for decoding as neuronal activity at zone entry and the test set as neuronal activity at each moment surrounding zone entry. The neural ensemble decoded the current zone reliably. *p*(zone) is the posterior probability indicating representation of a given zone entry as calculated by Bayesian decoding. (**f**) For **d**,**e**, the training set is zone entry and the test set is neuronal activity when the rat enters the zone, triggering the cue that signals the delay. Rat icon indicates that decoding is aligned to zone entry. Solid box indicates that the training set for the decoder is based on responses to zone entry. Open circles indicate reward locations. (**g**,**h**) *p*(reward) at each zone for OFC (**g**) and vStr (**h**), defining the training set for decoding as neuronal activity at reward delivery and the test set as neuronal activity at each moment surrounding zone entry. The neural ensemble at time of zone entry decoded the current reward type reliably. (**i**) For **g**,**h** the training set is the reward flavor and the test set is neuronal activity when the rat enters the zone, triggering the cue (tone). Rat icon indicates that decoding is aligned to zone-entry, as in **f**. Filled circles indicate that the training set is based on responses to reward-delivery, as in **c**. Dashed lines indicate zone location.

current site, the next site and the opposite site (**Fig. 2c**). All analyses were based on this categorization. All analyses used a leave-one-out approach so that the encounter being decoded was not included in the definition of the training set.

Both OFC (**Fig. 2a**) and vStr (**Fig. 2b**) were capable of reliably distinguishing between the current reward site (**Fig. 2c**) and the other sites (**Supplementary Fig. 8**). Shuffling the interspike intervals of the cells removed all of these effects. *p*(reward) and *p*(zone) calculated from shuffled data were consistently 0.14 (**Supplementary Fig. 8**).

Zone entry responses

Previous research has suggested that, in simple association tasks in which cues predict reward, both OFC and vStr cells respond to cues predictive of reward[12,15,24,28,31,34,35]. Both OFC and vStr neural ensembles distinguished the different zones both at the time of reward (**Fig. 2a–c**) and at the time of entering the zone (cue onset) (**Fig. 2d–f**) (single-cell differentiation, **Supplementary Figs. 6** and **7**; decoding differentiation, **Supplementary Figs. 8** and **9**). These representations were related; neural activity in OFC and vStr also predicted the reward type of the current zone during zone entry (cue onset) (**Fig. 2g–i** and **Supplementary Fig. 10**). Shuffling the interspike intervals eliminated these effects (**Supplementary Fig. 11**).

Both OFC and vStr responded strongly under conditions in which the rat determined the cost to be worth staying (**Fig. 3**) (for example, when the delay was below threshold; **Fig. 3a**,**b**). In contrast, neither structure represented expectations of reward under conditions in which the rat determined the cost to not be worth staying (that is, skips, when delay was above threshold; **Fig. 3c**,**d**). This suggests that these structures were indicating expected value, and predicting future actions. To directly test this hypothesis, we compared reward-related decoding when the rat encountered a delay near threshold (threshold ±2 s) and either stayed to sample the feeder (**Supplementary Fig. 12a**,**b**) or skipped to proceed to the next reward option (**Supplementary Fig. 12c**,**d**). When the rat stayed (waiting for a reward), both OFC and vStr increased their representations of the current reward at the time of zone entry. In contrast, when the rat skipped the current reward, neither OFC nor vStr reliably represented the current reward or zone. Shuffling the interspike intervals of the cells removed these effects (**Supplementary Fig. 13**).

Regret

Regret entails the recognition that one has made a mistake: that an alternative action would have been a better option to take[4,6]. As noted above, a regret-inducing situation requires that two properties be satisfied: (i) the undesirable outcome should be a result of the agent's previous action, and (ii) following the selection of an option, the outcome (value) of all options needs to be known, including the outcome (value) of the unselected options. Our task and behavior satisfies these conditions. Because the rats were time-limited on the Restaurant Row task, encountering a high-cost delay after not waiting through a low-cost delay means that skipping the low-cost delay was a particularly expensive missed opportunity.

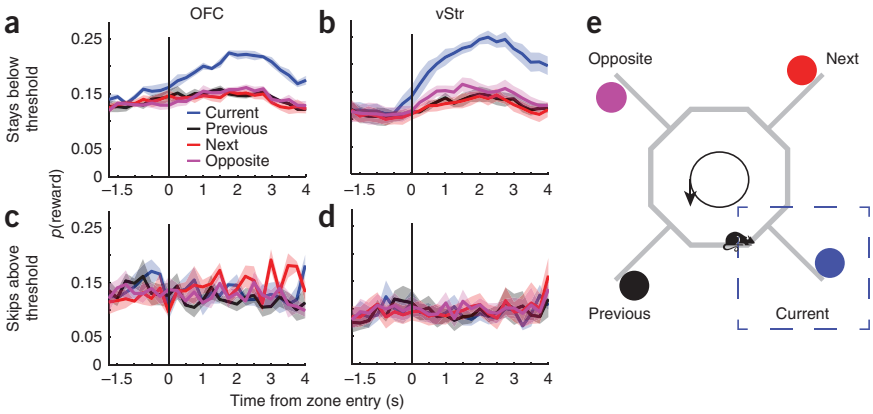


Figure 3 Representations of expected reward as a function of delay and threshold. To determine whether OFC and vStr signals predicted behavior at time of zone-entry, we measured *p*(reward) at each zone for all offers above and below the threshold for a given rat for a given flavor-reward site (shaded area, s.e.m.). (**a**,**b**) Low-cost offers in which the rat waited through the delay (distribution of current reward was determined to be significantly different, empirical cumulative distribution function, significant at $\alpha = 0.05$). (**c**,**d**) High-cost offers in which the rat skipped out and did not wait through the full delay. (**a**,**c**) OFC. (**b**,**d**) vStr. (**e**) This decoding operation was based on a training set at the reward but a test set at zone entry.

In the Restaurant Row task, a rat would sometimes skip an offer that was less than that rat's threshold for that flavor on that day and then encounter an offer at the subsequent site that was greater than that rat's threshold for that flavor on that day. Because the delay is a cost and value is matched (by definition) at threshold, this sequence is one in which the rat skipped a low-cost offer, only to find itself faced with a high-cost offer. From the economic and psychology literature, we can identify these sequences as potential 'regret-inducing' situations[4,6]. We can compare these conditions to control conditions that we would expect to induce disappointment rather than regret.

Literature suggests that during regret, there should be manifest changes in the rat's behavior and neurophysiology that reflect recognition of the missed opportunity, as well as subsequent behavioral choices that one might not have made normally. Theoretically, the key to regret is a representation of the action not taken[3,5,9,36,37]. This implies that there should be representations of the previous choice during the regret-inducing situations, particularly in contrast to control conditions that are merely disappointing.

Thus, we define a regret-inducing situation as one in which (i) the rat skipped a low-cost/high-value reward (delay less than measured threshold for that flavor for that day), and then (ii) the rat encountered a high-cost/low-value reward (delay greater than measured threshold for that flavor for that day). In this situation, the rat has made an economic mistake: if it had taken a different action (waited for that previous reward), it would have had a more valuable session. For consistency, we will refer to the opportunity in situation (i) as the previous zone or previous reward and the opportunity in situation (ii) as the current zone or current reward.

As noted above, one needs to differentiate regret-inducing sequences from sequences that are merely disappointing. By definition, a disappointing sequence occurs when one encounters a situation that is worse than expected, but not as a result of one's own agency. There are two controls that need to be taken into account, a control for the sequence of offers (control 1) and a control for the rat's actions (control 2).

To control for the sequence of offers, we took sequences in which the rat encountered the same sequence of offers, but took (stayed for) the first offer. This matched control should only induce disappointment (worse than expected, but not the fault of the rat)[7,8]. Control 1 differs from the regret-inducing situation only in that the rat took the previous offer rather than skipping it. In summary, control 1 was defined as situations in which the delay at the previous zone was below threshold and the rat waited for reward, followed by an encounter at the current zone such that delay was above threshold. In this situation, the rat did not make a mistake (as it waited for reward at the previous zone); the delay at the current zone was merely worse than the rat was willing to wait for, making the rat (presumably) disappointed. Control 1 controls for the sequence observed by the rat.

To control for the rat's actions, we took sequences in which the offer at the previous zone was greater than threshold (and skipped) and, again, the rat encountered a higher-than-threshold offer at the current zone. In this second control condition, the rat skipped the previous offer, but that was the 'correct' action to take, as the previous offer was above threshold. This second control condition should also induce disappointment because the rat has encountered two high-cost offers in a row. But this second control condition should not induce regret, because the rat's actions were consistent with its revealed preferences. Control 2 differs from the regret-inducing situation only in that the delay at the previous offer was above rather than below threshold. In summary, control 2 was defined as situations in which the delay at the previous zone was above threshold, followed by an encounter at the current zone such that delay was above threshold. In this situation, the rat did not make a mistake (since it skipped a high-cost delay at the previous zone), but the delay at the current zone was worse than the rat was willing to wait for, making the rat (presumably) disappointed. Control 2 controls for the reward sequence seen by the rat.

Potential regret and control instances were found within each session by comparing the delays at each of the zones to the threshold of that zone for that rat for that day. Regret instances and control instances were evenly distributed throughout each session across all rats. The distribution of the high-cost offers at the current zone did not differ between the potential regret-inducing sequences and matched controls (**Supplementary Fig. 14**).

Behaviorally, rats paused and looked backwards toward the previous option upon encountering a potentially regret-inducing sequence, but they did not do so in either control condition (**Fig. 4**). We identified pause-and-look events as points at which the rat's path showed high curvature and derived an orientation (see Online Methods). During potential regret-inducing sequences, rats were more likely to look backwards toward the previous option (**Fig. 4d**) than during either of the matched control conditions (**Fig. 4e**,**f**) ($P = 0.00042$, Watson's circular U test). In the first control condition (where the rat took a good offer and then encountered a bad offer), the rat tended to look toward the current spoke (zone) but then skip it and go on to the next zone. In the second control condition (where the rat encountered two bad offers in a row), the rat tended to look toward the next zone. Thus, there was a behavioral difference, implying that the rats recognized these three situations differently.

During potential regret instances, individual reward-responsive neurons in OFC and vStr showed activity patterns more consistent with the previous reward than the current one (**Fig. 5**). Neural activity peaked immediately after the start of the look back toward the previously

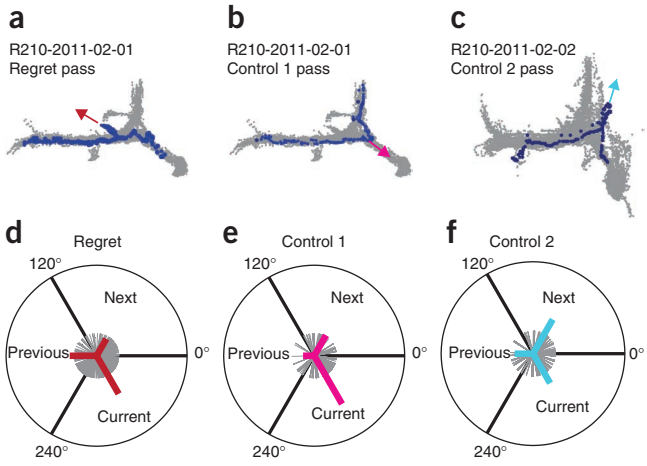


Figure 4 Behavioral responses in regret-inducing and control situations. All passes were rotated so as to align on entry into a current zone. Orientation was measured using the curvature measure as per the Online Methods. (**a**–**c**) Examples of approaches for each of the three conditions: regret-inducing, control 1 (same sequence but rat took previous option) and control 2 (two long delays in a row). Gray dots show all behavioral tracking samples from the example session. Blue dots show the current path taken in each example. The colors of the arrows correspond to the matching circular vector plots. Arrow directions indicate empirically determined curvature direction. In a regret-inducing example (**a**), when the rat entered the zone, he paused and looked backwards toward the previous zone. In a control 1 example (**b**), the rat looked toward the current reward spoke but proceeded on to the next zone. In a control 2 example (**c**), the rat looked toward the next zone but turned back toward the current reward. (**d**–**f**) Summary statistics. The first reorientation event was measured as per the Online Methods. Gray traces show all pausing reorientations over all instances in that condition. Heavy line shows vector average in each 120° arc. In regret-inducing conditions (**d**), rats tended to orient toward the previous zone or current spoke. In control 1 conditions (**e**), rats tended to orient only toward the current spoke. In control 2 conditions (**f**), rats tended to orient toward the next zone. The distributions in **d**–**f** were significantly different from each other (Watson's circular *U*; see text).

skipped, low-cost reward. To quantify these changes in representation during regret-inducing situations and disappointment-inducing controls, we examined the population dynamics using a Bayesian decoding algorithm. Population decoding analyses offer insight into the dynamics of neural population. Ensemble activity more accurately represents the dynamics of the entire population compared to that of a single cell. To determine the neural population representation during these situations, we measured the Bayesian representations of p(reward) and p(zone) from the ensemble including all cells.

While our first inclination was to look for representations of the missed reward, human subjects self-report that they regret actions taken or not more than they do missed outcomes[3,5,9,36]. We did find a weak representation of the missed reward (not significant; OFC, $P = 0.006$ after taking four potential decoding signals into account (multiple comparisons); vStr, $P = 0.0006$ after taking four potential decoding signals into account (multiple comparisons); see **Supplementary Fig. 15**). However, we found that there were strong representations of the previous decision point (p(zone)) that were significantly different from those of all other zones (outside the 95% confidence

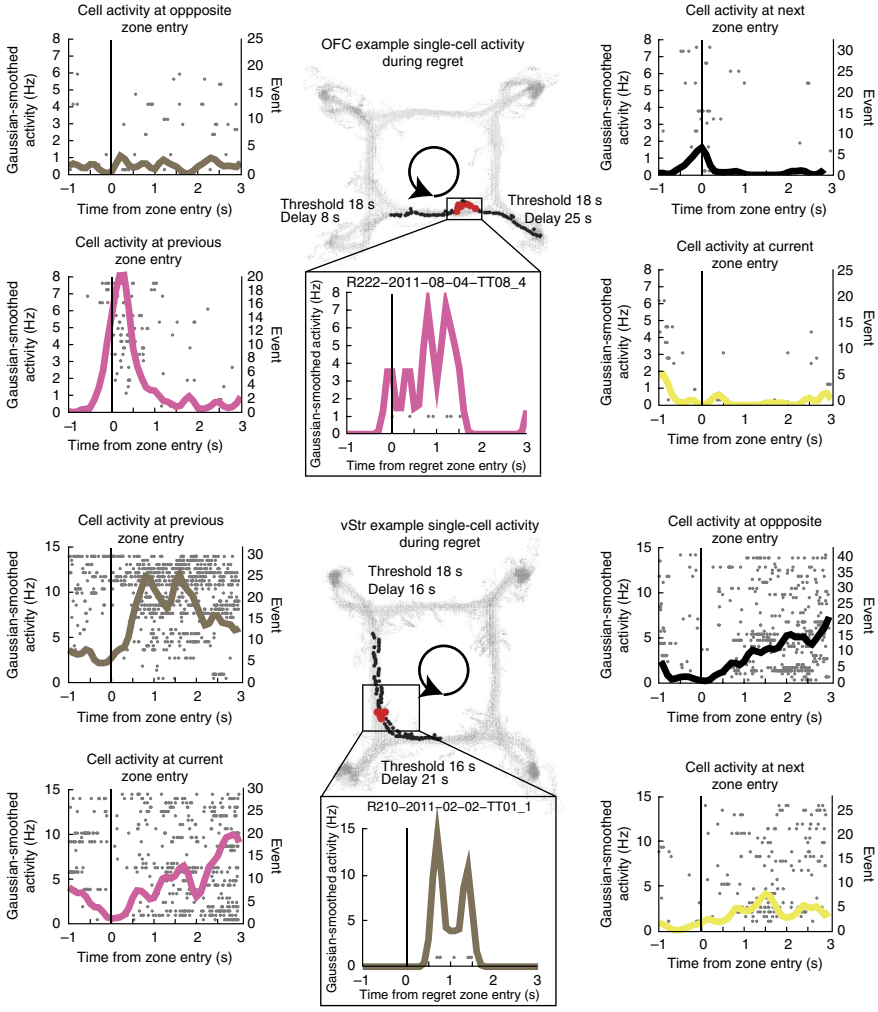


Figure 5 Single reward cells in OFC and vStr during regret-inducing situations. Top: OFC example cell during regret-inducing situation. Gray dots represent individual spikes. Solid colored lines indicate Gaussian-smoothed activity, Gaussian width $\sigma = 50$ ms. Black, unflavored pellets; pink, cherry flavored; yellow, banana flavored; brown, chocolate flavored. Black dots in the center panel represent position of the animal in this example lap during this instance. Red dots show position of the animal when the cell in question fired spikes. The rat traveled in a counterclockwise direction. The maze has been aligned so that the current zone is represented by the bottom right zone. This particular cell responded most to entry into the cherry reward zone, little to entry into the banana reward zone. When the rat skipped a low-cost cherry zone opportunity and encountered a high-cost banana zone opportunity, the rat looked back toward the previous reward, and the activity of the cell approximated that of the cherry zone-entry response. Bottom (display same as top panel): vStr example cell during a regret-inducing situation after skipping the chocolate reward zone and arriving at the cherry reward zone.

