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).

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.

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.

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).

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.

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.

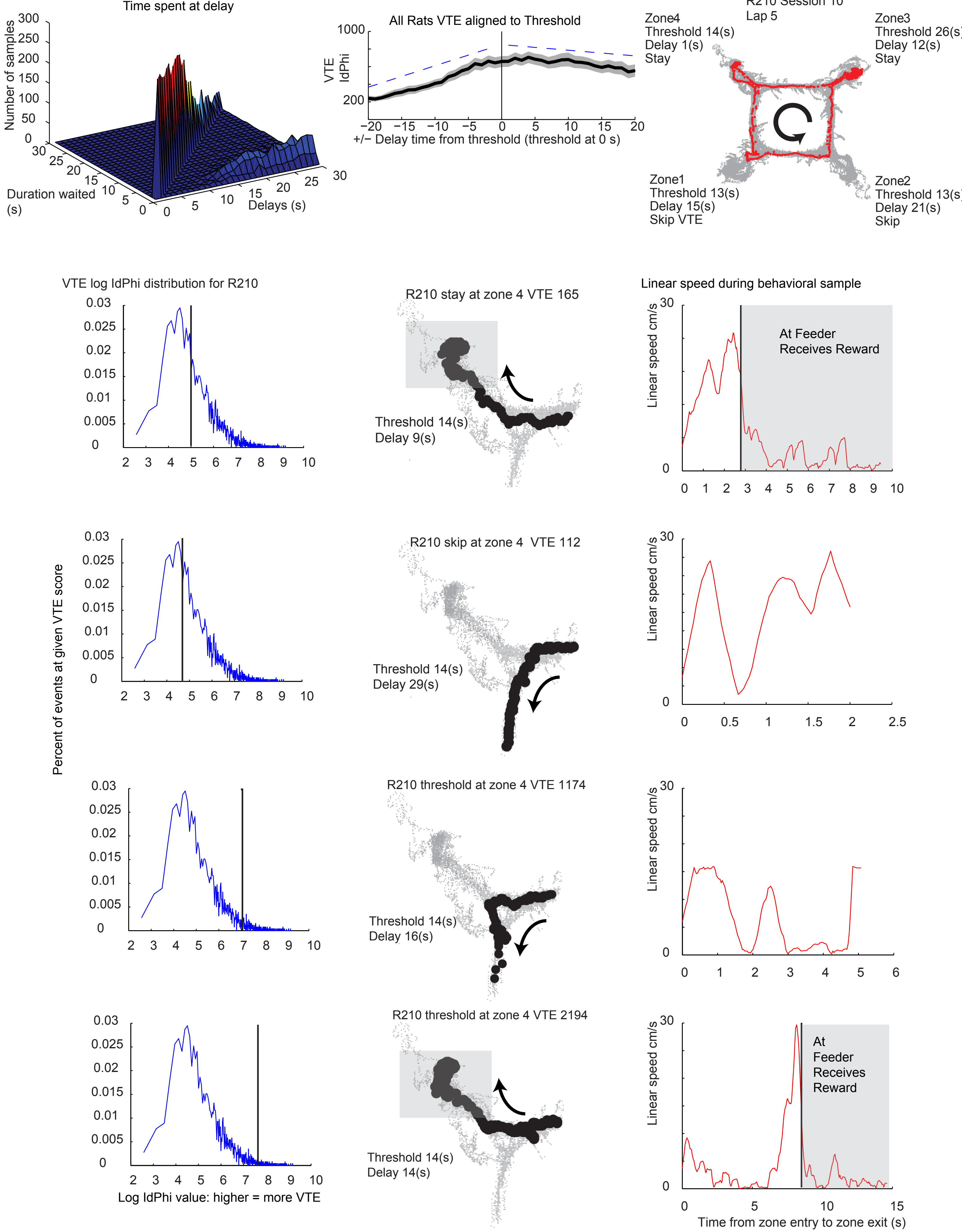
Supplemental Table 2

Summary of the decoding different conditions – Regret and Controls

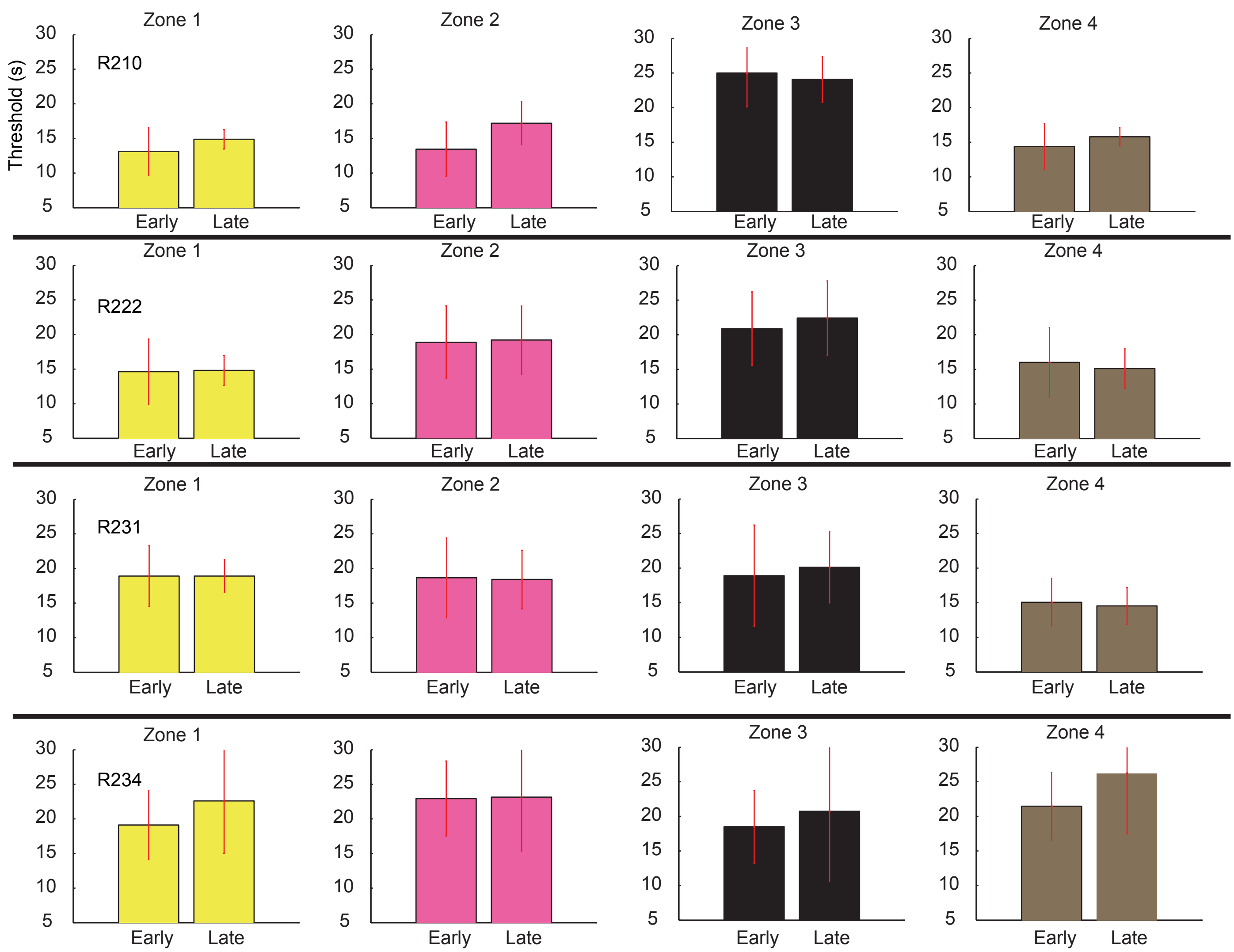
Condition	Offer at previous	Rat's action	Offer at current
Regret-inducing	Delay < threshold	SKIP	Delay > threshold
Control 1	Delay < threshold	STAY	Delay > threshold
Control 2	Delay > threshold	SKIP	Delay > threshold
Rejoice (Supplemental Fig S16)	Delay < threshold	SKIP	Delay < threshold
Rejoice (Supplemental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

