Behavioral/Cognitive

Orbitofrontal State Representations Are Related to Choice Adaptations and Reward Predictions

Thomas A. Stalnaker, Nishika Raheja, and Geoffrey Schoenbaum
Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, National Institute on Drug Abuse Intramural Research Program, Baltimore, Maryland 21224

Animals can categorize the environment into "states," defined by unique sets of available action-outcome contingencies in different contexts. Doing so helps them choose appropriate actions and make accurate outcome predictions when in each given state. State maps have been hypothesized to be held in the orbitofrontal cortex (OFC), an area implicated in decision-making and encoding information about outcome predictions. Here we recorded neural activity in OFC in 6 male rats to test state representations. Rats were trained on an odor-guided choice task consisting of five trial blocks containing distinct sets of action-outcome contingencies, constituting states, with unsignaled transitions between them. OFC neural ensembles were analyzed using decoding algorithms. Results indicate that the vast majority of OFC neurons contributed to representations of the current state at any point in time, independent of odor cues and reward delivery, even at the level of individual neurons. Across state transitions, these representations gradually integrated evidence for the new state; the rate at which this integration happened in the prechoice part of the trial was related to how quickly the rats' choices adapted to the new state. Finally, OFC representations of outcome predictions, often thought to be the primary function of OFC, were dependent on the accuracy of OFC state representations.

Key words: **cognitive map; odor; orbitofrontal; rat; single unit**

Significance Statement

A prominent hypothesis proposes that orbitofrontal cortex (OFC) tracks current location in a "cognitive map" of state space. Here we tested this idea in detail by analyzing neural activity recorded in OFC of rats performing a task consisting of a series of states, each defined by a set of available action-outcome contingencies. Results show that most OFC neurons contribute to state representations and that these representations are related to the rats' decision-making and OFC reward predictions. These findings suggest new interpretations of emotional dysregulation in pathologies, such as addiction, which have long been known to be related to OFC dysfunction.

Introduction

People are often faced with different contexts in which conflicting actions are appropriate. For example, using slang or coarse language in a relaxed environment with friends might be appropriate and lead to a positive outcome, such as building rapport; but using the same words at home while having dinner with one's parents might be inappropriate and have bad outcomes, such as disapproval or misunderstandings. Based on one's judgment of the situation, one could categorize social settings into two different "states": slang-appropriate and slang-inappropriate. By categorizing the environment into these different states, one could more easily decide at any point whether to use slang or more formal language. Such a strategy is widespread in human behavior.

Categorizing the environment into different "states," defined as unsignaled contexts associated with different sets of cue- or action-outcome contingencies, is also useful in reversal learning paradigms where, for instance, actions might lead to the opposite outcomes before versus after reversal. Separating the environment into a prereversal state and a postreversal state would be computationally efficient for animals trained in this setting, since it would help them more easily choose the action leading to their preferred outcome in each state, free of conflict with the opposing information (Gershman et al., 2010; Niv, 2019). This is true even if the state transition is unsignaled and can only be inferred

Received Apr. 1, 2020; revised Dec. 21, 2020; accepted Dec. 25, 2020.
Author contributions: G.S. and T.A.S. designed research; T.A.S. analyzed data; G.S. and T.A.S. edited the paper; G.S. and T.A.S. wrote the paper; T.A.S. and N.R. performed research; T.A.S. and N.R. wrote the first draft of the paper.
The authors declare no competing financial interests.
This work was supported by National Institute on Drug Abuse Intramural Research Program ZIA-DA000587. The opinions expressed in this article are the authors' own and do not reflect the view of the National Institutes of Health/Department of Health and Human Services. We thank the National Institute on Drug Abuse Intramural Research Program histology core for technical assistance in histology.
Correspondence should be addressed to Thomas A. Stalnaker at thomas.stalnaker@nih.gov or Geoffrey Schoenbaum at geoffrey.schoenbaum@nih.gov.
https://doi.org/10.1523/JNEUROSCI.0753-20.2020

from the changed contingencies. But how does the brain keep track of such inferred states? One proposal is that the orbitofrontal cortex (OFC) keeps a cognitive map of "state space," a representation of how the current environment is carved up into states, that is used by other areas to keep track of the current state, especially when that state is partially observable or hidden (Wilson et al., 2014). Although the dominant theory has long been that the central function of OFC in behavior is to represent the value or sensory features of expected outcomes (Gottfried et al., 2003; Padoa-Schioppa and Assad, 2006), more recently it has been noted that many of the characteristic effects of OFC lesions would result from an inability to keep track of the current state when it is not directly observable (Wilson et al., 2014). Under this proposal, the reason the OFC represents expected outcomes and reward value is that those quantities often help to define the current state space, especially in tasks commonly used experimentally (Niv, 2019).

Here we examined single-unit activity in 831 lateral OFC neurons recorded in rats performing a task with four distinct states, each consisting of a unique set of response-outcome contingencies. We made four broad predictions as to how neural activity should track states and relate to the rats' behavior across unsignaled state transitions. The first is that the current state should be decodable from the activity of a large proportion of OFC neurons. Second, representations of state in OFC neural activity should change across state transitions as animals integrate information identifying the new state. The new state information should first appear after new outcomes are delivered, because new outcomes would be the first evidence in our task that the action-outcome contingencies that define the state have changed. Third, if state information carried by OFC neurons is functionally important to behavior, the rate at which OFC state representations develop across state transitions should be related to the rate at which animals adapt to the new state. And fourth, other non-state information carried by OFC neurons should depend on state representations; specifically, we predicted that outcome-predictive information would depend on accurate OFC state representations. We tested these predictions by examining the information encoded in pseudo-ensembles of OFC single units using neural decoding algorithms. The results show that OFC ensembles hold strong state representations that are widely distributed across the OFC population, develop across state transitions, and are related to choice behavior and neural representations of expected outcomes.

