Systems/Circuits

Prefrontal Neurons Encode a Solution to the Credit-Assignment Problem

Wael F. Asaad,[1,2,3,4] **Peter M. Lauro,**[2,3] **János A. Perge,**[2] and **Emad N. Eskandar**[5,6]

[1]Department of Neurosurgery and [2]Department of Neuroscience, Brown University, Providence, Rhode Island 02912, [3]Brown University Alpert Medical School, Providence, Rhode Island 02903, [4]Norman Prince Neurosciences Institute, Rhode Island Hospital, Providence, Rhode Island 02903, [5]Department of Neurosurgery, Massachusetts General Hospital, Boston, Massachusetts 02114, and [6]Department of Neurosurgery, Harvard Medical School, Boston, Massachusetts 02115

To adapt successfully to our environments, we must use the outcomes of our choices to guide future behavior. Critically, we must be able to correctly assign credit for any particular outcome to the causal features which preceded it. In some cases, the causal features may be immediately evident, whereas in others they may be separated in time or intermingled with irrelevant environmental stimuli, creating a potentially nontrivial credit-assignment problem. We examined the neuronal representation of information relevant for credit assignment in the dorsolateral prefrontal cortex (dlPFC) of two male rhesus macaques performing a task that elicited key aspects of this problem. We found that neurons conveyed the information necessary for credit assignment. Specifically, neuronal activity reflected both the relevant cues and outcomes at the time of feedback and did so in a manner that was stable over time, in contrast to prior reports of representational instability in the dlPFC. Furthermore, these representations were most stable early in learning, when credit assignment was most needed. When the same features were not needed for credit assignment, these neuronal representations were much weaker or absent. These results demonstrate that the activity of dlPFC neurons conforms to the basic requirements of a system that performs credit assignment, and that spiking activity can serve as a stable mechanism that links causes and effects.

Key words: credit assignment; learning; monkey; population coding; prefrontal cortex; single neuron

---

Significance Statement

Credit assignment is the process by which we infer the causes of our successes and failures. We found that neuronal activity in the dorsolateral prefrontal cortex conveyed the necessary information for performing credit assignment. Importantly, while there are various potential mechanisms to retain a "trace" of the causal events over time, we observed that spiking activity was sufficiently stable to act as the link between causes and effects, in contrast to prior reports that suggested spiking representations were unstable over time. In addition, we observed that this stability varied as a function of learning, such that the neural code was more reliable over time during early learning, when it was most needed.

Introduction

Credit assignment is the process that links the outcomes of our choices with the responsible factors. For example, many of us

have had the experience of using trial and error to recall which passcode works in a particular circumstance. Upon discovering the successful entry, we may immediately review or mentally note that information to recall it the next time. This is a critical moment, because appropriately assigning credit for that success to the correct preceding event determines whether we have learned from this experience, or if we'll stumble through the same search process the next time as well. More generally, properly assigning credit for specific outcomes underlies our ability to infer causality and make sense of our environment.

Credit assignment is necessary for any form of associative learning, but it is more challenging when the causal environmental feature is ephemeral and so no longer present when the outcome is revealed (this is the temporal credit-assignment problem) or when multiple potentially relevant features are concurrently present (the structural credit-assignment problem).

The dorsolateral prefrontal cortex (dlPFC), whose neurons have been observed to respond selectively to relevant, attended information (Rainer et al., 1998a; Everling et al., 2002; Lebedev et al., 2004), as well as to maintain information over time (Fuster and Alexander, 1971; Funahashi et al., 1993; Miller et al., 1996; Rainer et al., 1998b; Markowitz et al., 2015), would seem likely to have the basic, necessary properties for solving the structural and temporal credit-assignment problems. Likewise, lateral PFC lesions are typically found to impair performance on tasks that have in common the requirement to use information about choice outcomes to guide future behavior (Dias et al., 1997; Parker and Gaffan, 1998; Mansouri et al., 2007; Rossi et al., 2009; Simmons et al., 2010; Rushworth et al., 2011; Kovach et al., 2012). These deficits often do not conform straightforwardly to a "working memory" narrative, but they are consistent with a potential role in credit assignment. Yet, the specific neuronal representations necessary for performing credit assignment, especially in its more challenging forms incorporating temporal or structural problems, have not been directly investigated in the PFC or elsewhere. This alternative framework may be useful to understand more precisely how these neuronal representations contribute to behavior, especially in light of growing neurophysiological and lesion evidence that PFC activity may not directly mediate simple working memory (Lara and Wallis, 2014, 2015; Pasternak et al., 2015). Therefore, to study the responses of individual dlPFC neurons and populations of dlPFC neurons during learning, we designed a task that invoked both the structural and temporal credit-assignment problems.