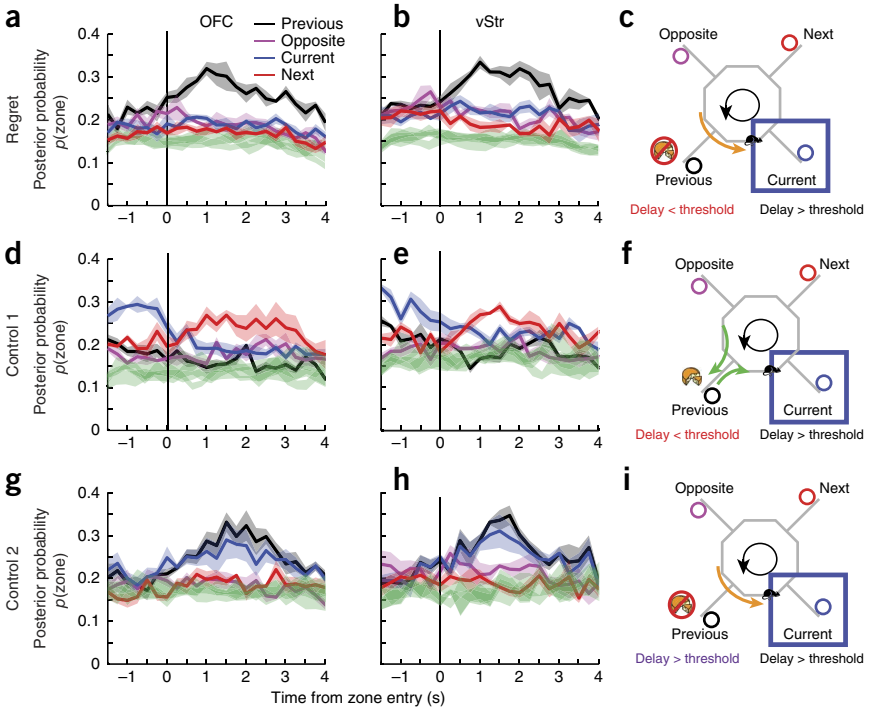


Figure 6 Neural representations in OFC and vStr represent the previous zone during behavioral regret instances. (**a**,**b**) In regret-inducing conditions, the *p*(zone) representation of the previous encounter was high after zone entry into the current zone for both OFC (**a**) and vStr (**b**) (shaded areas, s.e.m.). Green traces show decoding using shuffled inter-stimulus intervals. Decoding to the previous zone was significantly different from all other conditions, even after controlling for multiple comparisons (ANOVA; OFC, $P << 0.001$; vStr, $P << 0.001$; distribution significantly different as determined by empirical cumulative distribution function, significant at $\alpha = 0.05$). (**c**) The conditions being decoded in **a**,**b**: the rat has skipped the previous offer, even though the delay was less than threshold for that restaurant, and has now encountered a delay greater than threshold for the current restaurant. (**d**–**f**) In the control 1 condition, the *p*(zone) representation of the current zone increased until the rat heard the cue indicating a long delay, at which time the representation changed to reflect the next zone. In control 1, *p*(zone) representations of the current and next zones were significantly different from the other zones (ANOVA; vStr, $P << 0.001$; OFC, $P << 0.001$), although they were not different from each other after controlling for multiple comparisons (ANOVA; vStr, $P = 0.074$; OFC, $P = 0.619$). OFC (**d**), vStr (**e**) and cartoon indicating condition (**f**). (**g**–**i**) In the control 2 condition, the *p*(zone) representations of both the current and previous zones increased when the rat heard the cue indicating a long delay (compared to other zones, ANOVA; OFC, $P << 0.001$; vStr, $P << 0.001$). OFC (**g**), vStr (**h**) and cartoon indicating condition (**i**). Decodings to the current and previous zones in control 2 were not significantly different from each other (ANOVA; OFC, $P = 0.509$; vStr, $P = 0.268$).

interval as determined by empirical cumulative distribution function) (**Fig. 6a**–**c**). This differentiation of the previous zone was not observed in either control condition. In the first control (same sequence), both OFC and vStr demonstrated increased representations of the next zone (**Fig. 6d**–**f**). By definition, these control instances were high-cost encounters with the current reward site (for example, encountering a delay above threshold) and thus the rats were likely to skip them. In the second control condition (two bad offers), the representations of both the current and previous zones increased and were different from the representations of other rewards (**Fig. 6g**–**i**). However, this response was markedly different from that seen during potential regret instances, as the increase in representation of the previous zone could not be differentiated from the increase in representation of the current zone. Shuffling interspike intervals eliminated all of these effects (**Fig. 6**). Other, more positive situations (rejecting a low-cost previous offer and then encountering a low-cost offer or rejecting a high-cost offer and then encountering a low-cost offer) both led to strong representations of the current zone (**Supplementary Fig. 16**). In addition, when rats stayed for an above-threshold delay but then encountered a below-threshold delay (which could be described as a potential regret-inducing condition), we again found increased representations of the previous zone (**Supplementary Fig. 17**). The representations of the previous zone in this condition were smaller when compared to the previously described regret-inducing condition.

Thus, the rats showed different behaviors and different neurophysiological representations during regret-inducing situations, both of which reflected the information processing we would expect to see during regret. As noted above, an important role of regret in decision-making is that it changes subsequent decisions[37–39]. Consistent with this hypothesis, we found that rats were more likely to stay at the high-cost option in a regret-inducing situation than under either control condition (versus first control condition, $P = 0.01$; versus second control condition, $P = 0.06$; Wilcoxon, **Fig. 7a**). In addition, rats spent less time eating the food before proceeding on to the next reward site following regret-inducing situations as compared to non-regret situations (**Fig. 7b**). The handling time distributions were significantly different (regret condition versus all non-regret conditions, Wilcoxon, $P << 0.001$). After waiting for food through a long delay in a regret-inducing situation, rats rushed through eating and quickly went on to their next encounter.

The hypothesis that the neural representation of the previous zone reflects some information processing related to regret implies that

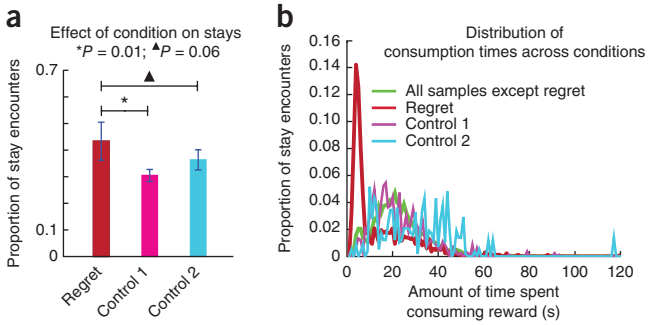


Figure 7 Behavioral changes following potential regret instances. (**a**) Comparing the proportion of stays to skips during each condition revealed that rats were more willing to wait for a reward following regret-inducing instances than control 1 instances (Wilcoxon, $*P = 0.01$) or control 2 instances (Wilcoxon, $▲P = 0.06$). (**b**) Rats spent less time consuming reward during regret than during non-regret instances. Typical handling time mean = 25.3 s, s.d. = 12.2 s; regret handling time mean = 15.2 s, s.d. 14.2 s. Control handling times were distributed the same as all non-regret handling times.

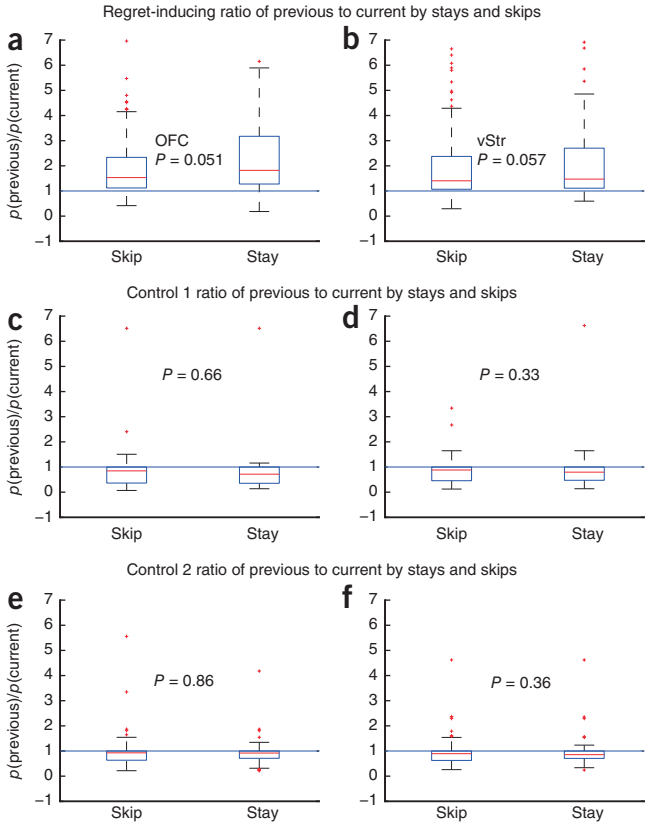


Figure 8 Behavioral and neurophysiological correspondences during regret. To determine whether the representations of previous reward were different when the rat chose to stay at the high-delay (high-cost) current zone, we measured the ratio between the *p*(zone) representation of the previous zone against the *p*(zone) representation of the current zone from 0 to 3 s following zone entry for all conditions in the event that the rat skipped or stayed. Each panel shows a box plot of the distribution of $p(\text{zone}_{\text{previous}})/p(\text{zone}_{\text{current}})$ ratios divided between stays and skips. Box limits are 25th and 75th percentiles, whiskers extend to data not considered outliers and outliers are plotted separately. (**a**) $p(\text{zone}_{\text{previous}})/p(\text{zone}_{\text{current}})$ ratios from OFC ensembles during regret-inducing conditions. (**b**) $p(\text{zone}_{\text{previous}})/p(\text{zone}_{\text{current}})$ ratios from vStr ensembles during regret-inducing conditions. (**c**,**d**) During control 1 conditions. (**e**,**f**) During control 2 conditions. Following regret inducing instances, when rats were more willing to wait for reward, $p(\text{zone}_{\text{previous}})$ was greater than $p(\text{zone}_{\text{current}})$.

there should be a relationship between that representation of the previous zone and the rat's subsequent actions. The hypothesis predicts that a stronger representation of the previous zone would lead to an increased likelihood of taking the high-cost (current) offer. To determine whether there was a relationship between a rat's willingness to take the high-cost offer and the neurophysiological representations, we compared the ratio of representations of the previous and the current zones and categorized these representations by stay or skip decisions at the current zone. This ratio was increased when the rat decided to stay, but only in the regret-inducing situations (**Fig. 8**). The ratio was unrelated to the decision to stay in the two control conditions. In regret-inducing situations, rats were more willing to stay on trials in which they showed an increased representation of the previous zone relative to the current zone.

DISCUSSION

Regret is the introspective recognition that a previously chosen action led to a less desirable outcome than an alternative action would have. The two keys to identifying regret are value and agency. The Restaurant Row task, in which rats made economic (value-related, cost-dependent) decisions allowed us to identify potentially regret-inducing situations. First, the Restaurant Row task was an economic task, in which rats revealed economic preferences just as human and nonhuman primates do[14,40,41]. Second, because the rats had a limited time budget, encountering a bad (above-threshold) offer after skipping a good (below-threshold) offer meant that the rat had missed an opportunity. By standard economic and psychological definitions, this sequence should induce regret[4,6,9]. We were able to identify two matched sets of controls that should induce disappointment but not regret: (1) situations in which the rat encountered a similar sequence of offers but took the previous low-cost option and (2) situations in which the rat encountered two above-threshold offers and skipped the previous high-cost option.

Our data indicate that behavioral and neurophysiological differences between the potential regret-inducing situations and the controls were consistent with a hypothesis that the rats were expressing something akin to human regret. During the regret-inducing situation, rats looked backwards toward the previous (missed) goal and the OFC and vStr were more likely to represent that previous goal. After it, rats were more likely to wait out the (current) high-cost offer, and they rushed through handling their reward when they did. Interestingly, we found that the neurophysiological representations of counterfactual information in the regret-inducing situation were more strongly related to the missed action (activity when the action was taken, measured by *p*(zone)) than to the missed outcome (activity when the reward was received, measured by *p*(reward)). This is consistent with data that humans express more regret about the actions taken (or not taken) than about the missed outcomes[3,5,9,36,37].

The Restaurant Row task had three features that made it particularly well suited to the identification of regret. First, it is an economic task on which rats reveal preferences. Second, the inclusion of four 'restaurants' allowed us to differentiate a general representation of other rewards from a specific representation of the mistaken choice. We found a clear and significant representation of the previous (lost) zone, but not the next or opposite zones. Third, the Restaurant Row task separates the choice of waiting (staying) or going (skipping) from reward receipt. This separation allowed us to differentiate the regret-induced representation of the previous (lost) reward (a small effect) from the regret-induced representation of the previous (mistaken) action (a large effect). Regret is more about the things you did or did not do than about the rewards you lost[5,9,36].

Previous evidence indicates that rats can combine information to form an expectation of a novel reward (imagining a particular outcome) and that both OFC[16,17,42] and vStr (if a model in the evaluation steps of the task exists) contribute to this process[23,24]. Our data indicate that violation of an expectation initiates a retrospective comparison (regretting a missed opportunity). As with the prospective calculation of expectation, this retrospective calculation of expectation influences future behavior: rats are more willing to wait for reward after a regret instance. These two processes, the act of imagining future outcomes and the process of regretting previous poor choices, are both necessary to modify future decisions to maximize reward. While some evidence suggests that OFC represents economic value[14], the representation of regret is more consistent with the hypothesis that OFC encodes the outcome parameters of the current, expected or imagined state[15–17,23]. The data presented here are also consistent with the essential role of OFC in proper credit assignment[43–45]. Previous studies have identified potential representations of the counterfactual could-have-been-chosen option in rats[25], monkeys[19] and humans[11]. In humans, representations of the value of the alternative outcome increase activity in OFC as measured by functional magnetic resonance imaging[1,11]. Abe and Lee[19] found that there were

representations of an untaken alternative option in monkey OFC on a cued decision-making task in which the alternative option that should have been taken was cued to the monkey after the incorrect decision.

The connectivity between OFC and vStr remains highly controversial, with some evidence pointing to connectivity[46–48] and other analyses suggesting a lack of connectivity[49,50]. The anatomical and functional mechanisms through which the OFC and vStr derive their representations of regret-related counterfactual information remains unknown. In addition, the analyses used here lack the temporal resolution necessary to determine any interactions between structures.

