

Theta-Band Phase Locking of Orbitofrontal Neurons during Reward Expectancy

Marijn van Wingerden, Martin Vinck, Jan Lankelma, and Cyriel M. A. Pennartz

Cognitive and Systems Neuroscience Group, Swammerdam Institute for Life Sciences–Center for Neuroscience, University of Amsterdam, 1090 GE Amsterdam, The Netherlands

The expectancy of a rewarding outcome following actions and cues is coded by a network of brain structures including the orbitofrontal cortex. Thus far, predicted reward was considered to be coded by time-averaged spike rates of neurons. However, besides firing rate, the precise timing of action potentials in relation to ongoing oscillations in local field potentials is thought to be of importance for effective communication between brain areas.

We performed multineuron and field potential recordings in orbitofrontal cortex of rats performing olfactory discrimination learning to study the temporal structure of coding predictive of outcome. After associative learning, field potentials were marked by theta oscillations, both in advance and during delivery of reward. Orbitofrontal neurons, especially those coding information about upcoming reward with their firing rate, phase locked to these oscillations in anticipation of reward. When established associations were reversed, phase locking collapsed in the anticipatory task phase, but returned when reward became predictable again after relearning. Behaviorally, the outcome anticipation phase was marked by licking responses, but the frequency of lick responses was dissociated from the strength of theta-band phase locking. The strength of theta-band phase locking by orbitofrontal neurons robustly follows the dynamics of associative learning as measured by behavior and correlates with the rat's current outcome expectancy. Theta-band phase locking may facilitate communication of outcome-related information between reward-related brain areas and offers a novel mechanism for coding value signals during reinforcement learning.

Introduction

Constructing predictive representations of future reward is essential for goal-directed behavior. Goal-directed behavior appears to be guided by a network of brain structures encoding the predicted outcome of cues and actions as a result of associative learning (Schoenbaum et al., 1998; Padoa-Schioppa and Assad, 2006; Schultz, 2006; Rolls et al., 2008). Such predictive representations have been found in the amygdala, prefrontal cortex, and ventral striatum and are a requirement for learning mechanisms that depend on the difference between actual and predicted reward (Rescorla and Wagner, 1972; Pennartz, 1997; Schultz et al., 1997; Sutton and Barto, 1998; Schultz, 2006). However, until now the neural coding of reward expectancy has been studied almost

exclusively in the domain of firing rates, whereas a number of issues in reward processing require considering its temporal organization. First, the effect that a reward-predictive representation in a given brain structure exerts on its target areas will depend on the temporal alignment of the firing of contributing neurons (Salinas and Sejnowski, 2000; Engel et al., 2001). Second, the temporal phasing of activity in two connected brain areas may modulate the efficacy of their communication (Womelsdorf et al., 2007). Third, reward-dependent associative learning mechanisms likely involve long-term synaptic modifications depending on the exact timing of presynaptic and postsynaptic activity (Levy and Steward, 1983; Markram et al., 1997; Lisman and Spruston, 2005; Cassenaer and Laurent, 2007).

We investigated the relationship between single-unit activity and local field potential (LFP) oscillations during an associative learning task in the orbitofrontal cortex (OFC), a prefrontal structure that is strongly connected with the amygdala, parahippocampal cortex, medial prefrontal cortex, and ventral tier of the basal ganglia, and is crucial for flexibly encoding reward-value representations and adjusting goal-directed behaviors (Bechara et al., 1994; Carmichael and Price, 1995; Ongür and Price, 2000; Fellows and Farah, 2003; Stalnaker et al., 2007; Schilman et al., 2008). Firing-rate changes of OFC neurons were shown to represent the predicted reward value of cues and actions, both at the single-cell and population level (Schoenbaum et al., 1998; Wallis and Miller, 2003; Padoa-Schioppa and Assad, 2006; van Duuren et al., 2008). Moreover, OFC representations can be updated when previously learned associations are no longer valid (Rolls et al., 1996;

Received Aug. 3, 2009; revised Feb. 10, 2010; accepted March 16, 2010.

This work was supported by the Netherlands Organization for Scientific Research–VICI Grant 918.46.609 (to C.M.A.P.) and the European Union Seventh Framework Programme FP7-ICT Grant 217148 (to C.M.A.P.). We would like to acknowledge the software tools or assistance provided by the following: Peter Lipa (University of Arizona, Tucson, AZ) for the use of BubbleClust; A. David Redish (University of Minnesota, Minneapolis, MN) for the use of MClust; Ruud Joosten and Laura Donga (Netherlands Institute for Neuroscience and University of Amsterdam, Amsterdam, The Netherlands) for help with rat surgeries; and Ed de Water, Theo van Lieshout, Cees van den Biggelaar, Johan Soede, Frans Pinkse, Hans Gerritsen, Ron Manuputy, Mattijs Bakker, and Wietze Buster (University of Amsterdam) for experimental setup and drive manufacturing. M.v.W. and C.M.A.P. designed the experiments; M.v.W. performed the experiments; M.v.W., M.V., and J.V.L. analyzed the data; and M.v.W., M.V., and C.M.A.P. wrote the paper.

Correspondence should be addressed to Marijn van Wingerden, Swammerdam Institute for Life Sciences–Center for Neuroscience, University of Amsterdam, P.O. Box 94216, 1090 GE Amsterdam, The Netherlands. E-mail: e.j.m.vanwingerden@uva.nl.

DOI:10.1523/JNEUROSCI.3860-09.2010

Copyright © 2010 the authors 0270-6474/10/307078-10\$15.00/0

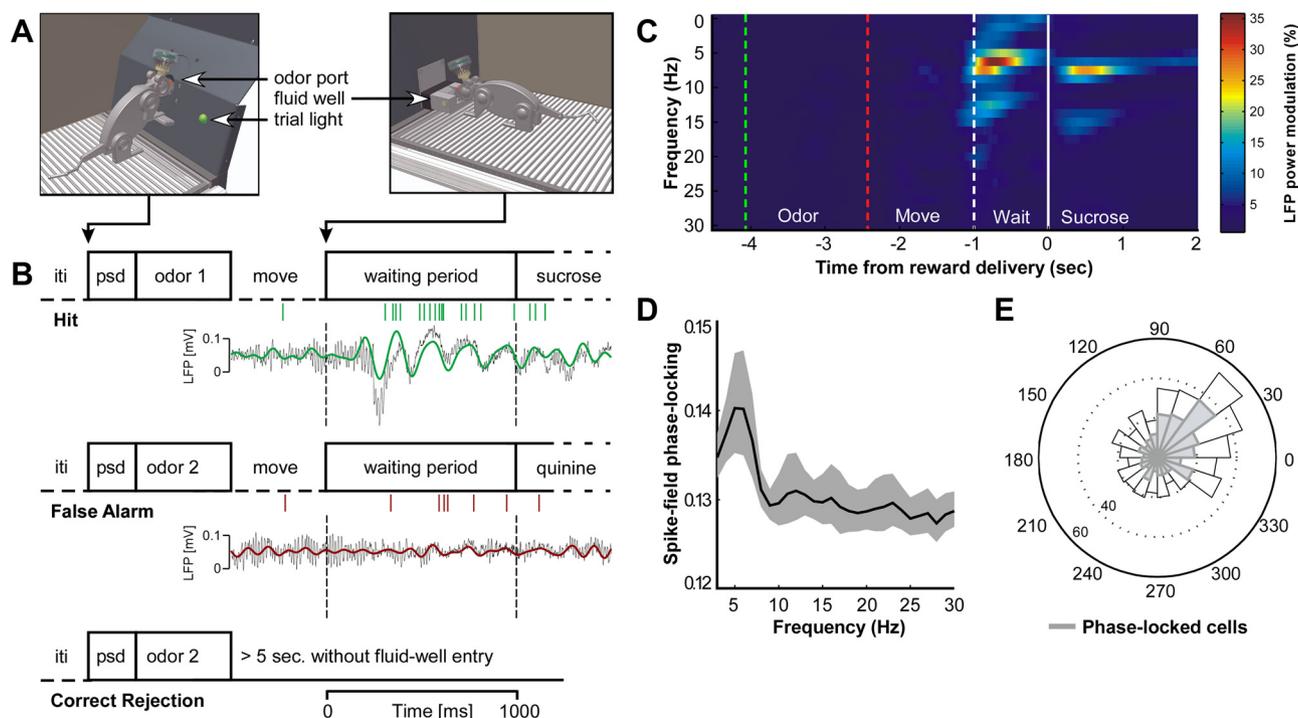


Figure 1. Task-related theta oscillations and phase locking. **A**, Operant chamber; impression of rat making odor poke (left) and fluid poke (right). **B**, Raw LFP segments recorded during different task periods (black) and bandpass-filtered LFP (4–12 Hz) (green trace, hit trial; red, false alarm trial) and spike trains from one unit, recorded on a different tetrode (vertical green/red bars). psd, prestimulus delay. **C**, LFP power modulation, synchronized in time on sucrose delivery. Pseudocolors indicate averaged percentage increase in LFP power compared to the intertrial interval period. Green dashed line, Median time of odor onset; red dashed line, median time of odor offset; white dashed line, start of waiting period; white solid line, sucrose delivery. Theta oscillations (9 Hz) found during odor sampling are not visible here due to synchronization on reward delivery. **D**, Spike–field phase-locking spectrum averaged across all task-segments and all units. Shaded regions indicate 95% bootstrapped confidence intervals on the mean. **E**, Polar histogram of mean phase values for all cells. Shaded part corresponds to number of significantly phase-locked units (270 of 525, Rayleigh test, $p < 0.05$). Mean phase, 11.3° . Distribution of phases deviated significantly from uniform distribution ($p < 0.001$).