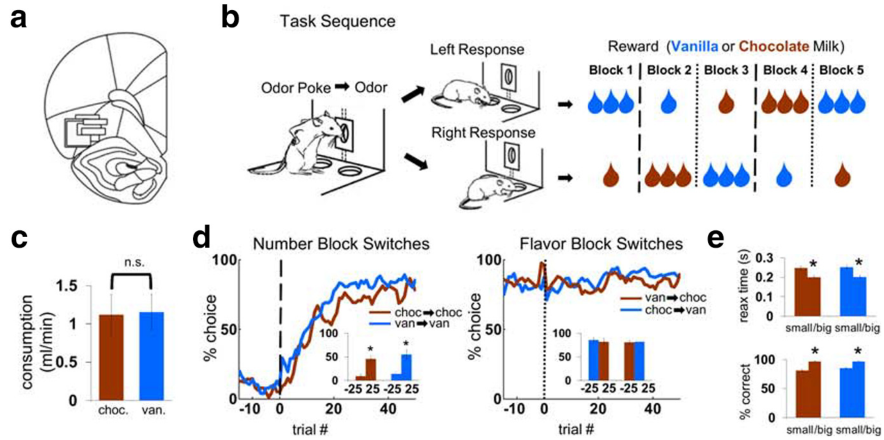

Figure 1. Recordings in OFC, odor-guided choice task, and behavioral results. ***a***, Recording sites in lateral OFC. Black boxes represent the approximate location from which recordings were made in each rat (in the left hemisphere). The width represents the estimated span of the electrode bundle (~1 mm), and the height represents the approximate extent of recording across all sessions. Bregma 2.8-3.6 mm. ***b***, Trial events and reward schedule. Trials started with odor delivery followed by choice for 1 or 3 drops of chocolate or vanilla milk. Two odors indicated forced choices, left or right; a third odor indicated free choice. Reward contingencies were stable across blocks of ~60 trials but switched in number of drops (dashed lines) or flavor (dotted lines) in four unsignaled transitions. The last four blocks always had the same four sets of action-outcome contingencies, but the order differed from day to day. ***c***, Results of a 10 min consumption test run in a separate group of rats ($t_{(10)} = 0.1$, $p = 0.93$). ***d***, Average trial-by-trial choice rates across number block switches (left), separately for 1 drop chocolate⟶3 drops chocolate compared with 1 drop vanilla⟶3 drops vanilla, and flavor block switches (right). Inset, Bar graphs compare choice rate in 25 trials before block switches versus 25 trials after block switches. ANOVA on difference in choice rates across transitions, with factors transition type and initial flavor; main effect of transition type ($F_{(1,92)} = 195.7$, $p < 0.001$), driven by significant changes across number transitions (planned contrast, $F_{(1,92)} = 445.9$, $p < 0.0001$), and insignificant changes across flavor transitions (planned contrast, $F_{(1,92)} = 1.3$, $p = 0.27$); no effect of initial flavor ($F_{(1,92)} = 0.0$, $p = 0.93$); no differences between vanilla-to-chocolate and chocolate-to-vanilla (planned contrast, $F_{(1,92)} = 2.3$, $p = 0.13$). ***e***, Average response latencies and accuracy on forced-choice trials. Within-subjects ANOVAs on reaction time and accuracy: main effects of reward number ($F_{(1,93)} = 62.2$, $p < 0.001$; $F_{(1,93)} = 182.3$, $p < 0.001$) but not flavor ($F_{(1,93)} = 0.3$, $p = 0.57$; $F_{(1,93)} = 5.3$, $p = 0.024$), nor interactions ($F_{(1,93)} = 0.1$, $p = 0.73$; $F_{(1,93)} = 5.1$, $p = 0.027$). $*p < 0.001$ vs. 1 drop condition.

Materials and Methods

Subjects. Six male Long-Evans rats weighing 175-200 g (~60 d old on arrival) were tested at the University of Maryland School of Medicine in accordance with School of Medicine and National Institutes of Health guidelines.

Surgical procedures and histology. Drivable electrodes consisting of bundles of eight 25-μm-diameter FeNiCr wires (Stablohm 675, California Fine Wire), were manufactured, electroplated with platinum to an impedance of ~300 kOhms, and implanted in the left OFC (3.0 mm anterior to bregma, 3.2 mm laterally, and, to begin, 4.0 mm ventral to the surface of the brain) in each rat. At the end of the study, the final electrode position was marked, the rats were killed with an overdose of isoflurane and perfused, and the brains were removed from the skulls and processed using standard techniques. All surgical procedures followed guidelines for aseptic technique.

Behavioral task. Recording was conducted in aluminum chambers fitted with a house light and a panel containing an odor port and two fluid wells (see Fig. 1*b*). The odor port was connected to an air flow dilution olfactometer to allow the rapid delivery of olfactory cues, and the fluid wells were connected to fluid delivery lines containing flavored milk (Nesquik brand chocolate or vanilla) diluted 50% with water. A custom C++ program controlled the house light and solenoid valves that delivered the odors and fluids; the program also recorded photobeam breaks at the odor port and fluid wells and lick detectors inside each fluid well.

Rats were trained before being implanted with electrodes and then retrained to work with the recording cable. The initial shaping phase gradually introduced all elements of the task (described below); thus, rats could learn the associative structure of the task over this period. Recording began when rats could complete five blocks of trials (at least 260 trials) with the cable.

Each of the 94 total recording sessions consisted of a series of self-paced trials organized into five blocks. Rats could initiate a trial by poking into the odor port while the house light was illuminated (for all analyses and figures, trial onset, or the end of the intertrial interval (ITI), was considered the time the house light was turned on). Beginning 500 ms after the odor poke, an odor would be delivered for 500 ms. The cessation of odor delivery served as a go-response indicating that rats could respond by poking at the left or right fluid well, after which fluid delivery would begin following a 500 ms delay. Three different instructive odors (which were initially neutral and then unchanged throughout training and testing) were used: two of these indicated fluid would only be available at the left well or right well (forced-choice left and forced-choice right, respectively); a third indicated fluid would be available at either well (free-choice). Trials were presented in a pseudorandom sequence such that the free-choice odor was presented on 7 of 20 trials and the left/right odors were presented in equal numbers ($\pm$1 over 250 trials).

Rewards were either one drop or three drops of chocolate or vanilla milk, with drop size ~0.05 ml and 500 ms between drops. Response-reward contingencies were consistent within blocks of trials, such that the same reward would be delivered for every correct response, either free- or forced-choice, according to whether it was at the left or right well. Upon each block transition, either the number of drops (1 or three drops) or flavor (chocolate or vanilla) would change on both sides while the other variable remained constant. These block transitions were not explicitly signaled and could not be predicted based on the exact number of trials. The first block, consisting of on average 43 $\pm$ 16 (SD) trials, was used to set the rats' expectations before the first transition. The length of the last four blocks varied nonsystematically at ~65 $\pm$ 11 (SD). The reward schedule was arranged so that the last four blocks always consisted of the same four sets of action-outcome contingencies (e.g., left-3 drops chocolate, right-1 drop vanilla), although the order of the blocks varied randomly from session to session. These four blocks defined four unique "states" of the task.

During testing, rats were limited to 10 min of *ad libitum* water each day, in addition to fluid earned in the task.

Flavor preference testing. In 6 rats from a separate experiment (same strain, source, and water restriction regimen), we compared consumption of the chocolate versus vanilla milk solution in two-bottle tests. All rats were tested for 10 total minutes, with the location of the bottles swapped every 30 s. Two rats were given five 2 min tests while the other 4 rats were given one 10 min test each.

Single-unit recording. Procedures were the same as described previously (Stalnaker et al., 2010). Wires were screened for activity daily; if no activity was detected, the rat was removed and the electrode assembly was advanced 40 or 80 μm. Otherwise, a session was conducted, and the electrode was advanced by at least 40 μm at the end of the session. Neural activity was recorded using Plexon Multichannel Acquisition Processor systems, interfaced with odor discrimination training chambers. Signals from the electrode wires were amplified and filtered by standard procedures described in previous studies. Waveforms (>2.5:1 signal-to-noise) were extracted from active channels and recorded with event time stamps sent by the behavioral program.

Data analysis. Units were sorted using Offline Sorter software from Plexon, using a template matching algorithm. Sorted files were then processed in Neuroexplorer to extract unit time stamps and relevant event markers. These data were subsequently analyzed in MATLAB.

Experimental design and statistical analysis. Behavioral data were analyzed by processing time-stamped events using MATLAB. Free-choice rates (see Fig. 1*d*) are the average percent of free-choice trials toward a particular side, scaled according to the overall number of rewarded trials (i.e., the scale includes both free- and forced-choice trials). For example, the free-choice rate in the last 25 trials before a block switch would take all free-choice trials occurring within the last 25 of any kind of trial and calculate the percentage of those trials toward the side delivering one drop (or three drops). Reaction time (see Fig. 1*e*) is the time from the cessation of odor delivery to withdrawal from the odor port, and percent correct is the percentage of forced-choice trials in which the rat chose the rewarded side.

For decoding analyses of neural data, MATLAB scripts and functions from the Neural Decoding Toolbox (www.readout.info) (Meyers, 2013) were modified as needed. Trial-aligned spike trains were constructed from each neuron by aligning raw spike trains to trial events (house light on, odor delivery start, odor port withdrawal, reward delivery start, house light off) and then concatenating them according to the average time between the events in the dataset. Then spikes were binned into sliding epochs of 500 or 1000 ms across the trial period. All rewarded trials were labeled according to the state, defined by the set of action-outcome contingencies available in the block from which the trial occurred (e.g., left-3 drops chocolate milk, right-1 drop vanilla milk). Then random pseudo-ensembles were repeatedly selected from all OFC neurons. For each decoding run, trials were randomly split into a test set, with one trial from each neuron in each state, and a training set, with as many remaining trials as possible such that an equal number of training trials were selected for each state for each neuron. Training sets were then averaged across trials for each neuron. Each test trial was classified according to which training set had the highest correlation coefficient with it across neurons. This test training split selection was repeated for each trial, and then the entire process was repeated between 100 and 1000 times, depending on the analysis. The overall percentage correct for each sliding bin is shown in the figures. Significance was assessed by randomly assigning trial labels 100 times, repeating the entire analysis each time, and from these 100 datasets constructing a null distribution of decoding percentages. For a bin to be defined as significantly greater than chance, it had to be part of a run of 5 consecutive bins in the upper 1% of the null distribution.