The Restaurant Row task introduced here allowed economic measures to identify potential regret-inducing situations, in which the rat made a decision that placed it in a less valuable situation. Because the task was time limited, any decision to wait for a reward decreased the amount of time available to receive future rewards. Human subjects self-report that they regret actions taken or not taken more than they do missed outcomes[3,5,9,36]. Intriguingly, during regret-inducing situations our decoding results showed strong representations of the previous zone entry, where the decision was made and the action taken (p(zone)), but weak and nonsignificant representations of the missed outcome (p(reward)). Most hypotheses suggest that the function of regret is a revaluation of a past opportunity that drives future behavioral changes[4,6]. After making a mistake and recognizing that mistake, rats were more likely to take a high-cost option and rush through the consumption of that less-valuable option.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank C. Boldt and K. Seeland for technical support and other members of the Redish laboratory for discussion. This work was supported by US National Institutes of Health grants T32 NS048944 (A.P.S.), T32 DA007234 (A.P.S.) and R01-DA030672 (A.D.R.).

AUTHOR CONTRIBUTIONS

A.P.S. and A.D.R. conducted the experiments, collected the data, performed the analysis and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Coricelli, G., Dolan, R.J. & Sirigu, A. Brain, emotion and decision making: the paradigmatic example of regret. *Trends Cogn. Sci.* **11**, 258–265 (2007).
2. Camille, N. *et al.* The involvement of the orbitofrontal cortex in the experience of regret. *Science* **304**, 1167–1170 (2004).
3. Gilovich, T. & Medvec, V.H. The experience of regret: what, when, and why. *Psychol. Rev.* **102**, 379–395 (1995).
4. Bell, D. Regret in decision making under uncertainty. *Oper. Res.* **30**, 961–981 (1982).
5. Landman, J. & Manis, J.D. What might have been: counterfactual thought concerning personal decisions. *Br. J. Psychol.* **83**, 473–477 (1992).
6. Loomes, G. & Sugden, R. Regret theory: an alternative theory of rational choice under uncertainty. *Econ. J.* **92**, 805–824 (1982).
7. Loomes, G. & Sugden, R. Disappointment and dynamic consistency in choice under uncertainty. *Rev. Econ. Stud.* **53**, 271–282 (1986).
8. Bell, D. Disappointment in decision making under uncertainty. *Oper. Res.* **33**, 1–27 (1985).
9. Landman, J. Regret: a theoretical and conceptual analysis. *J. Theory Soc. Behav.* **17**, 135–160 (1987).
10. Lee, D. Neural basis of quasi-rational decision making. *Curr. Opin. Neurobiol.* **16**, 191–198 (2006).
11. Coricelli, G. *et al.* Regret and its avoidance: a neuroimaging study of choice behavior. *Nat. Neurosci.* **8**, 1255–1262 (2005).
12. Schoenbaum, G., Chiba, A.A. & Gallagher, M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat. Neurosci.* **1**, 155–159 (1998).
13. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708 (1999).
14. Padoa-Schioppa, C. & Assad, J.A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
15. Jones, J.L. *et al.* Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* **338**, 953–956 (2012).
16. Takahashi, Y.K. *et al.* Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning. *Neuron* **80**, 507–518 (2013).
17. Wilson, R.C., Takahashi, Y.K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267–279 (2014).
18. Sul, J.H., Kim, H., Huh, N., Lee, D. & Jung, M.W. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* **66**, 449–460 (2010).
19. Abe, H. & Lee, D. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* **70**, 731–741 (2011).
20. Schoenbaum, G., Nugent, L.S., Saddoris, M.P. & Setlow, B. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport* **13**, 885–890 (2002).
21. Fellows, L.K. & Farah, M.J. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain* **126**, 1830–1837 (2003).
22. Rudebeck, P.H., Saunders, R.C., Prescott, A.T., Chau, L.S. & Murray, E.A. Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat. Neurosci.* **16**, 1140–1145 (2013).
23. McDannald, M.A., Lucantonio, F., Burke, K.A., Niv, Y. & Schoenbaum, G. Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J. Neurosci.* **31**, 2700–2705 (2011).
24. McDannald, M.A. *et al.* Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *Eur. J. Neurosci.* **35**, 991–996 (2012).
25. Steiner, A.P. & Redish, A.D. The road not taken: neural correlates of decision making in orbitofrontal cortex. *Front. Neurosci.* **6**, 131 (2012).
26. Cromwell, H.C. & Schultz, W. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* **89**, 2823–2838 (2003).
27. O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452–454 (2004).
28. Roesch, M.R., Singh, T., Brown, P.L., Mullins, S.E. & Schoenbaum, G. Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J. Neurosci.* **29**, 13365–13376 (2009).
29. van der Meer, M.A. & Redish, A. Covert expectation-of-reward in rat ventral striatum at decision points. *Front. Integr. Neurosci* **3**, 1–15 (2009).
30. Lavoie, A.M. & Mizumori, S.J. Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons of rats. *Brain Res.* **638**, 157–168 (1994).
31. Setlow, B., Schoenbaum, G. & Gallagher, M. Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron* **38**, 625–636 (2003).
32. Takahashi, Y.K. *et al.* The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* **62**, 269–280 (2009).
33. Takahashi, Y.K. *et al.* Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat. Neurosci.* **14**, 1590–1597 (2011).
34. Schoenbaum, G. & Eichenbaum, H. Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. *J. Neurophysiol.* **74**, 733–750 (1995).
35. Roitman, M.F., Wheeler, R.A. & Carelli, R.M. Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* **45**, 587–597 (2005).
36. Landman, J. *Regret: The Persistence of the Possible* (Oxford Univ. Press, 1993).
37. Connolly, T. & Butler, D. Regret in economic and psychological theories of choice. *J. Behav. Decis. Mak.* **19**, 139–154 (2006).
38. Aronson, E. The effect of effort on the attractiveness of rewarded and unrewarded stimuli. *J. Abnorm. Soc. Psychol.* **63**, 375–380 (1961).
39. Arkes, H.R. & Ayton, P. The sunk cost and Concorde effects: are humans less rational than lower animals? *Psychol. Bull.* **125**, 591 (1999).
40. Hare, T.A., Camerer, C.F. & Rangel, A. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* **324**, 646–648 (2009).
41. Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* **13**, 1292–1298 (2010).
42. Bray, S., Shimojo, S. & O'Doherty, J.P. Human medial orbitofrontal cortex is recruited during experience of imagined and real rewards. *J. Neurophysiol.* **103**, 2506–2512 (2010).
43. Walton, M.E., Behrens, T.E., Buckley, M.J., Rudebeck, P.H. & Rushworth, M.F. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* **65**, 927–939 (2010).
44. Noonan, M.P. *et al.* Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc. Natl. Acad. Sci. USA* **107**, 20547–20552 (2010).
45. Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E. & Behrens, T.E. Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**, 1054–1069 (2011).
46. Carmichael, S.T. & Price, J.L. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J. Comp. Neurol.* **363**, 615–641 (1995).
47. Carmichael, S.T. & Price, J.L. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* **363**, 642–664 (1995).
48. Ongür, D. & Price, J.L. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* **10**, 206–219 (2000).
49. Schilman, E.A., Uylings, H.B., Galis-de Graaf, Y., Joel, D. & Groenewegen, H.J. The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex. *Neurosci. Lett.* **432**, 40–45 (2008).
50. Mailly, P., Aliane, V., Groenewegen, H.J., Haber, S.N. & Deniau, J.M. The rat prefrontostriatal system analyzed in 3D: evidence for multiple interacting functional units. *J. Neurosci.* **33**, 5718–5727 (2013).

ONLINE METHODS

Animals. Four Fischer–Brown Norway rats aged 10–12 months at the start of behavior were used in this experiment. Rats were maintained at above 80% of their free-feeding weight. All experiments followed approved NIH guidelines and were approved by the Institutional Animal Care and Use Committee at the University of Minnesota. Four rats is a standard sample size for behavioral neurophysiology experiments measuring information processing in large neural ensembles. Each rat was from a different litter.

Experimental design. The Restaurant Row task consisted of a central ring (approximately 1 m in diameter) and four spokes leading off that ring (**Fig. 1a**). At the end of each spoke, a feeder (MedAssociates, St. Albans, VT) dispensed two 45-mg food pellets of a given flavor (banana, cherry, chocolate and unflavored (plain), Research Diets, New Brunswick, NJ). A given flavor remained at a constant spoke through the entire experiment. As the rats proceeded around the track, the rat's position was tracked from LEDs on the head via a camera in the ceiling. A spatial zone was defined for each spoke that included the complete spoke and a portion of the inner loop and aligned with the inner loop such that a rat could not miss a zone by running past it (boxes in **Fig. 1a**); that is, zones were defined such that a rat had to pass through the current zone before continuing on to the next zone. The zone's entries were separated by 90 degrees, and each one led to a potential reward location approximately 0.5 m from the entry point on the central, octagonal track. A trigger zone was defined so as to include a spoke and the portion of the ring nearby. Zones were primed in a sequential manner so that the rat ran in one direction around the loop. When the rat entered a primed zone, a tone sounded indicating the delay the rat would have to wait in that zone to receive food. Offered delays ranged from 1 s (identified by a 750-Hz tone) to 45 s (12 kHz). As long as the rat remained in the active zone, a tone sounded each second, decreasing in pitch (counting down by 250-Hz increments). If the rat left the zone, the tones stopped, and the next zone in the sequence was primed. In practice, rats waiting out a delay would proceed down the spoke and wait near the feeder; rats skipping a zone would proceed directly on to the next trigger zone. Each rat ran one 60-min session each day. During training, rats were allowed to run the task in any manner they saw fit. However, rewards were only available if they traveled through the zones sequentially, zone 1 to zone 2 to zone 3 to zone 4. If a rat traveled backwards, the rat would have to complete approximately three laps to prime the previous zone. Rats quickly learned that this behavior was not viable. Within 7 d, rats learned to travel in only one direction and to pass through each zone sequentially.

Rats were initially handled and accustomed to the different flavors as described previously[25]. Rats were shaped to the task in three stages. In the first stage, all offers were 1 s. Once rats ran 30 laps per session consistently, they progressed to the second stage. In the second stage, each offer was randomly chosen from 1 to 10 s (uniform distribution, independent between encounters). Again, once rats ran 30 laps per session consistently, they progressed to the third stage, in which they faced the full Restaurant Row task with offers selected between 1 and 30 s (uniform distribution). Two rats often waited out the full 30 s at some locations, so delays were increased for those rats to range from 1 to 45 s.

Once rats were completing at least 50 laps per session on the full Restaurant Row task, they were implanted with hyperdrives targeting the ventral striatum and orbitofrontal cortex. Rats were then reintroduced to the task until running well. Each day, rats were allowed to run for 60 min and often completed upwards of 70 entries per zone. Rats received all of their food on the track each day.

Control task (4 × 20). To confirm the economic nature of the Restaurant Row paradigm, two rats ran an additional task after completing all recordings. In this modified version, each rat ran one session per day that consisted of four blocks of 20 min per block. In each block, one reward site provided three food pellets (of its corresponding flavor), while the other three reward sites provided one food pellet (of their corresponding flavors). Delays ranged from 1 to 45 s (uniform distribution). Each of the four sites was the three-pellet site for one of the four blocks each day. Which site offered the greater reward in which block was pseudorandomly varied across days. Rats were removed to rest on a nearby flower pot for 60 s between each block.

Surgery. Rats were implanted with a dual-bundle 12 tetrode + 2 reference hyperdrive[25,29] aimed at the ventral striatum (6 tetrodes + 1 reference, M/L +1.8 mm, A/P +1.9 mm) and orbitofrontal cortex (6 tetrodes + 1 reference, M/L +2.5 mm, A/P +3.5 mm). For two rats, the two targets were left vStr and left OFC, while for two rats the two targets were right vStr and right OFC. Following surgery, tetrodes were turned daily until they reached vStr and OFC. Upon acquisition of large neural ensembles and a return to stable behavior on the maze, each rat ran a minimum of 10 recording days. Data reported here came from a total of 47 sessions distributed evenly over the four rats: R210, 12 sessions; R222, 12 sessions; R231, 13 sessions; R234, 10 sessions (**Supplementary Table 1**).

Data analysis. No data that met the inclusion criteria (as defined in the main text) were excluded. Analyses were automated and applied uniformly to all instances meeting the inclusion criteria. Data were analyzed on an encounter-by-encounter basis. Clusters were cut on a session-by-session basis; experimenters were blind to behavior when cutting clusters.

Behavior. *Threshold calculation.* At every encounter with a reward zone, the rat could wait through the delay or skip it and proceed to the next zone. If the rat chose to skip, it tended to do so quickly (**Supplementary Fig. 1**). Rats tended to wait for short delays and skip long delays, as expected (**Fig. 1**). To determine the threshold, we defined stays as 1 and skips as 0 and fit sigmoid functions of stay or skip as a function of delay using a least-squares fit (Matlab, MathWorks, Natick, MA). The threshold for 'above or below' calculations was defined as the midpoint of the sigmoid. We determined a threshold for each rat for each session for each zone. All preference data were measured during the task, and each rat demonstrated a different preference, indicated by the amount of time that rat was willing to wait for reward.

Identifying regret-inducing and control situations. On entry into a given ('current') zone, we defined the situation as regret-inducing if it met the following three conditions: (i) the offer at the previous zone was a delay <threshold for that previous zone for that rat for that session. (ii) The rat skipped the previous offer. (iii) The offer at the current zone was a delay >threshold for that current zone for that rat for that session.

The first control was defined using the same criteria as for regret-inducing situations, except that (ii) the rat took the previous offer. This control situation keeps the sequence of offers the same but controls for the rat's agency (choice).

The second control was defined using the same criteria as for regret-inducing situations, except that (i) the offer at the previous zone was a delay >threshold for that previous zone for that rat for that session. This control situation keeps the rat's choices the same but makes the choice to skip the previous option the correct one (see **Supplementary Table 2** for a summary of conditions).

Curvature. To identify the pause-and-look behavior, we measured the curvature of the path of the animal's head and identified the point of maximum curvature and the direction of that point. Curvature was measured through the following algorithmic sequence: the position of the head was measured at 60 Hz from the LEDs on the headstage via the camera in the ceiling, giving $\langle x,y \rangle$ coordinates, velocity $\langle dx,dy \rangle$ was calculated using the Janabi-Sharifi algorithm[51] and acceleration $\langle ddx,ddy \rangle$ was calculated by applying the Janabi-Sharifi algorithm to $\langle dx,dy \rangle$. Finally, the curvature[52] at each moment was defined as

$$\frac{(dx \times ddy) + (dy \times ddx)}{(dx^2 + dy^2)^{1.5}}$$

Neurophysiology. Cells were recorded on a 64-channel analog Cheetah-160 recording system (Neuralynx, Bozeman, MT) and sorted offline in MClust 3.5 (A.D.R.; current software available at http://redishlab.neuroscience.umn.edu/). For all sessions, the position of the rat was tracked via overhead camera viewing colored LEDs on the headstage.

Reward responsiveness. We are interested in determining how a cell modulates its activity during reward delivery. To measure this quantitatively, we compared the firing rate of the cell in the 3 s after reward delivery to 500 randomly selected 3-s intervals throughout the task. If a cell's firing rate is different (whether increased or decreased) during reward delivery, then it carries information about reward delivery. We can measure this change by determining whether the activity during the 3 s after reward delivery is significantly different than the bootstrap. Because

these distributions were not normal, we used a Wilcoxon test to calculate significance. Responsiveness to each reward site was calculated independently (see **Supplementary Table 1** for summary of cells per rat).