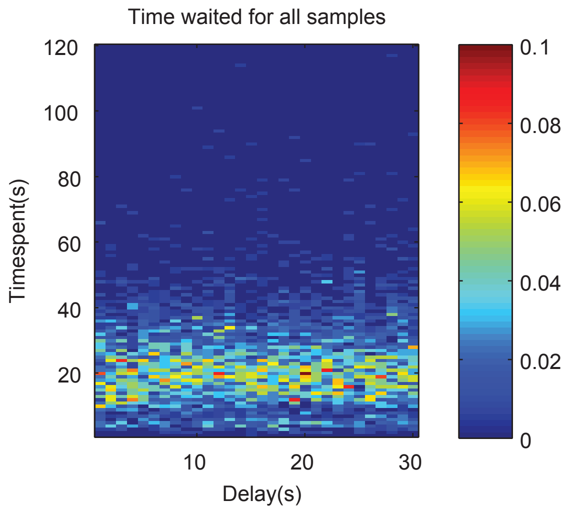
Supplemental Figure S1



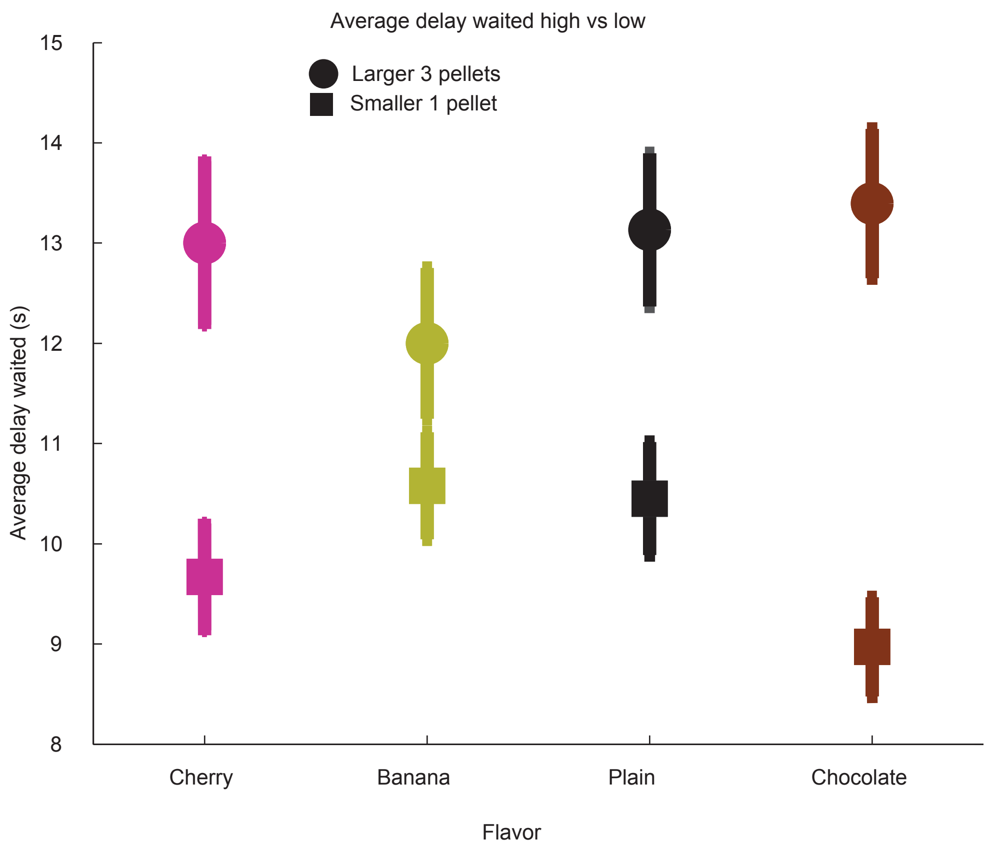
Supplemental Figure S2



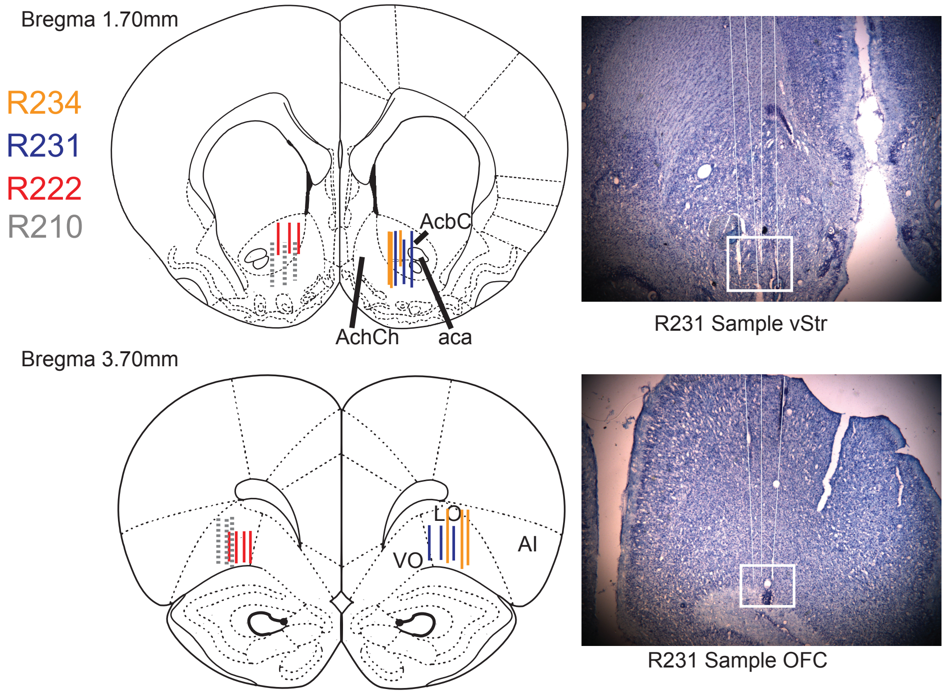
