Behavioral/Cognitive

Adaptive Encoding of Outcome Prediction by Prefrontal Cortex Ensembles Supports Behavioral Flexibility

Alberto Del Arco, ⑩**Junchol Park, Jesse Wood, Yunbok Kim, and Bita Moghaddam**

Department of Neuroscience and Center for Neuroscience, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

The prefrontal cortex (PFC) is thought to play a critical role in behavioral flexibility by monitoring action– outcome contingencies. How PFC ensembles represent shifts in behavior in response to changes in these contingencies remains unclear. We recorded single-unit activity and local field potentials in the dorsomedial PFC (dmPFC) of male rats during a set-shifting task that required them to update their behavior, among competing options, in response to changes in action– outcome contingencies. As behavior was updated, a subset of PFC ensembles encoded the current trial outcome before the outcome was presented. This novel outcome-prediction encoding was absent in a control task, in which actions were rewarded pseudorandomly, indicating that PFC neurons are not merely providing an expectancy signal. In both control and set-shifting tasks, dmPFC neurons displayed postoutcome discrimination activity, indicating that these neurons also monitor whether a behavior is successful in generating rewards. Gamma-power oscillatory activity increased before the outcome in both tasks but did not differentiate between expected outcomes, suggesting that this measure is not related to set-shifting behavior but reflects expectation of an outcome after action execution. These results demonstrate that PFC neurons support flexible rule-based action selection by predicting outcomes that follow a particular action.

Key words: cognitive flexibility; decision making; γ oscillations; outcome prediction; single-unit recording; working memory

---

Significance Statement

Tracking action– outcome contingencies and modifying behavior when those contingencies change is critical to behavioral flexibility. We find that ensembles of dorsomedial prefrontal cortex neurons differentiate between expected outcomes when action– outcome contingencies change. This predictive mode of signaling may be used to promote a new response strategy at the service of behavioral flexibility.

Introduction

Flexible rule-based changes in action selection is critical for optimal adaptation to changes in the environment (Miller and Cohen, 2001). This process, which is often referred to as set-shifting behavior, is modeled by tasks where organisms learn, by trial and error, that an action– outcome contingency has changed, and

that they need to change their behavior to obtain reward or avoid punishment. Across the mammalian species, set-shifting behavior depends on the functional integrity of the prefrontal cortex (PFC) subregions. In humans, damage to the PFC, or psychiatric disorders that involve PFC deficits, are associated with impaired set-shifting (Milner, 1963; Wilmsmeier et al., 2010). Lesions of the dorsolateral or medial PFC in monkeys (Dias et al., 1996) or its putative rat homolog, the dorsomedial PFC (dmPFC; Birrell and Brown, 2000), also impair set-shifting performance (Ragozzino et al., 1999; Stefani et al., 2003; Floresco et al., 2006, 2008; Darrah et al., 2008; Dalton et al., 2011; Park et al., 2016; Park and Moghaddam, 2017).

While most PFC recordings during set-shifting behavior have focused on rules or strategy encoding (Rich and Shapiro, 2009; Durstewitz et al., 2010; Rodgers and DeWeese, 2014; Bissonette and Roesch, 2015; Powell and Redish, 2016), less is known about the relationship between reward/outcome encoding and flexible behavior. We posited that adaptive encoding of expected rewards is critical for set-shifting performance. Although PFC neurons represent both received and predicted rewarding outcomes in monkeys (Watanabe, 1996; Mansouri et al., 2006; Seo et al., 2007;

Wallis, 2007; Histed et al., 2009; Wallis and Kennerley, 2010; Asaad and Eskandar, 2011; Donahue et al., 2013) and rodents (Narayanan and Laubach, 2008; Sul et al., 2010), a relationship between outcome-encoding neurons and flexible rule-based shifts in behavior has not been established.

To assess whether outcome-related activity in the dmPFC predicts set-shifting behavior, we recorded single-unit activity and local field potentials (LFPs) in the dmPFC of rats during a rule-based set-shifting task (Darrah et al., 2008) and a control task. In the "Rule" task, animals were rewarded after they guided their behavior according to two previously learned rules: an instrumental nose poke in a specific location (Side Rule) or a lit port (Light Rule). The rule was changed four times during each recording session, requiring rats to solve the new discrimination rule by trial and error based on the delivery or omission of reward. In the "No-Rule" task, trials were pseudorandomly rewarded and outcome delivery could not be predicted. As a consequence, animals could not shift their performance.

PFC neurons displayed outcome-predictive activity only during the Rule condition. This pattern emerged as animals adapted their behavior to the new rule, suggesting a link between the outcome-predictive activity and shift in action selection. γ Power increased during all action–outcome intervals in both tasks, suggesting that this measure reflects a general network activation during reward expectation but does not track behavioral flexibility.

Materials and Methods

Animals
Adult male Sprague Dawley rats weighing 300–360 g were pair-housed on a 12 h light/dark cycle (lights on at 7:00 P.M.). All experiments were performed during the dark phase when the animals are most active. The rats were placed on a mild food-restricted diet (15 g of rat chow per day) 2 days before starting behavioral experiments. All procedures were in accordance with the University of Pittsburgh's Institutional Animal Care and Use Committee and were conducted in accordance with the National Institute of Health *Guide for the Care and Use of Laboratory Animals*.

Experimental design
Animals were trained in the set-shifting task until they reached criterion and showed a stable performance (see Behavioral task). Then, through stereotaxic surgery, electrodes arrays were implanted in the medial PFC. One week after surgery, animals were retrained in the set-shifting task to criterion performance and recording sessions started (see Surgery and electrophysiology procedures). After completion of experiments, animals were anesthetized and their brains removed to confirm the placement of electrodes in the medial PFC (See Histology).

Behavioral task
Figure 1A depicts the behavior protocol. We designed the task to be performed under two outcome conditions (Stefani and Moghaddam, 2006) in the same session: Rule condition (sets 1–4), in which animals could guide their instrumental actions alternating between the two rules to receive a reward; and No-Rule condition (sets 5–8), in which animals' actions were pseudorandomly rewarded using a prefixed schedule regardless of their action.

The training protocol was similar to that described previously (Darrah et al., 2008). Rats were tested in an operant box (Coulbourn Instruments) containing a house light, a pellet magazine that could deliver food pellets (fortified dextrose, 45 mg; Bio-Serv) into a food trough, and three nose-poke holes arranged horizontally on the wall opposite the food trough.

Habituation and pretraining. Rats were handled in the vivarium for ≥3 d before habituation to experimental procedures. They were habituated to the operant chamber during two consecutive 20 min sessions. During these sessions, the house light was on and reward pellets were dispensed into the food trough at 30 s intervals, following consumption of the previous pellet. After habituation, rats performed four consecutive sessions of the two perceptual discrimination rules to be used in the

set-shifting task: "Light" and "Side" (see description below). The order of presentation of the discrimination rules was counterbalanced and the rats were reinforced with a single reward pellet for correct nose-poke responses, according to response discrimination rule. The sessions lasted 60 min (first and second sessions) and 90 min (third and fourth sessions) and the intertrial interval (ITI) was set at 5 s during the first two pretraining sessions and at 10 s during the last two pretraining sessions.

Rule task. After pretraining, the rats received daily training sessions on the set-shifting task (Fig. 1A). The task required the rats to shift their response patterns between two rules—the Light rule and the Side rule—each session, to receive a reward pellet. Each trial began when animals poked in the illuminated food trough. One second later, one of the three nose-poke cue lights was illuminated (cue onset). The Light rule required the rats to execute nose pokes into the illuminated cue hole, regardless of its location on the left, center, or right side of the chamber. The Side rule required animals to respond to a port in a designated spatial location (left, center, or right), regardless of which port was illuminated. Pseudo-randomization was used to ensure that the same cue light was never illuminated >2 times in a row.

In every trial, only correct responses were rewarded with a food pellet. There was a delay of 1 s between the response (instrumental poke) and the delivery or omission of the pellet. In both trials, correct and incorrect, the food trough was illuminated until the rats made a head entry into the food trough to end the trial. After a 10 s ITI, the food trough light was turned on again and the rats had to make a food trough poke to start a new trial. A trough poke to start every trial avoided premature responses before the cue onset.

We determined that animals shifted their behavior according to the new discrimination rule when they achieved 10 consecutive correct choices in a set (criterion). After reaching criterion, the response rule was immediately changed (extradimensional shift), requiring the rat to learn the new discrimination rule by trial and error based on the delivery or omission of reward. The task required the rats to reach the performance criterion four times (four sets) every session, resulting in three consecutive extradimensional shifts. The task was counterbalanced with eight possible sequences of extradimensional shifts (order and combination of rules) that were cycled in a systematic manner (for example, center–light–right–light; light–left–light–center; etc.). The rats were tested daily until stable performance was reached.