Bayesian decoding. We used a Bayesian decoding algorithm[53] with a training set defined by the neuronal firing rate at specific times of interest (250-ms window). Any decoding algorithm consists of three parts: (i) a training set of tuning curves that defines the expected activity as a function of the variable in question, (ii) a test set of spikes or firing rates and (iii) the posterior probability calculated from (i) and (ii). In this manuscript, we used two decoding processes: one in which the tuning curves were defined as the neural activity in the 3 s after reward delivery at the four reward locations [p(reward)], and one in which the tuning curves were defined as the neural activity in the 3 s after initial cue delivery (zone entry) [p(zone)]. When calculating p(zone), time after reward delivery was not included. This was only important for delays <3 s.

p*(reward).* Throughout the paper we refer to this measure as "p(reward)"; however, mathematically, it is p(reward | spikes). Assuming a uniform distribution of reward priors, this equation is

$$p(\text{reward}\,|\,\text{spikes}) = \frac{p(\text{spikes}\,|\,\text{reward}) \times 0.20}{p(\text{spikes})}$$

We defined the training set of p(spikes | reward) as the firing rate during the 3 s after a given reward delivery (for example, p(spikes | banana) etc.). To provide a control for unrelated activity, we also included a fifth condition in our calculation, the average firing rate during times the animal was not in any countdown zone. Thus, the training set consisted of five expected firing rates: firing rate after reward receipt (i) at banana, (ii) at cherry, (iii) at chocolate and (iv) at unflavored, plus a fifth control of expected firing rate (v) on the rest of the maze. Because of the inclusion of the fifth (average firing rate when not at reward) condition, the normalization factor is 0.20.

p*(zone).* Throughout the paper we refer to this measure as "p(zone)"; however, mathematically, it is p(zone | spikes). Assuming a uniform distribution of reward priors, this equation is

$$p(\text{zone}\,|\,\text{spikes}) = \frac{p(\text{spikes}\,|\,\text{zone}) \times 0.20}{p(\text{spikes})}$$

We defined the training set of p(spikes | zone) as the firing rate during the 3 s after entry into a given trigger zone (for example, p(spikes | banana zone) etc.). To provide a control for unrelated activity, we also included a fifth condition in our calculation, the average firing rate during times the animal was not in any trigger zone. Thus, the training set consisted of five expected firing rates: firing rate after zone entry (i) at banana, (ii) at cherry, (iii) at chocolate and (iv) at unflavored, plus a fifth control of expected firing rate (v) on the rest of the maze. Because of the inclusion of the fifth (average firing rate when not in any trigger zone) condition, the normalization factor is 0.20.

Calculating representations of previous, current, next and opposite. To average across passes between different rewards, we first calculated the posterior probability for a given question (for example, p(reward) or p(zone)) separately for each restaurant or zone. We then rotated the results on the basis of the zone or reward in question to define a current zone or reward (the one the rat is currently encountering), a previous zone or reward (the one the rat had just left), a next zone or reward (the one the rat would encounter next), and an opposite zone or reward.

By using ensemble decoding, we can effectively ask what recorded neurons are representing with the highest probability, taking into account both increases and decreases in firing rate. The ensemble reliably differentiated entries into the different zones as effectively as the different rewards (**Supplementary Figs. 8** and **9**). During normal behavior, the ensemble reliably represented the current zone on entry into it and the current reward on receipt of it.

51. Janabi-Sharifi, F., Hayward, V. & Chen, C.S.J. Discrete-time adaptive windowing for velocity estimation. *IEEE Trans. Control Syst. Technol.* **8**, 1003–1009 (2000).
52. Hart, W.E., Goldbaum, M., Cote, B., Kube, P. & Nelson, M.R. Measurement and classification of retinal vascular tortuosity. *Int. J. Med. Inform.* **53**, 239–252 (1999).
53. Zhang, K., Ginzburg, I., McNaughton, B.L. & Sejnowski, T.J. Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J. Neurophysiol.* **79**, 1017–1044 (1998).

doi:10.1038/nn.3740

Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task

Adam P. Steiner and A. David Redish

Supplemental Figure Captions

Supplemental Figure S1. Duration waited at delay and VTE behavioral summary.
First Row. In order to determine whether rats were waiting for a specific tone before leaving, we measured the time spent at each zone encounter over all rats, over all sessions. Graph shows number of seconds spent waiting as a function of the delay offer. A rat waiting out the entire delay would add into the x=y line; a rat leaving immediately would add into a cluster near the 0 duration waited. As can be seen in the histogram, rats tended to wait through the entire delay or leave after 3 seconds.

When rats encounter certain decisions, they sometimes pause and turn back and forth between the multiple options (defined as vicarious trial and error), as if deliberating between them (Muenzinger and Gentry, 1931; Muenzinger, 1938). In humans and other primates, a similar process can be seen in saccade-fixate-saccade (SFS) sequences (Padoa-Schioppa and Assad, 2006; Krajbich et al., 2010). Previous studies have found these VTE events to primarily occur during flexible (non-automated) behaviors (Muenzinger and Gentry, 1931; Johnson and Redish, 2007; van der Meer and Redish, 2009; Papale et al., 2012; Steiner and Redish, 2012), however, previous studies have not examined the relationship between VTE and decision difficulty.

Vicarious trial and error (VTE) was measured as the integrated absolute angular change in the orientation of motion of the head, as measured by sequences of head position samples (Papale et al., 2012; Steiner and Redish, 2012). This measure was calculated through a short algorithm sequence: first the position of the head <*x,y*> was sampled at 60 Hz via the Cheetah Neuralynx system. Change in head position <*dx,dy*> was calculated using the Janabi-Sharifi(Janabi-Sharifi et al., 2000) algorithm. Orientation of motion <*phi*>, was calculated as the arc-tangent of <*dx,dy*>. Change in orientation of motion was <*dphi*> was calculated by applying the Janabi-Sharifi algorithm to <*phi*>. VTE was measured as the sum of the absolute value of <*dphi*> over first two seconds of time after entering a zone <*IdPhi*>.

Rats running the task showed three clear behaviors on encountering a new spoke – they sometimes just ran down the spoke to sample the food-delivery site, they sometimes skipped the spoke, and they sometimes paused and expressed VTE at the decision-point. As noted above, sampling tended to occur when the delays were below the threshold that rat had for that flavor, while skips tended to occur when the delays were above threshold. We quantified VTE through a measure of the integrated angular velocity of the head position of the animal (Steiner and Redish, 2012). We found that VTE tended to occur at the threshold, decreasing dramatically when the delay was less than threshold (generally a sample), but also decreasing when the delay was greater than threshold (generally a skip), (blue dotted lines, bottom plot; Linear Regression, R^2=0.95, $p < 0.001$ pre threshold; R^2=0.76, $p < 0.001$ post threshold).

Nature Neuroscience: doi:10.1038/nn.3740

Several behaviors were typical when rats encountered a delay upon entering a zone. **Second Row**. If rats decided to stay, they generally proceeded to the reward site and waited until the tone counted down and reward was delivered (as indicated by the very low average speed for the remainder of the time in zone). On these passes VTE was typically quite low. **Third Row**. If the delay was above threshold, rats would often skip the zone relatively quickly (decrease in speed at 1 second followed by increasing speed after 2 seconds), spending little time in the current zone. VTE on these passes was typically low. **Fourth Row.** If rats encountered a close to threshold delay and chose to skip the reward, VTE remained high. Rats remained relatively stationary for a longer period of time (from 1 to 5 seconds) before finally locomoting and leaving the current zone for the next zone. **Fifth Row**. On close to threshold delays, rats demonstrated stronger VTE. If rats chose to sample the reward, they would proceed towards the feeder and wait through the remainder of the delay (early fluctuation in speed indicates high VTE, followed by decrease, near 0 cm/s speed indicates the rat has arrived at the feeder location where he remains until reward is received).

Supplemental Figure S2. Comparison of thresholds within session by rat. Thresholds were consistent within each session. If we compared the thresholds from the first half to the second half, no thresholds were significantly different between the first and second half of each session. Red bars represent the standard error.

Supplemental Figure S3. Overall food handling time. After consuming food, rats typically took 20-30 seconds before leaving the zone. This did not change as a function of the delay the rat had waited before receiving the food.

Supplemental Figure S4. Increasing the number of pellets increases the average delay waited. To determine if the rats took value into account when making decisions to stay or go (a key tenet in neuroeconomics (Montague and Berns, 2002, Padoa-Schioppa and Assad, 2006, Kable and Glimcher, 2007, Rangel et al., 2008)), two of the rats (R231 and R234) underwent an additional variation of the Restaurant Row task following completion on the unmodified version of the task.

In this modified version, sessions consisted of four 20 minute blocks. During each 20 minute block, one reward flavor site dispensed three food pellets rather than two pellets (i.e. 3x 45 mg), while the other sites only dispensed one food pellet (i.e. 1x 45 mg). The four blocks allowed us to have each site be the “more valuable” site for one block. The order was randomly determined each day. Delays were randomly selected, as in the original task. Each 20 minute block was followed by a one minute rest, during which time the rat was removed to a small flower pot to the side. Each rat ran one complete session of four blocks per day.

Rats were willing to wait longer for the larger reward (errors bars represent +/– standard error). This manipulation indicates that increasing the reward size increased the time rats were willing to wait, which implies that increasing reward size had more value, and that the rats were behaving economically.

Nature Neuroscience: doi:10.1038/nn.3740

There is no reason to expect the increase in the amount of time willing to wait for larger rewards to be linear. Subjective value depends upon the internally generated function for each reward (humans (Lichtenstein and Slovic, 2006, Krajbich et al., 2010), rats (Young, 1932, Berridge, 2009, Ahmed, 2010), primates (Padoa-Schioppa and Assad, 2006)). Because rats, like humans, have preferences, we would expect that different rewards would have different values. However, the only real way to measure a value is by the choices that occur within a given context. By measuring the revealed preferences for each flavor we are essentially determining the subjective value of each reward.

Rats discount hyperbolically (Mazur, 2001, Mazur and Biondi, 2009, Papale et al., 2012). In addition, it has been shown that rats' preference saturates as the number of pellets increases. Thus, the amount of time a rat will wait for 4 pellets is not twice the time a rat will wait for 2 pellets (Papale et al., 2012). We would not expect the value of 3 pellets to be exactly equal to 3x the value of 1 pellet. The amount a rat will consume at a given moment is not a linear relationship to the amount of food available. The time a rat would be willing to wait for 3 pellets should be greater than the time it would be willing to wait for 1 pellet. The time spent waiting for 3 pellets was larger than the average time spent waiting for 1 pellet

Supplemental Figure S5. Histology. Colored lines indicate where recordings for each tetrode in each rat began. Lines terminate where recordings were ended. Insets show example tracts and endpoints for tetrodes in OFC and vStr.

Supplemental Figure S6. Example reward-related cells from orbitofrontal cortex (OFC). Each super-panel (**a,b,c,d**) shows firing from a single cell. Within each super-panel, each subpanel shows that cell's response around the time of reward-delivery. The color of the trace indicates flavor (yellow = banana; black = plain/non-flavored; magenta = cherry; brown = chocolate) and the trace itself indicates the response (in spikes/sec) of the cell. Small dots indicate spikes on individual reward-delivery events. Trace shows average firing over all events, smoothed with a Gaussian window (sigma=50 ms). As can be seen in these examples, different cells responded differently (but reliably) to the different flavor-reward-sites.

Supplemental Figure S7. Example reward-related cells from ventral striatum (vStr). Display as in **Supplemental Figure S6**.

Supplemental Figure S8. Orbitofrontal (OFC) and ventral striatal (vStr) neural ensembles accurately tracked the rewarded flavor during reward receipt. Both OFC and vStr accurately tracked the rewarded flavor. Panels show the confusion matrices of the decoding. We calculated *p(Reward) @ Reward* for each flavor, using a leave-one-out approach to avoid the tautology. Note that, as per Methods, the decoding returns five values, for each of the four flavors plus the fifth "other" condition. **a,b,** The strong increase in the identity comparison implies separate representations of each flavor-reward-site. **c,d,** Shuffling the interspike intervals of the cells removes these representations.

Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S9. Orbitofrontal (OFC) and ventral striatal (vStr) neural ensembles differentiated cue signals at entry into the different zones. As in **Supplemental Figure S7**, panels show the confusion matrices of the decoding. We calculated *p(Zone) at Zone* for each zone, using a leave-one-out approach to avoid the tautology. Note that, as per Methods, the decoding returns five values, for each of the four flavors plus the fifth "other" condition. **a,b,** The strong increase in the identity comparison implies separate representations of each trigger zone. **c,d,** Shuffling the interspike intervals of the cells removes these representations.

Supplemental Figure S10. **Representations match between zone and reward.** To determine the relationship between cues and reward-related activity, we calculated the confusion matrices for the decoding for *p(Reward) at Zone*. **a,b**, The strong increase in the identity comparison implies matched representations between each reward and zone. **c,d**, Shuffling the interspike intervals of the cells removes these representations.

Supplemental Figure S11. Chance levels for decoding. To determine the chance level for the representations of *p(Reward) at Reward, p(Zone) at Zone,* and *p(Reward) at Zone,* we shuffled the interspike intervals. Shuffling the interspike intervals preserves the firing characteristics of the cells but disrupts their alignment to temporal events. Shuffling the interspike intervals for all cells during reward receipt produced a chance level of ~0.14 for all conditions.

Supplemental Figure S12. **Decoding close to threshold on skips and stays.** In order to determine whether orbitofrontal (OFC) and ventral striatal (vStr) signals predicted behavior differentially for similar offers, we measured *p(Reward) at Zone,* for all offers near threshold (delay within 2 seconds above or below threshold). **a,b**, Encounters in which the rat waited through the delay. **c,d**, Encounters in which the rat skipped out and did not wait through the full delay; **a,c**, OFC; **b,d**, vStr. Note that the current reward was better represented during stays than the other zones (**a,b**). In contrast, during skips, the current zone was not better represented; instead, the representations of the next zone began to appear after 2-3 seconds (**c,d**).

Supplemental Figure S13. Shuffled decoding close to threshold on skips and stays. Analysis of the same data shown in **Supplemental Figure S12**, but with interspike intervals shuffled. Shuffling interspike intervals removed all effects.

Supplemental Figure S14. **Matched samples for regret and control conditions.** It is important to ensure that the current delay offers made in the matched control encounters had the same distribution as the regret-inducing instances. Graph shows the cumulative distribution function (CDF) of the "current" offers included in each condition. The distributions were closely matched, indicating that any results seen (e.g. **Fig 5 Main Text**) were not a result of differences between the current offers. **a**, Regret-inducing vs. control 1. **b**, Regret-inducing vs. control 2. The thin lines on the empirical distribution plot represent the 95% confidence intervals (alpha = 0.05). Mann Whitney U tests indicated that the distribution of delays were not significantly different (vs control 1, **a**, p=0.20; vs control 2, **b**, p=0.11).

Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S15. ***p(Reward)*** **and** ***p(Reward)*** $_{shuffled}$**.** Under normal conditions the current reward is accurately represented. However during regret instances the current reward representations are drastically decreased. Instead neuronal firing rates more accurately represent the missed previous reward. The average decoding for the previous *p(Reward)* was different from the shuffled data (ANOVA $p << 0.001$ for vStr and $p << 0.001$). However, *p(Reward)* for OFC was not significant after controlling for multiple comparisons. *p(Reward)* for vStr was not significant after controlling for multiple comparisons.

Supplemental Figure S16. Additional conditions, in which the rat finds a below-threshold opportunity after skipping a previous delay. a–c If the first reward offer was lower than threshold and rats skipped then encountered a second reward lower than threshold, both OFC and vStr represent the current reward more accurately. This increase occurs immediately after the rat enters into the current, primed zone. This result is consistent with data indicating that OFC represents a given reward when a state paired with that reward has been entered (Wilson, 2014). Prior to entry into the current zone, there is no difference in the representations. **d–f** When the rats skipped a high-threshold, high cost delay and encountered a low cost delay, both OFC and vStr ensembles accurately represented the current reward.