Supplemental Figure S3



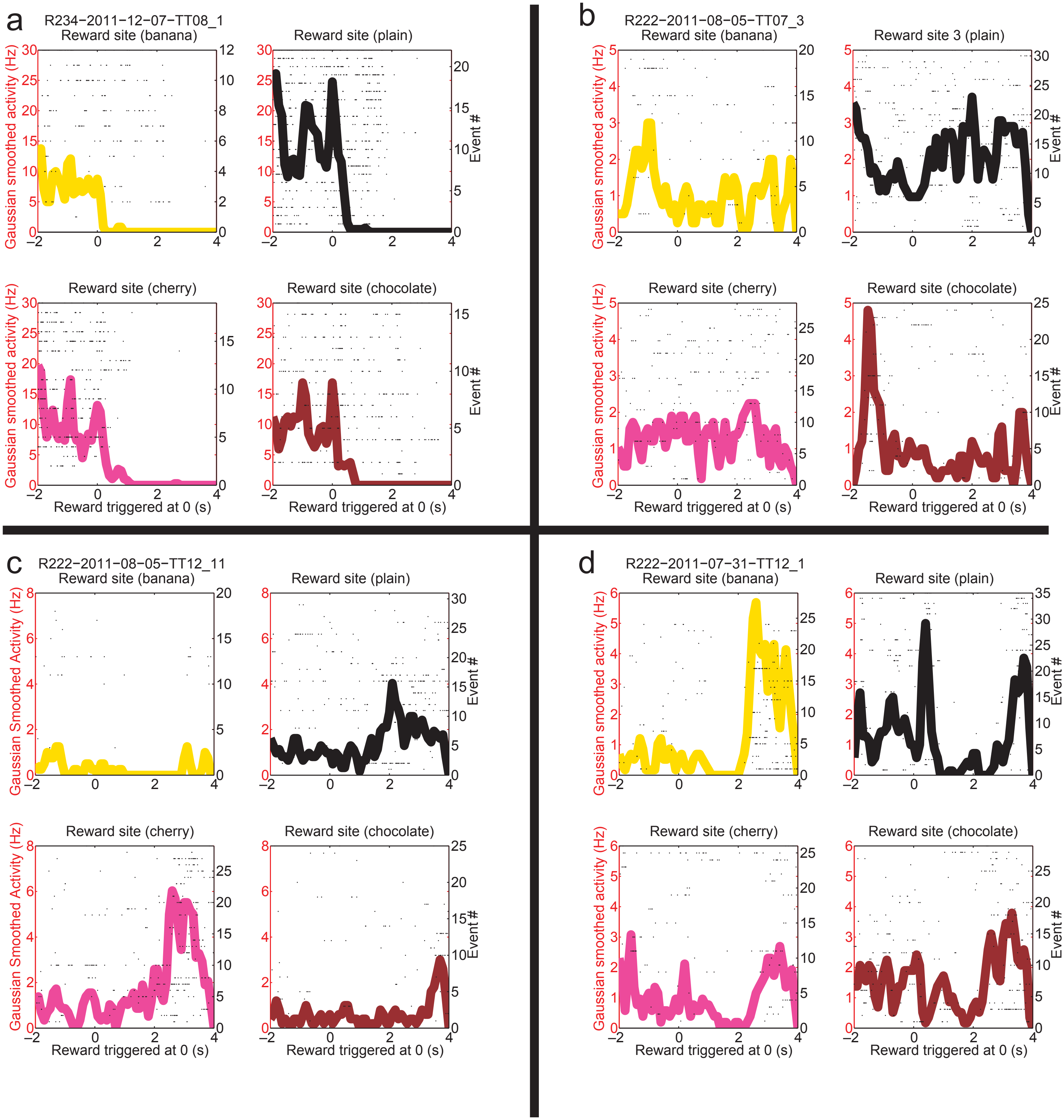
Supplemental Figure S4



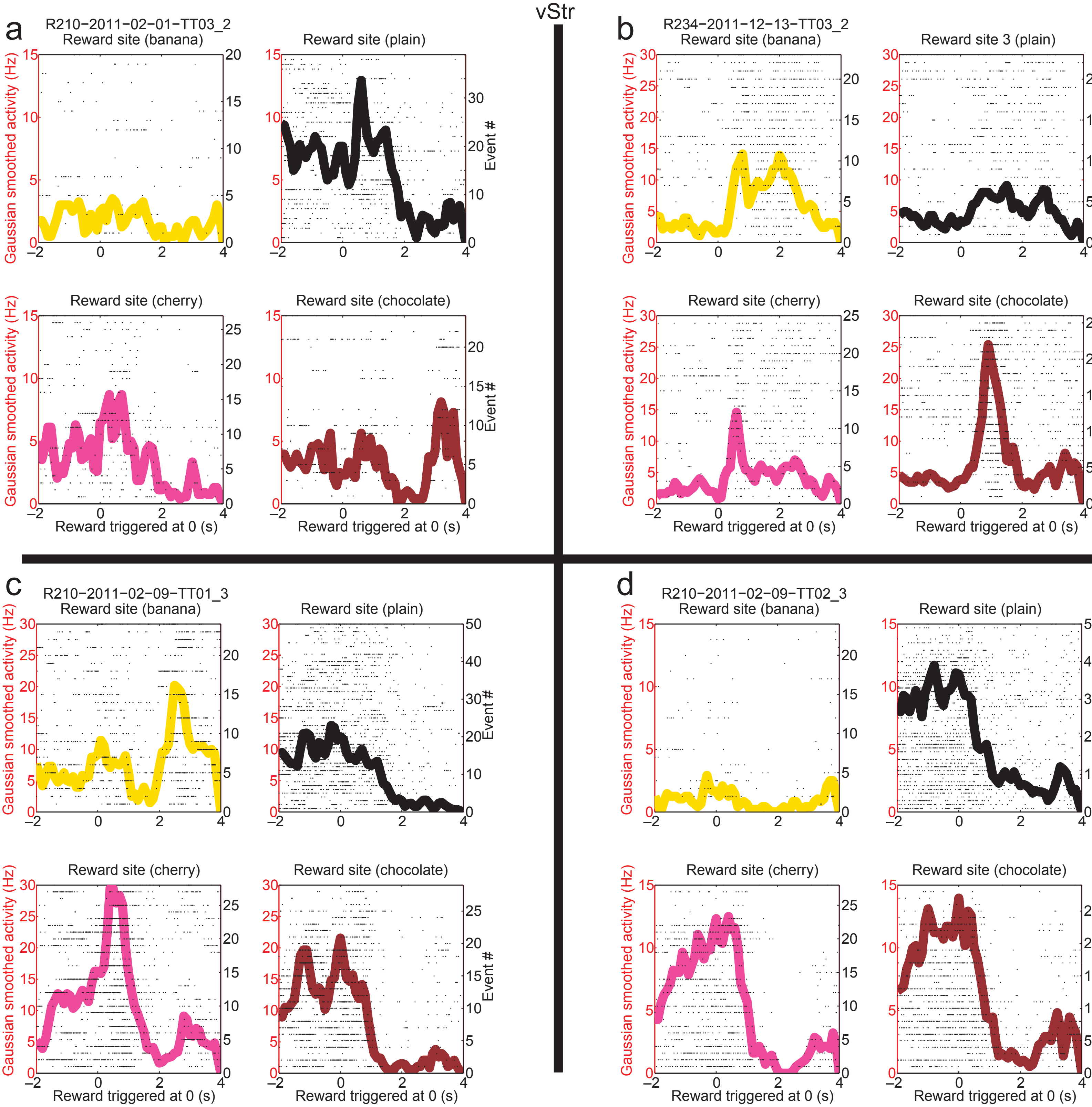
