



Published in final edited form as:

Nat Neurosci. 2015 February ; 18(2): 295–301. doi:10.1038/nn.3918.

Dynamic Routing of Task-relevant Signals for Decision Making in Dorsolateral Prefrontal Cortex

Christopher H. Donahue^{1,4} and Daeyeol Lee^{1,2,3}

¹Department of Neurobiology, Yale University School of Medicine, New Haven, CT 06510, USA

²Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, CT 06510, USA

³Department of Psychology, Yale University, New Haven, CT 06520, USA

Abstract

Neurons in the dorsolateral prefrontal cortex (DLPFC) encode a diverse array of sensory and mnemonic signals, but little is known about how this information is dynamically routed during decision making. We analyzed the neuronal activity in the DLPFC of monkeys performing a probabilistic reversal task where information about the probability and magnitude of reward was provided by the target color and numerical cues, respectively. The location of the target of a given color was randomized across trials, and therefore was not relevant for subsequent choices. DLPFC neurons encoded signals related to both task-relevant and irrelevant features, and task-relevant mnemonic signals were encoded congruently with choice signals. Furthermore, only the task-relevant signals related to previous events were more robustly encoded following rewarded outcomes. Thus, multiple types of neural signals are flexibly routed in the DLPFC so as to favor actions that maximize reward.

INTRODUCTION

While many studies emphasized the role of dorsolateral prefrontal cortex (DLPFC) in working memory^{1–6}, neurons in DLPFC also encode a highly heterogeneous set of signals, reflecting the stimulus features or cognitive processes related to the particular tasks performed by the animals^{7,8}. DLPFC neurons can encode both the attended and remembered locations^{9,10}, as well as experienced or expected outcomes^{11–13}, temporally discounted values^{14,15}, and other properties of rewards, such as their probabilities, magnitudes, and the effort necessary to acquire them¹⁶. While the heterogeneity of signals present in the DLPFC suggests that it may play an important role in integrating different sources of information to aid in decision-making, this region has also been shown to encode information that is irrelevant for the tasks that animals are trained to perform^{17,18}. This poses a problem as to

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

⁴Present Address: Gladstone Institute of Neurological Disease, San Francisco, CA 94158, USA

AUTHOR CONTRIBUTIONS

C.H.D and D.L. designed the experiments and wrote the manuscript. C.H.D. performed the experiment and analyzed the data.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

how signals that are most relevant in a particular context might gain priority over other less relevant signals.

We trained monkeys on a probabilistic reversal task where they had to estimate the reward probabilities based on their recent experience, and combine them with magnitude information that varied independently across trials. Reward probabilities were associated with the colors of the targets, while their spatial positions were randomized across trials. Therefore, the animals had to associate previous outcomes with the colors of chosen targets and ignore their spatial locations. Indeed, animals displayed a strong tendency to choose the target colors rewarded previously. We also found that rewards in the previous trial enhanced the encoding of task-relevant information. In addition, only the information about the task-relevant events in the previous trial was encoded congruently with the choice signals in the DLPFC. In contrast, reward did not influence the encoding of task-irrelevant signals. Thus, task-relevant and irrelevant signals in the DLPFC might be differentially integrated such that task-relevant information can bias behavior towards the selection of actions that maximize reward.

RESULTS

Behavioral effects of rewarded and unrewarded outcomes

The animals were trained on a probabilistic reversal task in which the magnitude of reward available from each target was independently varied across trials (Fig. 1a; see Online Methods). After achieving fixation on the central target, the animals were presented with a red target and a green target (target period). One of the target colors was associated with a high reward probability (80%) and the other was associated with a low reward probability (20%). After an additional delay (0.5 s), small yellow tokens appeared around each target indicating the magnitude of available reward (magnitude period). The animals were then required to shift their gaze towards one of the targets. The reward probabilities for red and green targets were fixed within a block and underwent un-signaled reversals so that the animals had to estimate them through experience.

The task required the animals to associate the colors of previously chosen targets (relevant information) with their outcomes and ignore previously chosen locations (irrelevant information), since target locations were randomized across trials. To examine how the animals were influenced by task-relevant versus task-irrelevant information, we examined the win-stay and lose-stay behavior in a color and spatial reference frame (Fig. 1b). We found that in 117/118 (99.2%) of the sessions there was a significant difference between win-stay and lose-switch behavior in terms of chosen colors. In contrast, none of the sessions showed a significant influence of previously chosen location. These results indicate that the animals correctly used task-relevant information from previous trials to make their choices.

To determine how the animals integrated all of the information required to make their decisions, we compared a series of reinforcement learning models¹⁹ that addressed three important issues (Supplementary Table 1; see Online Methods). First, we examined how rewarded and unrewarded outcomes differentially affected learning on subsequent trials.

Second, we examined whether the animals used information about the anti-correlated structure of reward probability reversals in this task. Finally, we examined whether the animals combined probability and magnitude information additively or multiplicatively. To systematically address these issues, we explored all possible models where these three factors were independently manipulated. The result was consistent across both animals. We found that the models with 2 learning rates in which target values were updated reciprocally (coupled) were superior in both monkeys. In both monkeys, the learning rates associated with rewarded trials was significantly larger than the learning rates associated with unrewarded trials (monkey O, $\alpha = 0.65$ and 0.05 ; monkey U, $\alpha = 0.64$ and 0.04 , for rewarded and unrewarded outcomes, respectively). These results suggest that the animals updated their value functions more after rewarded trials than when after unrewarded trials, and that this updating took into account the anti-correlated structure of reward probability reversals. We also found that the additive model (P+M) provided a better fit to the data than the model in which they were combined multiplicatively (EV) in every session for both monkeys. Thus, both animals might in fact have relied more on a suboptimal additive strategy rather than calculating expected values. When the performance of this additive model was evaluated using a leave-one-session-out cross-validation, the animal's choice could be predicted correctly in $91.0 \pm 2.3\%$ of the trials. For comparison, the clairvoyant model in which the animal always chooses the target with the higher expected value correctly predicted the animal's choices in $79.8 \pm 2.8\%$ of the trials.

DLPFC activity related to multiple features

Single-unit activity was recorded from a total of 226 neurons in the DLPFC of both monkeys (77 neurons in monkey O; 149 neurons in monkey U). We used a multiple linear regression model (see Online Methods) to characterize neural signals related to various events in current and previous trials and to identify how task-relevant (i.e. previously chosen color, previous outcome) and task-irrelevant features (previously chosen target location) are encoded. Overall, significant proportions of DLPFC neurons encoded both task-relevant and task-irrelevant features related to events from the previous trial (Fig. 2a). For example, during the interval surrounding the target onset (-250 to 250 ms relative to target onset), the proportion of neurons encoding information about the outcome in the previous trial (33.2%) and the previously chosen color (10.6%) were significantly greater than expected from chance (binomial test, $p < 10^{-3}$). Task-irrelevant information about the previously chosen location was even more strongly represented in the population than information about previously chosen color (28.3% vs. 10.6% ; 2 proportion z-test, $p < 10^{-5}$). Three examples of neurons encoding each of these features are shown in Figs. 2b–d.

Neurons in DLPFC also encoded multiple types of information presented in the current trial (Fig. 3a). For example, in the epoch following target onset (250 to 750 ms following target onset), 29.2% of neurons encoded the relative positions of the red and green targets. The neuron shown in Fig 3b increased its activity more when the red target appeared on the left than it appeared on the right (t-test, $p < 10^{-7}$). In the epoch following the onset of magnitude information (750 – 1250 ms following target onset), a significant fraction of neurons also encoded the animal's chosen color (15.0%), chosen target location (34.1%), and reward

magnitudes (31.9%). Single neuron examples for each of these signals are shown in Figs. 3c–e.