No-Rule task. In one session, following four sets of the Rule task (sets 1–4) as described above, animals were reinforced noncontingently for four sets (sets 5–8; Fig. 1A). Rats responded as detailed above, but rewards were given according to a predetermined pseudorandom schedule (no discrimination rules). Thus, during the noncontingent reward task, the rats performed 32 trials and received 17 rewards (15 omissions of rewards) every set. Although there were no rule shifts during the No-Rule task, the last 10 trials of every set were always rewarded (resembling the criterion of 10 consecutive correct responses before the rule changes in the Rule task). During the No-Rule task, trial outcomes did not provide information about rule sets.

Surgery and electrophysiology procedure
Upon completion of behavioral training (stable set-shifting performance), custom microelectrode arrays of eight polyamide-insulated tungsten wires (50 μm) were implanted under isoflurane anesthesia in the medial PFC (prelimbic; 3.0 mm anteroposterior, 0.7 mm dorsolateral, −4 mm dorsoventral from bregma) of rats (n = 10) as described previously (Park et al., 2016). The electrode array was secured onto the skull with dental cement using six screws as anchors. A silver wire was connected to one of the screws to be used as a ground. The rats were single-housed after electrode implantation. One week after surgery, animals were food-restricted again (15 g of rat chow per day) and were acclimated to the recording cable in the operant box for four 30 min sessions and retrained to criterion performance. Once set-shifting performance was stable and above criterion, recording sessions started.

Single units were recorded via a unity-gain field-effect transistor head stage and lightweight cabling, which passed through a commutator to allow freedom of movement within the test chamber (Neuro Biological Laboratories). Recorded single-unit activity was amplified at 1000× gain
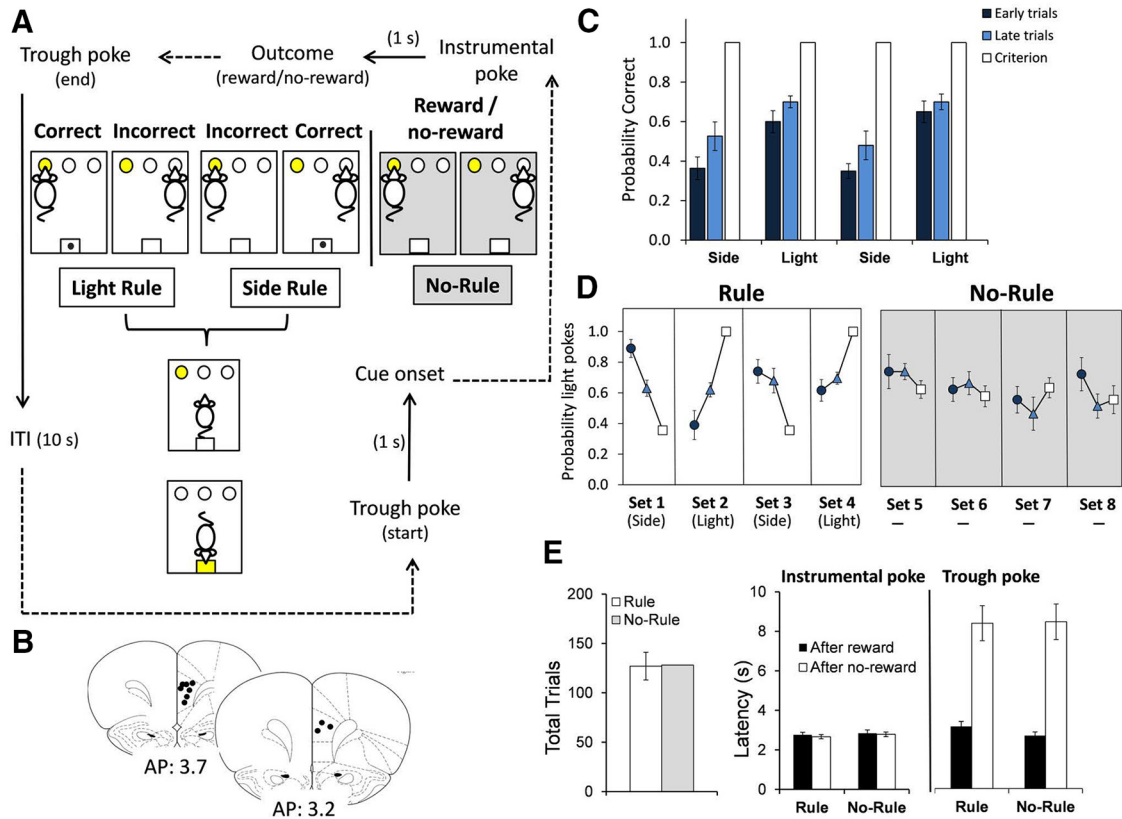
Figure 1. Set-shifting protocol, electrode placement, and a summary of behavioral data obtained during one recording session. ***A***, Animals were trained to perform three shifts (4 sets) between two rules: illumination (Light Rule) versus spatial location (Side Rule) at all sessions. After 10 s ITI, the light of the food trough turned on and animals were required to make a trough poke to start a new trial. Once cued, animals could poke into one of the three holes (left, center, or right), one of which was lit. During Light Rule sets, poking into the lit port (correct response) led to the delivery of a pellet (reward) with a 1 s delay. Poking into an unlit port (incorrect response) led to no reward. During Side Rule sets, the correct response depended on the spatial location (right in this illustration, but it could be left or center), regardless of which port was illuminated. During the No-Rule condition (shaded), pellets were delivered pseudorandomly regardless of the chosen port. After the instrumental poke, animals were required to make a trough poke to get the reward (in case of rewarded choice) and finish the trial to start the ITI. The rule changed without an explicit cue after 10 consecutive correct responses (criterion) and animals had to adapt to the new rule by trial and error. Dashed arrows indicate that time is not fixed. ***B***, Representation of electrode placement in the dmPFC (prelimbic area). ***C***, Bars (mean ± SEM; n = 10) show the probability of correct responses for every set (4 sets) during the set-shifting task. There were fewer correct responses in the early trials after the rule changed than in the late trials before reaching criterion (two-way ANOVA, $F_{(1,72)}$ = 8.28, p < 0.01). ***D***, Plot graphs (mean ± SEM; n = 10) show the probability of poking into lit ports during the two outcome conditions: the Rule condition, in which rewards depended on animals' choices; and the No-Rule condition (shaded), in which the delivery of rewards was not contingent on their actions. Animals changed their choices (increase poking probability), according to the set rule (indicated under the graph), during the Rule condition (two-way ANOVA, Light sets: $F_{(1,36)}$ = 5.46, p = 0.025; Side sets: $F_{(1,36)}$ = 5.36, p = 0.026), but not during the No-Rule condition (two-way ANOVA, 5–8 sets: $F_{(1,72)}$ = 1.04, p = 0.311). ***E***, Bars (mean ± SEM; n = 10) show that the number of total trials during the task (before reaching criterion) and the latency for instrumental actions did not change in Rule versus No-Rule conditions.

and analog bandpass filtered at 300–8000 Hz; LFPs were bandpass filtered at 0.7–170 Hz. Single-unit activity was digitized at 40 kHz, and LFPs were digitized at 40 kHz and downsampled to 1 kHz by Recorder software (Plexon). Single-unit activity was digitally high-pass filtered at 300 Hz, and LFPs were low-pass filtered at 125 Hz. Behavioral event markers from the operant box were sent to Recorder to mark events of interest (trough poke to start the trial, instrumental poke, trough poke to reward/no-reward). Single units were isolated in Offline Sorter (Plexon) using a combination of manual and semiautomatic sorting techniques as described previously (Sturman and Moghaddam, 2011).

Histology
After completion of the experiment, rats were anesthetized with chloral hydrate (400 mg/kg, i.p.) and perfused with saline and 4% buffered formalin. Brains were then removed and placed in 4% formalin. Brains were sectioned in coronal slices, stained with cresyl violet, and mounted to microscope slides. Electrode-tip placements were examined under a light microscope. Only rats with correct placements within the prelimbic PFC were included in electrophysiological analyses (Fig. 1B).

Analysis of electrophysiological data
Electrophysiological data were analyzed with custom-written scripts, executed in Matlab (MathWorks), along with the Chronux toolbox (http://chronux.org/). Multiple regression analyses were used to examine whether firing rate of units was predicted by one of the following: (1) the outcome of the previous and current trials (Outcome t-2, Outcome t-1, Outcome t, where t is current trial); (2) the set rule (Light/Side) and the spatial location (right/center/left); or (3) the interaction between set rule and spatial location. Regression coefficients generating p values <0.05 were considered significant. The analysis used all trials. We used 250 ms windows advancing in 50 ms steps for the −1 to 1 s duration around the trough poke to start the trial and for the −1 to 2 s around the instrumental action. To be considered outcome selective, a unit had to display a significant response in ≥4 of 5 bins (50 ms bin) for each of the 250 ms intervals in the following time windows: from 0 to 0.5 s (outcome-predictive) and from 1.25 to 1.75 s (outcome-responsive), time locked to the instrumental poke. χ^2 Tests were used to compare the proportion of units correlated with current outcomes under the two outcome conditions (Rule and No-Rule tasks) and during specific time windows (i.e., from 0 to 0.5 s and from 1.25 to 1.75 s, locked

to the instrumental poke). The same selectivity criteria and time windows were used to identify units selective for the set rule and spatial location, and their interaction.