Supplemental Figure S17. Posterior probability *p(Zone)* when the rat stayed for a delay at A > threshold at A and encounters a delay at B < threshold B. Both OFC and vStr ensembles increased their decoding to the previous reward. Representations of the previous zone were significant (OFC: ANOVA, $p << 0.05$; vStr: ANOVA, $p << 0.05$). However, these representations were not as strong compared to instances when the rat skipped a delay A < threshold at A and encountered a delay B > threshold B.

Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Table 1

Cellular yields by structure from each rat

Rat Number	OFC (number of cells)	vStr (number of cells)
R210	225	243
R222	329	43
R231	336	112
R234	235	61

Supplemental Table 1: Number of cells recorded from each structure by rat.

Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Table 2

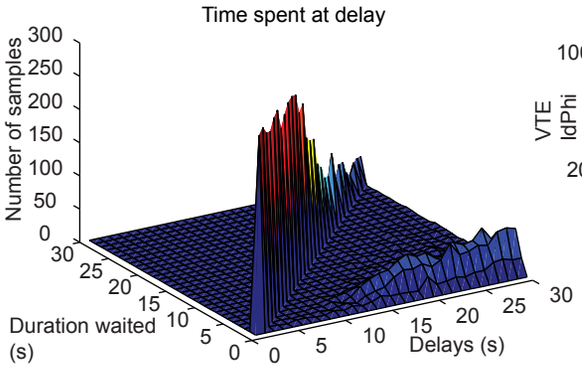
Summary of the decoding different conditions – Regret and Controls

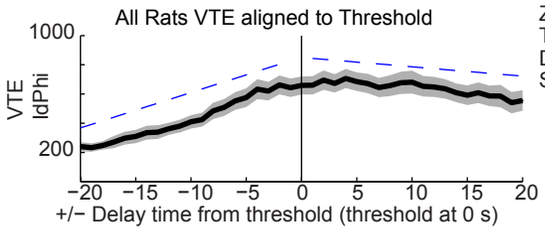
Condition	Offer at previous	Rat's action	Offer at current
Regret-inducing	Delay < threshold	SKIP	Delay > threshold
Control 1	Delay < threshold	STAY	Delay > threshold
Control 2	Delay > threshold	SKIP	Delay > threshold
Rejoice (Supplemental Fig S16)	Delay < threshold	SKIP	Delay < threshold
Rejoice (Supplemental Fig S16)	Delay > threshold	SKIP	Delay < threshold
High Stays (Supplemental Fig S17)	Delay > threshold	STAY	Delay < threshold

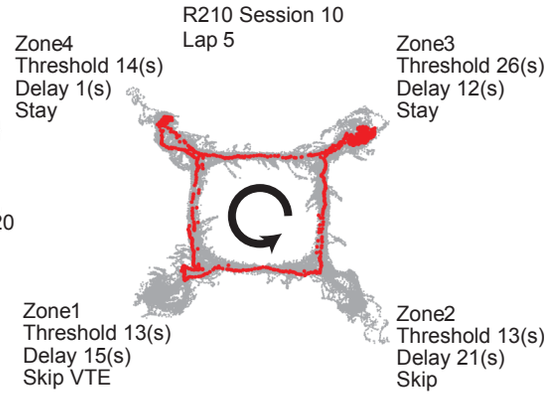
Supplemental Table 2: Summary of the different decoding conditions.

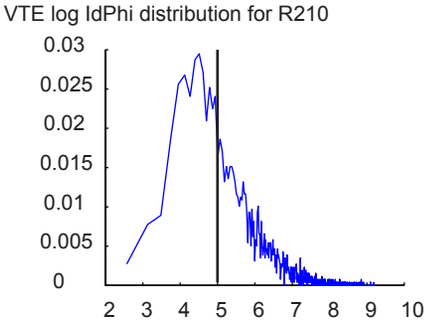
Nature Neuroscience: doi:10.1038/nn.3740