Next, we examined how signals related to the previous trial's events were integrated with information from the current trial. In particular, we tested how various task-related signals were combined so as to allow the animal to identify the high-value target location (HVL) in the current trial, namely the location of the target with the higher reward probability. To identify the HVL, the animals need to integrate three different pieces of information: the previously chosen color, the previous outcome, and the configuration of target positions on the current trial. During the target period, we found that 18.1% (binomial test, $p < 10^{-12}$) of the neurons encoded a 3-way interaction between these three factors, and therefore provided the specific information about the HVL (Fig. 4a). For example, the neuron in Fig. 4b increased its firing rate when the animal was rewarded in the previous trial and the previously chosen color was currently displayed on the right. In contrast, its firing rate decreased if the animal was rewarded in the previous trial and the previously chosen color was currently on the left. During the magnitude period, activity was significantly influenced by the HVL in 29 neurons (12.8%), and 16 neurons among them (55.2%) showed significant effects of both HVL and choice, suggesting that HVL and choice signals might co-exist in the DLPFC.

Neurons that integrated individual elements of the HVL signal were also present. For example, prior to the onset of magnitude cues (–250 to 250 ms relative to target onset), 10.2% of the neurons integrated signals related to the previously chosen color and previous outcome. The single neuron example shown in Fig. 4c reached its highest firing rate when the animal previously chose the red target and was rewarded. The integration and maintenance of such signals during the inter-trial interval might provide the animals with the information needed to determine the HVL once the target location information is presented. Since we defined the HVL as the 3-way interaction among the previously chosen color, previous outcome, and current target configurations, we also used a decoding analysis to test whether HVL signals could be decoded from a population of neurons with multiple main effects of these 3 variables (3 groups with different pairs of main effects and another group with all three main effects) or their 2-way interactions, but without the significant 3-way interaction (see Online Methods). When corrected for multiple comparisons, none of these 7 groups showed significant encoding of HVL signals, whereas as expected, HVL signals were reliably decoded from the neurons with significant 3-way interactions (t-test, $p < 10^{-5}$).

We found that information about the previous trial's chosen location and outcome were combined to indicate the previously rewarded location (PRL), even though this information was irrelevant in this task. In the epoch following target onset, 20.8% of the neurons encoded the PRL. The example neuron in Fig. 4d increased its activity more when the animal previously chose the right location and was rewarded. Thus, while a significant population of neurons in the DLPFC combined information at the single neuron level that was critical for performing this task (HVL), many DLPFC neurons still encoded conjunctions that were not useful to the animals (PRL).

Congruency of task-relevant information and choice signals

Given that task-relevant and irrelevant variables were both encoded in the DLPFC, we tested how each of these signals might be combined with the signals related to the animal's upcoming choice. We found that the task-relevant information (HVL) tended to be encoded in the same population of neurons that later encoded the animal's upcoming choice. Neurons encoding HVL during the target period were more likely to encode the animal's choice later during the magnitude period (24/41 neurons, 58.5%) than the neurons without HVL signals (53/185 neurons, 28.6%; χ^2 test, $p < 10^{-3}$). We next compared the regression coefficients for HVL defined in the epoch following target onset and the regression coefficients for upcoming choice defined in the epoch following the onset of magnitude cues. These regression coefficients were strongly correlated ($r = 0.55$, $p < 10^{-18}$), indicating that individual neurons encoded information about HVL and upcoming choice consistently (Fig. 5a). Namely, neurons that increased their firing rates when the HVL was displayed on the right also tended to increase their firing rates later in the trial when the animal made a rightward choice.

We also found that the DLPFC neurons encoding task-irrelevant information about the previously chosen target location (PRL) tended to encode the upcoming choice later in the trial. Neurons encoding PRL during the target period were more likely to encode the animal's upcoming choice during the magnitude period (24/47 neurons, 51.1%) than those without PRL signals (53 neurons, 29.6%; χ^2 test, $p < 0.01$). However, there was no significant correlation between regression coefficients for PRL and upcoming choice (Fig. 5b; $r = 0.01$; $p = 0.93$), and this was significantly smaller than the correlation for HVL and choice signals ($p < 10^{-10}$). These results revealed an important difference in the way that task-relevant and irrelevant information are encoded. Although both types of signals are present in the DLPFC, only task-relevant information was encoded congruently by the neurons encoding the animal's choice later. However, not all task-relevant information was treated in the same manner. We found that the information about the difference in reward magnitude was not encoded consistently by the choice-encoding neurons. During the magnitude period, the regression coefficients for the magnitude difference and choice were not significantly correlated ($r = 0.02$, $p = 0.78$). The regression coefficients for the magnitude difference was not correlated with those for the HVL estimated during the target period, either ($r = -0.02$, $p = 0.60$).

To further investigate how these signals were integrated in single neurons, we performed principal components analysis (PCA) on the regression coefficients (see Online Methods), and examined the weights assigned to the regression coefficients for the animal's upcoming choice, HVL, and PRL. If these signals mix independently, we would expect the first principal component (PC1) to capture variability in the population activity related to only the variable associated with the largest changes in firing rate. We found that PC1 captured shared variance related to both choice and HVL (Fig. 5c). The maximum value of the weights related to HVL was significantly higher than observed after shuffling the relationship among the regression coefficients ($p < 10^{-4}$), but this was not true for the weights related to the task-irrelevant PRL ($p = 0.11$). We repeated the same analysis on the absolute value of regression coefficients, and found that weights for both HVL and PRL were

significantly above the chance level (Fig. 5d; $p < 10^{-4}$). Thus, both HVL and PRL signals tended to combine in the same neurons that eventually signal the animal's choice, but only task-relevant HVL signals and choice were encoded in the same spatial reference frame.

Effects of reward on DLPFC activity

Since we observed a strong behavioral asymmetry following rewarded versus unrewarded trials, we tested whether the outcome of the previous trial also asymmetrically influenced the way in which different task-related variables are encoded in the DLPFC. Indeed, we found that DLPFC neurons tended to encode the current location of previously chosen color more robustly following a reward. For example, the neuron shown in Figure 6a increased its firing rate during the target period, when the previously chosen and rewarded color was presented on the right. In contrast, it decreased its firing rate when the previously chosen and rewarded color was presented on the left. However, when the previous trial was unrewarded, the position of previously chosen color did not strongly affect the activity of this neuron. These effects were present regardless of whether the animal eventually chose the leftward or rightward target. Thus, the information pertaining to the location of the previously chosen color was more robustly encoded when the previous trial was rewarded (see also Figure 4b). Furthermore, this neuron also increased its firing rate when the animal chose the rightward target in the current trial, suggesting that the HVL and choice were encoded in the same frame of reference as described above.

To investigate whether these effects were consistent across the population, we applied a linear discriminant analysis to decode the current location of the color chosen in the last trial separately according to whether this was rewarded or not (Fig. 6b). During the target period, we found that this information was encoded more robustly following a reward in both monkeys (paired t-test, $p < 10^{-12}$). The position of the previously chosen color could not be reliably decoded after unrewarded trials (paired t-test, $p = 0.52$). We also found that the average decoding accuracy for the color of the target chosen in the previous rewarded trial was significantly higher (0.52) than that in the unrewarded trial (0.50; paired t-test, $p < 0.05$). We next tested whether there were any differences in the ability to decode task-irrelevant information, namely previously chosen target location, following rewarded and unrewarded outcomes, and found that there was no significant difference (Fig. 6b; paired t-test, $p = 0.82$). In addition, we found a significant interaction between information type (location of previously chosen color vs. previously chosen location) and previous outcome (2-way ANOVA: $p = 0.01$).