The area under the receiver operating characteristic (ROC) curve was computed to assess discrimination between rewarded and unrewarded outcomes during the Rule and No-Rule tasks, and to compare discrimination during the early and late trials of the Rule task set. The ROC curve was also calculated to assess discrimination between light and side, right and no-right, center and no-center; and left and no-left. The area under the ROC (auROC) curve was bounded by 0 and 1, with more extreme ROC values indicative of greater discrimination, and 0.5 indicative of no discrimination, between conditions. auROC was obtained in 250 ms windows advanced in 50 ms steps. auROC curve of selective units was averaged and compared between the Rule and the No-Rule tasks. The auROC curve also was used to compare the selectivity of units during the early trials of the set (after the rule changed) and the late trials of the set (before criterion). The first/last 5 or 10 trials of every set were used to perform this analysis (5 or 10 depending on the number of trials of the set). Because some sets had <10 trials before criterion, there was some overlap (<15%) between the early and the late trials. Paired Student's t tests were used to find statistical significances in auROC curve population comparisons. A bootstrap analysis was performed to test whether ROC values of the Rule trials compared with the No-Rule trials, and early trials compared with late trials, were different by chance. For each unit, we shuffled the outcome for each trial and recalculated ROC areas 1000 times. Every time, shuffled ROC areas for the selective units during the Rule/No-Rule task, or early/late trials, were compared (paired t test).

Raw LFP signals from every trial were aligned to the instrumental action. Trials with clipping artifacts and trials with LFP values higher and lower than ±2.5 times the SD of the mean of the total signal were excluded. The power spectrum of every trial was calculated by fast Fourier transform using the Chronux function "mtspecgramc". This was done using a 500 ms moving window in 10 ms steps from −1 to 2 s locked to the instrumental poke. A multitaper approach was used because it improves spectrogram estimates when dealing with noninfinite time series data (Mitra and Pesaran, 1999). Each frequency bin (row) in the power spectrum was Z-score normalized to the average spectral power during a baseline period (a 3 s window of the ITI ending 1 s before the trough poke to start the trial). Rewarded and unrewarded trials were grouped in every animal to evaluate the effects of trial outcomes on the LFP power spectrum. Likewise, the early/late trials (see above) of every set were averaged to seek differences in the LFP power spectrum.

Statistical analyses
Behavioral performance data were analyzed with one-way and two-way ANOVAs (early/late trials × sets) and paired Student's t test (Rule vs
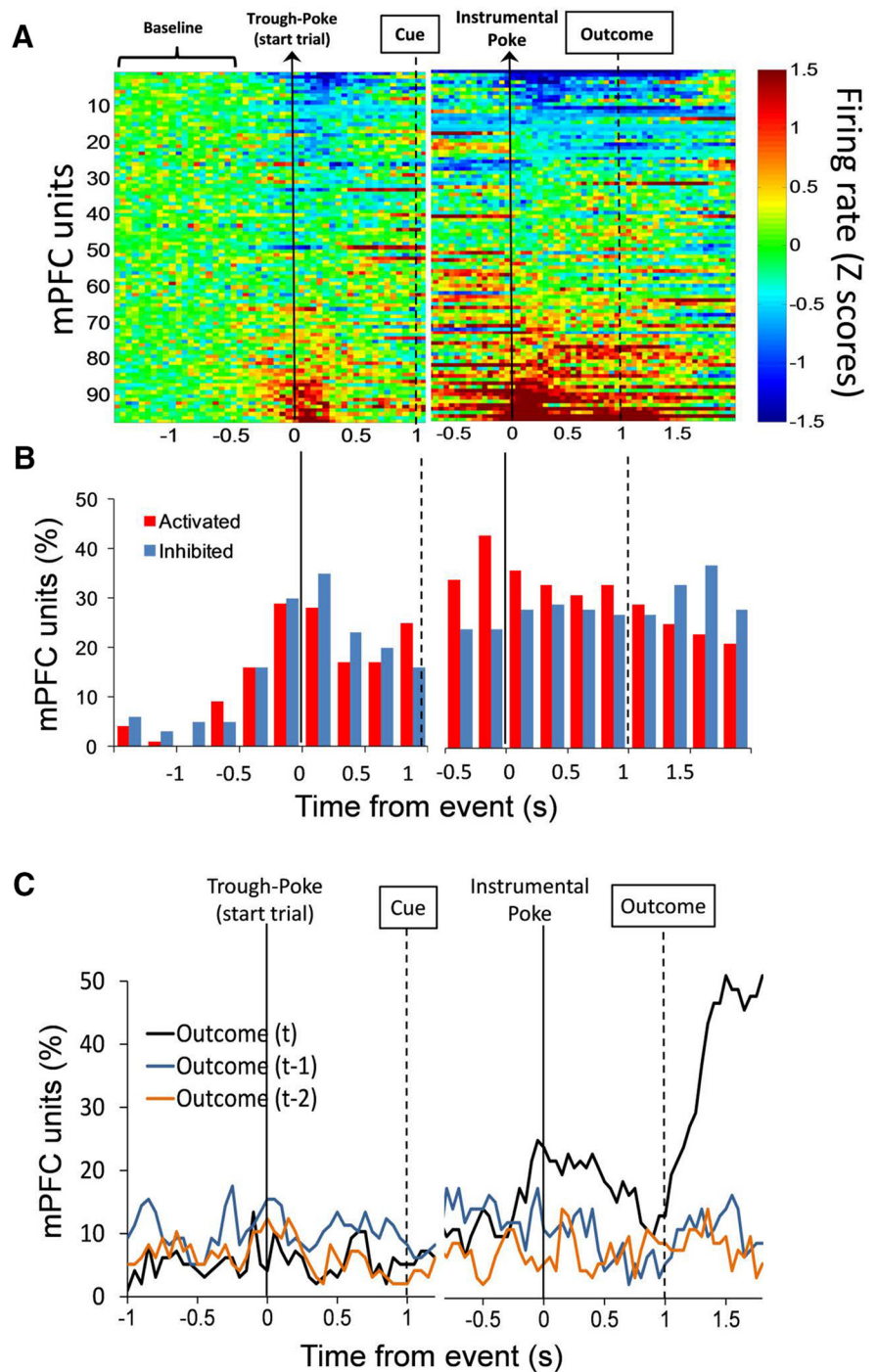

Figure 2. PFC single units change their activity and predict current trial outcomes during set-shifting performance. **A**, Heat plot represents the baseline-normalized firing rate for each unit ($n = 97$). Each row is the activity of an individual unit in 50 ms time bins aligned to corresponding task events and sorted from lowest to highest average normalized firing rate. **B**, Bars represent the time course of unit's activation and inhibition. The percentage of units was categorized as activated or inhibited over time based on whether their averaged activity by 250 ms time windows was significantly different from baseline activity. **C**, Time course showing the proportion of PFC units whose activity was significantly correlated to current outcome (t) and the previous two outcomes (t-1 and t-2) during set-shifting performance. A significantly higher proportion of units encoded current, but not previous, trial outcomes, both before ($\chi^2_{(2)} = 9.89$; $p < 0.01$, average across time bins from 0 to 0.5 s) and after ($\chi^2_{(2)} = 41.94$; $p < 0.001$, average across time bins from 1.25 to 1.75 s) the outcome. Arrows and thick lines indicate the timing of task events.

No-Rule). Single-unit activity data were analyzed with χ^2's and paired Student's t test to identify differences in the proportion of neurons and ROC values, respectively (early vs late trials and Rule vs No-Rule). LFP data were analyzed with two-way ANOVAs (Rule/No-Rule × early/late trials or rewarded/nonrewarded trials). $p < 0.05$ was

considered the cutoff for statistical significance. Data are represented as mean ± SEM.

Results

Behavioral performance

Animals showed a stable performance in the set-shifting task after training. The total number of trials and errors during the last six consecutive recording sessions were as follows: 159 ± 15 trials and 48 ± 6 errors; 167 ± 19 trials and 49 ± 8 errors; 143 ± 10 trials and 38 ± 3 errors; 187 ± 16 trials and 55 ± 7 errors; 176 ± 30 trials and 50 ± 7 errors; 167 ± 15 trials and 50 ± 6 errors ($F_{(1,5)} = 0.86$, $p = 0.512$, one-way ANOVA for total trials; $n = 10$ animals). Behavioral similarities and differences between Rule and No-Rule conditions are depicted in Figure 1. During the Rule condition, rule-based behavioral guidance was evident as the probability of correct responses increased from early to late trials within each set (Fig. 1C; $F_{(1,72)} = 8.28$, $p = 0.005$, two-way ANOVA, $n = 10$). Animals systematically changed their choice as a function of the rule set. Accordingly the probability of choosing the illuminated port increased or decreased across trials in the Light versus Side rule sets, respectively (Fig. 1D; Light sets: $F_{(1,36)} = 5.46$, $p = 0.025$; Side sets: $F_{(1,36)} = 5.36$, $p = 0.026$, two-way ANOVA). In the No-Rule condition, however, the probability of poking the illuminated ports did not significantly change across trials (Fig. 1D; 5–8 sets: $F_{(1,72)} = 1.04$, $p = 0.311$, two-way ANOVA). Thus, animals guided their behavior based on the task rule when outcomes were contingent upon action selection in the Rule condition.