Importantly, recent data suggest that neuronal activity in the lateral PFC dynamically evolves such as to severely limit the cross-temporal stability of representations (Meyers et al., 2008; Sigala et al., 2008; Barak et al., 2010; Stokes et al., 2013). This could be particularly problematic for the temporal credit-assignment problem, where a stable representation of relevant information over time is necessary so that reinforcement received when an outcome becomes apparent can be applied to the same neural ensemble that earlier signaled the causal features. However, that previous work involved tasks that did not require active learning or credit assignment at the time neurons were recorded and mostly relied on "pseudopopulations" of neurons (not all neurons considered by a particular cross-temporal decoder or correlation matrix were actually recorded simultaneously). Therefore, we focused on simultaneously recorded neural activity during a learning task requiring credit assignment, seeking to understand whether dlPFC activity can indeed provide the necessary information—including a stable representation over time—for solving the credit-assignment problem.

Materials and Methods

Subjects and behavior. Two male rhesus macaques, M1 and M2, performed the credit assignment and spatial tasks in a pseudorandom, block-wise, and interleaved fashion, with ≥4 repetitions of any particular block (correct cue or spatial location) in each individual session (Asaad and Eskandar, 2011). The task required animals to maintain central fixation as four cue objects were presented peripherally (500 ms), to hold fixation through a brief (1 s) delay, and then to make a choice by saccading to one of the four locations where the correct cue had appeared (Fig. 1). Generic visual feedback for correct or incorrect choices (a green circle or a red X) was then presented for 500 ms followed by automated juice reward if correct. Because, in the main credit-assignment task, this feedback did not reveal which cue had been at the selected location, a challenging credit-assignment problem needed to be solved for learning to occur. Animals learned which cue signaled the correct saccade target by trial and error. Once the correct cue was learned and the animal had

performed an additional ~40–50 trials, a new block was begun in which a different cue was designated correct, without any overt signal to the animal. In contrast to this cue-learning task, in the spatial version of this task, which the animals played on pairwise interleaved blocks, the correct choice was determined solely by its spatial position, and not by the identity of the cue that had previously appeared at that location. In both tasks, the arrangement of the cues varied randomly from trial to trial, but the same four cues were used throughout any individual session. Unique arrays of four cue pictures were chosen for each session from a pool of ~50 familiar stimuli (to limit potential perceptual learning about particular stimuli and to use session-unique combinations of those stimuli).

Because all trials within a session contained the same four pictures, an animal's attention to a particular cue was inferred through its correct performance on a trial. Although individual trials could occasionally be performed successfully by chance (25%), sustained high performance depended on generally consistent allocation of attention to the correct cue. Outcomes were sorted into four categories, as in our previous work (Asaad and Eskandar, 2011), based on the current and prior trial's outcome: unexpected correct, expected correct, unexpected incorrect, and expected incorrect. An outcome was considered unexpected if the prior trial's outcome differed from the current one. Monkeys M1 and M2 performed an average of 1019.1 and 963.5 correct trials (of 1281.1 and 1179.7 attempted trials without technical errors) per session over an average of 23.4 and 22.6 blocks, respectively. M1 performed 33 sessions and M2 performed 21 sessions.

The behavioral task was implemented in MonkeyLogic (Asaad and Eskandar, 2008; Asaad et al., 2013; www.monkeylogic.org). Eye position was monitored with an infrared tracking system (Iscan) at 120 Hz. Animals were always handled in accordance with National Institutes of Health policies and those of the Massachusetts General Hospital animal care and use committee.

Anatomy. To reconstruct recording locations, we relied upon a post-implant MRI of each animal with vitamin E-filled fiducial arrays marking specific recording trajectories. Each array was partitioned into discrete three-dimensional (3D) volumes using the 3dclust function of AFNI (Analysis of Functional Neuroimages) with a mask over the fiducial signals in the MRI (Cox, 1996). The resulting volumes were visualized in 3D with SUMA (AFNI Surface Mapper; Saad et al., 2004). The center coordinate of the array was calculated as the center of mass of a manually selected central fiducial volume within the array. The coordinates of adjacent, parallel trajectories extending in approximately the anterior–posterior and dorsomedial–ventrolateral directions were determined up to an 8 mm radius to encompass the full recording array span. These coordinates were then extended along the third (z-axis) dimension using the appropriate linear transformation matrices, then visualized in SUMA as final recording trajectories.