Supplemental Figure S5



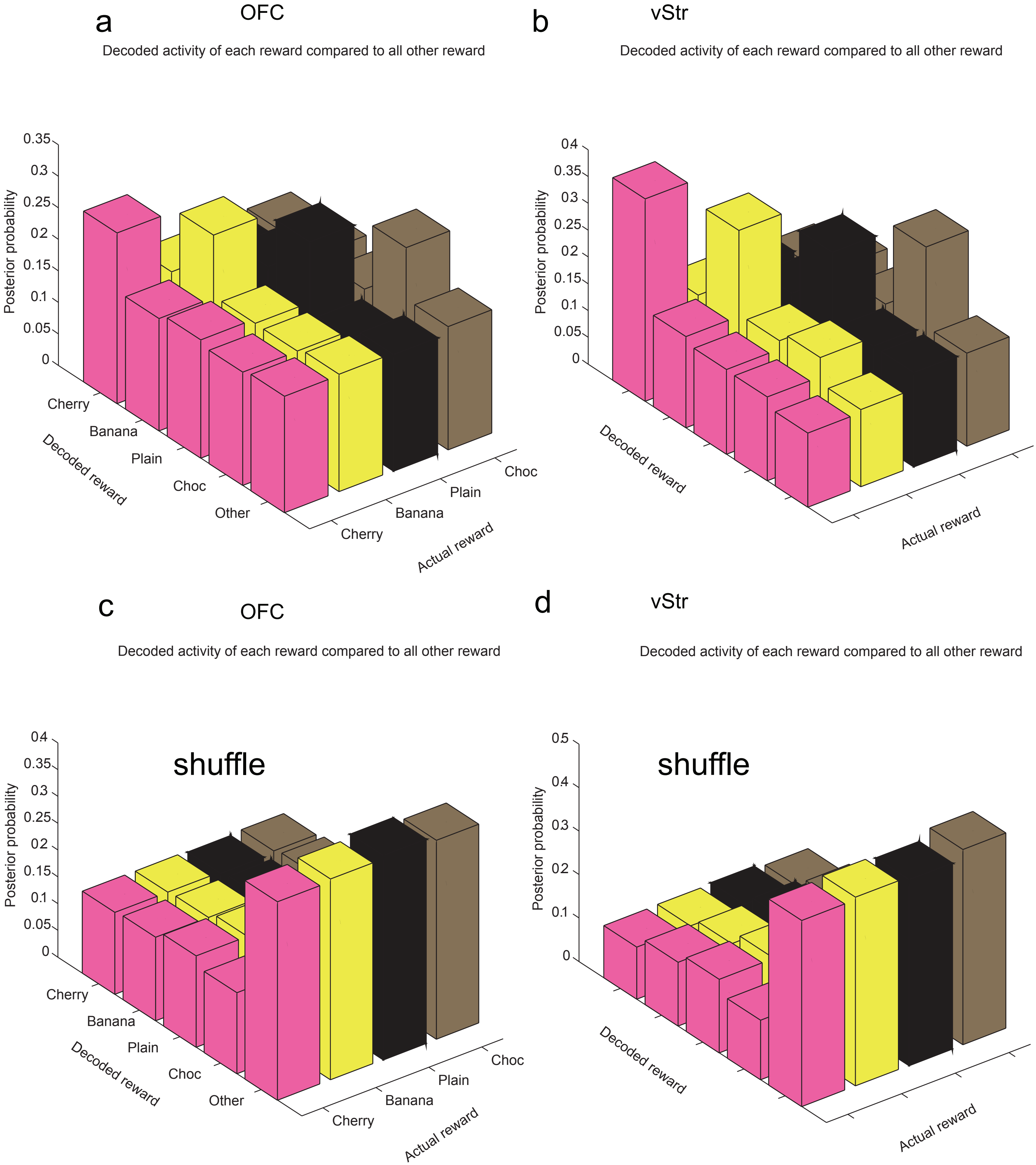
Supplemental Figure S6
OFC



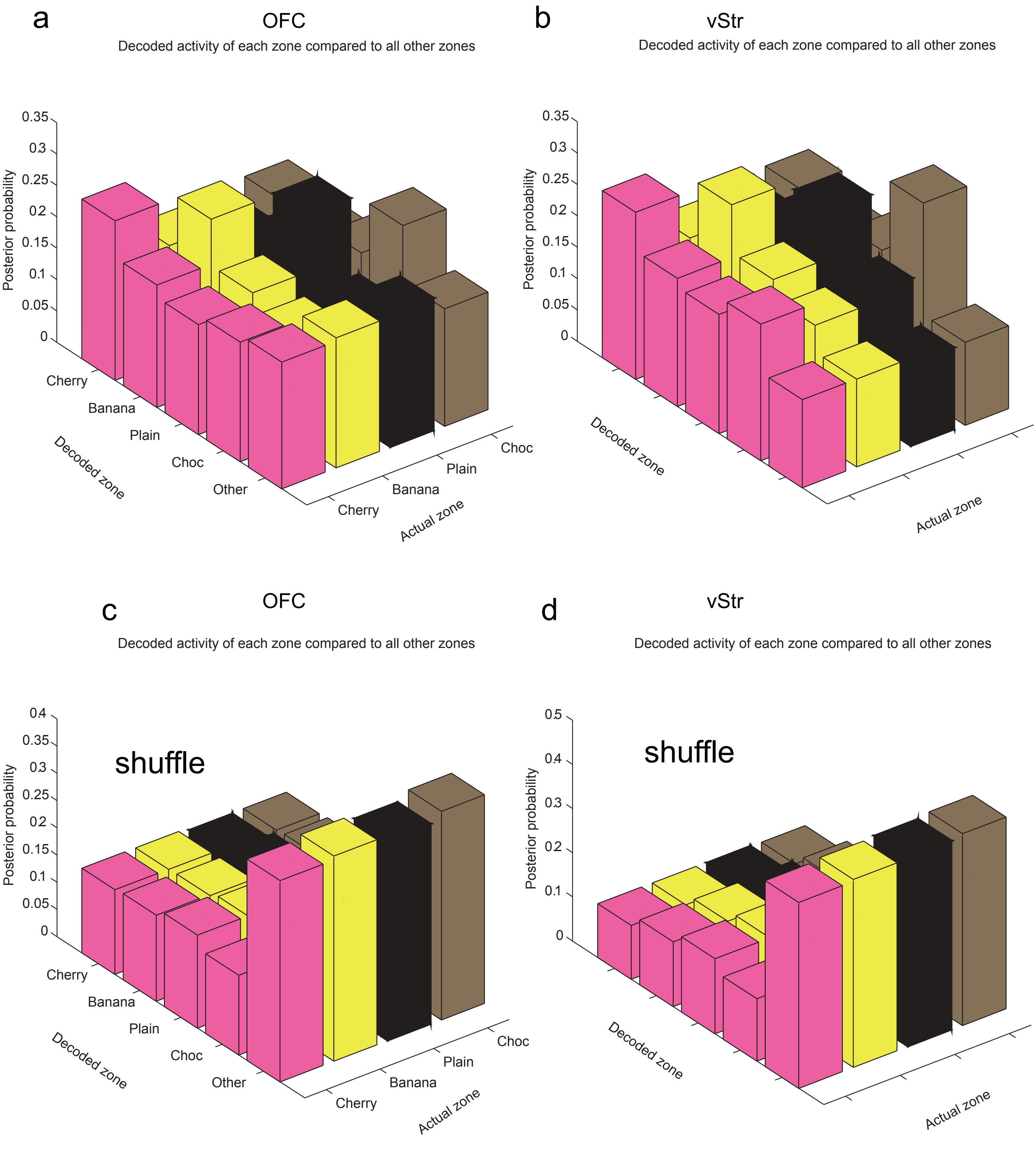
Supplemental Figure S7