The number of trials and the latency to execute an instrumental action and retrieve rewards did not differ in Rule versus No-Rule conditions (Fig. 1E). Animals performed 127 ± 14 trials (77 ± 9 rewarded correct; 50 ± 6 nonrewarded incorrect) before reaching criterion during the Rule condition; and 128 trials (68 rewarded and 60 nonrewarded) during the No-Rule condition. Reward delivery did not significantly change the latencies to make instrumental pokes (Rule: latency after rewarded trials, 2.76 ± 0.12 s; latency after nonrewarded trials, 2.66 ± 0.10 s, $t_{(9)} = 0.90$, $p = 0.389$; No-Rule: latency after rewarded trials, 2.85 ± 0.16 s; latency after nonrewarded trials, 2.79 ± 0.11 s, $t_{(9)} = 0.42$, $p = 0.681$; paired t test). The latency to poke in the food trough after nonrewarded responses was significantly longer than after rewarded responses in both conditions (Rule: latency after rewarded, 3.15 ± 0.25 s; latency after nonrewarded, 8.37 ± 0.88 s, $t_{(9)} = 6.76$, $p = 0.000$; No-Rule: latency after rewarded, 2.69 ± 0.18 s; latency after nonrewarded, 8.45 ± 0.90 s, $t_{(9)} = 6.89$, $p = 0.000$; paired t test).

Single-unit activity

PFC neurons predict and signal trial outcomes

Figure 2A shows the baseline-normalized (Z scores) firing rate of all single units ($n = 97$, two sessions, 10 rats; units per rat: 26, 12, 13, 11, 8, 7, 6, 5, 5, 4) time-locked to the initial poke to start the trial and the instrumental poke during the Rule condition. Many units were activated and inhibited at the time of the instrumental action and at the time of the outcome, when animals received feedback for their actions by delivery or omission of reward (Fig. 2B). A multiple regression analysis was performed to identify units that encoded the previous and/or current response outcome, i.e., units that significantly modulated their activity as a function of the response outcome in the previous and/or current trials. A significantly greater proportion of units encoded the current trial outcome during the interval between the action and the outcome (preoutcome period; $\chi^2_{(2)} = 9.89$; $p < 0.01$, averaged

Table 1. The number and percentage of medial PFC single selected units encoding each task variable during the set-shifting task; instrumental poke (0 – 0.5 s) and outcome (1.25–1.75 s)

	Instrumental poke	Outcome
Current outcome (t)	24 (25%)	48 (49%)
Most recent previous outcome (t-1)	10 (10%)	12 (12%)
Second most recent previous outcome (t-2)	4 (4%)	5 (5%)
Set Rule	12 (12%)	18 (19%)
Right	17 (17%)	18 (19%)
Center	14 (14%)	12 (12%)
Left	20 (21%)	21 (22%)
Rule × Right		
Side–Right	17 (18%)	14 (14%)
Light–Right	12 (12%)	10 (10%)
Rule × Center		
Side–Center	19 (20%)	10 (10%)
Light–Center	12 (12%)	17 (18%)
Rule × Left		
Side–Left	12 (12%)	12 (12%)
Light–Left	9 (9%)	9 (9%)

across time bins from 0 to 0.5 s), and after the outcome (postoutcome period; $\chi^2_{(2)} = 41.94$; $p < 0.001$, averaged across time bins from 1.25 to 1.75 s), compared with the proportions of units encoding previous trial outcomes during the same intervals (Fig. 2C). The units that encode current trial outcomes during the preoutcome period are referred to as outcome-predictive units. The units that encode current trial outcomes during the postoutcome period are referred to as outcome-responsive units. There were 24 of 97 (25%) outcome-predictive units and 48 of 97 (49%) outcome-responsive units according to selectivity criteria (see Material and Methods). In addition, we performed multiple regression analyses to identify selective units that encoded the set rule (Light vs Side), the spatial location (Right, Center, and Left) and the interaction set rule × spatial location. These data are summarized in Table 1.

To investigate whether predictive encoding and responsive encoding of the current outcome emerges selectively during rule-guided actions, we recorded neuronal activity during the Rule and the No-Rule conditions in the same session. The average baseline firing rate did not differ between the two conditions (Rule, 7.6 ± 1.3 Hz; No-Rule, 7.4 ± 1.3 Hz; 42 units). The heat plots of the overall response and the proportion of units responding to current outcomes during the Rule and No-Rule conditions are shown in Figure 3A–C. The outcome-predictive activity for rewarding and nonrewarding trials observed in the Rule condition was changed in the No-Rule condition, whereas the outcome-responsive activity was similar between the two conditions. The proportion of outcome-predictive units in the Rule condition was significantly greater than that of the No-Rule condition ($\chi^2_{(1)} = 5.12$; $p < 0.025$, average of time bins from 0 to 0.5 s; Fig. 3C). In contrast, the proportion of outcome-responsive units was not different between the two conditions ($\chi^2_{(1)} = 0.04$; $p > 0.50$, average of time bins from 1.25 to 1.75 s). There were 13 of 42 (31%) outcome-predictive units and 22 of 42 (52%) outcome-responsive units during the Rule condition according to selectivity criteria (see Material and Methods). These results indicate that significantly more PFC units represented impending outcomes when the reward was contingent on the animals' action, but not when actions were randomly reinforced.

To quantify each unit's discrimination of the current response outcome, we computed the auROC curves from distributions of firing rates in rewarded versus unrewarded trials (Wallis and
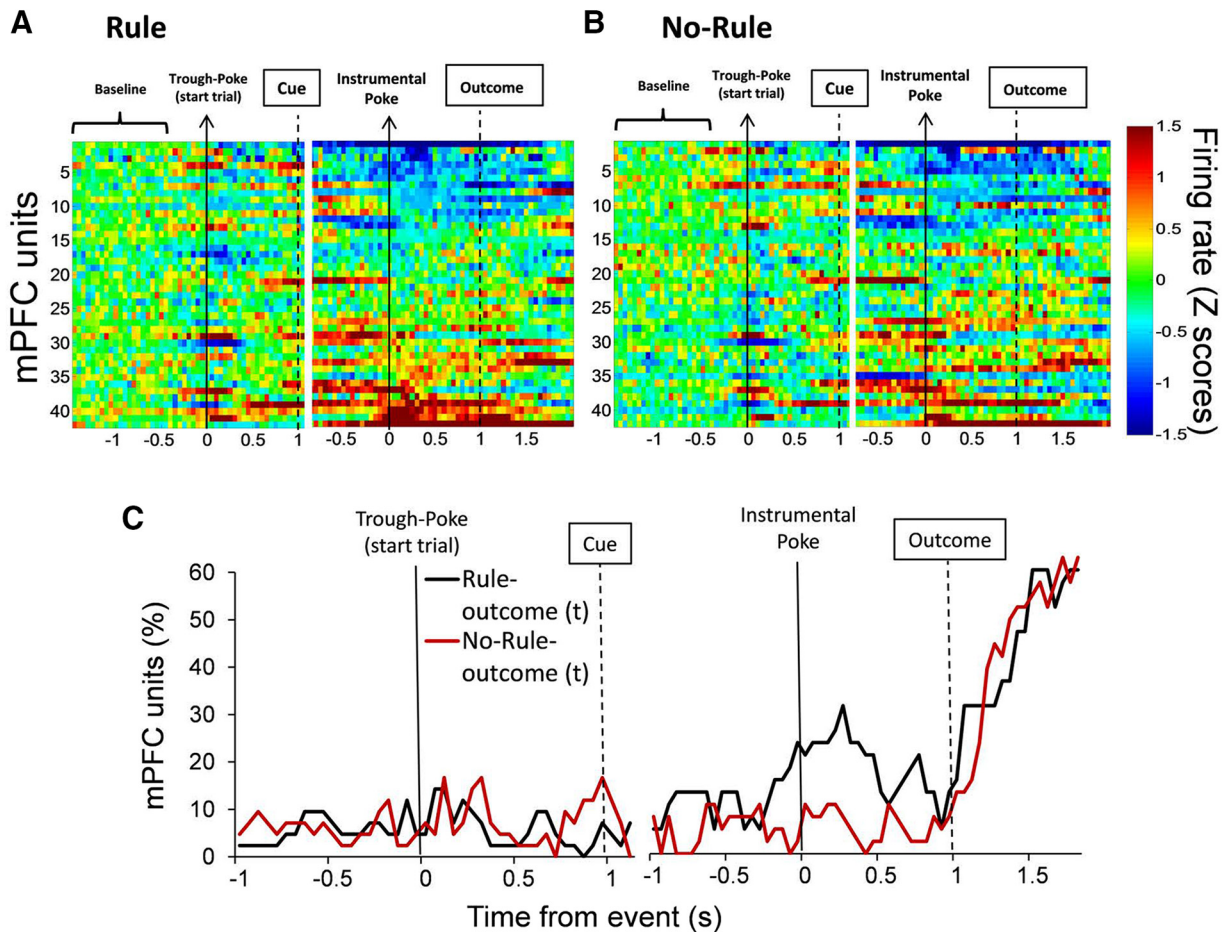
Figure 3. PFC single-unit firing rate is modulated by outcome conditions. ***A***, ***B***, Heat plots represent the baseline-normalized firing rate for each unit (*n* = 42) during the two outcome conditions, Rule (***A***) and No-Rule (***B***). Each row is the activity of an individual unit in 50 ms time bins aligned to corresponding task events and sorted from lowest to highest average normalized firing rate. ***C***, Time course showing the proportion of PFC units whose activity was significantly related to current (t) trial outcomes comparing Rule and No-Rule conditions. There were considerably more units anticipating the current trial outcome during the Rule compared with the No-Rule condition ($\chi^2_{(1)}$ = 5.12; $p < 0.025$, average across time bins from 0 to 0.5 s).