For the analysis in Figure 3, an ANOVA on state was run for each neuron across all correct trials using firing rate during light on to odor on (or, for the analysis in Fig. 3*b*, other epochs), and neurons were ranked according to the resulting *F* value. Then decoding was run for repeatedly selected pseudo-ensembles from each segment of the ranked population (>90%, 81%-90%, etc.). For the analysis in Figure 3*b*, *F* values were calculated for different epochs across the trial (last 1.5 s of ITI, 1.5 s before odor delivery, odor delivery to choice response, first 1.5 s of reward delivery, and 1.5 s after reward delivery to 1.0 s after house light off), and Pearson correlations were run between sets of *F* values from each pair of epochs.

For the multiple linear regression models of trial event factors, we used the MATLAB function "fitlm" to create the odor+direction+outcome model and "step" to add state as an additional predictor. We corrected for multiple comparisons in the odor+direction+outcome model by requiring 5 consecutive sliding bins with $p < 0.01$ and disregarding bins with fewer significant neurons than chance by χ^2 test. Because state and outcome are partially correlated (only two of the four possible outcomes occur in each state), state was considered to add significantly to the model when it had $p < 0.05$ and had a lower *p* value than outcome when each was added to the model, or when it had $p < 0.05$ when added after outcome to the model.

For the analysis shown in Figures 4 and 5, we constructed sets of training trials consisting of the last 20 trials of all blocks before transitions in the number of drops, and the last 20 trials of all blocks after transitions in the number of drops. Then we selected test trials in sliding 5-trial segments for Figure 4 (or 10-trial segments for Fig. 5*b*–*d*) from the end of the "before" training set to the beginning of the "after" training set. The decoding percentage was then the percentage of test trials classified as "before" versus "after." The heat plots were constructed by smoothing and converting the decoding percentages into a color scale. The significance levels shown on the scale (dotted lines) were calculated based on the null distribution constructed from the first set of test trials by randomly assigning trial labels 100 times, each time running the entire decoding analysis. For the analysis in Figure 5, we ranked drop number transitions according to the proportion of free-choices toward the 3-drop side in the first 10 trials after the transition. We took the top quartile of such transitions (highest choice rate toward the new 3-drop side) and the bottom quartile (lowest choice rate toward the new 3-drop side). Then we separately analyzed the neuronal ensembles recorded in each, matching the ensemble sizes used by the decoder in each condition. Because there were fewer neurons in these sets, we used 10-trial test sets and 1000 ms sliding epochs for this analysis to better smooth the data. For the analyses in Figure 5*c*, *d*, we included only forced-choice trials in the test-sets to avoid any bias in the rewards received in the top quartile versus bottom quartile test sets (because by definition, free-choice trials were more often toward the 3-drop side in the top quartile sets).

For the analysis shown in Figure 6, we constructed training sets for each state consisting of all rewarded forced-choice trials, excluding those occurring in the first 10 trials after drop number transitions. Test sets consisted of free-choice trials toward the 1-drop side or the 3-drop side, labeled according to the state, again excluding those occurring in the first 10 trials after drop number transitions. Ensemble sizes were again matched between the two conditions. This analysis tested whether a choice of the less preferred outcome (the 1-drop side) was associated with worse state representations, compared with state representations when the more preferred outcome was chosen. Forced-choice trials were chosen to construct the training sets because those trials were unbiased as to whether the preferred outcome was chosen. Significance levels were again assessed by randomly assigning state labels 100 times, each time repeating the entire analysis, to construct null distributions for each test set and each bin.

For the analysis shown in Figure 7, we first ran for each neuron an ANOVA on state using firing rate in the postreward epoch (1500 ms beginning 1500 ms after reward delivery) during the first 20 trials after block switches. The purpose of this was to find the neurons that were best at decoding state postreward right after switches, because neurons that were normally good at decoding state could provide a reliable readout of the trials on which state information was degraded in OFC as a whole. The 200 best-ranked neurons were then used to decode state. Then a yoked set of neurons was selected, consisting of all other neurons recorded in the same sessions as the state-decoding neurons. These neurons were used to decode the expected outcome. In order to ensure that this ensemble was as good at decoding expected outcome as possible, the worst 200 ranked neurons on an ANOVA on outcome prereward were excluded from this yoked set. Then we separated test trials into those on which the state had been decoded correctly by the first ensemble (the vast majority) and those on which the state had been decoded incorrectly by that ensemble, and separately calculated expected outcome-decoding percentages across sliding sets of epochs for each set of test trials. Significance was determined in the same way as in the other analyses. Because of how the yoking of the two ensembles worked, the final average ensemble size was 92 for state and 184 for outcome. We found that the qualitative pattern of results did not depend on exactly which neurons were selected for the state-decoding ensemble and the outcome-decoding ensemble, except that the state-decoding ensemble had to decode at a high percentage (>90%). For the decoding percentage as a function of ensemble size (see Fig. 7*c*,*d*), we calculated best fit parameters for a log function for each curve, and then calculated whether the parameters were significantly different using the nlinfit and nlparci MATLAB functions.

Results

We recorded 831 units from the lateral OFC (Fig. 1*a*) as rats performed an odor-guided choice task in which blocks of trials defined states that required the recall and application of conflicting associative information. A differently focused analysis of these units was previously published (Stalnaker et al., 2014), but here we sought to test whether OFC units held information about the current task "state." The task, illustrated in Figure 1*b*, consisted of a series of five blocks of trials presented across each of the 94 recording sessions. Each block was made up of a series of trials in which a particular milk reward (defined by the number of drops and the flavor) was delivered for a correct left response and a different milk reward for a correct right response. Unsignaled block switches resulted in changes to either the number of drops (one or three) or the flavor (vanilla or chocolate) of the rewards available at each well. Block switches were arranged such that each of the four unique sets of response-outcome contingencies (defining "states") were included in every session, although their order varied between sessions. There were four possible orders, one of which was randomly chosen for each session, such that the first switch was always a change in drop numbers, the second a change in flavors, the third a change in drop numbers, and the fourth a change in flavors. Each trial was initiated by turning on the house light, which signaled that the rat could make a nose poke. Nose pokes were followed after 500 ms by the delivery of an odor that signaled whether the trial would be free-choice (35% of trials), meaning that either well would yield reward, or forced-choice (65% of trials), meaning that only one of the two wells would yield reward. Free-choice trials served as a running index of which well the rats preferred at each point in the block, whereas forced-choice trials allowed outcomes at both wells to be experienced and neural data related to both outcomes to be analyzed. However, across both free- and forced-choice trials, the specific reward delivered for a correct left or right response in a particular block, and hence the state, was the same.

After initial shaping on the task, electrode bundles were implanted in lateral OFC. Rats' free choices tended to favor the 3-drop side. Across block transitions in which the number of drops changed, the free-choice rate switched from the 3-drop side on the old block to the 3-drop side on the new block over ∼20-25 trials on average, before reaching an asymptotic rate of ∼80%-90% (Fig. 1*d*). Rats' free-choice behavior across flavor transitions indicated no preference for chocolate- versus vanilla-flavored milk, as did a preference test run in a separate group of rats (Fig. 1*c*,*d*). On forced-choice trials, the overall error rate was very low (∼5%), but the preference for the 3-drop side was evident in significant differences in both the error rate and the latency to respond to the fluid well (Fig. 1*e*).

Do OFC ensembles represent the current state?

To test whether OFC ensembles contained information about the current state of the task, we constructed a simple decoding algorithm using all correct trials in the last four blocks of the sessions. Using an ensemble size of 25 OFC neurons, this algorithm decoded the current state above chance across all parts of the trials, including the ITI (Fig. 2*a*, left). When we calculated the state-decoding accuracy as a function of ensemble size, accuracy increased rapidly with more neurons, reaching an asymptote at nearly 100% for ensembles of ∼800 neurons across all parts of the trial (Fig. 2*b*). This shows that OFC ensembles tended to fire in distinct patterns according to the current state, such that a downstream recipient of OFC output could easily determine the current state, even when no external stimuli were being delivered (i.e., when the state was hidden).