The coordinate of intersection between a trajectory axis and dura was calculated by manually selecting where the reconstructed trajectory intersected the monkey's 3D MRI brain surface. Using the same transforms, the resulting coordinate was transformed to the trajectory space. The normal of the resulting coordinate was used as the z-axis value of the final dura–trajectory intersection coordinate to further align this coordinate along the trajectory axis. For subsequent electrode coordinate calculations, this point served as turn 0. Neuronal recording locations were described by their position along the array's anterior–posterior and dorsomedial–ventrolateral axes, and the number of 0.125 mm turns of the manual electrode microdrive. All recording locations were determined in the trajectory space, then transformed back to MRI native coordinates.

To visualize the anatomical location of each recording site, the monkey MRI brains underwent linear and nonlinear registration to the D99 macaque atlas using the script macaque_align.csh (Reveley et al., 2016). All MRI recording coordinates were transformed via this concatenated registration to the D99 atlas space. We display anatomical locations with a radius of 1 mm from the registered coordinates to account for error introduced throughout the imaging processing steps.

Electrophysiology. Individual dlPFC neurons were isolated using ≤16 acutely inserted microelectrodes through a permanent recording chamber implant. Neurons were selected for recording based solely upon their stability and signal-to-noise ratio, regardless of any behaviorally related
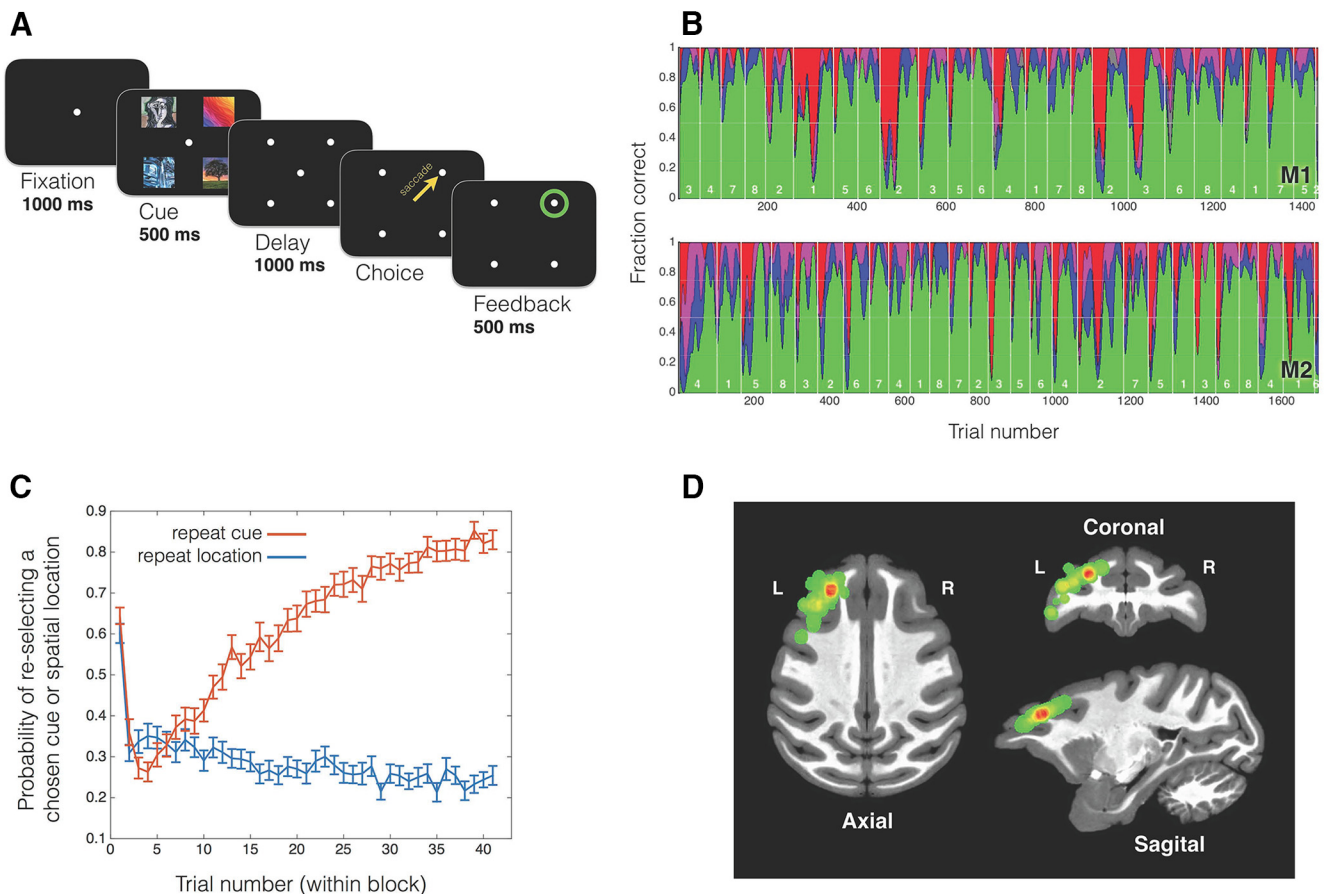
A