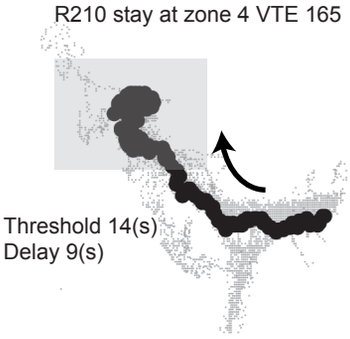
Supplemental Figure S1

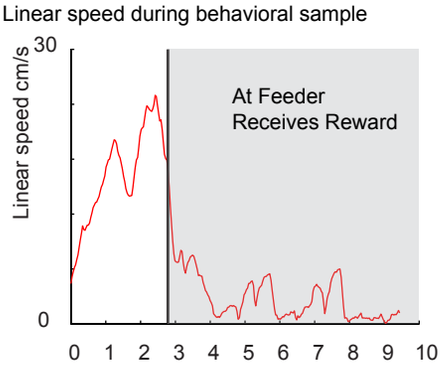


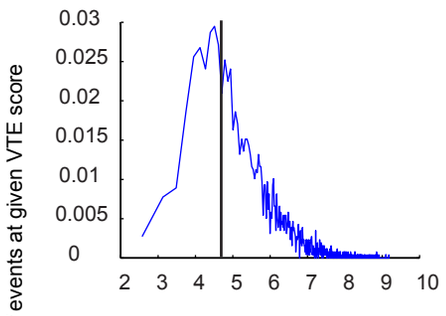


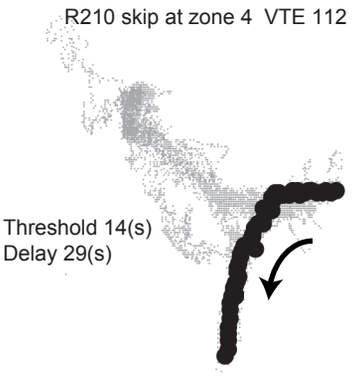


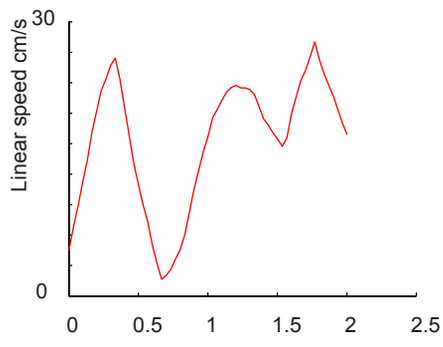


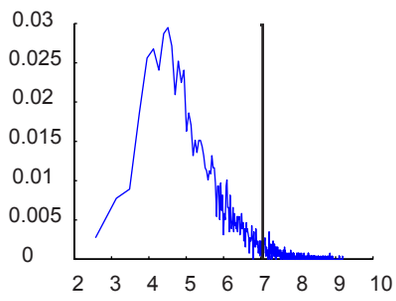


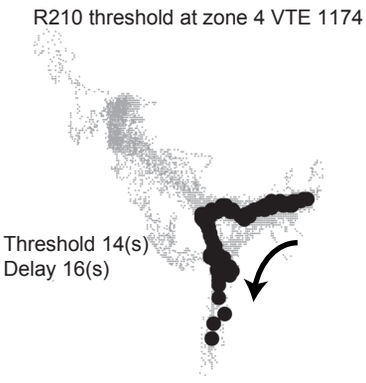


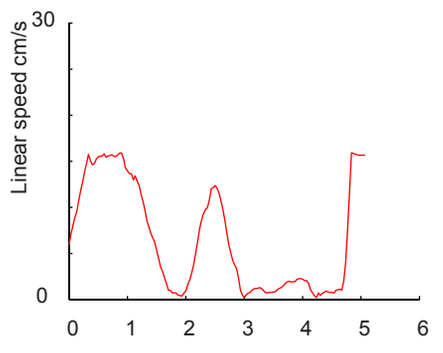


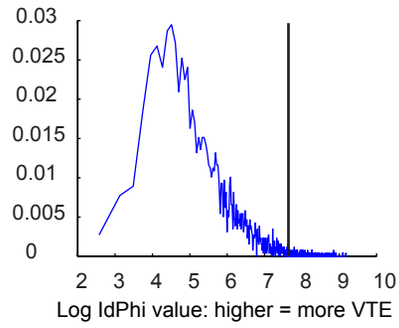


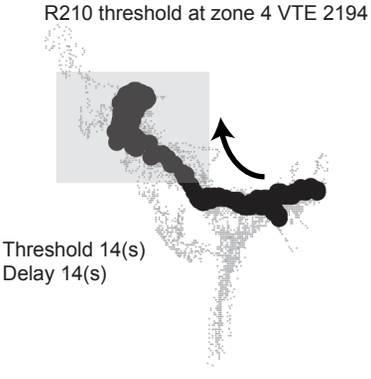


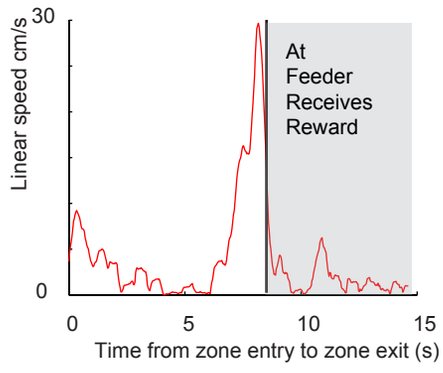






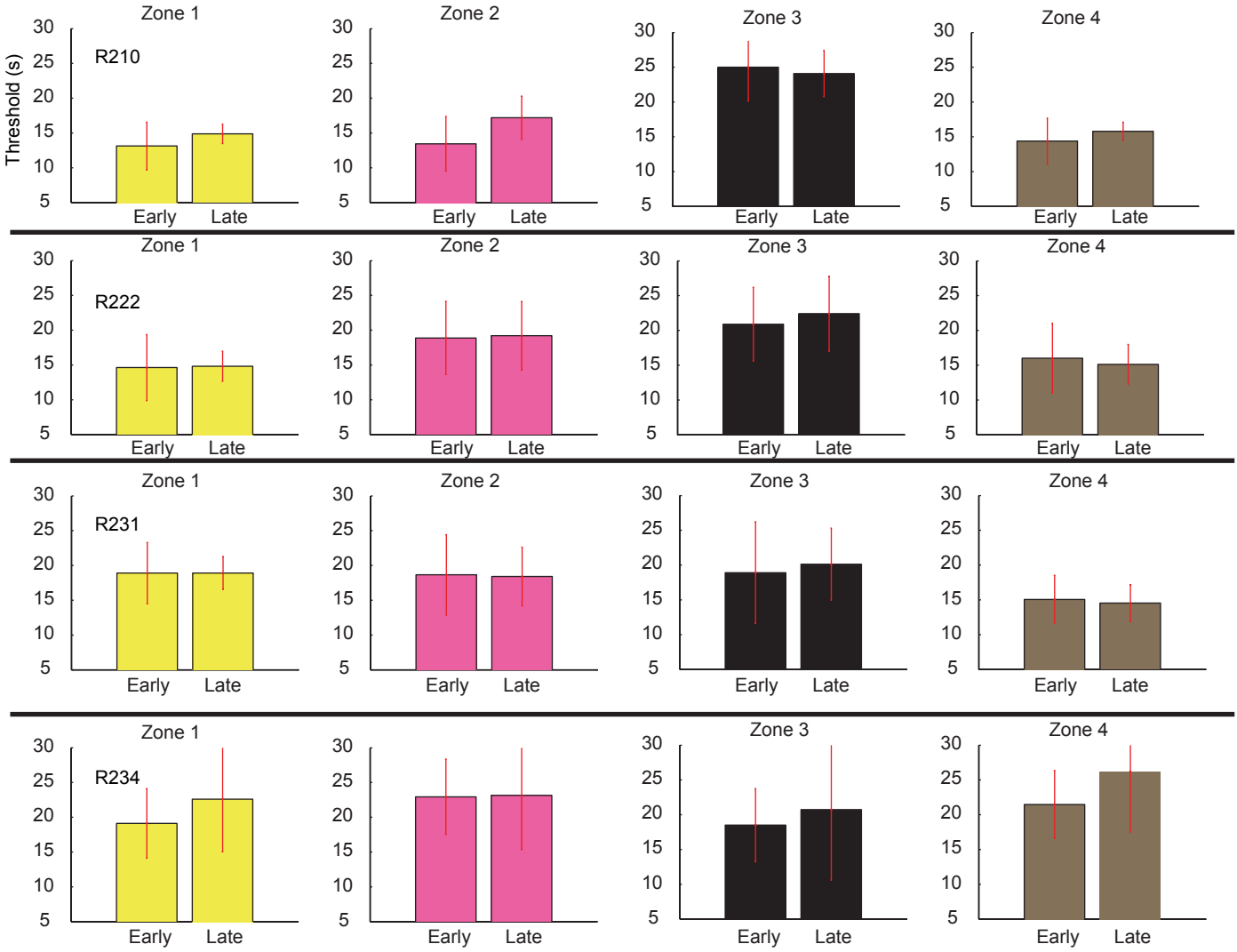




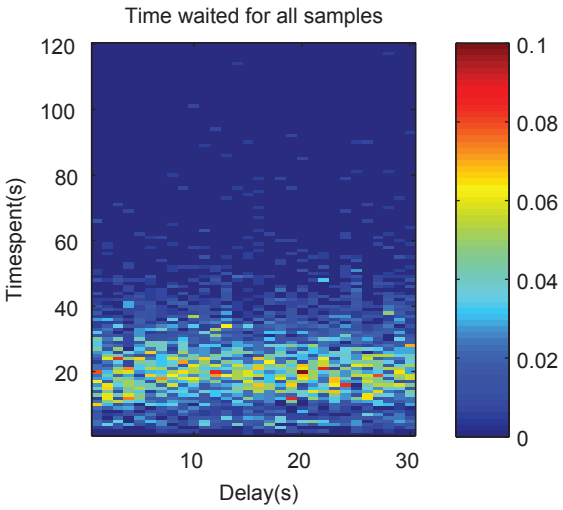


Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S2

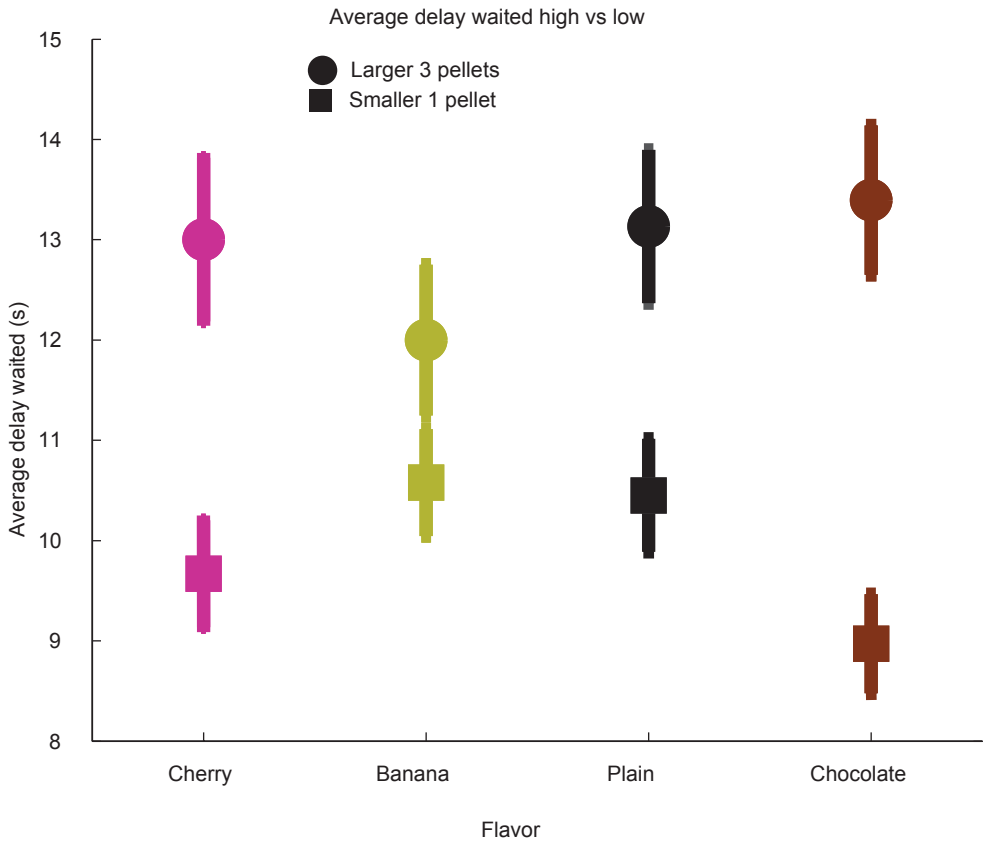


Supplemental Figure S3



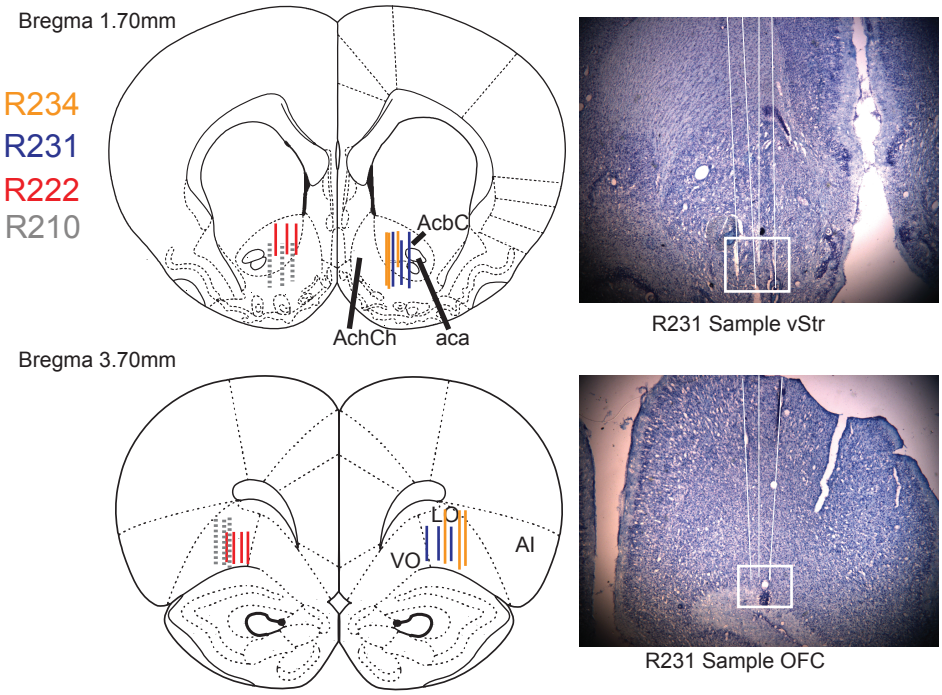
Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S4



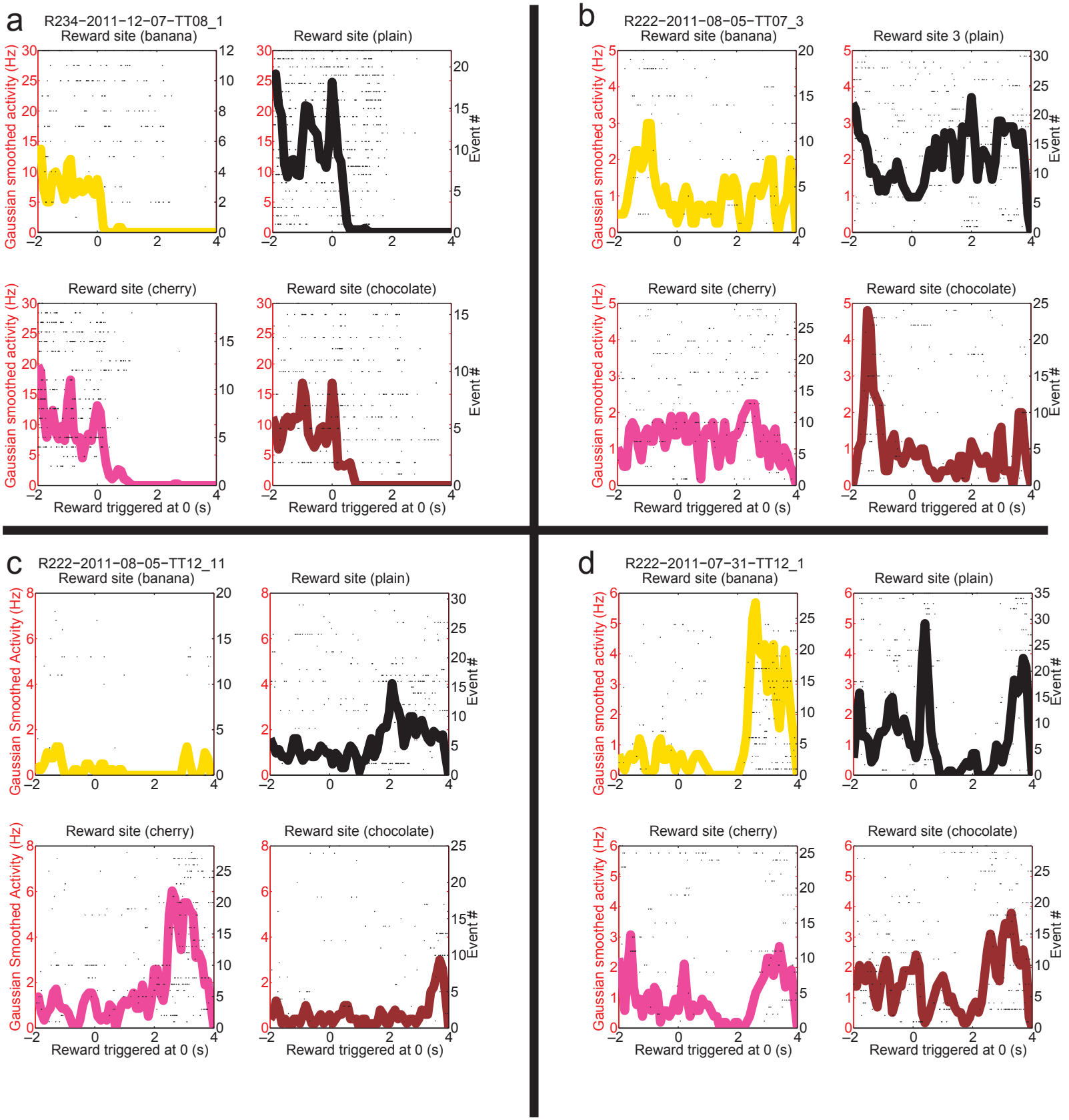
Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S5



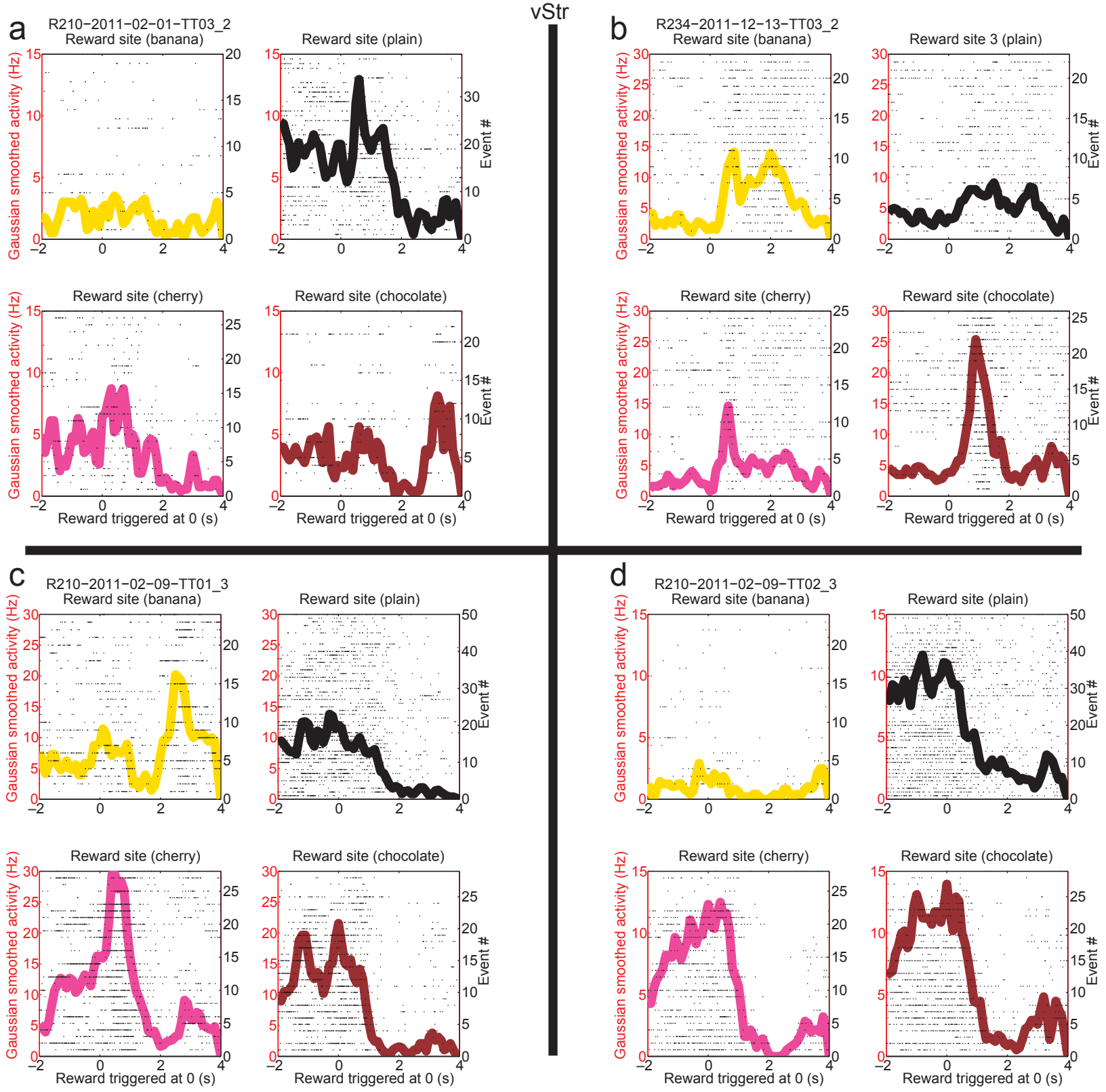
Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S6
OFC



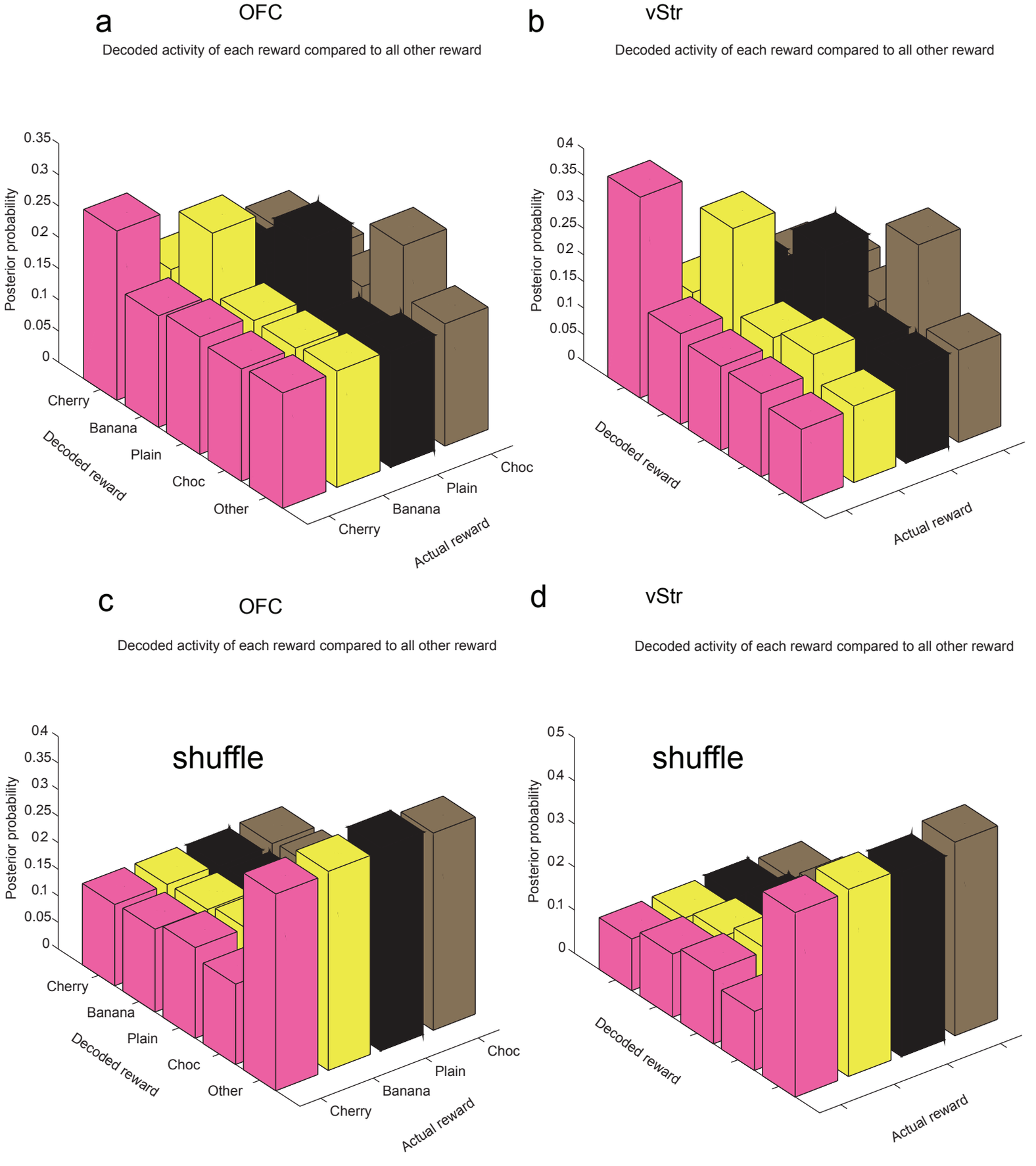
Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S7