We performed two control analyses to test whether state-decoding reflected an abstract representation of the state (Fig. 2*a*, right). In the first (Fig. 2*a*, top right), we trained the decoder on trials immediately after the outcome at the left well in each block was delivered, and we tested it on trials immediately after the outcome at the right well was delivered. Decoding in this condition was not different from decoding using all trials with matched training set sizes. Thus, state representations did not reflect a simple memory of which outcome had been delivered on the previous trial. Second, we trained the decoder on forced-choice trials and tested on free-choice trials (Fig. 2*a*, bottom right). Here decoding was not different from training on forced-choice and testing on held-out forced-choice trials, demonstrating that state representations generalize across trial type, consistent with an abstract representation of state.

We next asked how widespread this state information was across the OFC population. We ranked individual neurons according to each neuron's F value from an ANOVA run on firing rate using the factor "state" and then formed ensembles from subsets of the ranked population (e.g., >90%, >80%, etc). The average rank of the neurons making up these ensembles was highly correlated with the decoding accuracy of the ensembles, with the correlation line crossing the y axis above the significance

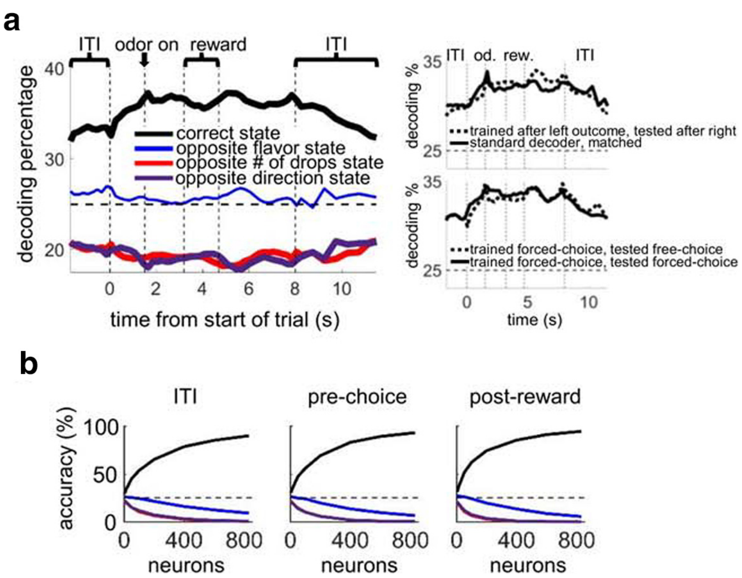

Figure 2. OFC pseudo-ensembles decode the current state accurately across all parts of trials. ***a***, Left, Black line indicates the percentage of trials classified to the correct state, using a 25 neuron pseudo-ensemble and a sliding 500 ms epoch aligned to different trial events and concatenated (the end of the ITI indicates the time the house light was turned on). Other lines indicate misclassification to the other possible states: blue represents the state with opposite flavor but the same number of drops (at both wells) as the correct state; red represents the state with the opposite number of drops but same flavor (at both wells) as the correct state; purple represents the state with the same two outcomes as the correct state, but at the opposite wells. Thick lines indicate significant difference from chance using a bootstrap distribution with shuffled labels ($p < 0.01$ for at least 5 consecutive bins). All rewarded trials were included in the decoder. Smaller panels on right of ***a*** represent two control analyses: Top, Training on trials immediately after outcome at left well, testing on trials immediately after outcome at right well. This shows that state representations do not reflect memory traces of the outcome delivered on the previous trial. Bottom, Training on forced-choice and testing on free-choice trials, showing that state representations generalize across trial type. Significant differences in control analyses were assessed at $p < 0.01$ by comparing decoding percentages as a function of ensembles sizes as described in Materials and Methods (nonsignificant for both). ***b***, State-decoding accuracy as a function of ensemble size using three 1000 ms epochs, with misclassification shown as in ***a***. Across all epochs, decoding accuracy approached 100%, with states differing only in the flavor of outcomes being the most likely to be misclassified. Dotted lines in all plots indicate chance level of accuracy (25%).

line for $p = 0.001$ (by bootstrap distribution with shuffled labels; Fig. 3). In addition, when we ran the state decoder with 1000 randomly selected 25 neuron ensembles, 99.4% had decoding significantly better than chance ($p < 0.001$, data not shown). Because we recorded neurons indiscriminately (without prescreening for task involvement), these analyses suggest that the vast majority of all OFC neurons can contribute to state representations held by small ensembles.

This analysis also revealed something unexpected. Although we ran the ANOVA to rank the neurons using a particular part of the trial (an epoch ending before the instructional odor was delivered to the rat; Fig. 3*a*, shading), the resulting ensembles were equally good at decoding the state across all other parts of the trial. This suggests that, although the average neuron held only a small amount of state information, it tended to hold that information across the whole trial. If, conversely, individual neurons had tended to specialize in one part of the trial and ensembles aggregated this information, our ranked ensembles would have been biased toward the epoch used to rank them. To confirm our initial impression, we tested the correlation between individual F values derived from different epochs across the trial (Fig. 3*b*). All epoch pairs showed significant correlations ($p < 0.001$) and high R^2 values, showing that, indeed, individual neurons held state information across entire trials. For example, 84% (402 of 480) of all OFC neurons with $p < 0.01$ for the state ANOVA using the preodor epoch also had $p < 0.01$ for the ANOVA using the post-reward epoch, with $R^2 = 0.38$ for the F value correlation between these two epochs. The average R^2 value across all epoch pairs was 0.29. Interestingly, the lowest correlations were observed between the reward period, in which OFC neurons tend to have strong phasic responsiveness, versus all other epochs (average $R^2 = 0.15$).

A related question is whether the OFC state representations seen in ensemble activity are built from representations of trial events carried by individual neurons, such as odor identity, choice direction, and outcome identity. We tested this idea by running ANOVAs on all neurons' firing rates in all time bins across all three trial event factors. After correcting for multiple comparisons, this yielded 434 of 831 neurons that were significant for at least one factor in one bin. Then we ran the state-decoding algorithm separately for the event-significant population versus the remaining event-nonsignificant population. As shown in Figure 3*d*, the event-significant population was significantly better at decoding state across the entire trial and ITI. However, the event-nonsignificant population still decoded state better than chance, showing that OFC state representation did not require strong event representations. Furthermore, the event-significant population decoded state better independently of the proportion of event-selective neurons in a particular bin (Fig. 3*d*, top, shaded circles). Indeed, the event-significant population was better at representing state even during the ITI and preodor-poke period, when there were no more event-selective neurons than chance. This suggests that state representations were not built from event representations, but rather that selecting the event-significant population distinguished the part of the OFC population that best represented both trial events and state. We further tested this idea by running a multiple linear regression model with factors odor, direction, and outcome, and then adding state to the model. In bins and neurons for which the model was significant, state added to the model 47% of the time, whereas in bins and neurons for which the model was not significant, state added to the model only 19% of the time. Thus, state as a predictive factor is largely separable from trial events and is more likely to be a strong predictor when trial events are also strong predictors.

How do OFC ensembles develop state representations when the state changes?

We next examined the development of OFC state representations across state transitions using a modified decoding algorithm that tested how well ensembles represented the previous versus the new state across a transition. This algorithm trained using one set of trials from the end of the previous block and a second set of trials from the end of the new block. Then it tested whether activity during sliding sets of trials in between more resembled one or the other of these two trial sets. The heat plot in Figure 4 illustrates the resulting data. Within a few trials of the state transition, epochs occurring after reward delivery were significantly