Miller, 2003). The discriminability of the impending outcome was significantly modulated by the presence versus absence of the action–outcome contingency. As shown in Figure 4*A*, the ROC area for the selective units was significantly higher during the preoutcome period in the Rule compared with the No-Rule condition ($t_{(12)}$ = 7.29, p = 0.000, paired t test, average of time bins from 0 to 0.5 s), but not during the postoutcome period ($t_{(21)}$ = 0.61, p = 0.540, paired t test, average of time bins from 1.25 to 1.75 s). The bootstrapped false-positive rate was α = 0.005. These results are consistent with the higher proportion of outcome-predictive units during the Rule condition (Fig. 3*C*) and extend those findings to show a higher discrimination to predict the response outcome in the presence of the rule defining the action–outcome relationship. Figure 4*B*,*C* shows one example neuron in which the outcome-predictive activity (spikes/s) for rewarding and nonrewarding trials observed in the Rule condition was changed in the No-Rule condition, whereas the outcome-responsive activity was similar between the two conditions. Figure 4*D*,*E* shows the net difference in the firing rate between rewarded and nonrewarded trials for every selective unit during the Rule and No-Rule conditions. The difference in the firing rate (average of all selective units) between rewarded and nonrewarded trials was significantly higher in the Rule compared with the No-Rule condition during the preoutcome period (13 units; $t_{(12)}$ = 2.78, p = 0.016, paired t test; Fig. 4*D*, inset), but not during the postoutcome

period (22 units; $t_{(21)}$ = 0.04, p = 0.964, paired t test; Fig. 4*E*, inset).

Outcome-predictive activity in the PFC anticipates behavioral shifts

The above results suggested that the outcome-predictive encoding in the PFC may reflect the switch to a different response strategy. To further examine this relationship, we tested whether outcome-predictive activity emerged as the animals achieved the rule shift across trials within each set. We found that the majority of outcome-predictive units showed enhanced discriminability in the late trials, before criterion was reached, as opposed to early trials, after the action–outcome contingency had been reset (Fig. 5*A*,*B*). The outcome-predictive selectivity was significantly greater in late trials compared with early trials ($t_{(23)}$ = 3.61, p = 0.001, paired t test, averaged across time bins from 0 to 0.5 s; Fig. 5*B*). The bootstrapped false-positive rate was α = 0.021. This effect seems to be consistent looking at rats individually (Table 2). The outcome-responsive selectivity did not differ between early versus late trials ($t_{(47)}$ = 0.53, p = 0.596, paired t test, averaged across time bins from 1.25 to 1.75 s).

We also tested whether the discriminability of units encoding for the set rule and the spatial location change in the late trials compared with that in the early trials. As shown in Table 3, there were no significant differences in the mean ROC values calcu-
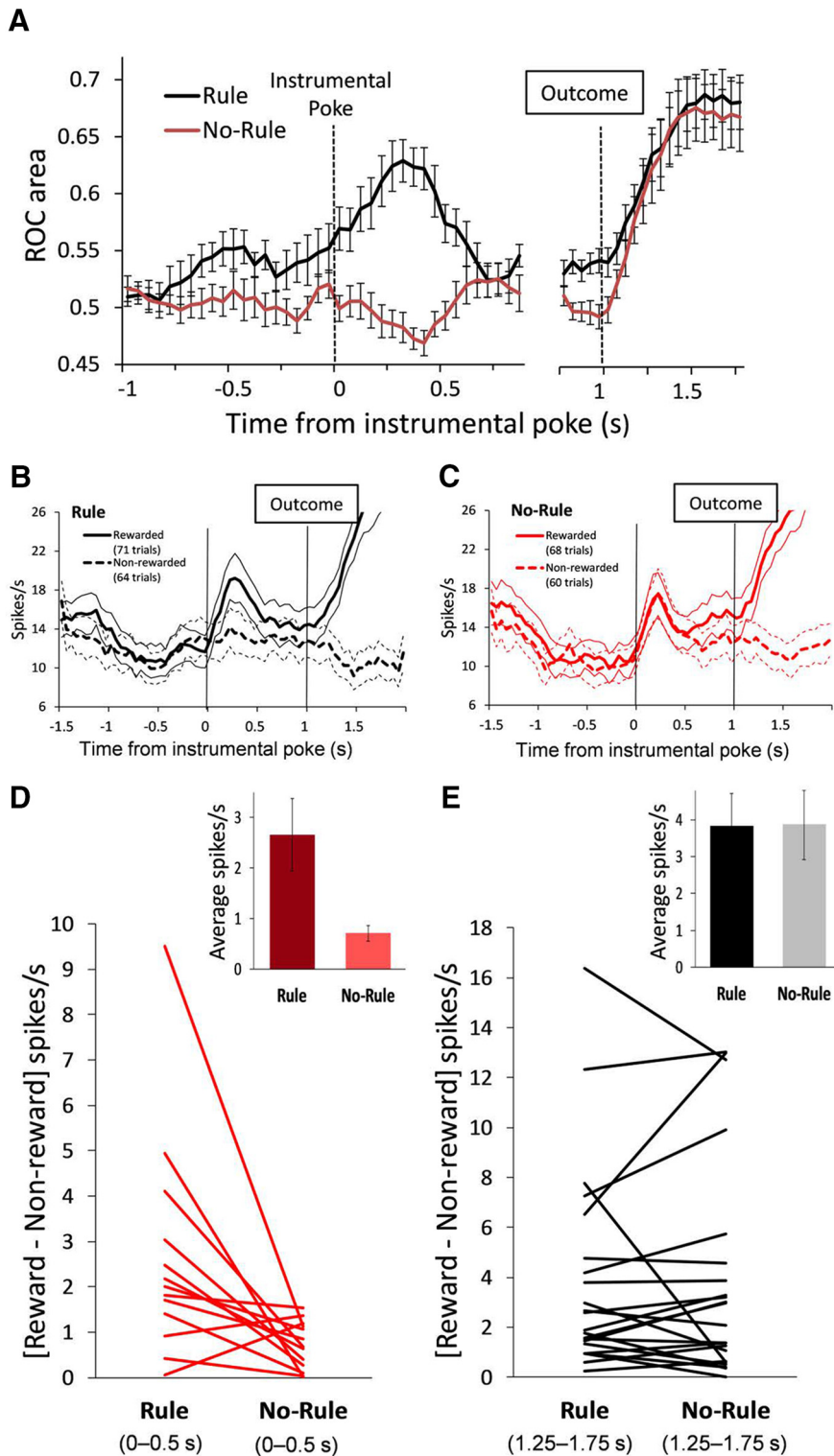
Figure 4. Outcome anticipation-related activity of PFC single units is modulated by the outcome condition. ***A***, Time course showing units' selectivity for current trial outcomes during the Rule condition compared with the No-Rule condition. Plots represent (mean ± SEM) the auROC curve of the population of selective units significantly correlated to current trial outcomes during the Rule condition. PFC units showed a greater selectivity to anticipate trial outcomes when these depended on animals' actions (Rule condition, paired *t* test, $t_{(12)} = 7.29$, $p = 0.000$, average across time bins from 0 to 0.5 s). The time axis is split to better represent the different number of selective units before (13 of 42 units) and after (22 of 42 units) the outcome. ***B***, ***C***, An example of a prefrontal single unit during the Rule (***B***) and the No-Rule condition (***C***). Units' activity anticipates current trial outcomes during the Rule, but not during the No-Rule, condition. Thick and dotted line plots represent the mean firing rates (spikes/s) of rewarded and nonrewarded trials, respectively. Thin line plots represent the corresponding ± SEM. Vertical lines indicate the instrumental poke and outcome time events. ***D***, ***E***, Net difference in the firing rate (spikes/s) between rewarded and nonrewarded trials for every selective unit during the Rule compared with the No-Rule condition, before (***D***; average from 0 to 0.5 s; 13 units) and after

lated for these variables. These results suggest that the outcome-predictive activity reflects the PFC neuronal representation of the new action–outcome contingency, since it arises as animals adapt their behavior to the set rule across trials.