Schoenbaum et al., 1998). If the relationship between single-unit activity and field oscillations is of importance for associative learning, we would expect to find changes in this relationship that follow the dynamics of learning, expressed preferentially by those cells whose activity is informative about upcoming task outcomes. In addition to testing this hypothesis, we also examined whether rhythmic LFP-locked spiking activity was associated with licking behavior, as this was found especially in advance of and following reward delivery.

Materials and Methods

Subjects

Three adult male Wistar rats (Harlan CPB), weighing 370–450 g at the time of surgery, served as subjects in these experiments. Before training and surgery, the rats were housed two to a cage on a reversed light/dark cycle (lights off: 7:00 A.M., lights on: 7:00 P.M.) with *ad libitum* food and water. After surgery, animals were housed individually in a transparent cage (40 × 40 × 40 cm), with other rats present in the climate-controlled colony room. During training and recording, rats were maintained on food restriction, with 5–15 g of food available from >0.5 h after training, depending on the amount of reward collected in the session, amounting to 90% free-feeding intake.

All experiments were conducted according to the National Guidelines on Animal Experiments and with approval of the Animal Experimentation Committee of the University of Amsterdam.

Behavioral training

Apparatus. Odor discrimination training was conducted in an operant chamber (56 × 30 × 40 cm L × W × H) equipped with an odor-sampling port and trial light on a front panel and a tray for delivery of fluids placed at the opposing wall (Fig. 1). The front panel was slanted at 45° (with respect to ground level) above the odor-sampling port to allow

unhindered nose poking into the odor port by implanted animals. Entries by the animals into the odor port and into the fluid well were recorded by photobeam interruptions and stored on a computer dedicated to behavioral data acquisition. Licking responses were detected by a separate photobeam situated within the fluid well. During recording sessions, behavioral events were synchronized with electrophysiological data acquisition running on a separate computer. Odors were delivered via separate glass vials and tubing to avoid mixing. Upon entering the application system, they were mixed 1:1 with clean air and released into the compartment directly behind the odor port by way of computer-controlled valves. Likewise, quinine and sucrose solutions were delivered to the fluid well via separate fluid lines and electronically controlled by valves (van Duuren et al., 2007). Training and task performance were devoid of human interference.

Training procedure. The animals were trained on a two-odor Go/No-Go discrimination task (Schoenbaum et al., 1998). After habituation to the chamber and pretraining, rats were confronted with odor discrimination problems. Each session, one novel odor was associated with reward (100 μ l of 15% sucrose solution), and a second novel odor with an aversive outcome (100 μ l of 0.01 M quinine solution). Training sessions consisted of blocks of 5 + 5 pseudorandomly ordered positive and negative odor trials. When a trial light was illuminated, rats could initiate a trial by poking their snout in the odor-sampling port. After a 500 ms delay, air flow through the odor-sampling port was switched from clean air to the selected clean air/odorant mixture. A correct nose poke in the odor port (wait for odor 500 ms, constituting a prestimulus delay, and sample odor for at least 750 ms) was indicated by the trial light turning off. After sampling, the rats could move over to the fluid well, into which they were required to make a nose poke for ≥ 1000 ms before the outcome (sucrose or quinine solution) was presented. We refer to the invariant 1000 ms delay period as the “waiting” or anticipatory period. This waiting period allows sampling of neural activity during the waiting pe-

riod that is not confounded by whole-body movement or novel sensory input (Schoenbaum et al., 1998; van Duuren et al., 2007). When the rats left the fluid well, an intertrial interval (ITI) of 10–15 s was observed before the next trial started. A correct rejection was scored if the rat refrained from entering the reinforcement tray for ≥ 5 s following sampling of the negative odor. Responses during the ITI had no programmed consequences, while premature responses (i.e., short pokes) during the odor-sampling or waiting period resulted in immediate termination of the current trial and the start of a new trial.

Rats were implanted after they reached behavioral criterion, viz., scoring $>85\%$ hits and correct rejections over a moving block of 20 trials. Following surgery, animals were retrained on the familiar odor pair with which initial training took place until performance was back at criterion level. Recording of neural activity started in the subsequent session. On each recording session, rats were confronted with a new odor pair. Furthermore, a reversal schedule of odor–outcome contingencies was appended to the session in case the rat reached criterion performance. To assess learning in the reversal phase, average performance over all sessions as a function of postreversal trial was compared against chance level using a binomial test with expected performance of 0.5 ($p < 0.05$).

Surgical procedures

Animals were anesthetized by intramuscular injection of 0.08 ml/100 g of Hypnorm (0.2 mg/ml fentanyl, 10 mg/ml fluanisone; VetaPharma), followed by 0.04 ml/100 g of Dormicum (5 mg/ml midazolam, s.c.; Roche), and mounted in a stereotaxic frame. Body temperature was maintained between 35 and 36°C. A microdrive, holding 14 individually moveable electrode drivers, was chronically implanted onto a craniotomy (diameter: 2 mm) in the left hemisphere dorsal to the OFC at 3.4–3.6 mm anterior and 3.0–3.2 mm lateral to bregma. The drivers were loaded with 12 tetrodes and 2 reference electrodes. Using dental cement, the drive was anchored to six stainless steel screws, one of which was positioned in the left parietal bone and served as ground. Immediately after surgery, all tetrodes and reference electrodes were advanced 0.8 mm into the brain. Next, the animal was allowed to recover for 7 d with *ad libitum* food and water, during which the 12 recording tetrodes were advanced in daily steps to the upper border of the OFC according to a standardized rat brain atlas (Paxinos and Watson, 2007). The reference electrodes were lowered to a depth of 1.2–2.0 mm and adjusted to minimize spiking activity on the reference channel.

After surgery, saline was injected subcutaneously (2 ml per flank), and pain relief was provided by 0.1 ml/100 g of presurgical weight of a 10% Finadyne (flunixin meglumine 50 mg/ml; Schering-Plough) solution administered in saline subcutaneously.

Electrophysiology

Using tetrodes (Gray et al., 1995), neural activity was recorded by a 64-channel Cheetah setup (Neuralynx). Signals were passed through a unity-gain preamplifier headstage and a 72-channel commutator (Dragonfly), amplified 5000 \times , and bandpass filtered between 600 and 6000 Hz for spike recordings. If a signal on any of the leads of a tetrode crossed a preset threshold, activity on all four leads was sampled at 32 kHz for 1 ms and stored for off-line analysis. Local field potentials recorded on all tetrodes were amplified 1000 \times , continuously sampled at 1874 Hz, and bandpass filtered between 1 and 475 Hz. Events in the behavioral task were coregistered and time stamped by the Cheetah system.

Data analysis

Isolation of single-unit activity. Spike trains were sorted to isolate single units using a semiautomated clustering algorithm (BubbleClust) followed by manual refinement using MClust. Automated and manual clustering of spikes was done using the waveform peak amplitude, area, squared amplitude integral, and the first three principal components. Clusters were accepted as single units when having no more than 0.1% of interspike intervals shorter than 2 ms.