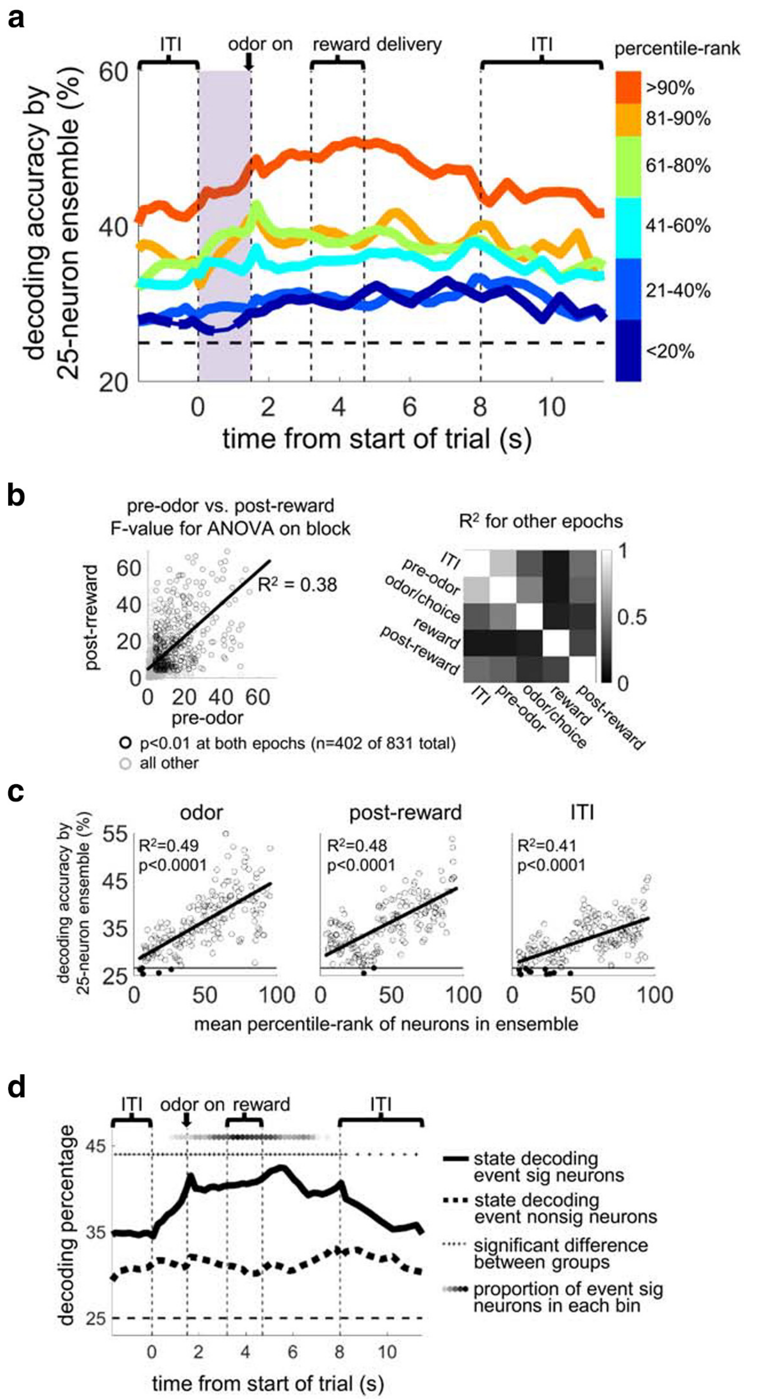

Figure 3. Individual OFC units maintain information about state across the trial, and state information is widely distributed across the OFC population. ***a***, The average accuracy of block decoding of 25 neuron OFC ensembles taken from ranked subpopulations. Neurons were ranked by the F values from an ANOVA using factor "state" run on each neuron's firing rate during the preodor period (shaded in purple). Ensembles from across the entire OFC population decoded state above chance across the whole trial, although they were ranked using firing rate only in the preodor epoch. Thick lines indicate significance versus chance at $p < 0.01$ by bootstrap for at least five consecutive bins. This suggests that individual units tended to maintain state information across the whole trial and that the majority of OFC neurons contributed to the state information observed in ensembles. ***b***, Scatter plot represents the correlation between the F values for block ANOVAs run on the prechoice epoch versus the postreward epoch across all OFC neurons. ***b***, Right, R^2 values for other pairs of epochs. ***c***, The linear correlation between mean percentile rank of the neurons in 200 25 neuron ensembles randomly selected from a segment of 10% of all the neurons sliding across the population from the top-ranked to the bottom-ranked neurons, versus their state-decoding accuracy for three 500 ms epochs across the trial. Horizontal line indicates the significance level for decoding accuracy at $p < 0.001$. Filled circles represent ensembles below the significance line, of which there were 9 of 200 for the odor epoch, 3 of 200 for the postreward epoch, and 14 of 200 for the ITI epoch. ***d***, Lines indicate state-decoding by 25 neuron ensembles selected from event-significant neurons (defined by ANOVA for factors odor, direction, or outcome in any bin) versus event-

←

nonsignificant neurons. Both groups decoded state better than chance and were significantly different from each other across all bins. Top, Shaded circles represent proportion of neurons significant for event ANOVAs in each bin, ranging from 0% (white) to 10% (black) of all possible significance tests across all neurons.

more similar to trials representing the new state, shown by dark blue patches on the heat plot. This suggests that the pattern of ensemble firing rates after reward almost immediately resembled that seen later in that same block, as opposed to the pattern occurring just a few trials previously at the end of the last block. However, in other parts of the trial, the pattern continued to resemble that in the previous state, as shown by the red and yellow patches in the heat plot. As trials in the new block proceeded (down the vertical axis), activity gradually came to resemble the new state pattern, shown by the heat plot changing from red to light blue to dark blue. By about trial 20, the new state pattern was represented uniformly across the whole trial. The timing of this development suggests that initial information about the state change arrives when an outcome is first observed to be different from in the previous block for that same response. Interestingly, it was not during reward delivery itself that the information first appeared in OFC ensembles, but ~1 s after the last reward drop delivery. This suggests that information about the number of drops must be integrated over time in ensemble patterns to represent the new state. Our analysis shows more trials are needed for this information to be integrated in earlier and later parts of the trial that are more distal from reward delivery.

We next tested whether the development of OFC state representations was correlated with the rate at which the rats adapted their choices to the new block. To do this, we split the OFC population into two groups according to whether neurons were recorded when rats switched quickly (top quartile of free-choice rate toward 3-drop side in the first 10 trials after the switch) versus when rats switched slowly (bottom quartile; Fig. 5*a*). This split was possible because there was considerable spontaneous variability in changes in the choice rate. On lower-quartile switches, rats perseverated by choosing the 1-drop side (formerly the 3-drop side) on 100% of free-choice trials in the first 10, whereas in the upper quartile they did so on only 28% of free-choice trials in the first 10. Then we looked separately at the development of state representations in OFC ensembles recorded in these two conditions (Fig. 5*b*). The overall pattern of decoding was roughly similar in the two conditions, recapitulating that evident in the overall population, with the new state pattern appearing first after reward

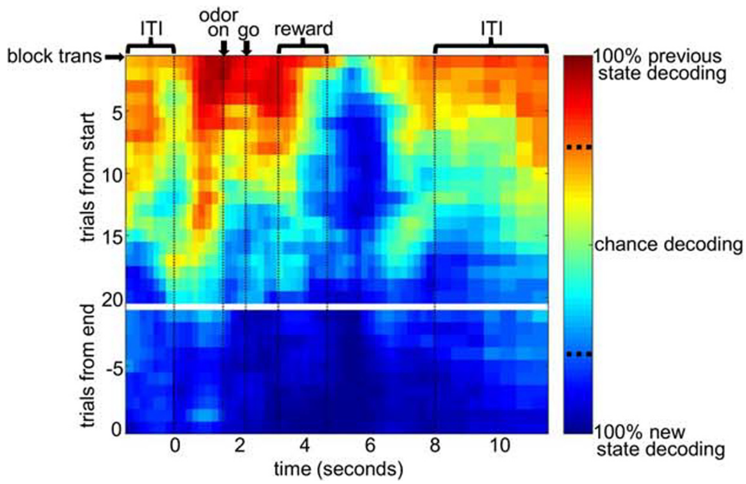

Figure 4. After a state transition, new state representations first emerge in OFC ensembles in the postreward epoch, after which it propagates to the other parts of the trial. Shown are results from a binary block decoder that tests how well OFC ensembles matched the old block pattern versus the new block pattern using sliding sets of 5 trials after a number block transition (see Materials and Methods). Spike trains were aligned to various events in the trial (shown with vertical dotted lines) and then concatenated and binned. Decoding percentages at each bin and set of trials were smoothed and converted into the color scale as shown for the heat plot. Red represents previous state-decoding. Blue represents new state-decoding. Vertical dotted lines on the color scale bar indicate significance levels for previous state or new state-decoding, based on a permutation test ($p < 0.05$, two-tailed).

and only gradually spreading to earlier and later parts of the trial. However, a marked difference appeared in the early parts of the trial, after the ITI ended but before the rat made the go response that indicated its choice. When choice rates were in the bottom quartile, activity here was strongly perseverative (dark red), whereas when choice rates were in the top quartile, decoding more quickly went to chance (green or light blue). We tested this more formally by analyzing decoding during a prechoice epoch in the two conditions (for this, we restricted our analysis to forced-choice trials to remove any bias in the kinds trials included in the top quartile vs bottom quartile datasets; Fig. 5*c*). This analysis confirmed the impression from the heat plots: When rats switched slowly, decoding for the new block was significantly below chance (reflecting the previous block) through about trials 12-15, whereas when rats switched quickly, decoding was significantly above chance. Later in the block, decoding in the two conditions became indistinguishable. When we ran a similar analysis using a postreward epoch, the neurons recorded on slowly switching blocks actually showed better decoding, but the differences were not significant (Fig. 5*d*). Because we only observed a significant difference in the prechoice part of the trial, it suggests the possibility that a failure to integrate new state information into OFC ensembles when rats are making choices could result in a failure to adapt choice behavior to the new contingencies.