B

C

D

Figure 1. Behavioral task, behavioral performance, and recording locations. *A*, Behavioral task. Animals first acquired, then held fixation for 1 s. This was followed by the appearance of a cue array consisting of four cue objects, randomly arranged, presented for 0.5 s. A 1 s "delay" interval followed. Then, the fixation spot disappeared and animals made their choice by saccading to the former location of one of the cues. Central fixation was maintained throughout the trial until the saccadic response. If the cue designated "correct" had appeared at the chosen location, generic positive feedback (a green circle) was presented for 0.5 s, followed by automated juice reward. If the chosen location had contained an "incorrect" cue, a red "X" was presented without subsequent reward. Once animals learned which cue marked the rewarded location and performed 40–50 further correct trials, a different cue was designated correct and they were required to relearn the correct cue using trial and error. The "spatial" task used the same sequence of cues and motor responses; however, a particular spatial location determined the correct response regardless of the cue that earlier appeared there. In that case, no credit assignment to the cues was necessary for learning. *B*, Behavioral performance from a typical session for each animal. Blocks, separated by vertical white lines, consisted of one feature (cue picture or spatial location) designated as correct. The white numbers identify the condition in that block (blocks 1–4 for each of the four cue pictures; blocks 5–8 for each of the spatial locations that could be designated correct). Blocks were interleaved in a pseudorandom fashion such that the cue-learning and spatial-learning tasks occurred in pairs and no individual block would be repeated within an eight-block cycle. Behavioral data were smoothed in a 10-trial boxcar average (green, correct choice; red, incorrect choice; pink, broken fixation; blue, early response; gray, no fixation to initiate trial). *C*, The behavioral strategy used by the animals. Here, rather than overall performance, the probabilities of repeating the immediately preceding choice of cue object (red) or spatial location (blue)—whatever those choices happened to be—are plotted as a function of trial number within a block (±SEs). Only the cue-learning blocks, in which the animals needed to learn which cue object marked the correct spatial location, are included. Note that these probabilities begin relatively higher due to the presence of preceding cue-learning or spatial-learning blocks. Animals relied relatively more on a spatial strategy (reselecting a particular spatial location) early during learning, then switched to a cue-based strategy (reselecting the location indicated by a particular cue) as learning progressed. *D*, Locations of neuronal recordings. The density of recording sites from both animals is projected onto a reference macaque brain and shown in the three standard planes (see Materials and Methods). Warmer colors indicate relatively more recordings performed at those locations. The visualized slices were selected to pass through the highest-density region.

responses. Signals were sorted into individual units using principal components analysis and individual waveform features (Plexon Offline Sorter, Plexon). Data analysis was performed in Matlab (Mathworks). Other results from this dataset have been previously reported (Asaad and Eskandar, 2011).

General data analysis. Analyses used a sliding 200 ms time window shifted in 25 ms steps across the trial. This time window was selected to match the typical dynamics of dlPFC neurons, which often show phasic responses of about this duration. Trials were divided into cue-triggered and feedback-triggered segments of activity because varying reaction times changed the precise interval between these events; we were most interested in neuronal activity evoked by these stimuli, and so precise alignment of neuronal activity to them was desirable. This approach is reflected in the figures containing a split x-axis (time). For the cross-temporal decoding analysis (see Figs. 6, 8), the two trial segments were concatenated for analysis and visualization. Because multiple compari-

sons were performed over time bins and/or neurons, all *p* values were subjected to the Benjamini–Hochberg procedure to limit the false discovery rate (Benjamini and Hochberg 1995), using a *q* level of 0.05.