Spike–LFP phase locking. All analyses were done with Matlab (The MathWorks) and using FieldTrip (<http://www.ru.nl/fcdonders/fieldtrip/>), an open source toolbox for the analysis of neurophysiological data. Briefly, for each spike of a given cell, we cut out a 1.0 s data segment of the LFPs

recorded simultaneously on the other tetrodes (such that we never included a spike–LFP pair from the same tetrode). Several methods exist to measure spike–LFP phase consistency (Fries et al., 1997; Jarvis and Mitra, 2001; Pesaran et al., 2002; Womelsdorf et al., 2006). The basis of all of these methods is to determine the phase of a spike or a spike train relative to an LFP trace in a particular frequency band. Here, we multiplied each unfiltered LFP data segment by a Hanning window and Fourier transformed the windowed data segment of length T , so that the spike-triggered LFP frequency (f) spectrum is given as follows:

$$X_i(f) = \sum_t^T w(t) x_i(t) e^{-2ift\pi}, \quad (1)$$

where $x_i(t)$ is the LFP time segment around the i th spike ($i = 1, 2, \dots, N$) and $w(t)$ the Hanning window. Using a 1.0 s data segment length set the Rayleigh frequency at 1 Hz, allowing a frequency resolution of 1 Hz. We determined the complex average spike-triggered LFP spectrum across the C different tetrodes as follows:

$$\overline{X_i(f)} = \frac{1}{C} \sum_{c=1}^C \frac{X_i^c(f)}{|X_i^c(f)|}. \quad (2)$$

Equation 2 ensures that the power of each LFP segment is ignored in the computation of the average spike phase. The spike phase is now simply given by $\theta_i = \arg(\overline{X_i(f)})$. We measured phase consistency by means of the spike–LFP phase-locking value, which is defined as the resultant vector length across all spikes N , as follows:

$$\text{PLV}(f) = \left| \frac{1}{N} \sum_{i=1}^N \frac{\overline{X_i(f)}}{|X_i(f)|} \right|. \quad (3)$$

The resultant vector length is a real number in the range of 0 (low phase consistency) to 1 (high phase consistency). The more spikes we use to obtain our statistical estimate of the resultant length, the more reliable it is. Spike–LFP phase locking is a biased measure with respect to the number of spikes that are entered in the computation. In the method used here (Womelsdorf et al., 2008) (cf. Lachaux et al., 2002), we controlled for the bias by always entering the same fixed number of spikes ($N = 50$) into Equation 3 when we compared between samples with a different number of elements. This was the case whenever we compared phase locking between trial types (see Figs. 2B,C, 3A,B, 4C), different task periods (see Fig. 2D), and different cell groups (see Fig. 4A,B). Only in the time-resolved estimate of phase locking (see Fig. 2A), we used a lower criterion of $N = 40$ spikes per cell because here the sliding time windows (500 ms) were half the size of regular LFP windows. We further reduced the statistical variance of the spike–LFP phase-locking estimate by means of a bootstrapping procedure. For every repetition, we drew a fixed number of spikes without replacement from all spikes in the sample. For each sample drawn, we determined the spike–LFP phase-locking value. Subsequently, we averaged these spike–LFP phase-locking values across all bootstrapped samples ($N = 5000$), producing an unbiased phase-locking value. Nonetheless, the statistical significance of spike–LFP phase locking can still be assessed across the entire sample of spikes, using the Rayleigh test ($p < 0.001$).

Time-resolved power spectra and timing of theta oscillation onset. For every tetrode, we isolated LFP data segments in 500 ms windows (setting the Rayleigh frequency resolution at 2 Hz) centered on the time point of interest, separated by time steps of 10 ms (see Fig. 3C,D, at 6 Hz). For Figure 1C, we used a window of 800 ms, yielding a frequency resolution of 1.25 Hz. Using Equation 1, the LFP power of the i th segment was determined as $|X_i(f)|$:

$$|X_i(f)| = \sum_t^T w(t) x_i(t) e^{-2ift\pi}. \quad (4)$$

To obtain the baseline-corrected LFP power $P_i(f)$, we first determined the LFP power in the intertrial interval in the same way as in the task period. We averaged the baseline LFP power across time segments and trials. We then defined the baseline-corrected trial LFP power by dividing the trial LFP power by the baseline LFP power.

Since the number of trials before and after reversal differs between sessions, the averaged power modulation shown in Figure 3, *C* and *D*, was only calculated for trials with $N \geq 5$ contributing sessions. To study the dynamics of theta oscillation onset around the odor–outcome reversal (see Fig. 3*E*), we calculated a center-of-mass (COM) index for the waiting period as follows:

$$\text{COM} = \frac{\sum_i^N (P_i(f)\tau_i)}{\sum_i^N P_i(f)}, \quad (5)$$

where N is the number of data segments between -2 and $+2$ s relative to reward delivery, and τ_i indicates the center of the LFP segment and ranges from -2 s to $+2$ s. To determine the trial-by-trial timing of the onset of the theta oscillation on hit trials (see Fig. 3*F*), we determined the first time point from -2 to $+2$ s relative to reward delivery, where the calculated theta power was $>50\%$ of the maximum theta power in that segment (threshold crossing). We pooled these values as a function of trial number relative to reversal across sessions and performed a regression analysis of theta onset versus trial number on the first 15 hit trials after reversal. To control for unequal trial numbers in the postreversal phase, we created a randomization distribution of correlation coefficients for this regression by shuffling trials from all sessions with more than four postreversal hit trials. We then compared the observed correlation coefficient against the randomization distribution of correlation coefficients to obtain the probability of finding that result. This statistical procedure ensured that any bias due to unequal postreversal trial numbers was removed.

Multiple regression analysis

To investigate the contribution of a type of behavioral correlate in firing rate to phase-locking values during a specific period, we coded the presence or absence of such a correlate per cell in a binary fashion. Next, we constructed a design matrix with firing-rate correlate presence as dummy independent variables to explain phase-locking values as a dependent variable, separately for each trial period. This resulted in a matrix of β -weights (correlation-coefficients) and associated p values (see Fig. 4*A, B*). As a control predictor, rat identity was entered, which yielded no additional explanatory power.

Results

We used tetrode arrays (Gray et al., 1995) to record single-unit activity and LFPs from the OFC in 17 sessions from three male Wistar rats. Before each session, the tetrodes were moved individually to optimize the number of cells recorded simultaneously. Histological verification of the tetrode endpoints and recording tracks showed that all recordings were performed between 3.2 and 4.2 mm anterior of bregma and confined to the ventral and lateral aspects of the OFC (supplemental Fig. S1*C–E*, available at www.jneurosci.org as supplemental material). Rats performed a two-odor discrimination task in which they were required, in each new session, to associate one novel odor (S+) to a positive trial outcome (sucrose solution) and a second novel odor (S−) to a negative outcome (quinine solution). All recordings were confirmed as having been made from OFC, as shown in the example in supplemental Figure S1*E* (available at www.jneurosci.org as supplemental material). This yielded a total of 525 well isolated single units (15–54 units per session, median: 29) in conjunction with LFP signals recorded simultaneously. In all analyses where pairs of single units and LFPs were used, the unit was recorded on a different tetrode than the LFP to avoid a possible frequency bias when the LFP around a spike was filtered.

Behavior

Following odor sampling, rats either made a Go decision, consisting of a locomotor response toward a fluid well followed by a

nose poke into the well and an immobile waiting period of 1 s before reinforcer delivery, or a No-Go decision, refraining from these responses (Fig. 1*A, B*). Rats required 86 ± 6 (mean \pm SEM) trials to reach the criterion of 85% correct Go/No-Go decisions over the last 20 trials, making on average 16 ± 2 false alarm responses (i.e., erroneous Go responses) (Fig. 1*B*). When the criterion was met, rats were exposed to a reversal schedule. Now, the odor previously associated with sucrose was coupled to quinine and vice versa. After reversal, rats quickly adjusted their behavior to the new odor–outcome contingencies, performing above chance level after on average 30 ± 4 trials (binomial test, $p < 0.05$). This provides evidence that the rats also learned the reversed odor–outcome contingencies. On hit trials, both before and after reversal, the mean reaction time (measured from odor offset to reward delivery) was negatively correlated with increasing trial number (both $p < 0.05$, Spearman's rank correlation). Licking behavior was analyzed separately (see below).