We also found that there was no difference in the way the neurons encoded upcoming choice information as a function of the previous outcome (paired t-test, $p = 0.69$; Fig. 6b). This confirmed that retrospective but not prospective task information was encoded more robustly following a rewarded outcome. Thus, task-relevant information (position of the color chosen in the previous trial) was encoded in a specific manner, becoming more robust following rewarded outcomes, while task-irrelevant information (previously chosen locations) was not influenced by the previous outcome. These results mirror the asymmetry in the behavior reported above, indicating that rewarded outcomes encouraged the animals to choose the

same color as in the previous trial and also enhanced the encoding of the corresponding information in the DLPFC.

Anatomical distributions

Neurons encoding HVL, choice, and PRL were broadly distributed throughout the DLPFC explored in this study (Supplementary Figure 1). We performed 1-way MANOVA to see whether neurons encoding these variables were differentially distributed in either the anterior-posterior or dorsal-ventral axis. While neither HVL nor PRL-coding neurons displayed a significant tendency to be located in any particular position (1-way MANOVA: $p=0.19$ and $p=0.68$, respectively), we found that neurons with choice-related activity tended to be located slightly more posteriorly ($p=0.03$). Neurons that conjunctively encoded both choice and HVL tended to be located more posteriorly and dorsally ($p<0.01$) while neurons that conjunctively encoded choice and PRL lacked any significant bias in their spatial distribution ($p=0.40$).

DISCUSSION

While neurons in the DLPFC encoded a diverse collection of signals related to both relevant and irrelevant aspects of the task, we found clear differences in how this information was integrated at the single neuron level. DLPFC neurons tended to encode information about the current position of the previously rewarded target color and the animal's upcoming choice congruently. In contrast, irrelevant information related to previously rewarded locations were encoded largely independently of upcoming choice locations. In addition, consistent with the animal's behavior, we found that only the signals relevant to the task were enhanced following a rewarded outcome. Thus, DLPFC might be a critical node in the network of brain areas that allows information from a variety of sources to get dynamically gated and transmitted to downstream neurons responsible for controlling the animal's behavior.

Behaviors related to the previous outcome

Although it is optimal to combine reward probabilities and magnitudes multiplicatively for calculating expected values, we found that both animals tested in this study used a heuristic additive strategy. While previous work has examined potentially different anatomical substrates for processing reward probability and magnitude information^{20–26}, relatively little is known about how these signals are combined to calculate expected value. While our results indicate that reward probability and magnitude signals are combined in the DLPFC, the animals did not use this information optimally. It is possible that the task design in this study did not provide sufficiently strong incentive to calculate the expected value of reward accurately, especially given that the animals also had to estimate reward probabilities.

We also found that rewarded outcomes had a much greater influence on the animals' subsequent behavior than unrewarded outcomes. This asymmetry has been reported across a wide range of behavioral paradigms including a competitive game^{13,27} and a paired associate task²⁸ in primates, as well as in a free choice task using probabilistic outcomes in

rodents²⁹. This suggests that rewarded and unrewarded outcomes might contribute differently to the neural process of updating the animal's decision-making strategies.

Lastly, we found that animals tended to update chosen and unchosen reward probabilities reciprocally in our task. This has potential implications for model-free versus model-based reinforcement learning^{19,30–32}. In simple model-free reinforcement learning, animals update only the action values associated with chosen actions. In our task, the objective reward probabilities were anti-correlated for the two targets. Exploiting knowledge of this structure would allow the animal to improve the speed and accuracy of learning by reciprocally adjusting the value of both targets. Different striatal and cortical circuits have been associated with model-free and model-based reinforcement learning^{30–35}. In addition, the prefrontal cortex might also play an important role in arbitrating the competition between multiple learning algorithms^{36,37}. However, how prefrontal activity related to previous outcomes might be utilized to adjust the estimates of multiple reward probabilities simultaneously is not known, and should be investigated further.

Integration of task-relevant signals in DLPFC

Decision-making in natural environments often requires animals to integrate various types of information over multiple timescales^{38–40}. Indeed, persistent activity often carries information about actions and outcomes from previous trials in dynamic decision-making tasks^{13,39}, as well as in tasks that require the use of abstract response strategies depending on events in the previous trial^{41,42}. Neurons in the DLPFC are also capable of integrating information across multiple sensory modalities⁴³. In the present study, we found that many DLPFC neurons encode signals related to task-irrelevant events in previous trials. Similarly, recent studies have shown that some DLPFC neurons encode task-irrelevant features of sensory stimuli presented previously¹⁸. Such encoding of task-irrelevant signals may be important for the behavioral flexibility often attributed to prefrontal cortex, since stimulus-outcome or action-outcome contingencies can change unexpectedly in a dynamic environment. In particular, residual coding of task-irrelevant information in the DLPFC might facilitate the detection of changes in the animal's environment. Consistent with this possibility, DLPFC neurons often encode task-irrelevant features during perceptual decision making^{17,44}. For example, in a recent study⁴⁴, monkeys were required to integrate noisy sensory evidence related to the color or motion direction of stimuli during a perceptual decision-making task. The relevant sensory evidence was selectively integrated in the prefrontal cortex along the axis in the state space corresponding to the animal's choice. In addition, the degree to which the neural activity in the state space deviates from the choice axis increased with the strength of sensory signals. The results in the present study suggest that during value-based decision-making task, task-irrelevant information from memory, such as the animal's choices and their outcomes in preceding trials, might be encoded similarly in the DLPFC. In particular, DLPFC signals related to task-relevant information were integrated better with the choice signals than signals related to task-irrelevant information. Therefore, similar mechanisms might exist in the prefrontal cortex for gating and selection of task-relevant information regardless of whether such information originates from sensory stimuli or previous experience. Coding of task-relevant signals in the DLPFC might change adaptively during learning to promote behaviors leading to more favorable

outcomes in a particular context. Such flexible coding might not be unique to the DLPFC, since the directional tuning of neurons in the supplemental eye field (SEF) have been previously shown to shift while animals learn novel conditional visuomotor relationships^{45–47}.

We also found that DLPFC signals related to the task relevant information, but not those related to task-irrelevant information, was selectively enhanced following a positive outcome. This tendency has been reported across a diverse set of brain regions, including DLPFC^{27,28}, SEF and lateral intraparietal area²⁷, striatum²⁹, and hippocampus⁴⁸. Interestingly, the nature of the behavioral task may dictate precisely what types of signals are enhanced. For example, we have previously shown that signals related to past actions but not upcoming actions were encoded more robustly following rewards in a computer-simulated competitive game²⁷, while another study found similar effects related to signals encoding upcoming actions²⁸. In the present study, we found that signals related to neither past nor upcoming actions were influenced by the previous outcome. Instead, a more complex combination of signals corresponding to the location of the color chosen in the previous trial was enhanced following a rewarded outcome. This particular variable (HVL) indicated the position of the target with a high reward probability, and therefore was relevant for the task in the present study. DLPFC neurons with HVL signals tended to encode the animal's choice once the information about the reward magnitude became available. Therefore, HVL signals might reflect the animal's tentative choice that can be revised later. However, two lines of evidence suggest that HVL and choice signals are not equivalent. First, HVL signals were asymmetrically affected by the outcome of the animal's choice in the previous trial such that information about the location of the previously chosen color was encoded in the DLPFC only after the rewarded trial. Second, even after the onset of magnitude cues, DLPFC neurons encoded HVL and choice signals concurrently, suggesting that the HVL signals were not merely temporary choice signals. Instead, reward-dependent enhancement of task-relevant signals in the DLPFC might be a general mechanism for learning to optimize the outcomes of the animal's behavior.