LFP activity

The power of low-γ-band oscillations (30–50 Hz) increased during the 1 s interval between the action and the outcome (Fig. 6*A*). There was no difference in the low-γ power (30–50 Hz) between Rule and No-Rule conditions ($F_{(1,20)} = 2.24$, $p = 0.149$, two-way ANOVA). Also, the γ power (30–50 Hz) was not significantly different between rewarded versus nonrewarded trials (Fig. 6*B*; $F_{(1,20)} = 0.09$, $p = 0.762$, two-way ANOVA) or early versus late trials (Fig. 6*C*; $F_{(1,20)} = 0.07$, $p = 0.794$, two-way ANOVA) in both conditions. These data suggest that low-γ power in the medial PFC may reflect task engagement, but not shifting behavior.

Discussion

Tracking action–outcome contingencies and modifying behavior when those contingencies change is critical to behavioral flexibility. We find that dmPFC neurons differentiate between expected outcomes when action–outcome contingencies change. This mode of outcome prediction signaling may be used to promote the exploitation of a new response strategy and ultimately facilitate behavioral flexibility.

Set-shifting performance and outcome-related activity in the PFC

To better understand how outcome representation by dmPFC neurons related to behavioral shifts, we used an operant set-shifting task in which (1) there was no external cue to signal rule changes because adaptation to the rule set depended on trial and error based on the delivery and omission of rewards; and (2) rats performed several shifts back and forth between two perceptual dimensions in the same recording session. Consistent with previous work (Watanabe, 1996; Mansouri et al., 2006; Seo et al., 2007; Wallis, 2007; Narayanan and Laubach, 2008; Histed et al., 2009; Sul et al., 2010; Wallis and Kennerley, 2010; Asaad and Eskan-

---

(***E***; average from 1.25 to 1.75 s; 22 units) the outcome. Insets show that the difference in the firing rate (average of selective units) between rewarded and nonrewarded trials is significantly higher in the Rule compared with the No-Rule condition before ($t_{(12)} = 2.78$, $p = 0.016$, paired *t* test), but not after, the outcome.
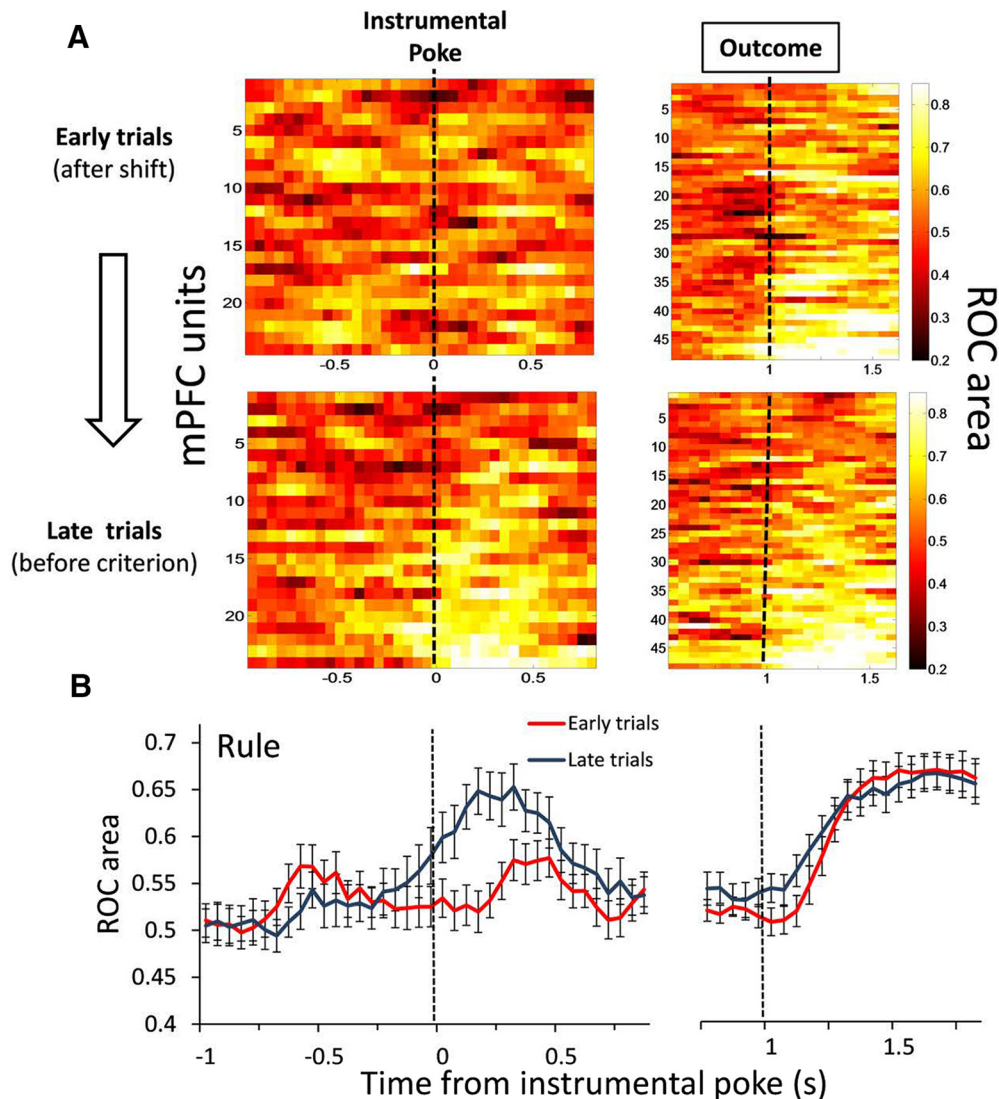
Figure 5. The selectivity of PFC neurons to anticipate trial outcomes is modulated by animals' adaptation to the set rule. *A*, Graphs represent the time course of current trial-outcome selectivity during set-shifting performance (Rule task) comparing the early trials after the shift (top) to the late trials before reaching criterion (middle). Heat plots represent the ROC area of the population of selective units significantly correlated to the current outcome by grouping the early trials and late trials of sets. Each row is the ROC area of a single unit sorted from lowest to highest ROC values. *B*, Temporal profile of the averaged ROC area (mean ± SEM) of the population of units represented in the heat plots. The selectivity to anticipate current trial outcomes was greater during the late trials of the set compared with the early trials (paired *t* test, $t_{(23)} = 3.61$, $p = 0.001$, average of time bins from 0 to 0.5 s). The time axis is split to better represent the different number of selective units before (24 of 97 units) and after (48 of 97 units) the outcome. Dashed vertical lines indicate the time of the instrumental poke and outcome events.

Table 2. Selected outcome-predictive units and mean auROC curve (mean ± SEM) per rat during the set-shifting task; instrumental poke (0 – 0.5 s), comparing early and late trials of the sets

Rat (units)	Instrumental poke	
	Early trials	Late trials
Ad31 (7)	0.56 ± 0.03	0.59 ± 0.05
Ad32 (2)	0.53 ± 0.03	0.59 ± 0.01
Ad33 (8)	0.53 ± 0.02	0.64 ± 0.04
Ad34 (2)	0.60 ± 0.07	0.59 ± 0.08
Ad35 (1)	0.51	0.65
Ad36 (2)	0.61 ± 0.06	0.69 ± 0.06
Ad38 (1)	0.61	0.70
Ad39 (1)	0.47	0.70

dar, 2011; Donahue et al., 2013), we found that, during performance of the set-shifting task, dmPFC neurons encoded information about current outcome, i.e., signaling reward delivery or omission, as well as impending outcomes.

We next addressed whether outcome encoding is related to behavioral shifts due to changes in action–outcome contingencies. This was addressed in two ways. First, we compared outcome-related activity in the dmPFC during the early trials of the set, after the rule was changed, to the late trials of the set, before animals changed their behavior (reached criterion). PFC neurons predicted outcomes more accurately during the late trials of the set than during the early trials of the set, suggesting that they anticipate the shift. Second, we recorded during a control task—No-Rule condition—in which actions resulted in random delivery of rewards. Behavioral performance adapted to the set rule during the Rule condition but did not change during the No-Rule condition. The proportion of neurons that anticipated future trial outcomes was significantly higher during the Rule condition, when animals adapted their behavior to the rules, than during the No-Rule condition. Moreover, according to ROC analysis, the discrimination of neurons to predict trial outcomes was significantly more accurate during the Rule compared with

Table 3. Mean auROC curve (mean ± SEM) for all the selected units encoding each task variable during the set-shifting task; instrumental poke (0 – 0.5 s) and outcome (1.25–1.75 s) time events, comparing early and late trials of the sets