Analysis of oscillatory activity

We observed increments in theta power (4–12 Hz) in the LFPs during the waiting period of each trial, when the rat anticipated reward, and during reward consumption, but not during movement (Fig. 1*B, C*). The characteristics of these oscillations are different from the well known theta rhythm recorded from hippocampus, which has been linked to locomotion and spatial and episodic memory formation (O'Keefe and Recce, 1993; Skaggs et al., 1996; Buzsáki, 2005). The LFP reflects the activity of a large number of neurons and could be confounded by volume conduction of signals from other brain areas. In addition to the tetrode recordings, we performed recordings in two additional animals in the same behavioral paradigm using linear silicon probes with electrodes spanning the dorsoventral axis of the OFC (supplemental Fig. S1*C, D*, available at www.jneurosci.org as supplemental material). Current-source density analysis of the recorded LFPs showed the presence of a sink-source pair in the OFC, indicating that at least part of the recorded theta oscillatory activity is locally generated (supplemental Fig. S1*A, B*, available at www.jneurosci.org as supplemental material). Next, we examined whether theta-band rhythmicity was also visible in the firing patterns of single neurons recorded from the OFC. We determined the phase of each spike relative to the LFP components for frequencies ranging from 3 to 30 Hz. The normalized sum of these single-cell phase vectors has a mean angle and a resultant length between 0 and 1. The resultant length of the vector reflects how consistently a cell fires to a phase of the theta oscillation, with a spike–field phase-locking value of 0 indicating no consistency in phases, and a value of 1 indicating full consistency of spike phases (for a detailed explanation, see Materials and Methods). In a majority of OFC units (51%), spikes were significantly phase locked to theta oscillations, with a maximum strength of locking at 6 Hz (Fig. 1*D*).

Theta-band phase locking in relation to behavioral task periods

Based on the prominent theta-band activity during the waiting period, we hypothesized that OFC theta oscillations may contribute to processing of information related to outcome expectancy. To address this issue, we investigated theta-band phase locking during this period as a function of time (Fig. 2*A*). We calculated time–frequency representations of phase locking using symmetric segments of 500 ms centered on each time point, in steps of 10 ms. We found that the strength of spike–field phase locking steeply rose during the waiting period, with a peak at 6 Hz, and

rapidly collapsed at reward delivery (Fig. 2A). To contrast phase locking across different event-related time windows, we defined the following task periods in each trial: the odor-sampling period (0.5 s window starting at odor onset), the movement period (a variable-length window from odor offset until 0.2 s before fluid well entry), the waiting period (a 0.8 s window from 0.2 s after fluid well entry until fluid delivery), and the fluid delivery period (a 2 s window starting at fluid delivery). For each cell, a task-period-specific spike–LFP phase-locking spectrum was determined using the sample of spikes across all applicable trials within the task period under consideration. To remove the spike-number-related bias when comparing between different windows, we applied the bootstrapping procedure using 50 spikes per cell for each condition. Subsequently, we averaged the spike–LFP phase-locking spectrum across all units considered for a given task period.

This analysis confirmed that the strongest theta-band phase locking occurred during the waiting period (Fig. 2B–D). Importantly, the increment in phase locking during the waiting period was specific to the expectation of a positive outcome: phase locking was significantly higher during the waiting period before sucrose than during the waiting period before quinine and intertrial interval fluid-well visits (Fig. 2B). It was also significantly higher than phase locking during both the sucrose and quinine delivery periods (Fig. 2C,D) [all comparisons: Wilcoxon's matched-pairs signed-rank (WMPSR) test between values at 6 Hz, unless noted otherwise]. These observations confirm that theta-band phase locking does not occur merely as a consequence of immobility or a general state of outcome expectancy, but relates specifically to expectancy of reward due to a learned association. Second, the enhanced phase locking during reward anticipation does not appear to be a trivial consequence of enhanced theta power, because the reward delivery period was marked by elevated theta power but a low degree of phase locking when all cells were considered (Figs. 1C, 2C).

To investigate whether theta-band rhythmicity is also present in the local spiking activity per se, we used the high-pass-filtered (600–6000 Hz) multiunit spike train instead of the LFP to calculate phase locking of units (see supplemental material, available at www.jneurosci.org). Again, we found spectrally specific phase locking during the waiting period that was significantly higher in the waiting period before sucrose delivery than in the waiting period before quinine delivery ($p < 0.01$ WMPSR) (supplemental Fig. S3A, available at www.jneurosci.org as supplemental material). In addition, there was significantly more theta-band rhythmicity in the spiking output of the OFC during the waiting period before sucrose than both the waiting period before quinine and the intertrial interval ($p < 0.05$, one-tailed t test) (supplemental Fig.

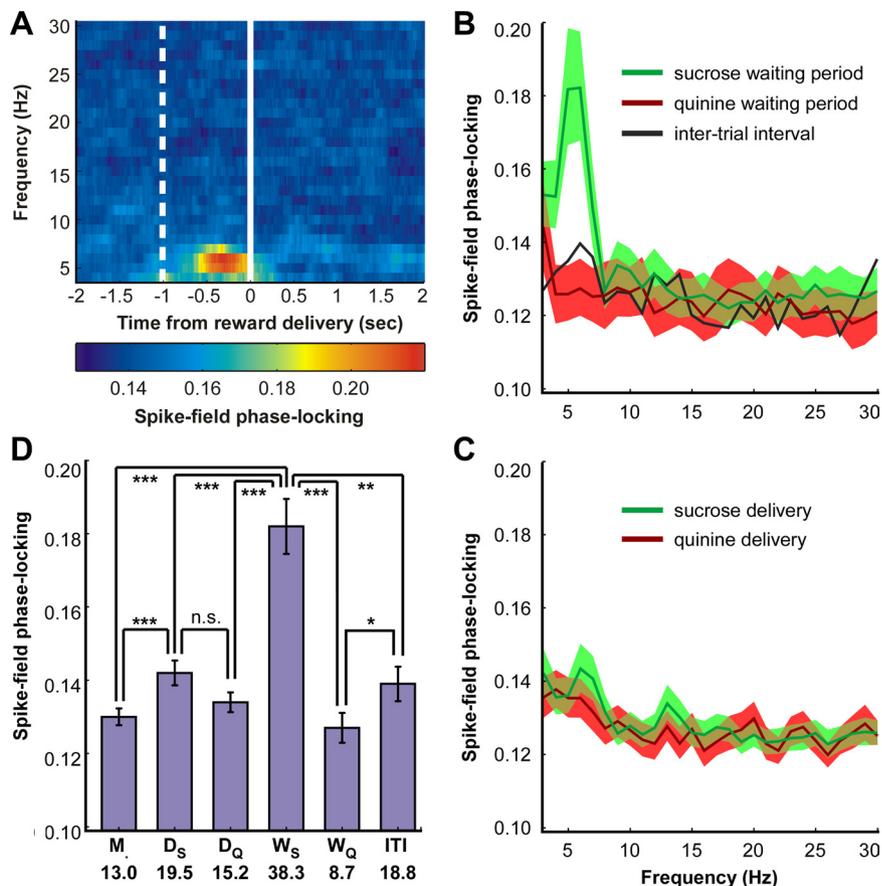


Figure 2. Modulation of spike–field phase locking during reward expectation and delivery. **A**, Time course of average (across all cells) phase locking centered on reward delivery. Phase locking was calculated in 500 ms symmetric windows centered on each time point. Solid line, Reward delivery; dashed line, start of waiting period. **B**, Average spike–field phase-locking spectra during waiting period, separated for sucrose, quinine delivery, and during (unrewarded) intertrial interval fluid pokes. Shaded regions indicate 95% bootstrapped confidence intervals on the mean. **C**, Same as in **B**, but now during fluid delivery period. **D**, Bar plot of phase-locking values at 6 Hz (mean \pm SEM) for different task periods. M, Movement; D_s , sucrose delivery; D_q , quinine delivery; W_s , waiting period before sucrose; W_q , waiting period before quinine. Numbers indicate percentages of cells significantly phase locked (Rayleigh's test $p < 0.05$). Comparisons, *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$, Wilcoxon's matched-pairs signed-rank test.

S3B–D, available at www.jneurosci.org as supplemental material). Both results confirm the presence of outcome-dependent theta-band rhythmicity in OFC spiking activity during the anticipatory period.

Spectral analysis of the oscillatory activity recorded during different trial periods indicated that besides the prominent theta-band (peak at 6 Hz) activity found during the waiting and reward delivery periods (Fig. 1C), another, weaker type of theta-band activity was present during the odor-sampling period, peaking at 9 Hz. Because the focus of this paper is on reward expectancy outside stimulus periods, theta-band phase locking during the odor-sampling period will not be further discussed below.