In conclusion, these results suggest that despite the heterogeneity of signals represented in DLPFC, task-relevant and task-irrelevant signals might be processed differently. In addition, heterogeneous activity patterns in the DLPFC may actually provide computational advantages by broadening the possible range of information available in the DLPFC⁴⁹. This diversity may enable a substrate for representations to be updated flexibly based on changing task demands^{49,50}.

ONLINE METHODS

Animal preparation

Two male rhesus monkeys (O and U, body weight 9–11 kg) were used. Their ages at the time of recording were 5.8 years (monkey O) and 5.1 years (monkey U). Monkey O had been previously trained on a manual joystick task, but had not been used for electrophysiological recordings prior to this experiment. Monkey U had not been used for any prior experiments. Both animals were socially housed throughout these experiments. Eye movements were monitored at a sampling rate of 225 Hz with an infrared eye tracker

(ET49, Thomas Recording, Germany). All procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at Yale University.

Behavioral task

The animals fixated a small white square at the center of a computer monitor to initiate each trial (Fig. 1a). Following a 0.5 s fore-period, two peripheral targets were simultaneously presented along the horizontal meridian. One of the targets was green, while the other target was red, and their positions were randomized across trials. After a 0.5 s interval (target period), a set of small yellow tokens were presented around each target, indicating the magnitude of available reward (magnitude period). The central fixation cue was extinguished after a random interval ranging from 0.5 to 1.2 s, chosen according to a truncated exponential distribution with a mean of 705 ms. The animals were then free to shift their gaze towards one of the two peripheral targets. After fixating the chosen target for another 0.5 s, the animals received visual feedback. A red or green ring around the chosen target indicated that the animals would be rewarded, while a gray or blue ring indicated that they were not to be rewarded. An additional delay of 0.5 s (feedback period) separated the onset of the feedback ring from reward delivery. In trials where the animals were rewarded, they received the magnitude of juice associated with the chosen target. Each token corresponded to one drop of juice (0.1 ml).

The reward magnitudes associated with each target color were drawn from the following 10 possible pairs: {1,1}, {1,2}, {1,4}, {1,8}, {2,1}, {2,4}, {4,1}, {4,2}, {4,4}, {8,1}. Each magnitude pair was counter-balanced across target locations, yielding 20 unique trial conditions (10 magnitude combinations \times 2 target locations). At the beginning of a session, one trial condition was randomly selected without replacement from the pool of 20. After completing 20 trials, the set of trial conditions was replenished to assure even sampling.

The red and green targets were associated with different reward probabilities, which were fixed within a block of trials and alternated across blocks that consisted of 20 or 80 trials. One target color was associated with a high reward probability (80%), and the other was associated with a low reward probability (20%). These reward probabilities underwent un-signaled reversals across blocks so that the animals had to learn them through experience. Target reward probabilities were anti-correlated such that the reversals for the two targets always occurred simultaneously. We did not find any systematic differences in either animal's behavior or in neural activity related to different block lengths. Therefore, the data from all blocks were combined in the analyses.

Analysis of behavioral data

The optimal strategy in this task is to multiply estimated reward probabilities with magnitudes to calculate expected values. To investigate how estimated reward probabilities and magnitudes were combined, we first tested a series of models to empirically estimate reward probabilities. Specifically, we examined variations of reinforcement learning (RL) models¹⁹ in which the reward probabilities were updated using either one learning rate used for all outcomes, or two learning rates that were allowed to differ depending on whether the outcome was rewarded or unrewarded. In the RL model with a single learning rate (RL1),

$$P_c(t+1)=P_c(t)+\alpha \cdot [R(t)-P_c(t)],$$

where $P_c(t)$ is the estimated reward probability of the chosen target on trial t , α is the learning rate, and $R(t)$ is the trial's outcome on trial t (1 if rewarded, 0 otherwise). Similarly, in the RL model with double learning rates (RL2),

$$P_c(t+1)=P_c(t)+\{\delta(R(t)-1) \alpha_{\text{rew}}+\delta(R(t))\alpha_{\text{unr}}\} [R(t)-P_c(t)],$$

where α_{rew} is the learning rate associated with rewarded outcomes, and α_{unr} is the learning rate associated with unrewarded outcomes. $\delta(x)=1$, if $x=0$, and 0 otherwise.

We also tested models to assess whether the animals might have had knowledge that the reward probabilities of the targets were anti-correlated. That is, we tested models where only the reward probability of the chosen target, $P_c(t)$, was updated (uncoupled), against models where the chosen target was updated as above, while the reward probability for the unchosen target was determined as $= 1 - P_c(t)$.

Using these estimates of reward probabilities, we proceeded to test whether they were combined multiplicatively with the reward magnitude, as would be expected if the animal made its choice according to the expected value of reward. To achieve this aim, we directly compared several variations of models in which probabilities and magnitudes were combined either additively or multiplicatively, or both. In the additive model (P+M),

$$\text{logit } p(\text{Red})=\beta_0+\beta_1 \Delta P_{\text{RG}}(t)+\beta_2 \Delta M_{\text{RG}}(t),$$

where $P_{\text{RG}}(t) = P_{\text{red}}(t) - P_{\text{green}}(t)$ is the difference in estimated reward probabilities of the red and green targets, and $M_{\text{RG}}(t) = M_{\text{red}}(t) - M_{\text{green}}(t)$ is the difference in magnitudes of the targets. In the multiplicative or expected-value model (EV),

$$\text{logit } p(\text{Red})=\beta_0+\beta_1 \Delta \text{EV}_{\text{RG}}(t),$$

where $p(\text{Red})$ denotes the probability of choosing the red target, and $\text{EV}_{\text{RG}}(t) = P_{\text{red}}(t) \times M_{\text{red}}(t) - P_{\text{green}}(t) \times M_{\text{green}}(t)$ is the difference in estimated expected values of the targets on trial t . The full model (P+M+EV) combined additive and multiplicative components:

$$\text{logit } p(\text{Red})=\beta_0+\beta_1 \Delta P_{\text{RG}}(t)+\beta_2 \Delta M_{\text{RG}}(t)+\beta_3 \Delta \text{EV}_{\text{RG}}(t).$$

Maximum likelihood estimates for model parameters were estimated using the `fminsearch` function in Matlab (Mathworks, Inc. MA) by combining trials across all sessions. Since the models differed in the number of free parameters, we used the Bayesian information criterion (BIC) for model comparison which penalized the use of additional parameters. We

tested all 12 combinations of models based on the three different factors described above, namely, the number of learning rates (RL1 or RL2), whether the animals exploited the correlation in reward probabilities (coupled or uncoupled), and how probability and magnitude information was combined (P+M, EV, or P+M+EV). These models were separately fit to the behavioral data in each monkey. We also performed a leave-one-out cross-validation to compare the performance of the models across individual sessions (Supplementary Table 1).