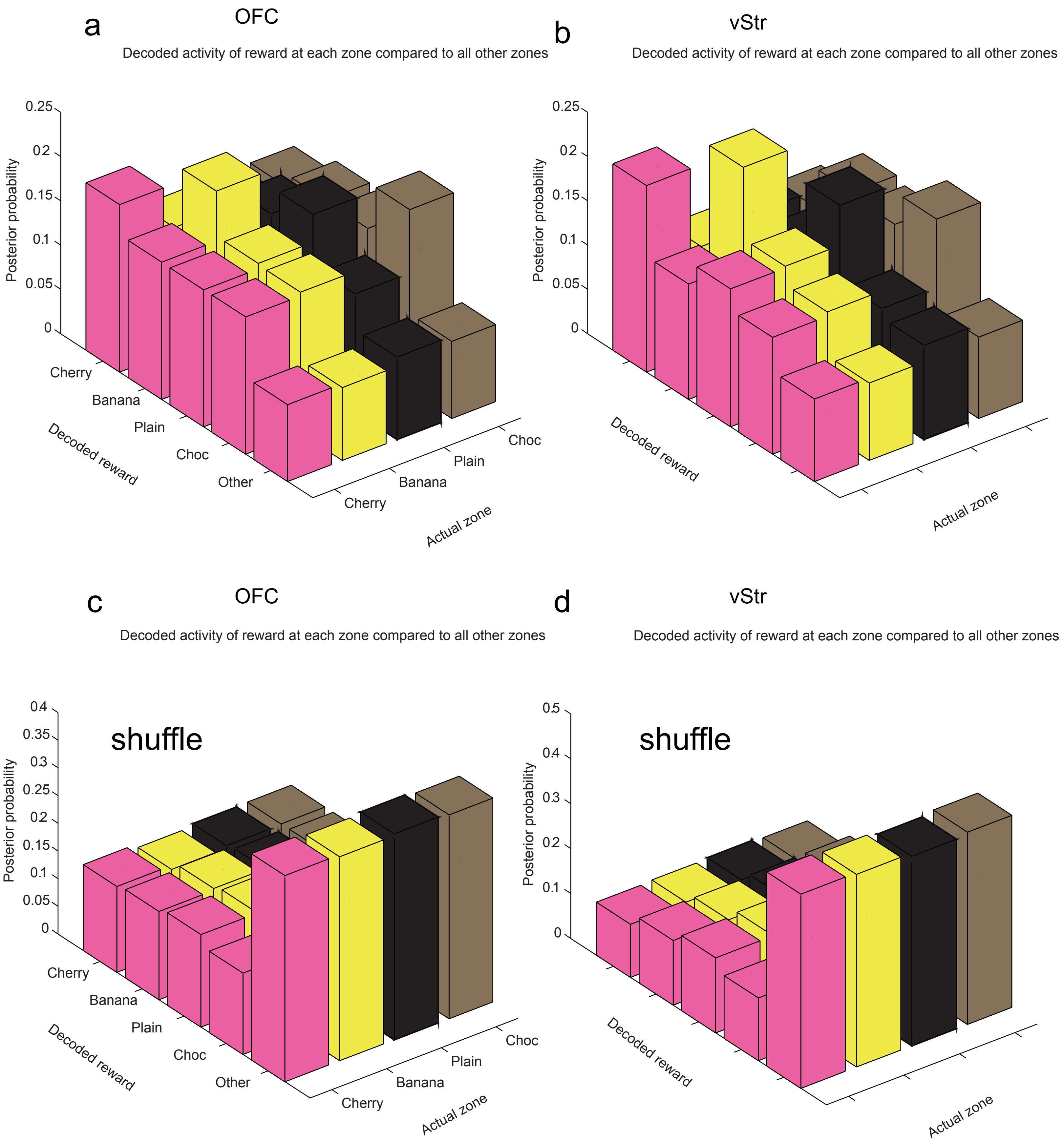
Supplemental Figure S8



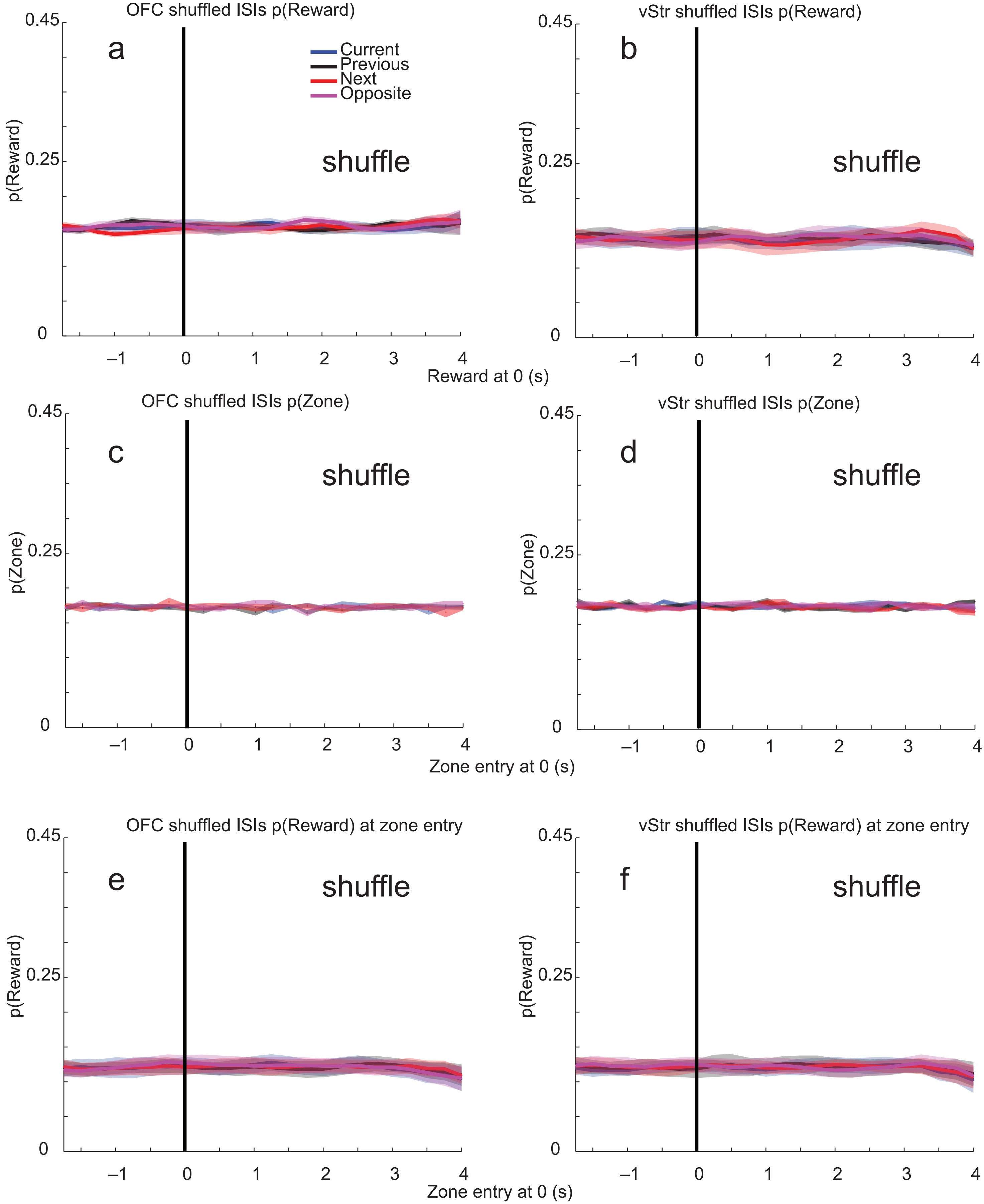
Supplemental Figure S9



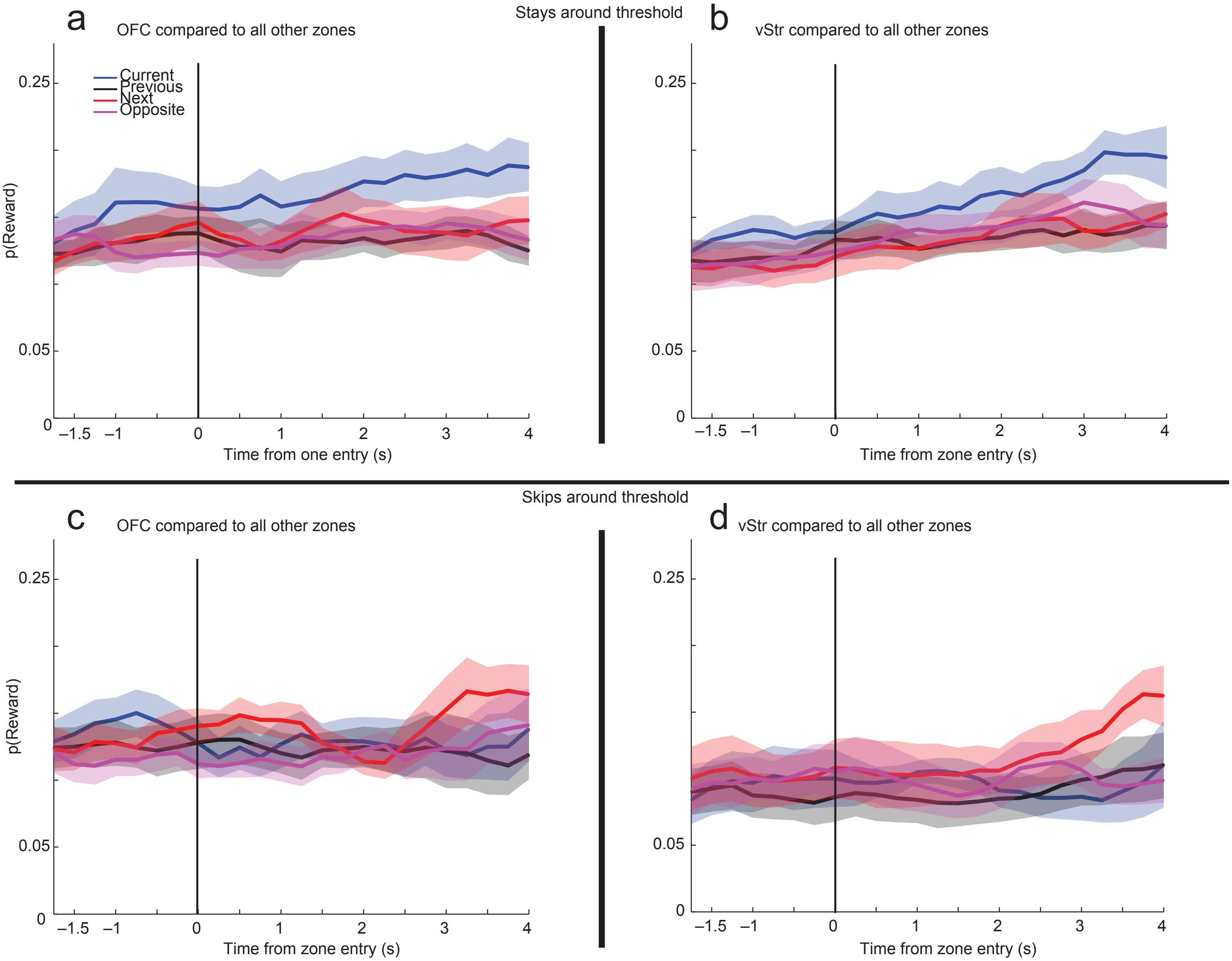