Strength of theta-band phase locking correlates with learning

We further investigated the relationship between theta-band phase locking and flexible learning. Because neural representations of reward expectancy are thought to be updated when an animal acquires or adjusts odor-reward associations (Rolls et al., 1996; Schoenbaum et al., 1998; Padoa-Schioppa and Assad, 2006), we hypothesized that theta-band rhythmicity is relevant for modification of these representations. To determine changes in spike–LFP phase-locking value across different stages of sessions in which odor-reward representations were modified, as expressed by changes in the behavior of the animal, we divided

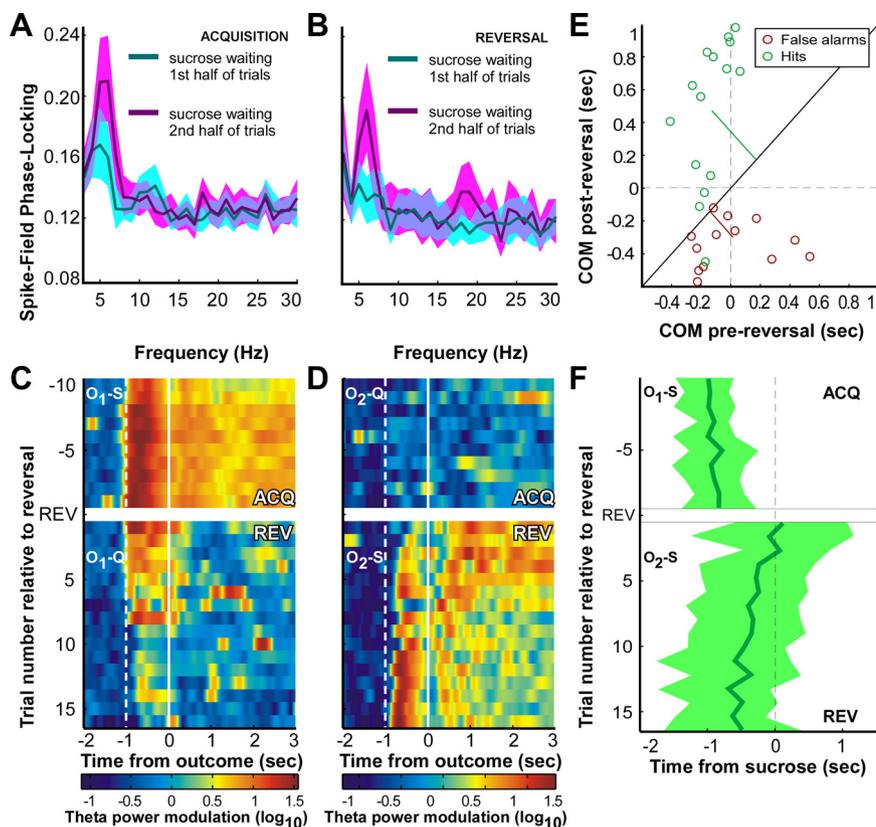


Figure 3. Learning-related changes in phase locking and LFP. *A*, Spike–field phase-locking spectra for prereversal phase, separated in early (cyan) and late (magenta) trials. *B*, Same as *A*, but now for reversal phase. *C*, Trial-by-trial modulation of theta power (6 Hz) relative to reversal. Pseudocolor scale indicates theta power modulation compared to baseline, averaged over contributing sessions (see Materials and Methods). Before reversal, odor 1 predicts sucrose (O₁-S). After reversal, odor 1 predicts quinine (O₁-Q). Dashed white line, Start of waiting period; solid white line, fluid delivery; ACQ, acquisition; REV, reversal. *D*, Same as *C*, but now synchronized on fluid delivery following odor 2. *E*, Hit/false alarm trials show opposite patterns for theta power COM (see Materials and Methods). Dashed lines indicate moment of fluid delivery. Red and green lines indicate mean shifts. *F*, Trial-by-trial timing of theta onset on hit trials in relation to reversal. Green line indicates mean first time point in time window where instant theta power reached >50% of maximum theta power on that trial. Shading indicates SD.

each session into four quadrants: the first and second half of the trials during the acquisition of odor-reward contingencies (pre-reversal period) (Fig. 3*A*), and the first and second half of the trials after the reversal of odor-reward contingencies (postreversal period) (Fig. 3*B*). In support of our hypothesis, we observed a significant increase in phase locking during the waiting period from the first to second half of the prereversal period ($p < 0.001$) (Fig. 3*A*). During the postreversal period, we first observed a significantly lower value of theta-band phase locking during the waiting period on early trials compared to the second half of acquisition ($p < 0.001$). As the animal acquired the reversed odor–outcome contingencies, theta-band phase locking under the sucrose rewarded condition returned to prereversal levels. During the second half of the reversal period, theta-band phase locking in this period was significantly higher compared to the first half ($p < 0.001$) (Fig. 3*B*). These findings demonstrate that theta-band phase locking during the waiting period does not reflect a fixed appetitive/aversive outcome dichotomy, but is related to the rat’s current outcome expectancy.

Trial-by-trial timing of LFP theta power modulation

Generally, spike–field phase-locking methods cannot be used on a trial-by-trial level because the number of spikes per trial is too low. To study the learning-related changes in theta-band oscillatory activity on a trial-by-trial basis, we instead examined the

evolution of LFP theta power across trials in relation to fluid delivery. Typically, on acquisition hit trials, the onset of the theta power increment was situated in the waiting period (Fig. 3*C*, top). On hit trials immediately following reversal, reward delivery is unexpected and the onset of the theta power increment appeared in the reward delivery phase (Fig. 3*D*, bottom), even though the behavioral sequence leading up to sucrose delivery is very similar. As the rat acquired the reversed odor-reinforcer contingencies, the onset of the theta power increment gradually returned to the waiting period (Fig. 3*D*, bottom). For false alarm trials, an opposite pattern was observed. During acquisition, the theta power increment was generally much weaker (Fig. 3*D*, top). On early reversal false alarm trials, the onset of the theta power increment was situated in the anticipatory period (Fig. 4*C*, bottom), resembling acquisition hit trials. Because these false alarm trials followed sampling of the previously positive odor, the elevated theta power may reflect the not yet updated outcome expectancy pattern. On subsequent false alarm trials, theta power gradually subsided, in line with the continued coupling of negative outcome to this odor. To quantify these observations across sessions, we calculated the center-of-mass (where “mass” is theta power) timing relative to fluid delivery, averaged across acquisition trials versus early reversal trials and separately for hit and false alarm trials (Fig. 3*E*, green and red). Indeed, the relation between COM timing

before and after reversal for hits and false alarms is opposite in almost all sessions. On hit trials, from acquisition (Fig. 3*C*, top) to early reversal (postreversal hit trial 1–5) (Fig. 3*D*, bottom) the COM shifted from 0.072 ± 0.033 s to 0.670 ± 0.033 s ($p < 0.001$), while on false alarm trials, the COM shifted in the other direction from 0.022 ± 0.075 s to -0.106 ± 0.042 s ($p < 0.001$). To specify this shift on a trial-by-trial basis, we defined the onset of the theta oscillation as the first moment when theta power crossed a threshold set at 50% of the maximum theta power on that trial. Regression analysis across sessions showed that, in the first 15 hit trials postreversal, theta onset shifted systematically to earlier moments with increasing trial number (slope: -0.066 s/trial, $p < 0.05$) (Fig. 3*F*). These results confirm that theta-band oscillatory activity mirrors the dynamic updating of outcome expectancies during learning, although behaviors closely associated with expectancy also need to be taken into account (see below).