Neurophysiological recording

Activity of individual neurons in the DLPFC was recorded extracellularly (left hemisphere in both monkeys) using a 16-channel multielectrode recording system (Thomas Recording, Germany) and a multichannel acquisition processor (Plexon, TX). Based on magnetic resonance images, the recording chamber was centered over the principal sulcus and located anterior to the genu of the arcuate sulcus (monkey O, 4 mm; monkey U, 10 mm). All neurons selected for analysis were located anterior to the frontal eye field, which was defined by eye movements evoked by electrical stimulation in monkey O (current < 50 uA). The recording chamber in monkey U was located sufficiently anterior to the frontal eye field, so stimulation was not performed in this animal. Each neuron in the dataset was recorded for a minimum of 320 trials (77 and 149 neurons in monkeys O and U, respectively), and on average for 518.8 trials (SD = 147.2 trials). We did not preselect neurons based on activity, and all neurons that could be sufficiently isolated for the minimum number of trials were included in the analyses.

Analysis of neural data

Linear regression analyses—A linear regression model was used to investigate how each neuron encoded both task-relevant and task-irrelevant information from the current and previous trials. Some of this information relied on the integration of simpler, more basic elements. For example, task-irrelevant information about the location of the rewarded target in the previous trial required knowledge of both the previously chosen location and previous outcome. Therefore, to analyze these higher-order features, we used the following multiple linear regression model which included simpler potentially confounding terms:

$$y = \beta_0 + \beta_1 C_{LR}(t) + \beta_2 C_{LR}(t-1) + \beta_3 R(t-1) + \beta_4 POS_{RG}(t) + \beta_5 M_R(t) + \beta_6 M_L(t) + \beta_7 C_{RG}(t) + \beta_8 C_{RG}(t-1) + \beta_9 R(t-1) \times POS_{RG}(t) + \beta_{10} C_{RG}(t-1) \times POS_{RG}(t) + \beta_{11} C_{RG}(t-1) \times R(t-1) + \beta_{12} PRL(t) + \beta_{13} HVL(t).$$

where $C_{LR}(t)$ is the position of the chosen target on trial t (1 if right, -1 otherwise), $R(t)$ is the outcome on trial t (1 if rewarded, -1 otherwise), $POS_{RG}(t)$ is the positions of the red and green targets on trial t (1 if red is on the right, -1 otherwise), $M_R(t)$ and $M_L(t)$ are the magnitudes of rewards available from the right and left targets on trial t , respectively, and $C_{RG}(t)$ is the color chosen by the animal on trial t (1 if chose red, -1 if chose green). In this model, task-irrelevant information related to the previously rewarded location, abbreviated hereafter as $PRL(t)$, is given by the interaction between the previously chosen location and previous outcome, namely $PRL(t) = C_{LR}(t-1) \times R(t-1)$. Task-relevant information corresponding to the high-value target location, $HVL(t)$, is defined as the current position of

the target color that was rewarded in the previous trial. This corresponds to a 3-way interaction between the previously chosen color, previous outcome, and current color locations, namely, $HVL(t) = C_{RG}(t-1) \times R(t-1) \times POS_{RG}(t)$. To test whether signals related to reward magnitude and choice are encoded consistently across different DLPFC neurons, we have also test the same regression model in which the two terms related to reward magnitude (M_R and M_L) were replaced by their difference ($M = M_R - M_L$).

Correlation and principal components analyses—As described in the Results, partially overlapping populations of neurons encoded task-relevant information related to HVL and task-irrelevant information related to PRL. To investigate how these populations were related to the neurons that encoded the animal's choice location, we examined the correlation between the regression coefficients calculated in non-overlapping 500 ms time bins. The use of non-overlapping bins was necessary for avoiding biases related to correlations between the regressors themselves. Therefore, the regression coefficients corresponding to HVL and PRL were estimated for the epoch following target onset (250 to 750 ms after target onset) while the regression coefficients for choice were taken from the period following magnitude onset (750 to 1250 ms relative to target onset).

We performed principal components analysis (PCA) to examine this further. PCA is an orthogonal linear transform used to find the axes that maximize the variance explained in the data. Since we found that the regression coefficients corresponding to HVL were strongly correlated with choice signals, it suggests that a mixture of signals related to HVL and choice may be present in the first principal component (PC1). Therefore, we constructed a matrix based on the regression coefficients derived from each variable. Rows were made up of each neuron, and columns were comprised of each variable's regression coefficient as a function of time. We applied PCA to this matrix and plotted the time course of loadings for PC1 corresponding to choice, HVL, and PRL. To assess the statistical significance of observed loadings revealed by this analysis, we randomly shuffled labels corresponding to the identities of different neurons independently for different variables 10,000 times. Next, we determined where the maximum value corresponding to each regression coefficient was relative to the shuffled distribution. Values corresponding to $\alpha=0.05$ were used to determine whether the loadings for the PC1 were statistically significant or not.

We tested whether the regression coefficients for magnitude difference (M) during the magnitude period were correlated with those for the HVL during the target period. We also examined how the regression coefficients for magnitude difference and choice were correlated during the magnitude period. However, the value of correlation coefficient between regression coefficients can be biased due to the correlation between the regressor themselves, when they are estimated from the same data. Therefore, to test the statistical significance of the correlation between the regression coefficients estimated for the same time interval, we randomly split the trials into two halves and examined the correlation coefficient between the regression coefficients estimated from two separate groups of trials.

Decoding analysis—To test whether signals related to the position and color of the target chosen by the animal was enhanced after a reward or not, we applied a linear discriminant analysis with 5-fold cross-validation separately for each neuron. As in our previous study²⁷,

the training set was balanced according to the current location of the previously chosen target, the upcoming choice location, and the previous outcome (yielding 8 different categories). In order to assure an unbiased classifier, we randomly removed trials until each category had the same number of trials. We first used the classifier to see whether we could extract information about HVL using the firing rates of neurons in the current trial for previously rewarded and unrewarded trials separately. We applied a 500 ms sliding window advanced in 50 ms increments from 500 ms before target onset to 1500 ms after target onset. Trials were randomly assigned to 5 different subgroups (for 5-fold cross validation). Each subgroup served as a test set, and the combined trials from the 4 remaining subgroups served as a training set which was used to find the firing rate boundary that maximally separated rightward and leftward categories. We then calculated how well we could classify the trials in the test set given this firing rate boundary. The process was repeated for each subgroup and the results averaged to obtain a time course for decoding accuracy for each neuron. Population means were constructed by averaging over all neurons. The same decoding procedure was repeated for decoding PRL and the animal's upcoming choice. In the process of balancing the decoder, the number of trials available was not sufficient (<5 trials in one category) for a small fraction of neurons (10/226) to allow for proper cross validation. We removed these neurons from the analysis and ran the decoder on the remaining 216 neurons that survived this criterion (77/77 neurons in monkey O; 139/149 neurons in monkey U).

Statistical analysis

Binomial tests (one-tailed) were used to determine whether the fraction of neurons encoding a particular signal was significantly greater than chance (5%). Paired t-tests (two-tailed) were used to test whether means for firing rates or decoding accuracy significantly differed between experimental conditions. χ^2 -tests were used to test whether neurons independently encoded choice and HVL or PRL. The statistical significance of specific regressors in the multiple linear regression analysis was assessed with a two-tailed t-test. Data distribution was assumed to be normal but this was not formally tested.

All neurons that could be isolated for at least 320 trials were included in this dataset, and there was no blinding for data collection or analysis. Trials were pseudorandomly presented such that one trial condition was randomly chosen without replacement from a pool of 20 possible conditions, balancing across target magnitudes and positions. Following successful completion of all trial conditions, the set was replenished to assure even sampling. The same size in the present study was not based on any statistical test, but is comparable to that of our previous studies^{14,35}.

A Supplementary Methods Checklist is available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Mark Hammond and Patrice Kurnath for their technical support, and Zhihao Zhang for his help with the experiment. This study was supported by the National Institute of Health (R01 DA029330 and T32 NS007224).