Supplemental Figure S10



Supplemental Figure S11

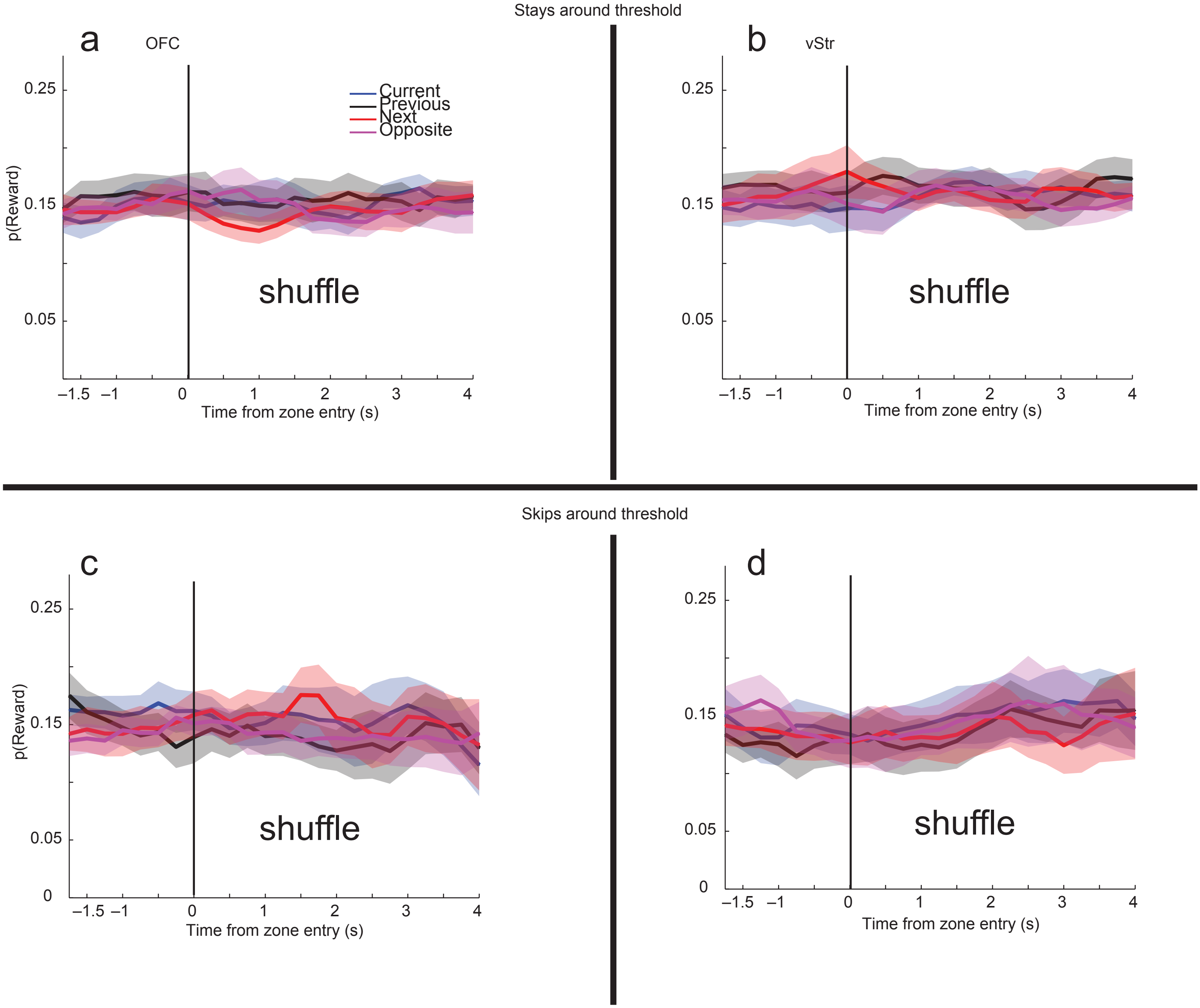


Supplemental Figure S12



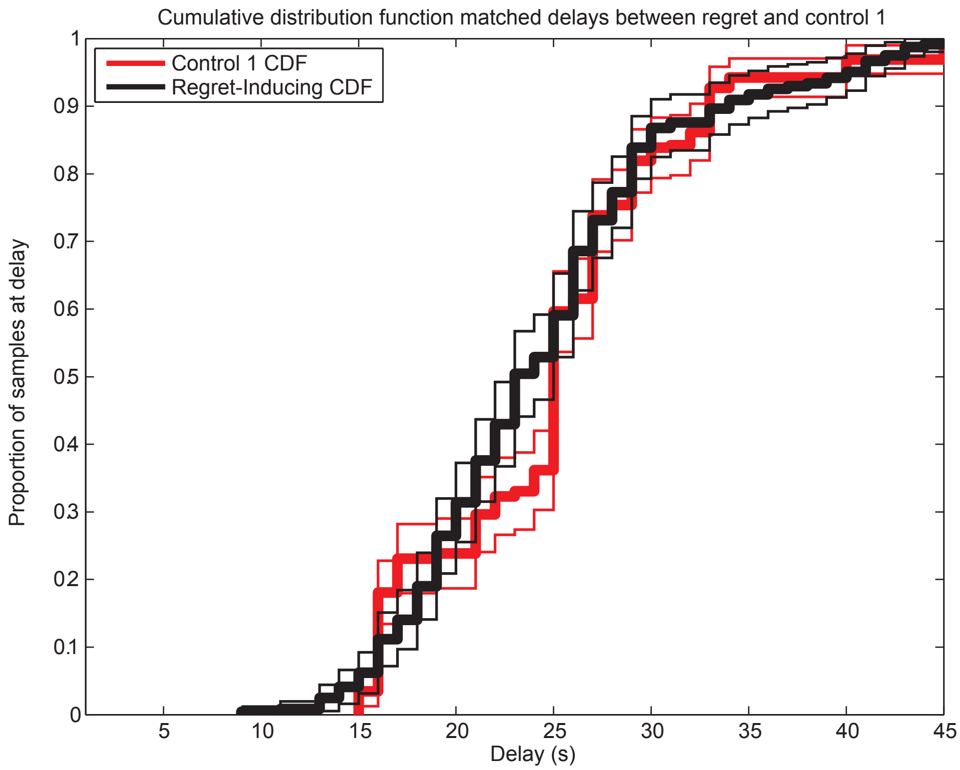