Phase locking by neuronal groups with different task-related firing patterns

So far, we have pooled all neurons regardless of their firing-rate selectivity for particular task phases and behavioral events. However, OFC cells are known to exhibit a wide range of firing-rate modulations in relation to task events, and a subset of OFC cells shows firing-rate sensitivity to actual and predicted reward

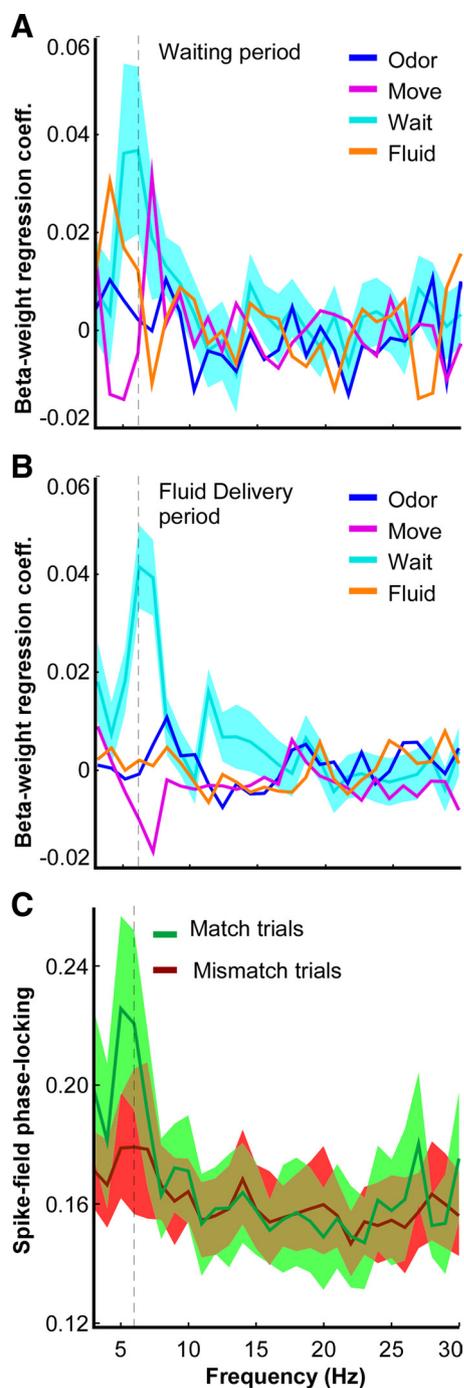


Figure 4. Relationship between behavioral firing-rate correlates and spike-field phase locking for different trial periods. **A**, Regression coefficients for the four categories of possible behavioral correlates. Only the waiting period correlate is a significant positive predictor of phase-locking values during the waiting period (tested at 6 Hz). Shaded regions indicate SEM. **B**, As in **A**, but now for the fluid delivery period. **C**, Phase-locking spectra separate for trials in which the outcome decoded from the neuron's firing rate matched (green) or failed to match (red) the actual outcome. Shaded regions indicate 95% bootstrapped confidence intervals on the mean. There was a significant difference at 6 Hz ($p < 0.01$, Wilcoxon's matched-pairs signed-rank test).

(Schoenbaum et al., 1998; Wallis and Miller, 2003; Padoa-Schioppa and Assad, 2006; van Duuren et al., 2008, 2009). Because strong theta-band phase locking during the waiting period becomes a reliable predictor of reward with learning, we predict that this property is predominantly expressed by those cells that show outcome selectivity in their firing rates. To investigate this,

Table 1. Distribution of behavioral firing-rate correlates across the recorded population of cells

None	Odor sampling	Movement period	Waiting period	Fluid delivery
[324]/0.62				
Odor sampling	43 (26) [88]/0.17	14 (6) 0.03	9 (7) 0.02	10 (8) 0.02
Movement period	14 (4) [72]/0.14	36 (12) 0.07	9 (5) 0.02	4 (0) 0.01
Waiting period	9 (2) [64]/0.12	9 (6) 0.02	24 (12) 0.05	9 (6) 0.02
Fluid delivery	10 (9) [62]/0.12	4 (2) 0.01	9 (5) 0.02	28 (20) 0.06
Triple	OD_WT_FD 6/0.01	OD_MV_FD 2/0.01	OD_MV_WT 4/0.01	MV_WT_FD 3/0.01

Distribution of 256 significant task-related correlates from a sample of 201 out of 525 units. Rows show categories of task-related correlates, columns specify single and double correlates in absolute numbers and fractions of total cell number. The numbers in brackets in the leftmost column indicate total number of units with a behavioral correlate in that category (including double and triple correlates, thus amounting to more than 256 correlates). Numbers in parentheses indicate the number of cells in correlate category that significantly discriminate between odors or outcomes. OD, Odor sampling; MV, movement period; WT, waiting period; FD, fluid delivery.

we classified cells according to the task periods in which they showed firing-rate responses that were significantly different from baseline in four basic groups (Table 1; supplemental Fig. S2, available at www.jneurosci.org as supplemental material). Often, cells exhibited firing rate correlates to more than one task period. To investigate the relationship between behavioral firing-rate correlates and phase locking, we performed a multiple linear regression analysis to predict phase-locking values from the presence of one or more correlates per cell.

Of the four possible behavioral firing-rate correlates, only the presence of a firing-rate correlate during the waiting period for a given cell (W-cell) was a strong and positive predictor of theta-band phase locking during the waiting period at 6 Hz (β -weight regression coefficient \pm SE = 0.038 ± 0.019 , $p < 0.05$ two-sided t test) (Fig. 4A). Though it might seem obvious that the group of cells with a firing-rate correlate during the waiting period should also show the highest phase-locking value, this is not a trivial finding since for each cell, a fixed number of spikes was used to calculate phase-locking values, eliminating a bias toward higher phase-locking values due to high firing rate. Strikingly, when we restricted this analysis to the reward delivery period (Fig. 4B), only the W-cell category was positively correlated with phase-locking values during reward consumption (β -weight \pm SE = 0.041 ± 0.007 , $p < 0.001$). These findings suggest that the increase in theta-band phase locking is dynamically expressed in anticipation of and during actual delivery of reward by a select group of neurons that also express expected outcome information through differential firing rates.

Trial-by-trial phase locking per unit predicts outcome

As phase-locking values cannot be meaningfully defined per trial, we could not determine whether phase-locking values were a better predictor of upcoming outcome than single-cell firing rates. However, we could select the group of trials in which a single cell's firing rate suggested one outcome, when actually the other outcome was coming up (mismatch trials) and contrast the phase-locking values for this group with the values for the group of trials in which the firing-rate of a single cell correctly predicted the upcoming outcome (match trials). In this contrast, we included those cells that significantly predicted the upcoming outcome (using a Bonferroni correction for the total number of cells; see Materials and Methods). To decode the predicted outcome per cell per trial, we determined the mean firing-rate response in

the waiting period for both quinine and sucrose delivery. Next, we used a leave-one-out decoding algorithm to predict for a single-cell trial response during the waiting period whether the actual fluid delivered was sucrose or quinine. This was determined by calculating the Euclidean distance of the single-cell's firing rate on that trial to the mean firing rate in the training set of quinine and sucrose trials (the training set was defined as the set of trials remaining after removal of the single trial under study). Next, we pooled the phase-locking values per condition (match/mismatch) over cells (Fig. 4C). We hypothesized that the phase-locking value would be lower on mismatch trials than on match trials. As expected, we observed a higher phase-locking value (0.22 ± 0.02) for decoded trials with a correct prediction than for decoded trials with an incorrect prediction (0.18 ± 0.01 , $p < 0.01$) (Fig. 4C). Thus, when the firing rates of single neurons correctly reflect the upcoming outcome, these cells are more strongly locked to the ongoing theta oscillation, possibly enhancing their effect on target cells. Conversely, when the outcome prediction associated with a unit's firing pattern fails to match the upcoming outcome, the unit's spikes on that trial are less phase locked.

Theta-band oscillatory activity in relation to licking behavior

Thus far, theta-band phase locking has been examined in the context of reward expectancy. In rats, expectancy can be tightly coupled to direct motor consequences such as anticipatory licking (Neafsey et al., 1986; Gutierrez et al., 2006, 2010). We found no systematic relationship between individual lick responses and theta activity in LFP traces (i.e., individual licks were not systematically associated with theta oscillations). However, based on the observation by Gutierrez et al. (2006) that OFC spiking activity may correlate with clusters of lick responses, we checked whether the frequency of licking was associated with theta-band phase locking under the respective trial conditions (supplemental Fig. S4, available at www.jneurosci.org as supplemental material). The intensity of licking was higher during the sucrose delivery period than during the sucrose waiting period, yet phase locking was selectively strong for the waiting but not the delivery period (compare Fig. 2D). Moreover, licking activity was significantly more frequent during the quinine-anticipation period as compared to visits to the fluid well during intertrial intervals, yet the phase locking was weaker during this anticipation period (compare Fig. 2D). Although these findings do not preclude any relationship between theta-band activity and licking (Gutierrez et al., 2010), they do suggest that theta-band phase locking is unlikely to be linked directly to the vigor or intensity of rhythmic motor activity expressed as licking.