References

1. Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol.* 1989; 61:331–349. [PubMed: 2918358]
2. Lara AH, Wallis JD. Executive control processes underlying multi-item working memory. *Nat Neurosci.* 2014; 17:876–883. [PubMed: 24747574]
3. Romo R, Brody CD, Hernández A, Lemus L. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature.* 1999; 399:470–473. [PubMed: 10365959]
4. Constantinidis C, Franowicz MN, Goldman-Rakic PS. The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nat Neurosci.* 2001; 4:311–316. [PubMed: 11224549]
5. O'Scalaidhe SP, Wilson FA, Goldman-Rakic PS. Areal segregation of face-processing neurons in prefrontal cortex. *Science.* 1997; 278:1135–1138. [PubMed: 9353197]
6. Rao SC, Rainer G, Miller EK. Integration of what and where in the primate prefrontal cortex. *Science.* 1997; 276:821–824. [PubMed: 9115211]
7. Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci.* 2001; 24:167–202. [PubMed: 11283309]
8. Tanji J, Hoshi E. Role of the lateral prefrontal cortex in executive behavioral control. *Physiol Rev.* 2008; 88:37–57. [PubMed: 18195082]
9. Lebedev MA, Messinger A, Kralik JD, Wise SP. Representation of attended versus remembered locations in prefrontal cortex. *PLoS Biol.* 2004; 2:e365. [PubMed: 15510225]
10. Messinger A, Lebedev MA, Kralik JD, Wise SP. Multitasking of attention and memory functions in the primate prefrontal cortex. *J Neurosci.* 2009; 29:5640–5653. [PubMed: 19403830]
11. Watanabe M. Reward expectancy in primate prefrontal neurons. *Nature.* 1996; 382:629–632. [PubMed: 8757133]
12. Leon MI, Shadlen MN. Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron.* 1999; 24:415–426. [PubMed: 10571234]
13. Barraclough DJ, Conroy ML, Lee D. Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci.* 2004; 7:404–410. [PubMed: 15004564]
14. Kim S, Hwang J, Lee D. Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron.* 2008; 59:161–172. [PubMed: 18614037]
15. Kim S, Cai X, Hwang J, Lee D. Prefrontal and striatal activity related to values of objects and locations. *Front Neurosci.* 2012; 6:108. [PubMed: 22822390]
16. Kennerley SW, Dahmubed AF, Lara AH, Wallis JD. Neurons in the frontal lobe encode the value of multiple decision variables. *J Cog Neurosci.* 2009; 21:1162–117. [PubMed: 19403830]
17. Lauwereyns J, et al. Responses to task-irrelevant visual features by primate prefrontal neurons. *J Neurophysiol.* 2001; 86:2001–2010. [PubMed: 11600657]
18. Genovesio A, Tsujimoto S, Navarra G, Falcone R, Wise SP. Autonomous encoding of irrelevant goals and outcomes by prefrontal cortex neurons. *J Neurosci.* 2014; 34:1970–1978. [PubMed: 24478376]
19. Sutton, RS., Barto, AG. *Reinforcement Learning: An Introduction.* MIT Press; Cambridge, Massachusetts, USA: 1998.
20. Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. *Science.* 2005; 307:1642–1645. [PubMed: 15761155]
21. Knutson B, Taylor J, Kaufman M, Peterson R, Glover G. Distributed neural representation of expected value. *J Neurosci.* 2005; 25:4806–4812. [PubMed: 15888656]
22. Yacubian J, et al. Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J Neurosci.* 2006; 26:9530–9537. [PubMed: 16971537]
23. Tobler PN, Christopoulos GI, O'Doherty JP, Dolan RJ, Schultz W. Neuronal distortions of reward probability without choice. *J Neurosci.* 2008; 28:11703–11711. [PubMed: 18987206]
24. Christopoulos GI, Tobler PN, Bossaerts P, Dolan RJ, Schultz W. Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *J Neurosci.* 2009; 29:12574–12583. [PubMed: 19812332]

25. Venkatraman V, Payne JW, Bettman JR, Luce MF, Huettel SA. Separate neural mechanisms underlie choices and strategic preferences in risky decision making. *Neuron*. 2009; 62:593–602. [PubMed: 19477159]
26. Berns GS, Bell E. Striatal topography of probability and magnitude information for decisions under uncertainty. *NeuroImage*. 2012; 59:3166–3172. [PubMed: 22100418]
27. Donahue CH, Seo H, Lee D. Cortical signals for rewarded actions and strategic exploration. *Neuron*. 2013; 80:223–234. [PubMed: 24012280]
28. Histed MH, Pasupathy A, Miller EK. Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. *Neuron*. 2009; 63:244–253. [PubMed: 19640482]
29. Ito M, Doya K. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J Neurosci*. 2009; 29:9861–9874. [PubMed: 19657038]
30. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*. 2005; 8:1704–1711. [PubMed: 16286932]
31. Hampton AN, Bossaerts P, O’Doherty JP. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci*. 26:8360–8367.
32. Lee D, Seo H, Jung MW. Neural basis of reinforcement learning and decision making. *Annu Rev Neurosci*. 2012; 35:287–308. [PubMed: 22462543]
33. Glascher J, Daw N, Dayan P, O’Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010; 66:585–595. [PubMed: 20510862]
34. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans’ choices and striatal prediction errors. *Neuron*. 2011; 69:1204–1215. [PubMed: 21435563]
35. Abe H, Lee D. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron*. 2011; 70:731–741. [PubMed: 21609828]
36. Lee SW, Shimojo S, O’Doherty JP. Neural computations underlying arbitration between model-based and model-free learning. *Neuron*. 2014; 81:687–699. [PubMed: 24507199]
37. Seo H, Cai X, Donahue CH, Lee D. Neural correlates of strategic reasoning during competitive games. *Science*. 2014; 346:340–343. [PubMed: 25236468]
38. Asaad WF, Rainer G, Miller EK. Neural activity in the primate prefrontal cortex during associative learning. *Neuron*. 1998; 21:1399–1407. [PubMed: 9883732]
39. Curtis CE, Lee D. Beyond working memory: the role of persistent activity in decision making. *Trends Cog Sci*. 2010; 14:216–222.
40. Bernacchia A, Seo H, Lee D, Wang XJ. A reservoir of time constants for memory traces in cortical neurons. *Nat Neurosci*. 2011; 14:366–372. [PubMed: 21317906]
41. Genovesio A, Brasted PJ, Mitz AR, Wise SP. Prefrontal cortex activity related to abstract response strategies. *Neuron*. 2005; 47:207–320.
42. Genovesio A, Brasted PJ, Wise SP. Representation of future and previous spatial goals by separate neural populations in prefrontal cortex. *J Neurosci*. 2006; 26:7305–7316. [PubMed: 16822988]
43. Fuster JM, Bodner M, Kroger JK. Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature*. 2000; 405:347–351. [PubMed: 10830963]
44. Mante V, Sussillo D, Shenoy KV, Newsome WT. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*. 2013; 503:78–84. [PubMed: 24201281]
45. Chen LL, Wise SP. Neuronal activity in the supplementary eye field during acquisition of conditional oculomotor associations. *J Neurophysiol*. 1995; 73:1101–1121. [PubMed: 7608758]
46. Chen LL, Wise SP. Evolution of directional preferences in the supplementary eye field during acquisition of conditional oculomotor associations. *J Neurosci*. 1996; 16:3067–3081. [PubMed: 8622136]
47. Chen LL, Wise SP. Conditional oculomotor learning: population vectors in the supplementary eye field. *J Neurophysiol*. 1997; 78:1166–1169. [PubMed: 9307145]
48. Singer AC, Frank LM. Rewarded outcomes enhance reactivation of experience in the hippocampus. *Neuron*. 2009; 64:910–921. [PubMed: 20064396]

49. Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S. The importance of mixed selectivity in complex cognitive tasks. *Nature*. 2013; 497:585–590. [PubMed: 23685452]
50. Duncan J. An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci*. 2001; 2:820–829. [PubMed: 11715058]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

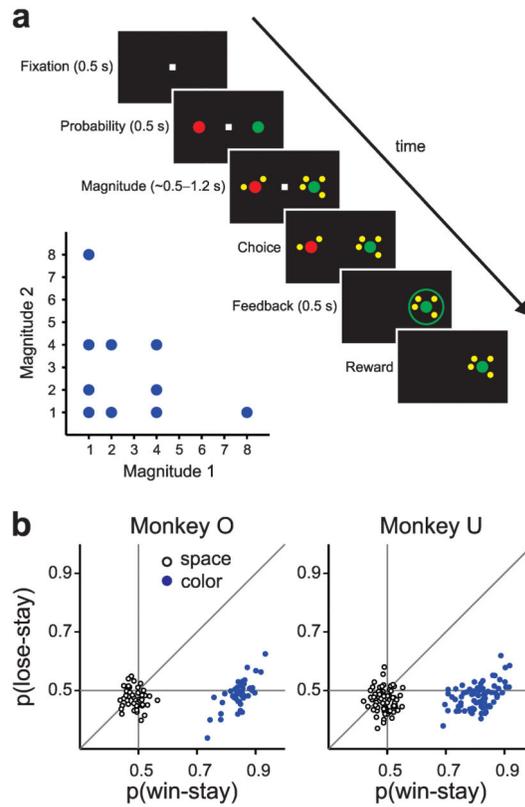
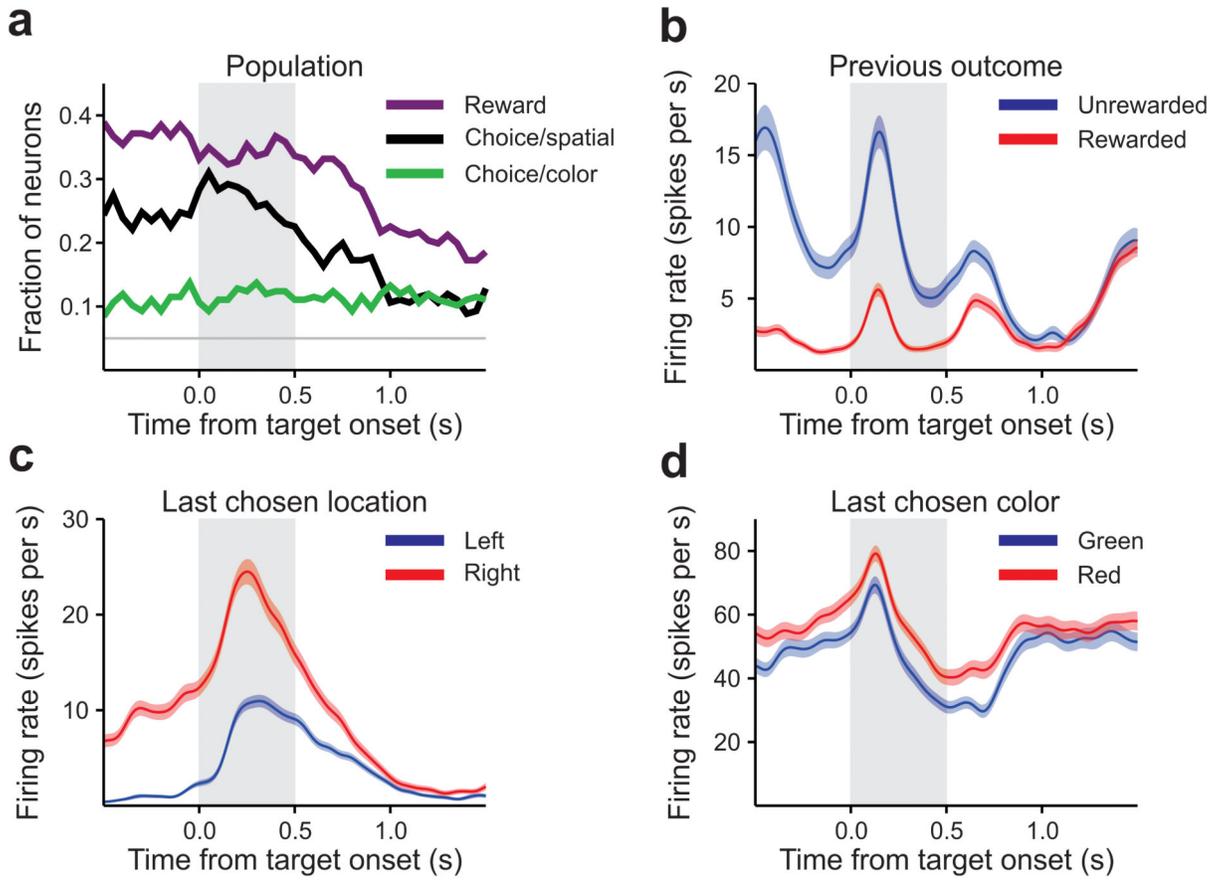


Figure 1. Behavioral task and performance. (a). Probabilistic reversal task, and magnitude combinations used (inset). (b) The proportion of trials in which the animal chose the same target color or location as in the previous trial after the previous choice was rewarded (win-stay) or unrewarded (lose-stay) ($n=45$ sessions in monkey O and 73 sessions in monkey U).

**Figure 2.**

Population summary and single neuron examples for activity related to events in the previous trial. (a) Fraction of neurons significantly encoding outcomes, chosen locations, and chosen colors in the previous trial ($n=226$ neurons; 77 and 149 neurons from monkeys O and U, respectively). (b) Example neuron showing effect of outcome in the previous trial. (c) Example neuron showing effect of previously chosen location. (d) Example neuron showing effects of previously chosen color. Gray background, target period; shaded areas, \pm SEM.

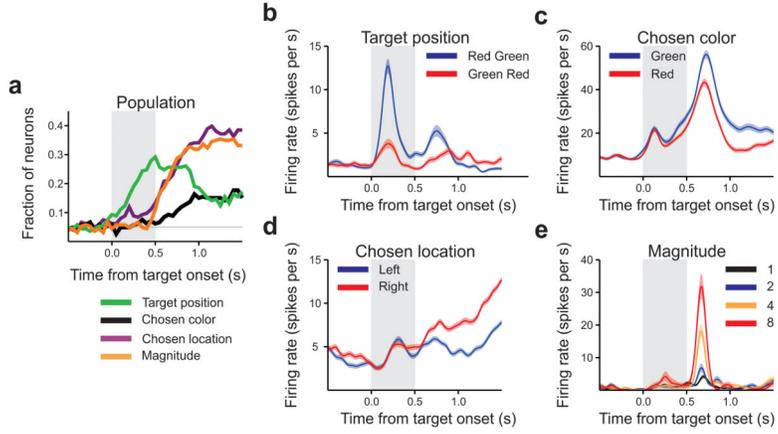


Figure 3. Population summary and single neuron examples for activity related to events in the current trial. (a) Fraction of neurons significantly encoding the chosen target location, chosen color, target color positions, and magnitudes in the current trial (n=226 neurons). (b) Example neuron showing effect of target position. (c) Example neuron showing effect of chosen color. (d) Example neuron showing effect of chosen location. (e) Example neuron encoding reward magnitude of the rightward target. Same format as in Fig. 2.