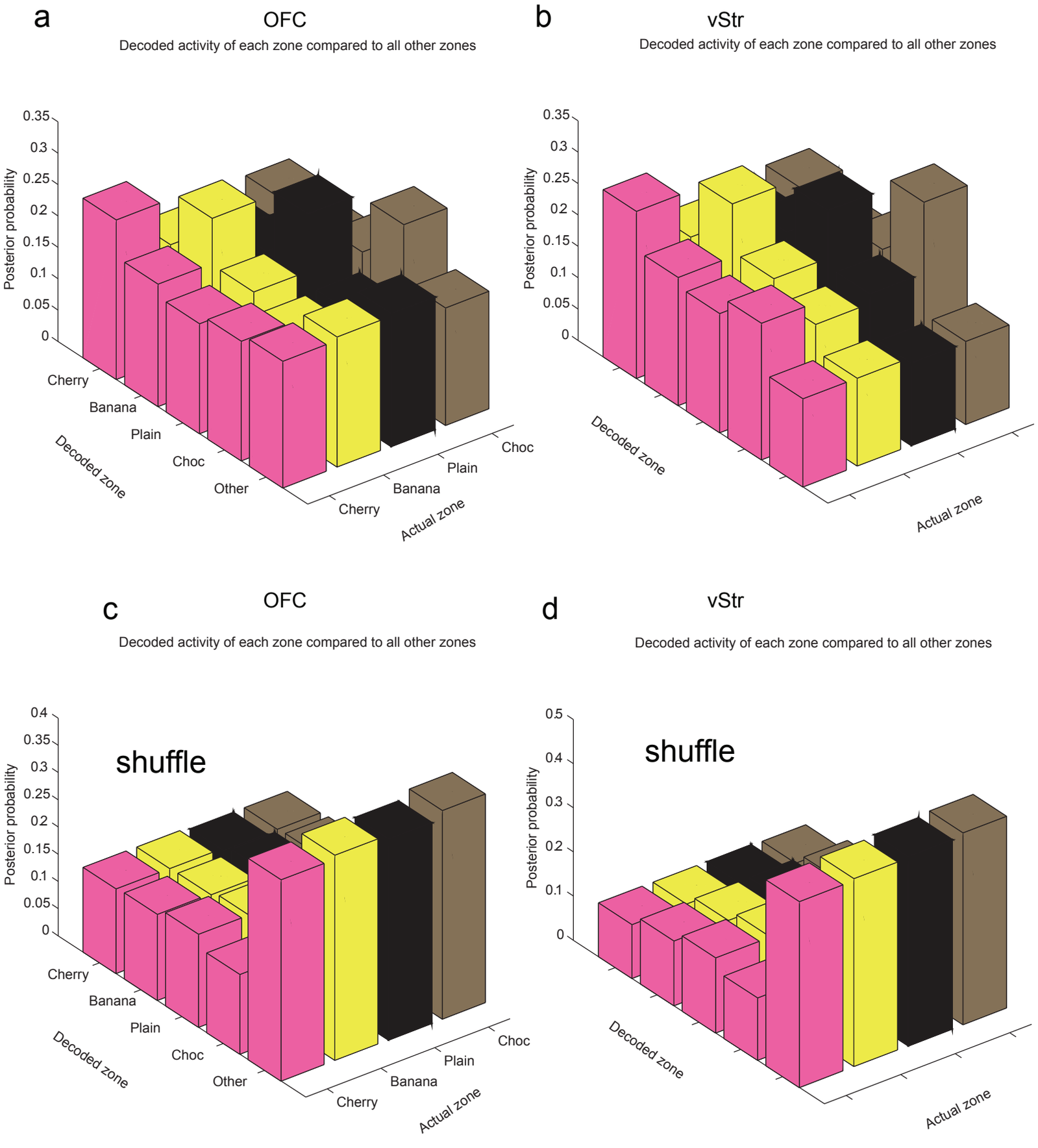
Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S8



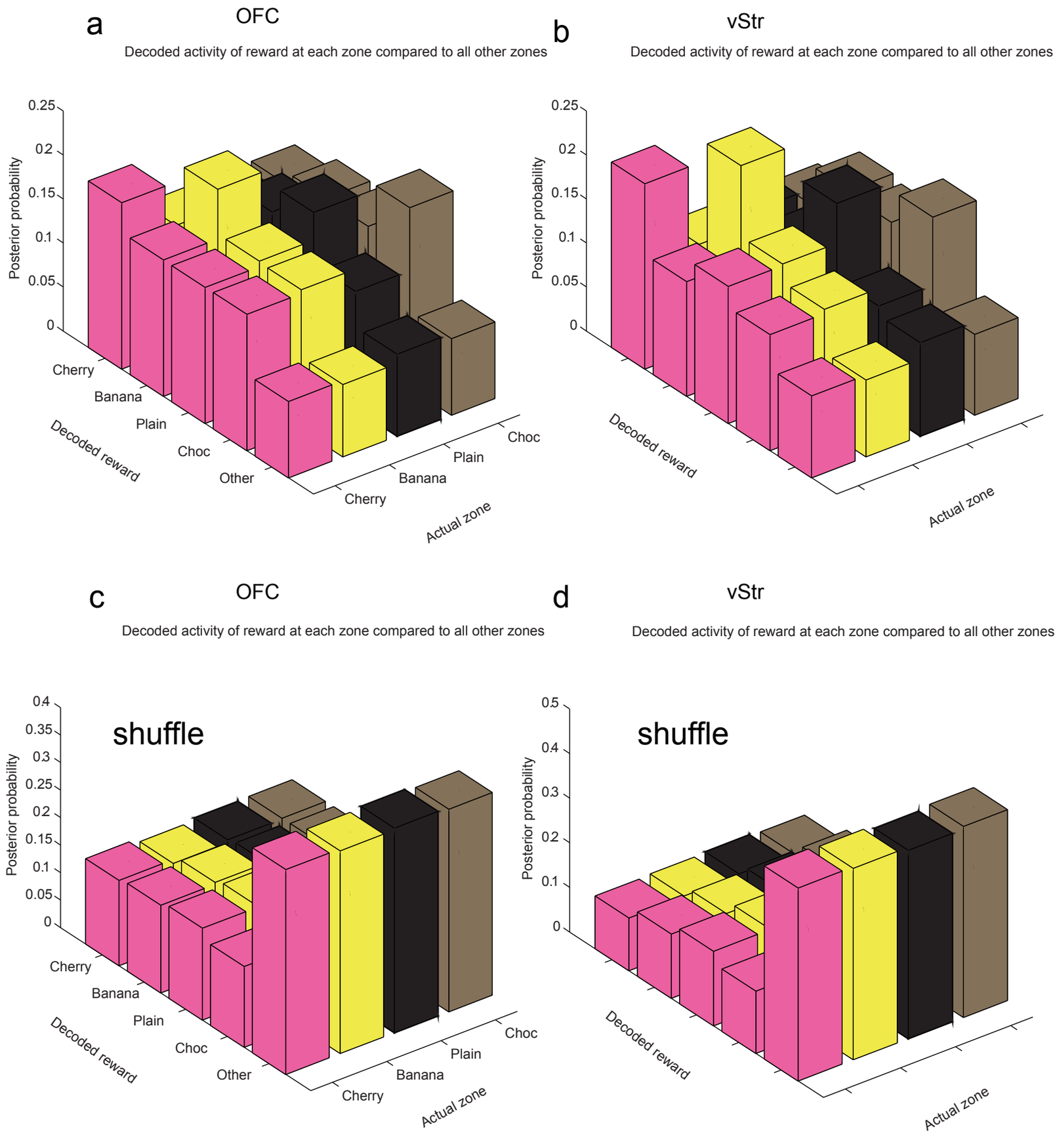
Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S9



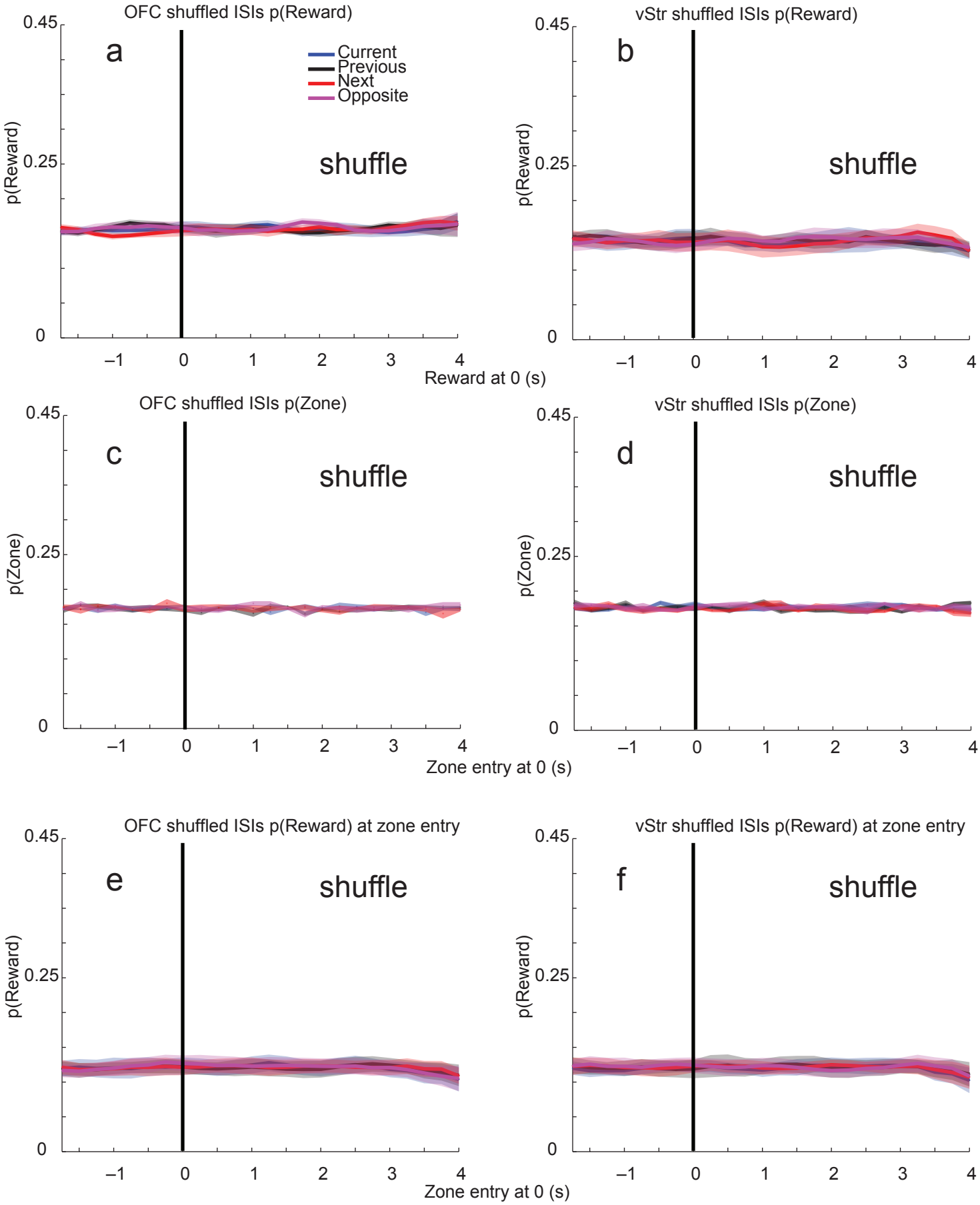
Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S10



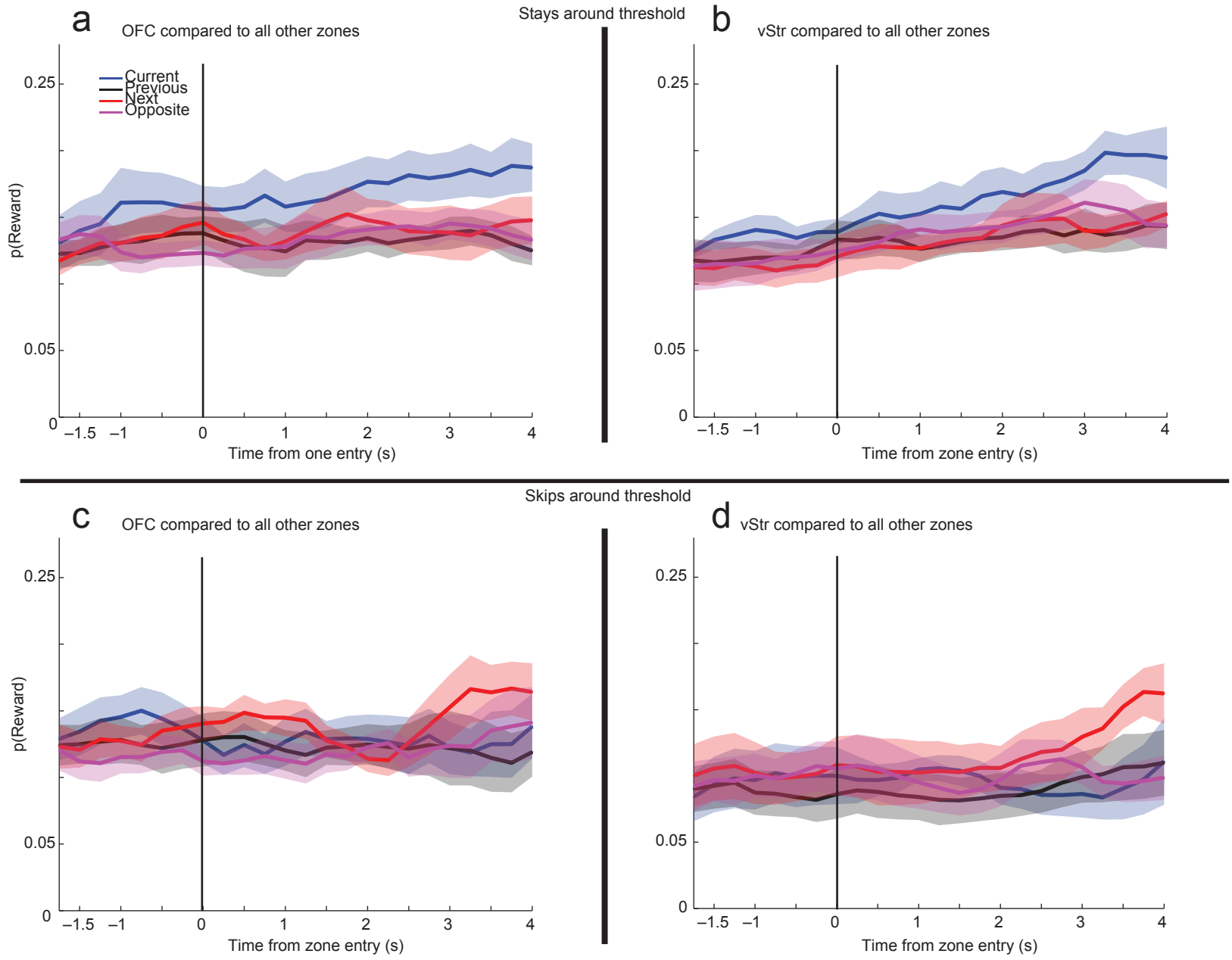
Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S11