Discussion

To summarize, our results indicate that OFC neurons engage in strong theta-band phase locking in anticipation of reward, mirroring the animal's current expectation of positive outcome as measured by behavioral Go responses. Specifically, cells whose firing rate was selectively modulated in anticipation of reward showed the strongest phase locking to the ongoing LFP theta oscillation during reward anticipation and delivery periods. When, on reversal of odor–outcome contingencies, established outcome expectancies were violated, theta-band phase locking declined (Fig. 3B) and the onset of theta oscillations following the previously unrewarded odor now appeared in the reward delivery period, coinciding with delivery of unexpected reward (Fig. 3D, F). As the animals acquired the reversed odor–outcome con-

tingencies, the onset of theta oscillations shifted back over trials to earlier moments in the waiting period (Fig. 3D, bottom).

Although the exact anatomical origin of LFP theta oscillations recorded from the OFC is not known, the present data suggest that theta oscillations are, at least in part, locally generated, while other brain areas may also contribute. First, theta-band rhythm was not detectable during locomotion toward the reward site, in contrast to what is expected for a hippocampal source (Fig. 1C). Second, we found a strong alignment of theta-band spike phases per cell (Fig. 1D) as well as between single-cell mean phases (Fig. 1E), suggesting that rhythmic spiking and synaptic activity at the population level contribute to the recorded LFP theta oscillations. Third, the results of the current source density analysis indicate the presence of a sink-source couple within the OFC, suggesting a local source of theta oscillations (supplemental Fig. S1A–C, available at www.jneurosci.org as supplemental material), although not precluding other, more remote sources. Finally, when we substituted the multiunit spiking activity for the LFP as a reference to calculate phase locking and when we calculated the frequency components of the autocorrelograms of composite multiunit spike trains (cMUAs), spectrally specific differences in theta-band oscillatory activity between behavioral conditions remained present (supplemental Fig. S3, available at www.jneurosci.org as supplemental material). Multiunit activity originates from a much smaller tissue volume than the LFP (Mitzdorf, 1987; Logothetis, 2003); therefore, spike–MUA phase locking and cMUA autocorrelation are measures local to the OFC. Thus, our results demonstrate that during reward expectation, the spiking output of the OFC not only shows theta-band rhythmicity, but also phase locks to locally recorded neural mass activity.

In addition to correlating with reward expectancy, the question arises whether theta-band phase locking is also related to licking behavior (Gutierrez et al., 2006, 2010). Unlike licking, which occurred across both waiting and delivery periods, theta-band phase locking was selectively enhanced only during the sucrose anticipation period (Fig. 2D). Theta-band phase locking and the frequency of lick responses were dissociated between sucrose anticipation and consumption periods (supplemental Fig. S4A, available at www.jneurosci.org as supplemental material). Moreover, during the quinine anticipation period in false alarm trials, the frequency of lick responses was higher than in intertrial intervals, whereas phase locking was weaker (supplemental Fig. S4B, available at www.jneurosci.org as supplemental material), again indicating a dissociation. Thus, phase locking does not appear to correlate primarily to licking behavior, but a more global linkage certainly may exist because reward expectancy can be considered a primary drive to initiate licking. Synchronized firing across OFC, amygdala, and agranular insular cortex may arise in conjunction with licking-related activity (Gutierrez et al., 2010), which however does not contradict a concurrent role in reward expectancy coding. Such an interleaving of rhythmic movement and information coding has been noted before in hippocampal, somatosensory, and olfactory structures, where theta activity correlates with locomotion, whisking, and respiratory rhythms, respectively (Vanderwolf, 1969; Lisman, 2005; O'Keefe and Burgess, 2005; Kleinfeld et al., 2006; Verhagen et al., 2007).

The present findings may have significant implications for our understanding of neural systems mechanisms for processing and acquiring reward information in the brain. Analysis of reward processing has thus far focused on firing-rate modulation by prediction or delivery of reward (Schultz et al., 1997; Schoenbaum et

al., 1998; van Duuren et al., 2008), and thus the temporal structuring of reward-related firing patterns remained unknown. Phase locking of OFC firing, as described here, may facilitate the temporal alignment of OFC output with firing in target areas, such as the amygdala, striatum, and neocortical sensory-associational areas (Carmichael and Price, 1995; Schilman et al., 2008), and enhance the impact of this synchronized output during selective phases of the oscillation (Salinas and Sejnowski, 2000; Engel et al., 2001). This impact may well depend on phase consistency between theta oscillations in OFC and in target areas themselves (Chrobak and Buzsáki, 1998; Hyman et al., 2003; Womelsdorf et al., 2007; Schroeder and Lakatos, 2009). Theta-band phase locking may thus facilitate the communication between areas involved in reward-dependent learning.

Recently, Schoenbaum et al. (2007) suggested that outcome-expectancy representations in OFC may serve to update and modulate learned associations stored in amygdala and related structures by comparing expected to actual outcomes. We observed a sustained theta-band activity during reward expectancy and delivery, expressed by elevated theta power and phase locking specific to outcome-selective cells. Previously, neural activity in amygdala and mPFC was shown to synchronize in the theta band (Pape et al., 1998; Siapas et al., 2005; Paz et al., 2008). We hypothesize that the synaptic updating process in OFC and related circuits underlying modification of stored associations may rely on spike-timing-dependent plasticity (STDP) (Levy and Steward, 1983; Cassenaer and Laurent, 2007) during coherent theta-band activity, which could be boosted by repetitive licking (Gutierrez et al., 2006). Under this hypothesis, the firing patterns phase locked to OFC theta oscillations convey a synchronized reward expectancy signal to target areas, which may or may not be continued into the period of the actual outcome, with significant consequences for learning. For example, when the outcome fails to meet the predicted signal (as in reversal learning), theta oscillatory activity in the delivery period collapses (Fig. 3C). Although this collapse was observed for theta-band LFP power, it can be safely assumed that the synchronization in spiking output is reduced at the same time, which may lead to a net weakening of synapses between OFC and target cells, consistent with STDP principles (Abbott and Nelson, 2000; Hayashi and Igarashi, 2009). Predictive synchronized firing after sampling of the other odor, previously coupled to quinine but now to reward, is regenerated at the moment of unpredicted reward (Fig. 3D,F), which may selectively strengthen another set of synapses onto target cells. During early reversal, the onset of oscillatory activity shifted to progressively earlier moments in the trial, in broad agreement with models of temporal difference (TD) learning (Schultz et al., 1997; Sutton and Barto, 1998). This class of reinforcement learning (RL) models posits the existence of a signal representing the error in reward prediction, as in classic conditioning theory (Rescorla and Wagner, 1972). This signal shifts backwards in time across learning trials toward the earliest reward-predicting event available to the subject. Despite this global similarity, our findings add two novel elements. First, the theta rhythm may implement a timing or “clocking” system that provides temporal structure to value signals. Such a timing system may fill in an important gap in RL models that have been underconstrained as to how time steps may be neurally implemented (cf. Schultz et al., 1997). Second, in contrast to the transient nature of error signals, theta-band phase-locked activity conveys value signals in a sustained manner as they come to precede reward delivery. Thus, this type of value signal may complement an early and transient reward prediction error signal emitted by mesencephalic dopa-

mine neurons after learning (Schultz et al., 1997). As compared to the quantities encoded in TD-learning and other RL models, the theta oscillatory activity in OFC resembles a value function $V(t)$, reflecting current reward predictions through time (cf. Pennartz, 1997; Niv and Schoenbaum, 2008). Because the theta oscillatory signal in OFC is expressed as a neural mass signal, it may exert vast and overriding effects on activity in its target structures. Although this hypothesis awaits further investigation, the current findings provide strong support for a role of the OFC in the temporal coordination of signaling expected and actual rewards, partly in conjunction with its role in rhythmic licking.

Future experiments may test the causal role of OFC theta oscillatory activity in the acquisition and reversal of learned associations, for instance by local infusions of substances that were shown to alter oscillations in other brain structures [e.g., GABA_B and muscarinic receptor antagonists (Leung and Shen, 2007) as well as blockers of I_h (Glasgow and Chapman, 2008)]. Second, further experiments may examine whether information on negative outcome predictions is also temporally structured by way of oscillatory activity, and how positive predictions about non-ingestible rewards are coded. Aversion- or punishment-related information may be coded outside the region of OFC currently studied, for instance in anterior cingulate cortex (Johansen et al., 2001) or nondopaminergic cell groups of the VTA (Ungless et al., 2004). Another possibility explaining the lack of theta-band phase locking during Go responses before quinine is that false alarm responses are generated by habit systems and in the absence of a clear negative expectancy, which if present would lead the animal to refrain from making a response that will be punished. Classical conditioning studies with an aversive outcome may help resolve this matter.