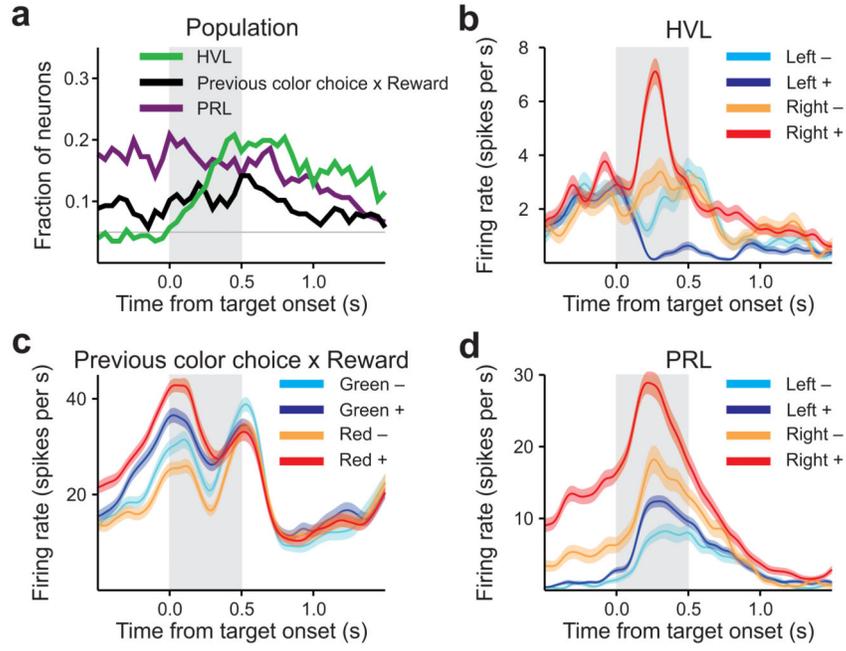


Figure 4. Population summary and single neuron examples related to interaction effects. (a) Fraction of neurons showing effects of previously rewarded location (PRL), previously chosen color and outcome, and their interaction corresponding to the high-value location (HVL) (n=226 neurons). (b) Example neuron showing encoding of HVL. (c) Example neuron encoding previously chosen color \times previous outcome interaction. (d) Example neuron encoding PRL. Same format as in Fig. 2.

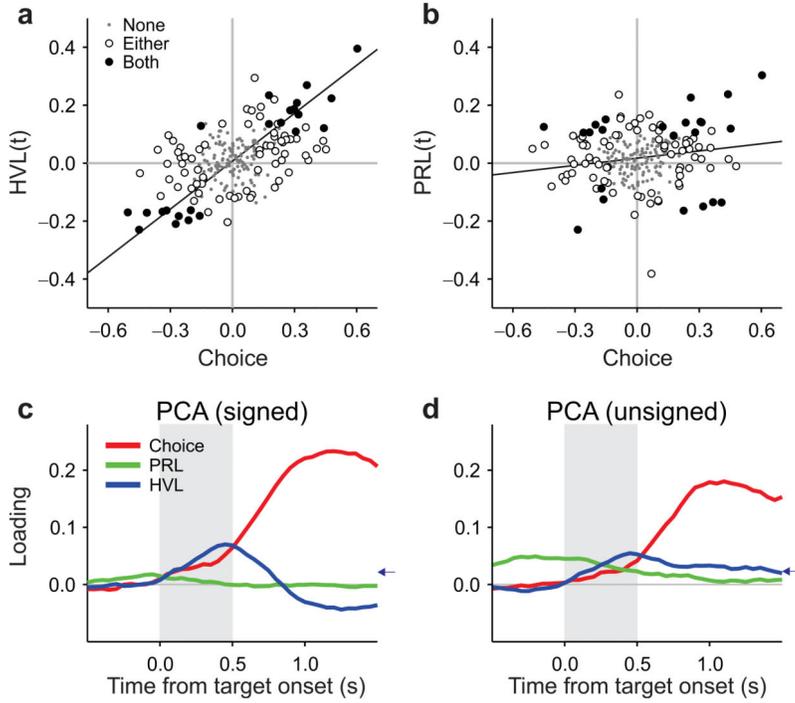


Figure 5. Congruent coding of HVL and choice. (a) Relationship between regression coefficients for HVL versus choice. (b) Relationship between regression coefficients for PRL versus choice (n=226 neurons). (c) Time course of first principal component (see Online Methods) related to choice, HVL, and PRL using signed regression coefficients. (d) Same result for unsigned regression coefficients. The arrow indicates chance levels based on 10,000 random shuffles ($p < 0.05$).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

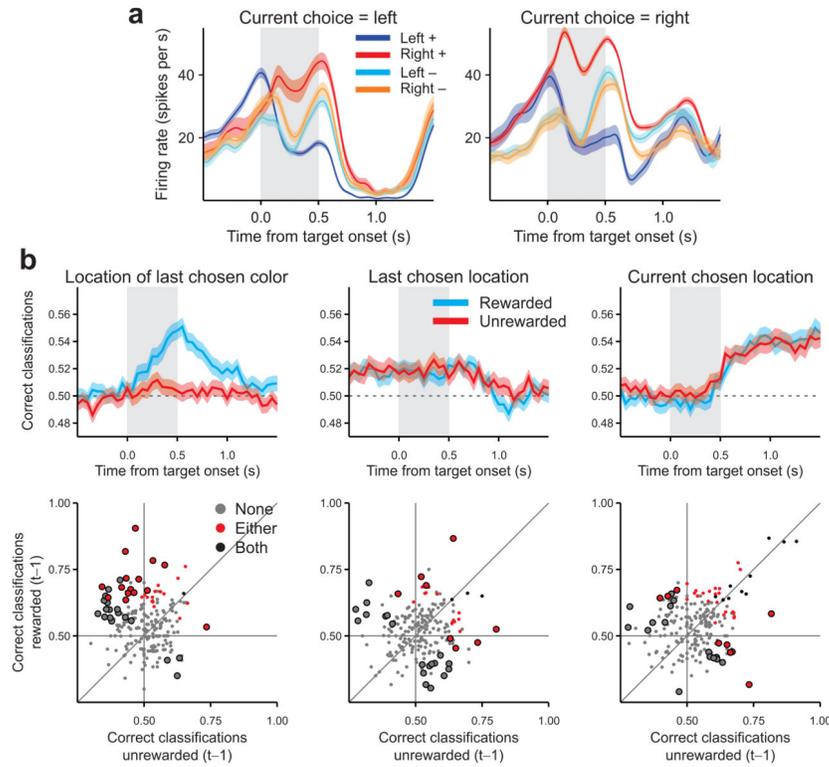


Figure 6. Effects of reward on task-relevant and irrelevant signals in the DLPFC. (a) Example neuron encoding the high-value location (HVL), shown separately for leftward and rightward choices in the current trial. (b) Decoding accuracy for the current location of the previous chosen color, previous chosen location, and current chosen location, shown separately for previously rewarded and unrewarded trials (n=216 neurons that survived our criterion for cross-validation; see Online Methods). Shaded areas, \pm SEM. The scatterplots in the bottom display the fraction of correct classifications for each neuron during the target period. Colors indicate whether neither, one, or both decoding analyses applied to previously rewarded and unrewarded were significantly above chance ($p < 0.05$), and large symbols indicate that the difference between rewarded and unrewarded trials was statistically significant (z-test, $p < 0.05$).