Does the above analysis mean that OFC state representations are directly driving choice behavior? Here we tested this idea by comparing state-decoding when rats chose the preferred outcome (the 3-drop side) versus the antipreferred outcome (the 1-drop side). We excluded trials right after state transitions, when the preferred side would be ambiguous. If OFC state representations were directly driving choices, we would expect to see better representations of state when rats chose the preferred side after this transition period. However, as shown in Figure 6, state-decoding was indistinguishable in the two conditions. In combination with the previous analysis, this suggests that, rather than directly driving choices, OFC state representations may only be critical to judgment as to the current state right after state transitions, and other parts of the brain are able drive the choice thereafter.

Do OFC representations of state drive other information held by OFC ensembles?

The theoretical advantage of categorizing the environment into states is that it would allow the appropriate rules or contingencies to be recalled in conflicting situations. In our task, for example, knowing the current state would allow the brain to better predict what outcome to expect based on a chosen action. We tested whether this was true within the OFC network by selecting OFC ensembles that were usually correct at decoding the current state after reward delivery immediately after block switches. Because this was the epoch when OFC ensembles were most accurate (see Figs. 4, 5), we reasoned that, in the rare cases when these ensembles were incorrect, it would mean that state information might be widely degraded in OFC. If state information was degraded after reward delivery, then that information would have no opportunity to recover on the following trial until reward was delivered. We predicted, therefore, that OFC ensembles would be unable to decode the expected outcome on trials following trials with degraded state information after reward delivery. We tested this idea by using the remaining OFC ensembles (excluding the neurons used to identify trials with degraded state information) to decode the expected outcome. As shown in Figure 7, we found that, when the state had been misrepresented previously, these independent ensembles (matched for the same sessions) were unable to decode the outcome better than chance until reward delivery started. In contrast, when the state had been represented accurately (which happened on the majority of trials), these independent ensembles decoded the outcome better than chance well before reward delivery. This analysis suggests that outcome predictive information is derived from the correct representation of the state in OFC ensembles. As a control, we ran the same analysis using trials later in the block, when the state would be well established. Here we found equally good outcome-decoding regardless of whether the state had been misrepresented on the previous trial. This result suggests that, when the current state is unambiguous, outcome predictions do not depend specifically on the accuracy of OFC state representations.

Discussion

Animals can divide the world into "states" that define the rules or associations appropriate for governing behavior. This strategy is computationally efficient when appropriate behaviors conflict between different situations because it allows animals to switch behaviors without needing to relearn the rules each time they change (Wilson et al., 2014). Here we trained rats on a task consisting of four states presented in an order that changed daily. States were not signaled by any external cue in addition to the sets of rewards available at the two fluid wells, which changed at each block transition and continued throughout the following block of trials. We found that the vast majority of neurons recorded in the lateral OFC held sufficient state information to contribute to ensemble representations of the current state at any point in time. State was even well represented during ITIs when

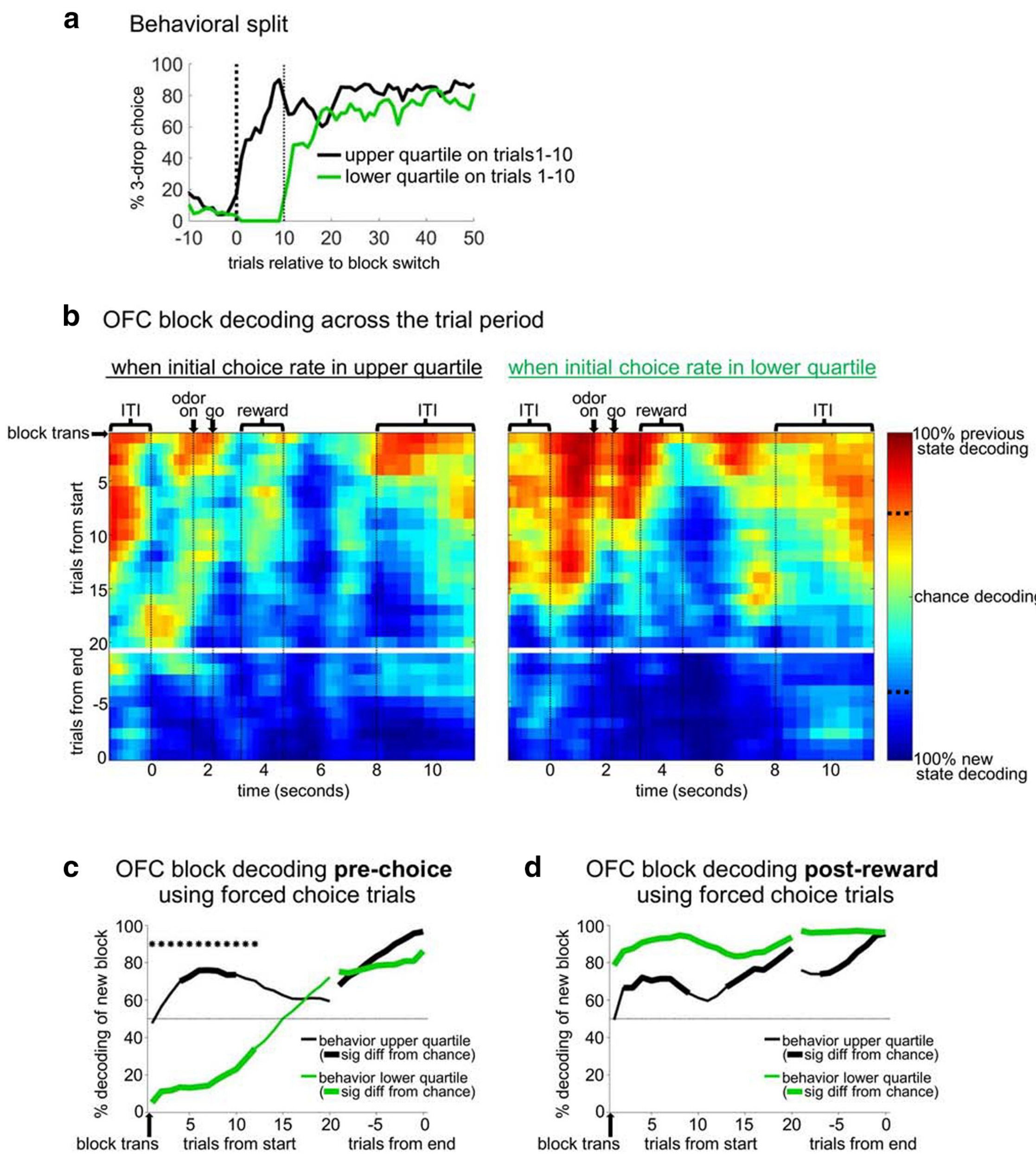

Figure 5. Faster behavioral adaptation to a state transition is associated with better OFC state-decoding during the prechoice part of the trial. ***a***, Number block transitions were split into the top and bottom quartiles of all transitions based on the free-choice rate in the first 10 trials after the transition. In bottom quartile switches, rats perseverated, choosing what was previously the 3-drop side, even though it had switched to deliver a single drop. Conversely, in upper quartile switches, rats immediately began switching to choose the new 3-drop side. We then ran the binary state decoder using only OFC neurons recorded across either the top or bottom quartile switches (***b-d***). ***b***, The resulting heat plots show that, in the prechoice part of the trial, neurons recorded during upper quartile switches adapted quickly to the new state code, whereas those recorded in lower quartile switches strongly encoded the old state for 10-15 trials after the transition. Line figures illustrated this effect for a prechoice epoch (***c***) or a postreward epoch (***d***). These latter analyses only included correct forced-choice trials, meaning that rewards received were not different in the two conditions. Thick lines indicate significance relative to chance. ***c***, *Significant differences between the conditions, both by permutation tests ($p < 0.05$ for five consecutive bins).