	Instrumental poke		Outcome	
	Early trials	Late trials	Early trials	Late trials
Current outcome (t)	0.55 ± 0.01	0.63 ± 0.02[a]	0.65 ± 0.01	0.64 ± 0.01
Most recent previous outcome (t-1)	0.58 ± 0.02	0.57 ± 0.02	0.56 ± 0.02	0.58 ± 0.02
Second most recent previous outcome (t-2)	0.57 ± 0.01	0.56 ± 0.03	0.58 ± 0.02	0.49 ± 0.04
Set rule	0.57 ± 0.02	0.58 ± 0.02	0.60 ± 0.02	0.57 ± 0.02
Right	0.58 ± 0.03	0.55 ± 0.01	0.60 ± 0.02	0.59 ± 0.02
Center	0.57 ± 0.02	0.55 ± 0.02	0.56 ± 0.02	0.54 ± 0.01
Left	0.57 ± 0.01	0.58 ± 0.01	0.58 ± 0.01	0.58 ± 0.01

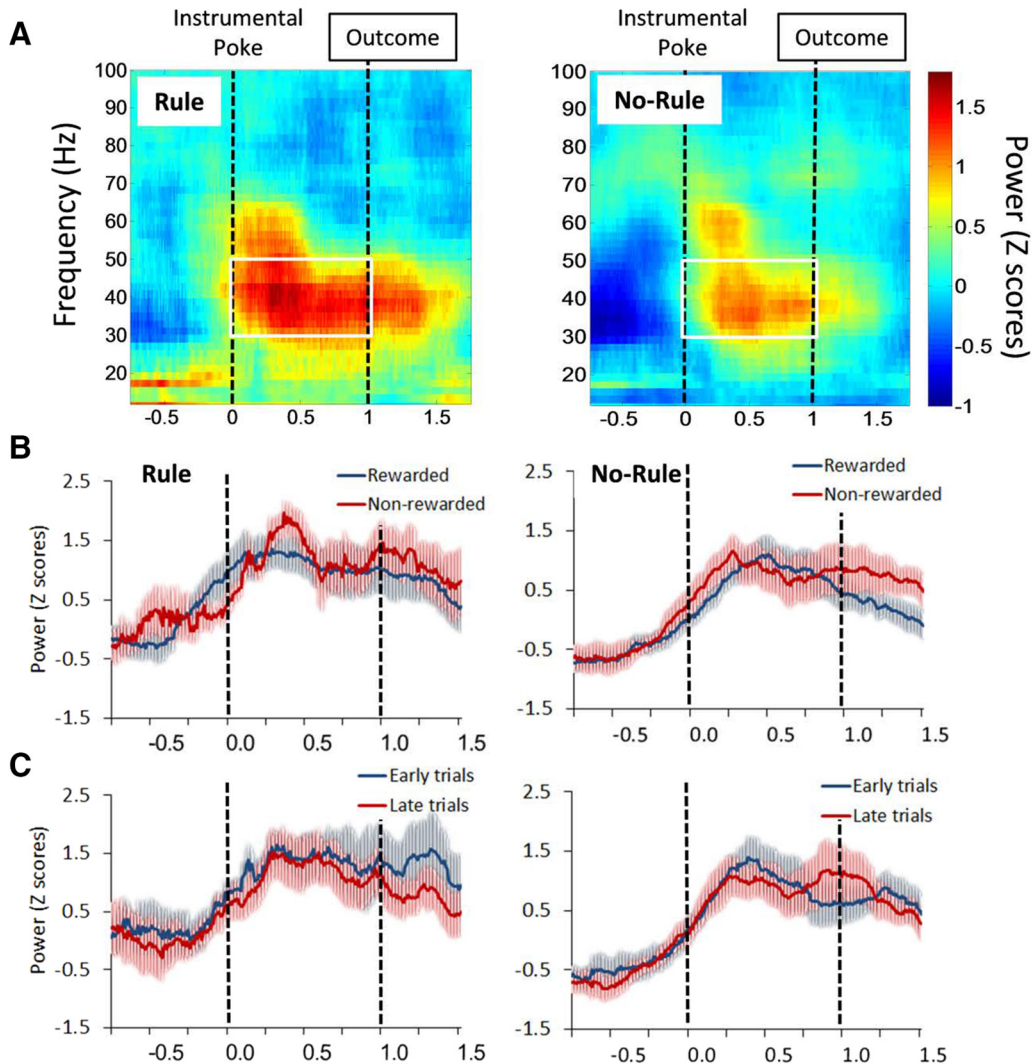[a]$p < 0.001$, paired Student's t test.

Figure 6. Changes in LFP oscillation power during set-shifting performance. **A**, Heat plots represent the time course of normalized LFP power locked to the instrumental poke during both outcome conditions, Rule and No-Rule. **B**, **C**, Plots represent the time course of normalized LFP power (averaged across 30 –50 Hz frequency bands; area inside the white square) during both conditions in rewarded compared with nonrewarded trials (**B**) and early trials after the shift compared with late trials before criterion (**C**). Low-γ power increases between the action and the outcome in both outcome conditions but it is not significantly modulated by the outcome of current trials or trial order. Shaded area represents the SEM. Dashed vertical lines indicate the time of the instrumental poke and outcome events.

the No-Rule condition. Importantly, outcome-responsive activity did not differ between rewarded and unrewarded conditions. Collectively, these results indicate that outcome-predictive, but not outcome-responsive, activity in the dmPFC is associated with animals updating their response strategy according to changes in set rules.

How is the predictive activity generated in the dmPFC? Previous studies have shown that neurons in another PFC region, the

orbitofrontal cortex, encode upcoming rewards (or their probability) when these rewards are associated with external cues predicting them (Watanabe, 1996; Stalnaker et al., 2007; Wallis, 2007; Simmons and Richmond, 2008; van Duuren et al., 2009). This is unlikely to be the case here because in our task animals did not have external cues associated with trial outcomes (reward or omission of reward) and were exposed to all competing options

in every trial. Instead, outcome-predictive activity in the PFC was more likely related to internally generated information guided by the memory of recent behavior choices and resulting outcomes. These would involve such constructs as working memory and/or accumulated confidence (reduced uncertainty) for action selection (Kepecs et al., 2008; Mainen and Kepecs, 2009; Karlsson et al., 2012). Along the same lines, the PFC can encode information about previous outcomes and goals not directly related to the performance of the task, which may ultimately lead to the implementation of new behavioral strategies (Genovesio et al., 2014; Donahue and Lee, 2015; Schuck et al., 2015). In this context, it can be hypothesized that the absence of anticipatory activity in the PFC promotes searching behavior (exploration), whereas generation of this activity promotes repetition of successful behavior (exploitation), thus facilitating behavioral shifts. The outcome prediction-related activity would, therefore, be part of a more complex prefrontal circuit involved in the top-down regulation of behavioral flexibility, in which other areas of the brain, such as the striatum, hippocampus, thalamus, brainstem (dopamine and noradrenaline projecting neurons), and lateral habenula are also involved (Aston-Jones and Cohen, 2005; Floresco et al., 2006, 2009; Ragozzino, 2007; Bissonette et al., 2013; Janitzky et al., 2015; Kawai et al., 2015).

Unlike outcome-predictive activity, the outcome-responsive activity in the dmPFC did not differentiate between Rule and No-Rule conditions or early/late trials of the set. In fact, there were no differences on how outcome-responsive neurons discriminate reward delivery during either outcome conditions or trial order. These results suggest that the outcome-responsive signal is not associated with rule changes. This may be expected with our set-shifting task because a single trial (rewarded or nonrewarded) did not necessarily mean that the rule had changed.

γ-Band oscillations are not related to behavioral updating
An increase in LFP activity was observed selectively in the γ-band range between action execution and outcome delivery. This signal, however, did not differentiate between the type of expected outcomes (reward or no reward) nor did it differ based on the currently valid rule. Previous studies in the PFCs of monkeys have shown changes in LFP oscillations associated with abstract rule representation (Buschman et al., 2012) and feedback signals (Quilodran et al., 2008; Rothé et al., 2011) during behavioral adaptation. A recent mouse study also suggested that γ rhythms in the PFC are related to cognitive inflexibility (Cho et al., 2015). Our data, however, indicate that γ oscillations are not directly related to set-shifting behavior but reflect expectation of an outcome after action execution.

γ Oscillations in the PFC are associated with attention and stimulus detection (Benchenane et al., 2011). Performing the set-shifting task on a trial-by-trial basis requires maintaining the association between choices and outcomes during both rewarded and unrewarded trials. The increase in medial PFC γ oscillations found in the present study may be critical for local circuit processing of the selected actions linking them to their outcomes across multiple trials. In line with this possibility, previous studies have shown increases of γ oscillations (from 45 to 60 Hz) in the PFC associated with instrumental learning (Yu et al., 2012) and independent of correct and incorrect choices during the learning of a discrimination task (van Wingerden et al., 2010).

In conclusion, dmPFC ensembles use outcome prediction as a signal to promote shifting to a behavior that is currently successful at generating reward. This signal is independent of γ oscillatory activity, which selectively represents outcome expectation independent of previous or currently valid action–outcome contingencies.