References

- Abbott LF, Nelson SB (2000) Synaptic plasticity: taming the beast. *Nat Neurosci* 3 [Suppl]:1178–1183.
- Bechara A, Damasio AR, Damasio H, Anderson SW (1994) Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50:7–15.
- Buzsáki G (2005) Theta rhythm of navigation: link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus* 15:827–840.
- Carmichael ST, Price JL (1995) Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol* 363:642–664.
- Cassenaer S, Laurent G (2007) Hebbian STDP in mushroom bodies facilitates the synchronous flow of olfactory information in locusts. *Nature* 448:709–713.
- Chrobak JJ, Buzsáki G (1998) Operational dynamics in the hippocampal-entorhinal axis. *Neurosci Biobehav Rev* 22:303–310.
- Engel AK, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci* 2:704–716.
- Fellows LK, Farah MJ (2003) Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain* 126:1830–1837.
- Fries P, Roelfsema PR, Engel AK, König P, Singer W (1997) Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proc Natl Acad Sci U S A* 94:12699–12704.
- Glasgow SD, Chapman CA (2008) Conductances mediating intrinsic theta-frequency membrane potential oscillations in layer II parasubicular neurons. *J Neurophysiol* 100:2746–2756.
- Gray CM, Maldonado PE, Wilson M, McNaughton B (1995) Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *J Neurosci Methods* 63:43–54.
- Gutierrez R, Carmena JM, Nicolelis MA, Simon SA (2006) Orbitofrontal ensemble activity monitors licking and distinguishes among natural rewards. *J Neurophysiol* 95:119–133.
- Gutierrez R, Simon SA, Nicolelis MA (2010) Licking-induced synchrony in

- the taste-reward circuit improves cue discrimination during learning. *J Neurosci* 30:287–303.
- Hayashi H, Igarashi J (2009) LTD windows of the STDP learning rule and synaptic connections having a large transmission delay enable robust sequence learning amid background noise. *Cogn Neurodyn* 3:119–130.
- Hyman JM, Wyble BP, Goyal V, Rossi CA, Hasselmo ME (2003) Stimulation in hippocampal region CA1 in behaving rats yields long-term potentiation when delivered to the peak of theta and long-term depression when delivered to the trough. *J Neurosci* 23:11725–11731.
- Jarvis MR, Mitra PP (2001) Sampling properties of the spectrum and coherence of sequences of action potentials. *Neural Comput* 13:717–749.
- Johansen JP, Fields HL, Manning BH (2001) The affective component of pain in rodents: direct evidence for a contribution of the anterior cingulate cortex. *Proc Natl Acad Sci U S A* 98:8077–8082.
- Kleinfeld D, Ahissar E, Diamond ME (2006) Active sensation: insights from the rodent vibrissa sensorimotor system. *Curr Opin Neurobiol* 16:435–444.
- Lachaux JP, Lutz A, Rudrauf D, Cosmelli D, Le Van Quyen M, Martinerie J, Varela F (2002) Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence. *Neurophysiol Clin* 32:157–174.
- Leung LS, Shen B (2007) GABAB receptor blockade enhances theta and gamma rhythms in the hippocampus of behaving rats. *Hippocampus* 17:281–291.
- Levy WB, Steward O (1983) Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience* 8:791–797.
- Lisman J (2005) The theta/gamma discrete phase code occurring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus* 15:913–922.
- Lisman J, Spruston N (2005) Postsynaptic depolarization requirements for LTP and LTD: a critique of spike timing-dependent plasticity. *Nat Neurosci* 8:839–841.
- Logothetis NK (2003) The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci* 23:3963–3971.
- Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275:213–215.
- Mitzdorf U (1987) Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex. *Int J Neurosci* 33:33–59.
- Neafsey EJ, Bold EL, Haas G, Hurlley-Gius KM, Quirk G, Sievert CF, Terberry RR (1986) The organization of the rat motor cortex: a microstimulation mapping study. *Brain Res* 396:77–96.
- Niv Y, Schoenbaum G (2008) Dialogues on prediction errors. *Trends Cogn Sci* 12:265–272.
- O'Keefe J, Burgess N (2005) Dual phase and rate coding in hippocampal place cells: theoretical significance and relationship to entorhinal grid cells. *Hippocampus* 15:853–866.
- O'Keefe J, Recce ML (1993) Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3:317–330.
- Ongür D, Price JL (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 10:206–219.
- Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature* 441:223–226.
- Pape HC, Paré D, Driesang RB (1998) Two types of intrinsic oscillations in neurons of the lateral and basolateral nuclei of the amygdala. *J Neurophysiol* 79:205–216.
- Paxinos G, Watson C (2007) *The rat brain in stereotaxic coordinates*, Ed 6. Amsterdam: Academic/Elsevier.
- Paz R, Bauer EP, Paré D (2008) Theta synchronizes the activity of medial prefrontal neurons during learning. *Learn Mem* 15:524–531.
- Pennartz CMA (1997) Reinforcement learning by Hebbian synapses with adaptive thresholds. *Neuroscience* 81:303–319.
- Pesaran B, Pezaris JS, Sahani M, Mitra PP, Andersen RA (2002) Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat Neurosci* 5:805–811.
- Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical conditioning II: current research and theory* (Black AH, Prokasy WF, eds), pp 64–99. New York: Appleton-Century-Crofts.
- Rolls ET, Critchley HD, Mason R, Wakeman EA (1996) Orbitofrontal cortex neurons: role in olfactory and visual association learning. *J Neurophysiol* 75:1970–1981.
- Rolls ET, McCabe C, Redoute J (2008) Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cereb Cortex* 18:652–663.
- Salinas E, Sejnowski TJ (2000) Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *J Neurosci* 20:6193–6209.
- Schilman EA, Uylings HB, Galis-de Graaf Y, Joel D, Groenewegen HJ (2008) The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex. *Neurosci Lett* 432:40–45.
- Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat Neurosci* 1:155–159.
- Schoenbaum G, Saddoris MP, Stalnaker TA (2007) Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Ann N Y Acad Sci* 1121:320–335.
- Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32:9–18.
- Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Siapas AG, Lubenov EV, Wilson MA (2005) Prefrontal phase locking to hippocampal theta oscillations. *Neuron* 46:141–151.
- Skaggs WE, McNaughton BL, Wilson MA, Barnes CA (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6:149–172.
- Stalnaker TA, Franz TM, Singh T, Schoenbaum G (2007) Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments. *Neuron* 54:51–58.
- Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction*. Cambridge, MA: MIT.
- Ungless MA, Magill PJ, Bolam JP (2004) Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 303:2040–2042.
- van Duuren E, Escámez FA, Joosten RN, Visser R, Mulder AB, Pennartz CM (2007) Neural coding of reward magnitude in the orbitofrontal cortex of the rat during a five-odor olfactory discrimination task. *Learn Mem* 14:446–456.
- van Duuren E, Lankelma J, Pennartz CMA (2008) Population coding of reward magnitude in the orbitofrontal cortex of the rat. *J Neurosci* 28:8590–8603.
- van Duuren E, van der Plasse G, Lankelma J, Joosten RN, Feenstra MG, Pennartz CM (2009) Single-cell and population coding of expected reward probability in the orbitofrontal cortex of the rat. *J Neurosci* 29:8965–8976.
- Vanderwolf CH (1969) Hippocampal electrical activity and voluntary movement in the rat. *Electroencephalogr Clin Neurophysiol* 26:407–418.
- Verhagen JV, Wesson DW, Netoff TI, White JA, Wachowiak M (2007) Sniffing controls an adaptive filter of sensory input to the olfactory bulb. *Nat Neurosci* 10:631–639.
- Wallis JD, Miller EK (2003) Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18:2069–2081.
- Womelsdorf T, Fries P, Mitra PP, Desimone R (2006) Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature* 439:733–736.
- Womelsdorf T, Schoffelen J-M, Oostenveld R, Singer W, Desimone R, Engel AK, Fries P (2007) Modulation of neuronal interactions through neuronal synchronization. *Science* 316:1609–1612.
- Womelsdorf T, Vinck MA, Bosman C, Oostenveld R, Desimone R, Fries P (2008) Putative excitatory and inhibitory neurons synchronize at different phases of the gamma cycle in visual area V4 of awake monkeys. *Front Syst Neurosci Conf Abstr Comput Syst Neurosci* 2008. www.cosyne.org/wiki/Poster2008-2.