rats were not engaged in the task. Importantly, these state representations in OFC did not reflect memory traces of what outcome had been delivered on the previous trial, and they generalized across free- and forced-choice trials. Further the state representations were largely dissociable from representation of individual trial events, which were also present, as described in a previous publication (Stalnaker et al., 2014); state was represented by portions of the neural population that did not distinguish trial events, and within individual neurons activity related to state and activity related to trial events coexisted. Thus, the state representations were not simply an ensemble property secondary to individual neurons encoding the defining characteristics of trials in a given block (odor, reward availability). Instead, the state representations are well positioned to promote the emergence of activity specific to the trial events in particular blocks, as if setting the stage for neural activity encoding the proper associations.

Similar representations have been reported previously in rat, nonhuman primate, and human OFC (Wilson et al., 2014; Bradfield et al., 2015; Saez et al., 2015; Schuck et al., 2016; Schuck and Niv, 2019). Here we extend these findings by showing that individual OFC neurons tend to maintain state representations across time. This suggests that an abstract state representation is not a property that emerges from multiple different

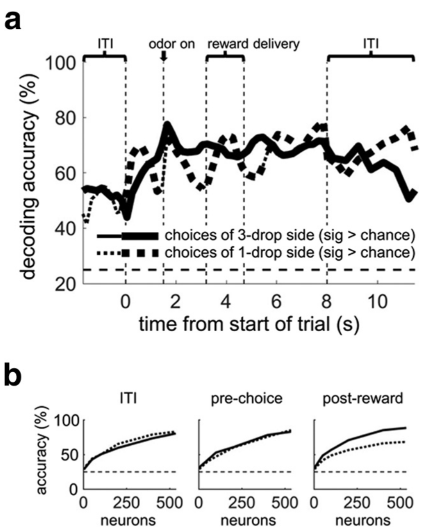

Figure 6. The accuracy of state-decoding is not related to whether rats accurately chose the side with the preferred outcome. ***a***, State-decoding accuracy across the trial divided into appropriate choices (free-choices of the 3-drop reward; solid line) or inappropriate choices (free-choices of the 1-drop reward; dotted line), excluding the first 10 trials after changes in drop number. All neurons on which enough inappropriate choices were made were included ($n = 535$). ***b***, Line plots represent decoding accuracy as a function of ensemble size in both conditions during the ITI, prechoice, and postreward epochs.

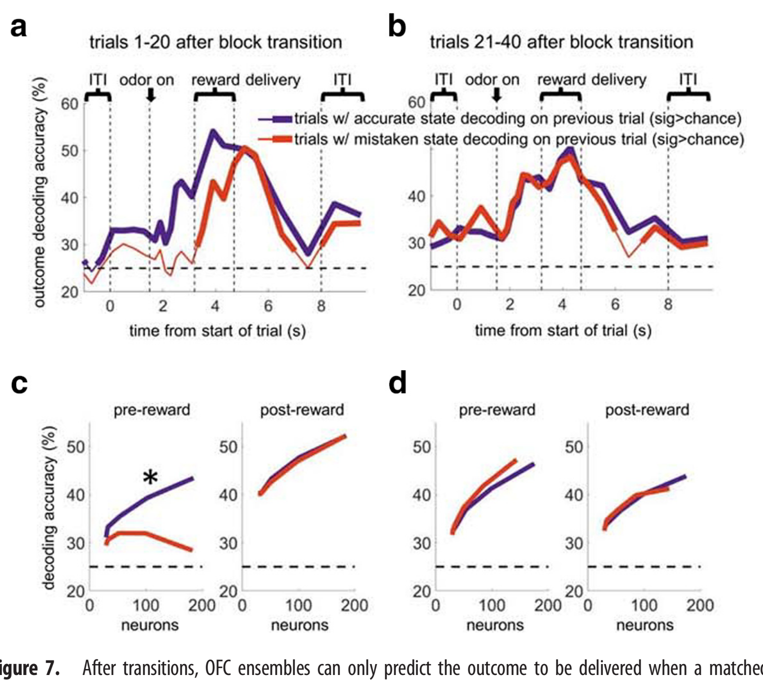

Figure 7. After transitions, OFC ensembles can only predict the outcome to be delivered when a matched OFC ensemble had decoded the state accurately after reward delivery on the previous trial. ***a***, Outcome-decoding accuracy of OFC ensembles when a matched ensemble recorded in the same sessions accurately (blue) or inaccurately (red) decoded the state on the previous trial, using the first 20 trials after block transitions. ***b***, A parallel analysis using trials 21-40 after block transitions. The state-decoding epoch was always a 1500 ms epoch beginning 1500 ms after the initial reward delivery time stamp on the previous trial. ***c***, ***d***, Same analysis as a function of outcome-decoding ensemble size. All rewarded trials were included, except those in which the previous trial was itself not rewarded. $*p < 0.01$ comparing the estimated parameter values for the best fit of the two functions.

representations in different neurons but is explicitly represented by variance in single-unit activity in OFC. Additionally, we show that the state representations in OFC are related to correct behavior, particularly in the initial trials after a state transition, and that recall of the appropriate state by OFC ensembles at the end of one trial predicts the recall of appropriate outcome expectations by OFC neurons on the subsequent trial, as if the state representation is setting the stage for which rules will be activated.

It is worth noting that what we have demonstrated is just one type of state that might be represented even in our simple task; our result does not preclude the representation of other hierarchically ordered states, either higher or lower in order than the states as we have defined them. For example, in our experimental design, because there were only four possible block orders, there was a higher-order state structure consisting of the sequence of the lower-order states. Such information might also be represented in OFC; however, because each session consisted of only a single instance of this higher-order "state," we were unable to test whether rats or OFC neurons also tracked them. But regardless, there would be utility in keeping track of the lower-order states, defined by the available outcomes, to know what action to choose or which well at which to expect the preferred outcome.

Our findings generally support the proposal that OFC is a critical node in the brain's representation of a "cognitive map." A cognitive map is analogous to a spatial map in that it organizes knowledge about the relationships between stimuli occurring in a particular setting (Wilson et al., 2014; Behrens et al., 2018; Niv, 2019). In this case, the map would represent that in one state, when a correct left response yields three drops of chocolate milk, a correct right response would yield one drop of vanilla milk, and so on for each of the four possible states. The state would therefore be akin to an implicit context that links the available actions and their probable outcomes during a particular block of trials. Under this idea, the representation of state that we observed in OFC neural activity might be conceived of as a pointer indicating the belief about the current state, facilitating the recall of context-appropriate contingencies locally and in distal structures.

OFC state representations changed quickly after changes in the reward contingencies. That the state representation first changes after reward delivery suggests that reward size and flavor, which in combination with the preceding response define the state, provides sensory evidence as to the current state that OFC neurons integrate over time. The gradual integrative process can be interpreted as the confidence in the current state, akin to decision confidence, which is also represented in OFC (Lak et al., 2014; Hirokawa et al., 2019). As more evidence is accumulated over repeated trials, the confidence as to the current state spreads to parts of the trial in which the reward identity must be remembered for longer. The mechanism for this integrative process is unknown. One speculative hypothesis would be that it depends on dopamine input. Dopaminergic prediction errors are known to occur at reward delivery after block transitions in this task and others like it (Hollerman and Schultz, 1998; Takahashi et al., 2017), and the magnitudes of prediction errors are theoretically important to distinguish whether a new state has been entered or the old state is changing (Gershman et al., 2010; Niv, 2019).