References

Asaad WF, Eskandar EN (2011) Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. J Neurosci 31:17772–17787. CrossRef Medline

Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci 28:403–450. CrossRef Medline

Benchenane K, Tiesinga PH, Battaglia FP (2011) Oscillations in the prefrontal cortex: a gateway to memory and attention. Curr Opin Neurobiol 21:475–485. CrossRef Medline

Birrell JM, Brown VJ (2000) Medial frontal cortex mediates perceptual attentional set shifting in the rat. J Neurosci 20:4320–4324. Medline

Bissonette GB, Roesch MR (2015) Neural correlates of rules and conflict in medial prefrontal cortex during decision and feedback epochs. Front Behav Neurosci 9:266. CrossRef Medline

Bissonette GB, Powell EM, Roesch MR (2013) Neural structures underlying set-shifting: roles of medial prefrontal cortex and anterior cingulate cortex. Behav Brain Res 250:91–101. CrossRef Medline

Buschman TJ, Denovellis EL, Diogo C, Bullock D, Miller EK (2012) Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. Neuron 76:838–846. CrossRef Medline

Cho KK, Hoch R, Lee AT, Patel T, Rubenstein JL, Sohal VS (2015) Gamma rhythms link prefrontal interneuron dysfunction with cognitive inflexibility in dlx5/6+/− mice. Neuron 85:1332–1343. CrossRef Medline

Dalton GL, Ma LM, Phillips AG, Floresco SB (2011) Blockade of NMDA GluN2B receptors selectively impairs behavioral flexibility but not initial discrimination learning. Psychopharmacology (Berl) 216:525–535. CrossRef Medline

Darrah JM, Stefani MR, Moghaddam B (2008) Interaction of N-methyl-D-aspartate and group 5 metabotropic glutamate receptors on behavioral flexibility using a novel operant set-shift paradigm. Behav Pharmacol 19:225–234. CrossRef Medline

Dias R, Robbins TW, Roberts AC (1996) Dissociation in prefrontal cortex of affective and attentional shifts. Nature 380:69–72. CrossRef Medline

Donahue CH, Lee D (2015) Dynamic routing of task-relevant signals for decision making in dorsolateral prefrontal cortex. Nat Neurosci 18:295–301. CrossRef Medline

Donahue CH, Seo H, Lee D (2013) Cortical signals for rewarded actions and strategic exploration. Neuron 80:223–234. CrossRef Medline

Durstewitz D, Vittoz NM, Floresco SB, Seamans JK (2010) Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. Neuron 66:438–448. CrossRef Medline

Floresco SB, Magyar O, Ghods-Sharifi S, Vexelman C, Tse MT (2006) Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. Neuropsychopharmacology 31:297–309. CrossRef Medline

Floresco SB, Block AE, Tse MT (2008) Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. Behav Brain Res 190:85–96. CrossRef Medline

Floresco SB, Zhang Y, Enomoto T (2009) Neural circuits subserving behavioral flexibility and their relevance to schizophrenia. Behav Brain Res 204:396–409. CrossRef Medline

Genovesio A, Tsujimoto S, Navarra G, Falcone R, Wise SP (2014) Autonomous encoding of irrelevant goals and outcomes by prefrontal cortex neurons. J Neurosci 34:1970–1978. CrossRef Medline

Histed MH, Pasupathy A, Miller EK (2009) Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. Neuron 63:244–253. CrossRef Medline

Janitzky K, Lippert MT, Engelhorn A, Tegtmeier J, Goldschmidt J, Heinze HJ, Ohl FW (2015) Optogenetic silencing of locus coeruleus activity in mice impairs cognitive flexibility in an attentional set-shifting task. Front Behav Neurosci 9:286. CrossRef Medline

Karlsson MP, Tervo DG, Karpova AY (2012) Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. Science 338:135–139. CrossRef Medline

Kawai T, Yamada H, Sato N, Takada M, Matsumoto M (2015) Roles of the lateral habenula and anterior cingulate cortex in negative outcome mon-

itoring and behavioral adjustment in nonhuman primates. Neuron 88: 792–804. CrossRef Medline

Kepecs A, Uchida N, Zariwala HA, Mainen ZF (2008) Neural correlates, computation and behavioural impact of decision confidence. Nature 455: 227–231. CrossRef Medline

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

Mansouri FA, Matsumoto K, Tanaka K (2006) Prefrontal cell activities related to monkeys' success and failure in adapting to rule changes in a Wisconsin card sorting test analog. J Neurosci 26:2745–2756. CrossRef Medline

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167–202. CrossRef Medline

Milner B (1963) Effects of different brain lesions on card sorting: the role of the frontal lobes. Arch Neurol 9:90–100. CrossRef

Mitra PP, Pesaran B (1999) Analysis of dynamic brain imaging data. Biophys J 76:691–708. CrossRef Medline

Narayanan NS, Laubach M (2008) Neuronal correlates of post-error slowing in the rat dorsomedial prefrontal cortex. J Neurophysiol 100:520–525. CrossRef Medline

Park J, Moghaddam B (2017) Impact of anxiety on prefrontal cortex encoding of cognitive flexibility. Neuroscience 345:193–202. CrossRef Medline

Park J, Wood J, Bondi C, Del Arco A, Moghaddam B (2016) Anxiety evokes hypofrontality and disrupts rule-relevant encoding by dorsomedial prefrontal cortex neurons. J Neurosci 36:3322–3335. CrossRef Medline

Powell NJ, Redish AD (2016) Representational changes of latent strategies in rat medial prefrontal cortex precede changes in behaviour. Nat Commun 7:12830. CrossRef Medline

Quilodran R, RothéM, Procyk E (2008) Behavioral shifts and action valuation in the anterior cingulate cortex. Neuron 57:314–325. CrossRef Medline

Ragozzino ME (2007) The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. Ann N Y Acad Sci 1121:355–375. CrossRef Medline

Ragozzino ME, Detrick S, Kesner RP (1999) Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. J Neurosci 19:4585–4594. Medline

Rich EL, Shapiro M (2009) Rat prefrontal cortical neurons selectively code strategy switches. J Neurosci 29:7208–7219. CrossRef Medline

Rodgers CC, DeWeese MR (2014) Neural correlates of task switching in prefrontal cortex and primary auditory cortex in a novel stimulus selection task for rodents. Neuron 82:1157–1170. CrossRef Medline

Rothé M, Quilodran R, Sallet J, Procyk E (2011) Coordination of high gamma activity in anterior cingulate and lateral prefrontal cortical areas during adaptation. J Neurosci 31:11110–11117. CrossRef Medline

Schuck NW, Gaschler R, Wenke D, Heinzle J, Frensch PA, Haynes JD, Rever-

beri C (2015) Medial prefrontal cortex predicts internally driven strategy shifts. Neuron 86:331–340. CrossRef Medline

Seo H, Barraclough DJ, Lee D (2007) Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. Cereb Cortex 17:110–117.

Simmons JM, Richmond BJ (2008) Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. Cereb Cortex 18:93–103. CrossRef Medline

Stalnaker TA, Roesch MR, Calu DJ, Burke KA, Singh T, Schoenbaum G (2007) Neural correlates of inflexible behavior in the orbitofrontal-amygdalar circuit after cocaine exposure. Ann N Y Acad Sci 1121:598–609. CrossRef Medline

Stefani MR, Moghaddam B (2006) Rule learning and reward contingency are associated with dissociable patterns of dopamine activation in the rat prefrontal cortex, nucleus accumbens, and dorsal striatum. J Neurosci 26:8810–8818. CrossRef Medline

Stefani MR, Groth K, Moghaddam B (2003) Glutamate receptors in the rat medial prefrontal cortex regulate set-shifting ability. Behav Neurosci 117: 728–737. CrossRef Medline

Sturman DA, Moghaddam B (2011) Reduced neuronal inhibition and coordination of adolescent prefrontal cortex during motivated behavior. J Neurosci 31:1471–1478. CrossRef Medline

Sul JH, Kim H, Huh N, Lee D, Jung MW (2010) Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. Neuron 66:449–460. CrossRef Medline

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. CrossRef Medline

van Wingerden M, Vinck M, Lankelma JV, Pennartz CM (2010) Learning-associated gamma-band phase-locking of action–outcome selective neurons in orbitofrontal cortex. J Neurosci 30:10025–10038. CrossRef Medline

Wallis JD (2007) Neuronal mechanisms in prefrontal cortex underlying adaptive choice behavior. Ann N Y Acad Sci 1121:447–460. CrossRef Medline

Wallis JD, Kennerley SW (2010) Heterogeneous reward signals in prefrontal cortex. Curr Opin Neurobiol 20:191–198. CrossRef Medline

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. CrossRef Medline

Watanabe M (1996) Reward expectancy in primate prefrontal neurons. Nature 382:629–632. CrossRef Medline

Wilmsmeier A, Ohrmann P, Suslow T, Siegmund A, Koelkebeck K, Rothermundt M, Kugel H, Arolt V, Bauer J, Pedersen A (2010) Neural correlates of set-shifting: decomposing executive functions in schizophrenia. J Psychiatry Neurosci 35:321–329. CrossRef Medline

Yu C, Fan D, Lopez A, Yin HH (2012) Dynamic changes in single unit activity and gamma oscillations in a thalamocortical circuit during rapid instrumental learning. PLoS One 7:e50578. CrossRef Medline