Supplemental Figure S13

Shuffled ISIs: decoding around threshold

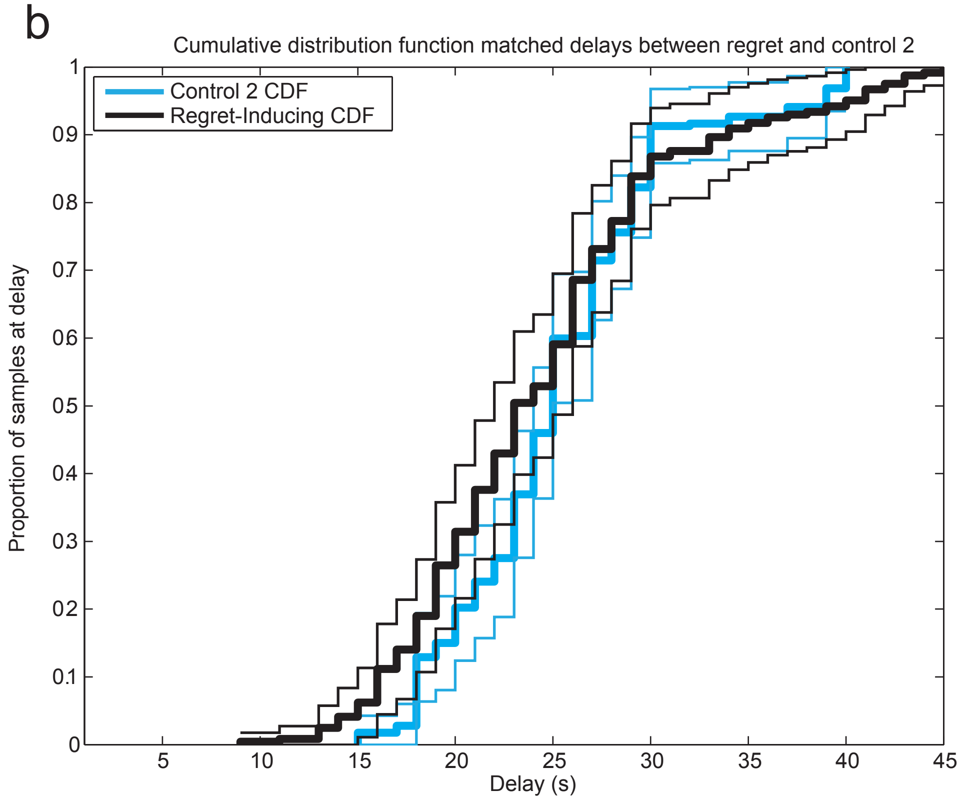


Supplemental Figure S14