OFC state representations were also related to rats' choice behavior, specifically after changes in the reward contingencies early in each block. When rats were slow to adapt free-choice behavior on a change in the number of drops delivered at both wells, OFC state representations in the prechoice period strongly represented the previous state in the first 10-15 trials after the transition, whereas when rats quickly adapted free-choices to the new contingencies, OFC represented the new state in the prechoice period during those trials. This relationship was only true in the prechoice period; it did not hold later in the trial, where state representations were unrelated to choices, nor was it true later in blocks, when rats' behavior was more stable, suggesting that they were more certain of the current state. In this later period of the trial blocks, OFC state representations were equally accurate whether rats chose accurately (i.e., the preferred reward) or inaccurately (i.e., the antipreferred reward). These findings suggest that OFC state representations may be most important for accurate decision-making right after state transitions when the state is ambiguous. This contrasts with what we have previously observed in cholinergic interneurons in dorsomedial striatum, in which the current state was misrepresented when rats chose inaccurately throughout blocks (Stalnaker et al., 2016). Interestingly, when the ipsilateral OFC was lesioned, this relationship between striatal state representations and choice was eliminated (Stalnaker et al., 2016). Together, these results suggest a circuit in which OFC activity sets the current belief about the state after switches, a setting that striatal circuitry (Bradfield and Balleine, 2017; Sharpe et al., 2019), or other interconnected areas, such as amygdala (Saez et al., 2015), are able to maintain thereafter to directly drive choices.

Knowing the current state would be useful because it would allow outcomes to be better predicted based on chosen, or potentially chosen, actions. It has long been known that, across species, OFC has prominent representations of the value and sensory features of predicted outcomes (Gottfried et al., 2003; Padoa-Schioppa and Assad, 2006; Mainen and Kepecs, 2009; Wallis and Kennerley, 2011; Rudebeck and Murray, 2014; Rich and Wallis, 2016; Rudebeck et al., 2017). If OFC state representations are being used to generate outcome predictions, representations of the expected outcome, within OFC or elsewhere in the brain, should depend on an accurate representation of the current state. We tested this prediction within OFC by examining encoding of the expected outcome by OFC ensembles when the state had been accurately or inaccurately decoded after reward delivery on the previous trial. We found that, when the state had been misrepresented right after state transitions, outcome-decoding was inaccurate (at chance levels) on the next trial until the outcome was delivered, at which time it quickly rose to match high and significant levels of outcome-decoding when the state had been accurately represented on the previous trial. This relationship provides evidence that OFC has circuitry that combines knowledge of the current state with knowledge of the chosen action to predict the outcome when the state is ambiguous. However, later in blocks, when the current state was presumably unambiguous, outcome predictive decoding was independent of the accuracy of OFC state representations. This suggests that other parts of the brain are maintaining state representations at that time, and conversely that OFC state representations are only necessary when the state is ambiguous. It has recently been proposed that more medial parts of OFC represent the expected outcome, while more lateral OFC represents the state (Bradfield et al., 2015; Bradfield and Hart, 2020). In the current study, we did not record in medial OFC, but our findings do not negate that possibility, since the neural correlates of expected outcomes observed might be retrieved from medial areas. More broadly, the current results are consistent with a recent proposal that representations of expected outcome or reward value would only be observed in OFC to the extent that they are part of state representations (Niv, 2019). As such, understanding the mechanisms through which OFC represents and uses cognitive maps may have important implications for understanding OFC dysfunctions in addiction and other affective disorders related to misevaluating expected outcomes (Goldstein et al., 2007; Lucantonio et al., 2012; Bernardi and Salzman, 2019).

References

Behrens TE, Muller TH, Whittington JC, Mark S, Baram AB, Stachenfeld KL, Kurth-Nelson Z (2018) What is a cognitive map? Organizing knowledge for flexible behavior. Neuron 100:490–509.

Bernardi S, Salzman CD (2019) The contribution of nonhuman primate research to the understanding of emotion and cognition and its clinical relevance. Proc Natl Acad Sci USA 116:26305–26312.

Bradfield LA, Balleine BW (2017) Thalamic control of dorsomedial striatum regulates internal state to guide goal-directed action selection. J Neurosci 37:3721–3733.

Bradfield LA, Hart G (2020) Rodent medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space. Neurosci Biobehav Rev 108:287–294.

Bradfield LA, Dezfouli A, van Holstein M, Chieng B, Balleine BW (2015) Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. Neuron 88:1268–1280.

Gershman SJ, Blei DM, Niv Y (2010) Context, learning, and extinction. Psychol Rev 117:197–209.

Goldstein RZ, Tomasi D, Rajaram S, Cottone LA, Zhang L, Maloney T, Telang F, Alia-Klein N, Volkow ND (2007) Role of the anterior cingulate and medial orbitofrontal cortex in processing drug cues in cocaine addiction. Neuroscience 144:1153–1159.

Gottfried JA, O'Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science 301:1104–1107.

Hirokawa J, Vaughan A, Masset P, Ott T, Kepecs A (2019) Frontal cortex neuron types categorically encode single decision variables. Nature 576:446–451.

Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1:304–309.

Lak A, Costa GM, Romberg E, Koulakov AA, Mainen ZF, Kepecs A (2014) Orbitofrontal cortex is required for optimal waiting based on decision confidence. Neuron 84:190–201.

Lucantonio F, Stalnaker TA, Shaham Y, Niv Y, Schoenbaum G (2012) The impact of orbitofrontal dysfunction on cocaine addiction. Nat Neurosci 15:358–366.

Mainen ZF, Kepecs A (2009) Neural representation of behavioral outcomes in the orbitofrontal cortex. Curr Opin Neurobiol 19:84–91.

Meyers EM (2013) The neural decoding toolbox. Front Neuroinform 7:8.

Niv Y (2019) Learning task-state representations. Nat Neurosci 22:1544–1553.

Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. Nature 441:223–226.

Rich EL, Wallis JD (2016) Decoding subjective decisions from orbitofrontal cortex. Nat Neurosci 19:973–980.

Rudebeck PH, Murray EA (2014) The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. Neuron 84:1143–1156.

Rudebeck PH, Saunders RC, Lundgren DA, Murray EA (2017) Specialized representations of value in the orbital and ventrolateral prefrontal cortex: desirability versus availability of outcomes. Neuron 95:1208–1220. e1205.

Saez A, Rigotti M, Ostojic S, Fusi S, Salzman CD (2015) Abstract context representations in primate amygdala and prefrontal cortex. Neuron 87:869–881.

Schuck NW, Niv Y (2019) Sequential replay of nonspatial task states in the human hippocampus. Science 364:eaaw5181.

Schuck NW, Cai MB, Wilson RC, Niv Y (2016) Human orbitofrontal cortex represents a cognitive map of state space. Neuron 91:1402–1412.

Sharpe MJ, Stalnaker T, Schuck NW, Killcross S, Schoenbaum G, Niv Y (2019) An integrated model of action selection: distinct modes of cortical control of striatal decision making. Annu Rev Psychol 70:53–76.

Stalnaker TA, Calhoon G, Ogawa M, Roesch MR, Schoenbaum G (2010) Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. Front Integr Neurosci 4:12.

Stalnaker TA, Cooch NK, McDannald MA, Liu TL, Wied H, Schoenbaum G (2014) Orbitofrontal neurons infer the value and identity of predicted outcomes. Nat Commun 5:3926.

Stalnaker TA, Berg B, Aujla N, Schoenbaum G (2016) Cholinergic interneurons use orbitofrontal input to track beliefs about current state. J Neurosci 36:6242–6257.

Takahashi YK, Batchelor HM, Liu B, Khanna A, Morales M, Schoenbaum G (2017) Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. Neuron 95: 1395–1405.e1393.

Wallis JD, Kennerley SW (2011) Contrasting reward signals in the orbitofrontal cortex and anterior cingulate cortex. Ann NY Acad Sci 1239:33–42.

Wilson RC, Takahashi YK, Schoenbaum G, Niv Y (2014) Orbitofrontal cortex as a cognitive map of task space. Neuron 81:267–279.