Nature Neuroscience: doi:10.1038/nn.3740

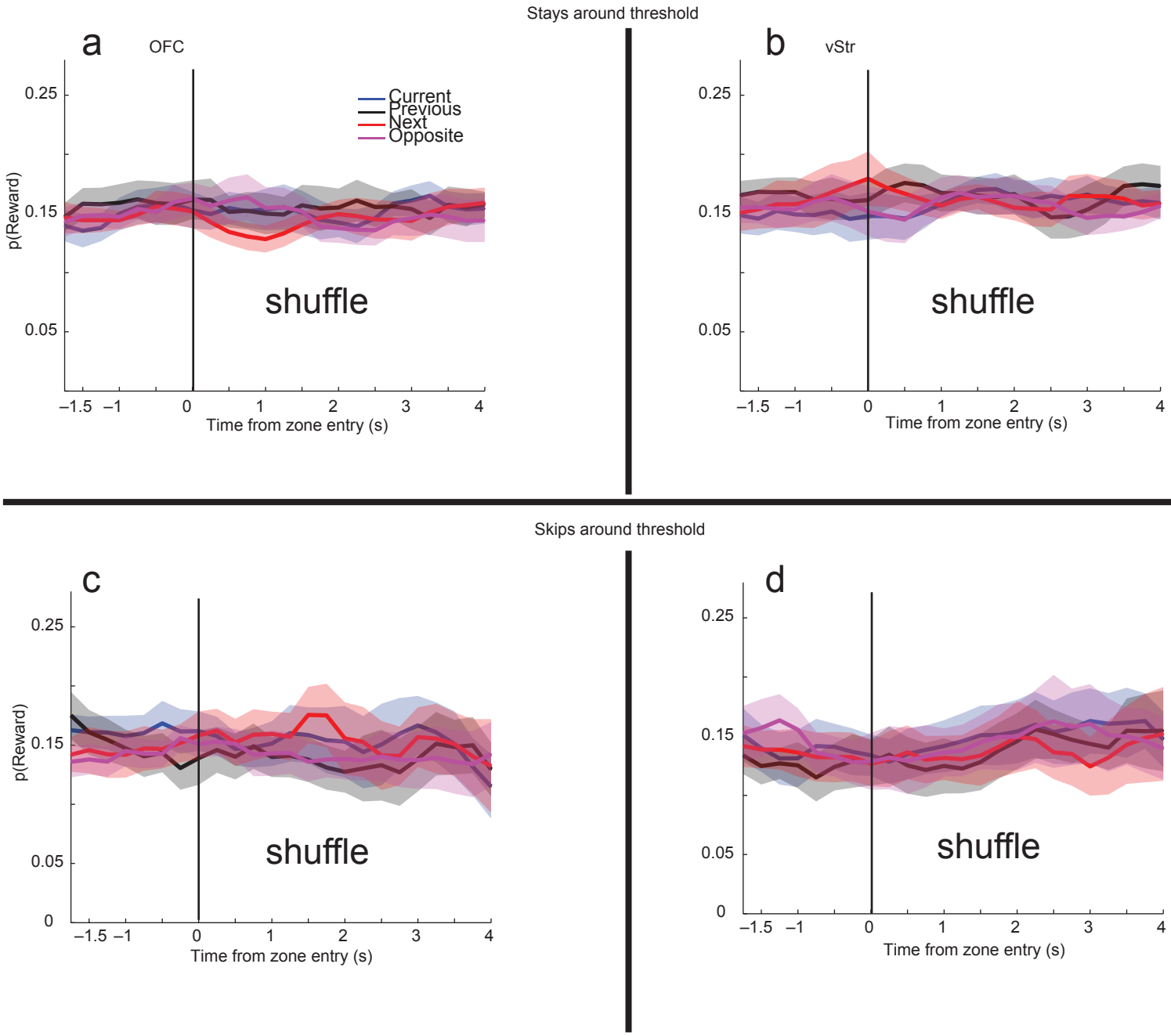
Supplemental Figure S12



Nature Neuroscience: doi:10.1038/nn.3740

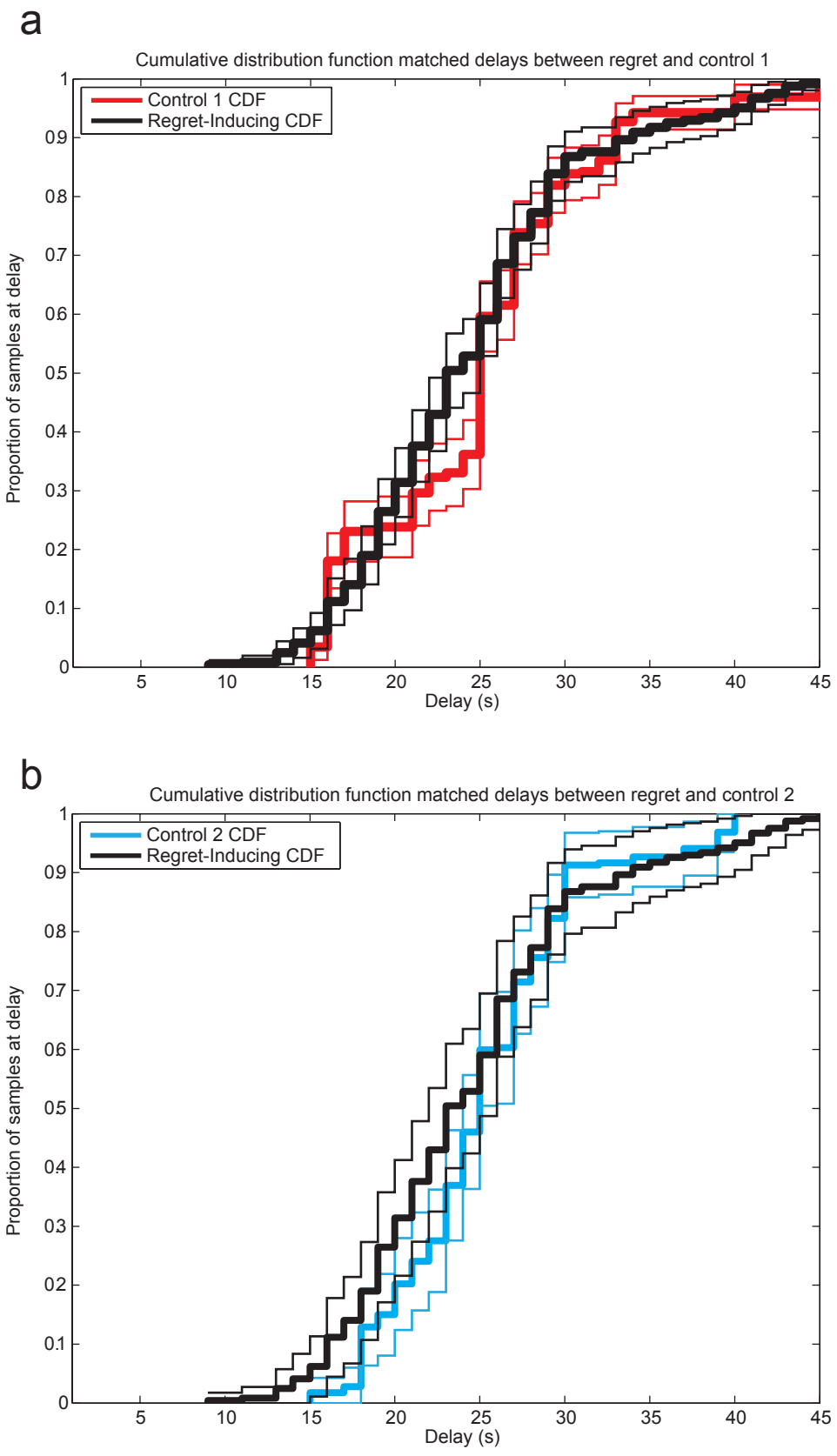
Supplemental Figure S13

Shuffled ISIs: decoding around threshold

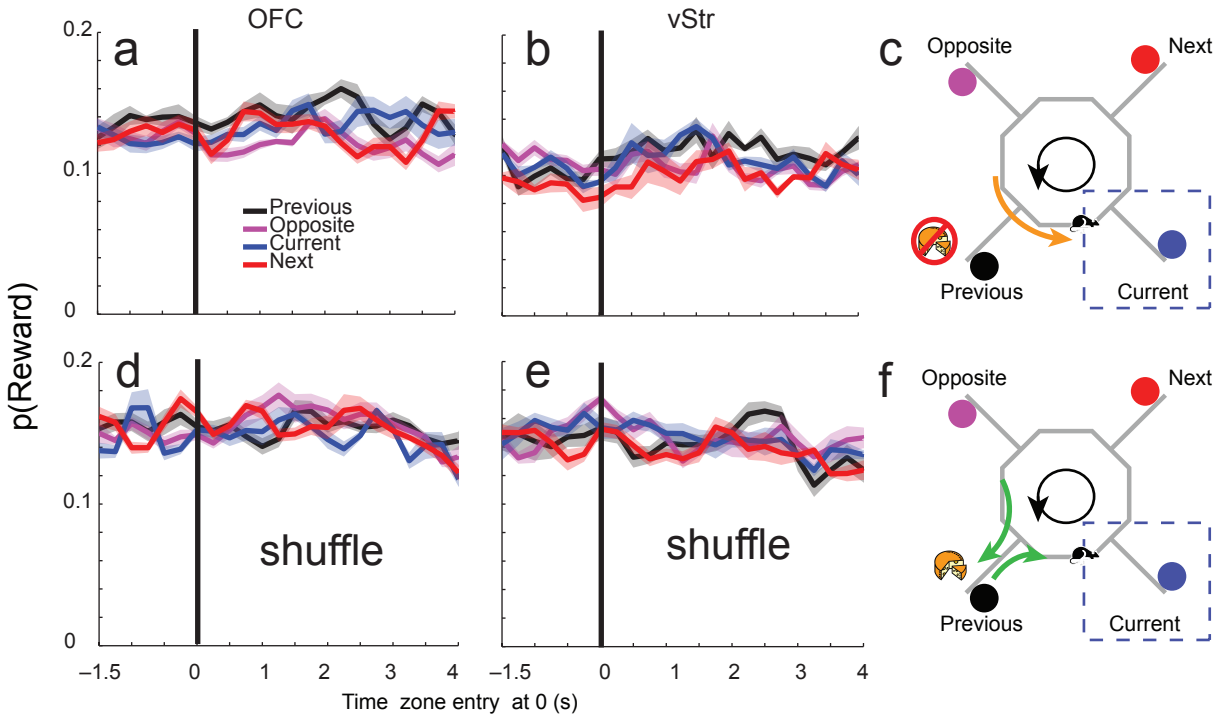


Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S14

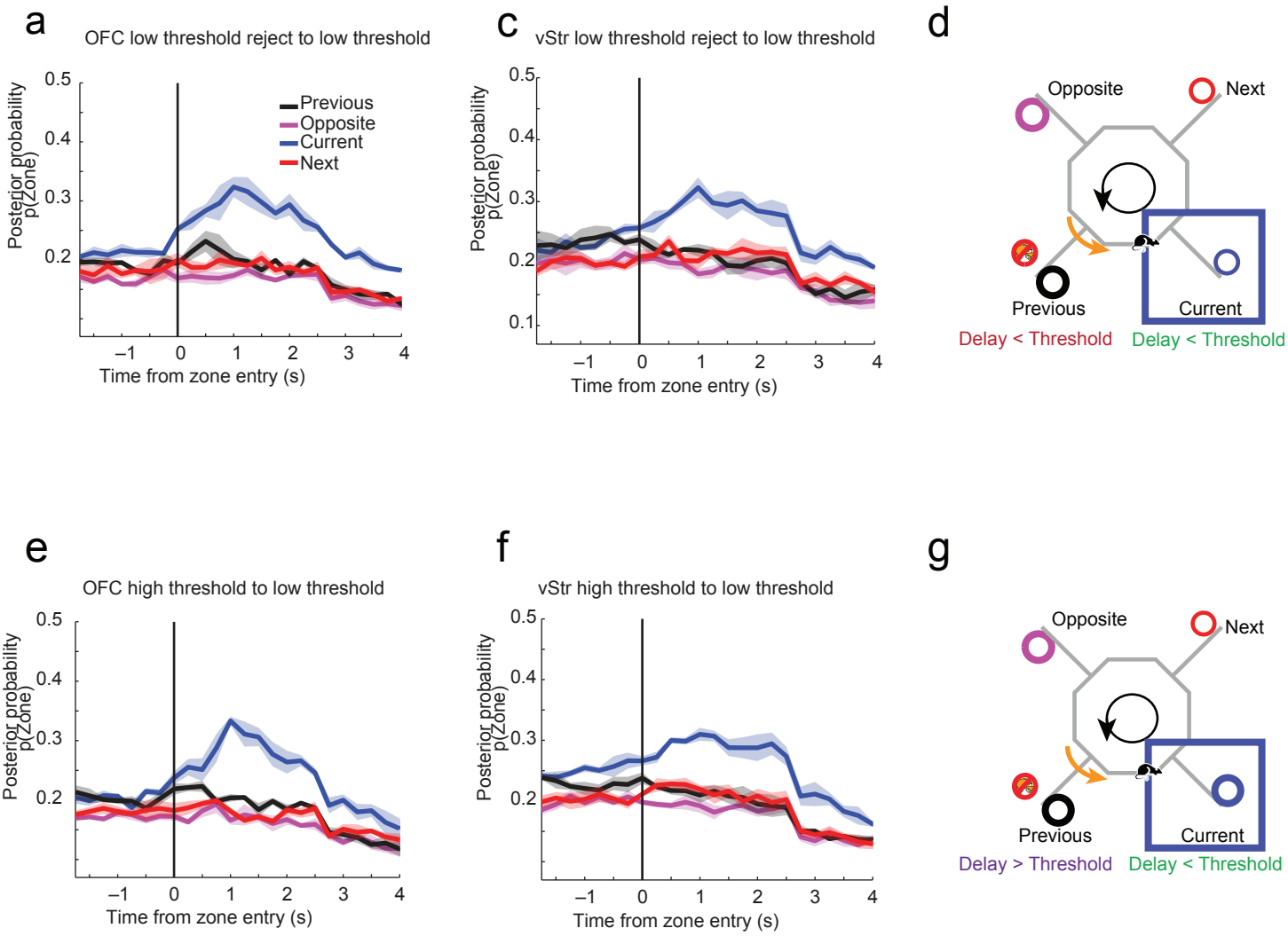


Supplemental Figure 15



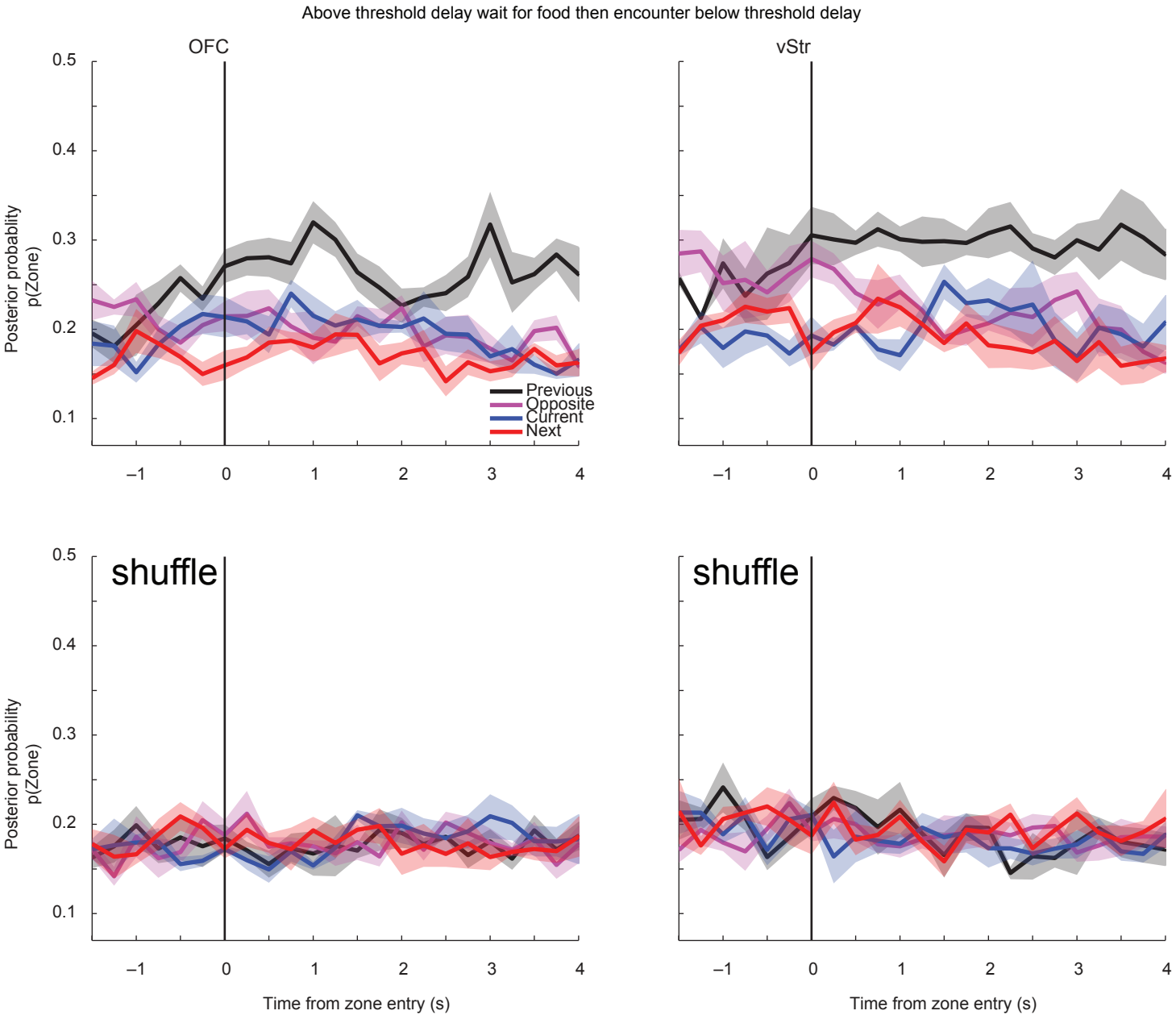
Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S16



Nature Neuroscience: doi:10.1038/nn.3740

Supplemental Figure S17



Nature Neuroscience: doi:10.1038/nn.3740