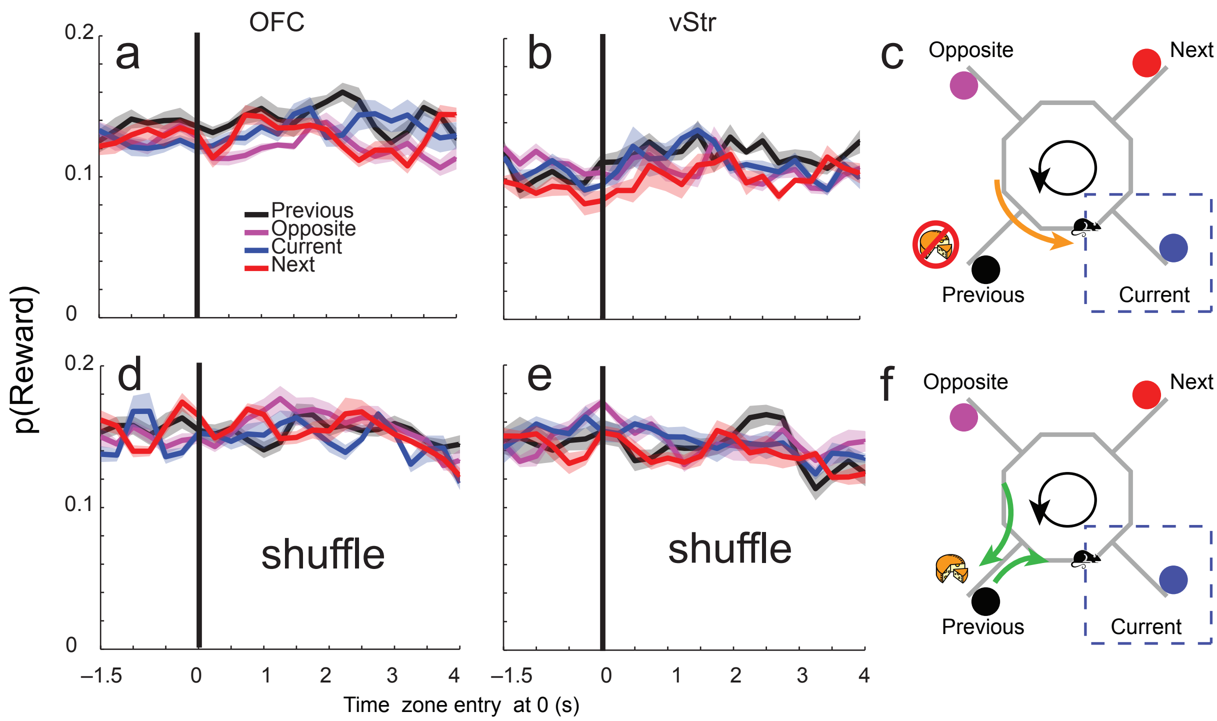
a



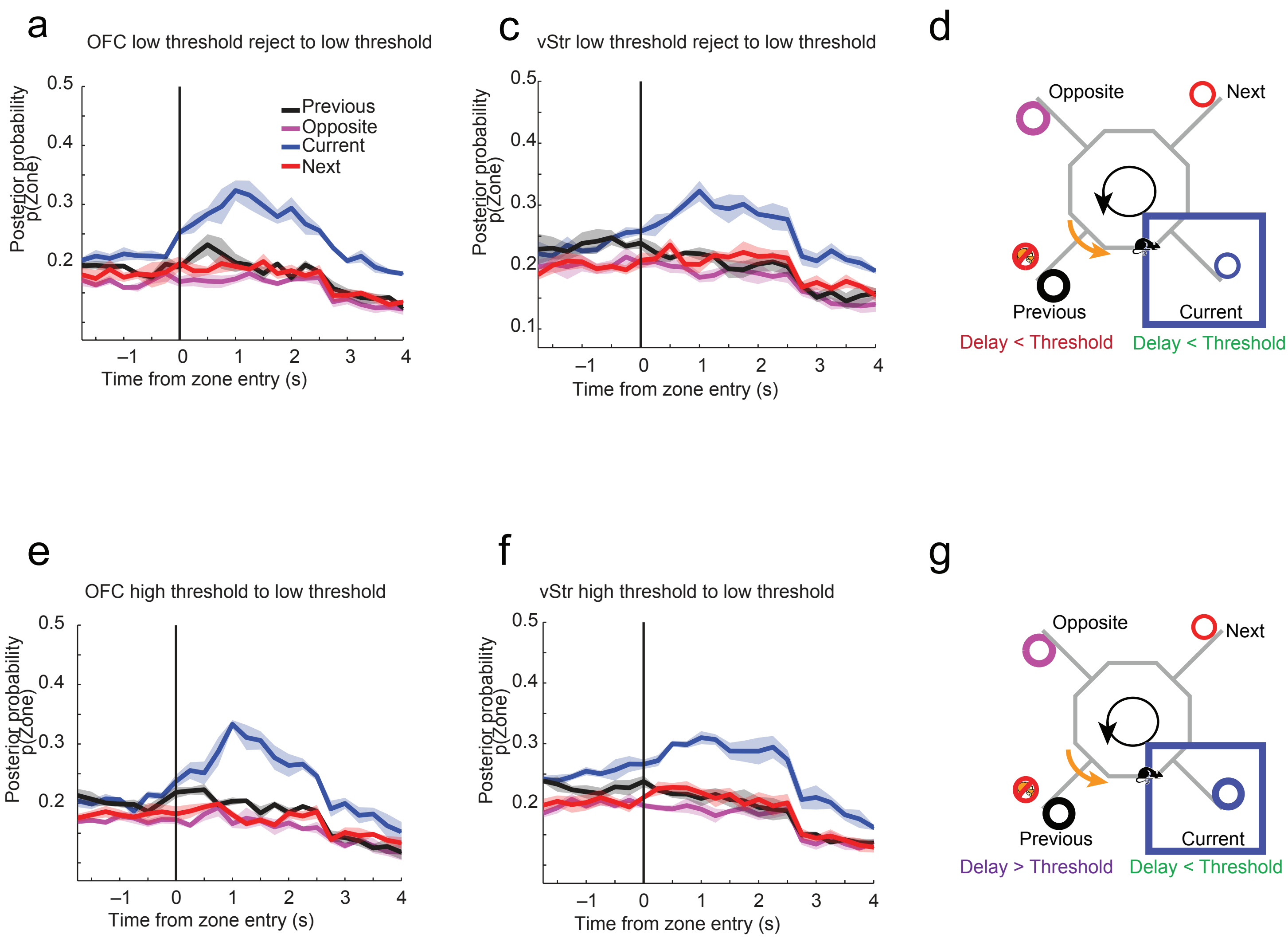
b



Supplemental Figure 15



Supplemental Figure S16



Supplemental Figure S17