Because cue selectivity is determined by comparing spike rates across blocks of trials, there is the potential for slow changes in baseline activity, whether mechanical or physiological, to influence this measure. Therefore, although information about the cue objects could persist between trials (because the same cue remained correct across an entire block), we conservatively repeated some analyses (see Results) while excluding any neuron that showed significant cue selectivity during a 500 ms baseline epoch (from −750 to −250 ms before cue array onset). Specifically, if a neuron exhibited cue selectivity during any 200 ms bin within that interval, it was eliminated from the reanalysis. Assessing spatial selectivity during the cue-learning task was not as susceptible to this concern because the spatial choice varied trial by trial rather than block by block.

Neural information. The amount of information neuronal responses (R) conveyed about task features (S) was quantified using an information theoretic approach. Specifically, the information (in bits) about a feature (cue or outcome) evident in a neuron's activity was calculated as the probability of observing a particular neuronal response, r (number of spikes in the examined time bin) given a particular stimulus, s [individual feature exemplar: cues (A, B, C, or D) or (unexpected correct, expected correct, unexpected incorrect, or expected incorrect) outcomes], summing over neuronal responses ($r_i \epsilon R$) and feature exemplars ($s_j \epsilon S$) as follows:

$$\sum_j \sum_i p(r_i|s_j) \times \log_2\left(\frac{p(r_i|s_j)}{p(r_i)}\right)$$

The probability of a response given a particular stimulus was normalized by the log ratio comparing the probability of observing that specific neuronal response given a stimulus divided by the probability of observing that response over all trials. The maximum amount of information that could be conveyed about a feature given four exemplars was two bits.

The significance of the information conveyed was assessed using a bootstrap method in which spike rates were randomly shuffled across the four feature exemplars followed by recalculation of the information metric. This process was repeated 1000 times to obtain a distribution of information values and the p values of the actual (nonshuffled) data were obtained by comparison of the observed information value to this distribution using a one-tailed test.

To understand how much information was conveyed by neurons as a function of learning, we determined how neural information varied during learning in those neurons that carried significant information about the cue objects or spatial locations. Specifically, we recalculated the information metric in five-trial segments of data sorted with respect to the achievement of learning criterion (four trials in a row correct, because the probability of observing that number of consecutive correct choices by chance was $0.25^4 < 0.01$). Because differences in the number of trials used for any particular analysis will change the available entropy, the neural information calculated here is not directly comparable to that in other analyses (as above). More importantly, differences in the number of trials at different stages of learning within this analysis can produce false trends that represent simply the varying entropy. Therefore, we compared the information obtained in this analysis to shuffled controls in which the assignments of trials to cue objects or spatial locations were randomized, maintaining consistent entropy with respect to the original, nonshuffled calculation, at each stage of learning.

Population decoding. To determine the amount of information a population of neurons conveyed about the correct cue, we used a simple linear decoder (Gochin et al., 1994; Victor and Purpura, 1996; Samengo, 2002) to classify trials according to the spike rates of simultaneously recorded neurons. Individual recording sessions contained between 2 and 29 dlPFC neurons. The analytic approach began by creating four mean population activity vectors for the n neurons recorded in a single session in R^n space, with one vector for each of the four cues, considering only correctly performed trials when that cue was designated correct. These four mean vectors were then compared against neural activity vectors derived from individual trials (not drawn from the training set of trials). Each trial was classified as belonging to the group represented by the mean vector having the shortest Euclidean distance to it. Classification accuracy was simply the proportion of trials in which neuronal activity most closely matched the correct cue rather than the other cues, based upon this n-dimensional distance metric. Monte Carlo cross-validation was performed over 100 iterations, each using an 80:20 (training/testing) split of the data, yielding 100 accuracy scores. This procedure was repeated for each time bin. The stability of neuronal population representations over time was assessed by training and testing the data using different time bins (Meyers et al., 2008; Stokes et al., 2013). Here, above-chance performance resulted when a classifier trained at one time could successfully be applied to activity observed at a different time within the trial, reflecting some degree of similarity in the neuronal representations. For this analysis, neuronal spike rates were smoothed with a

Gaussian kernel ($\sigma = 25$ ms) before taking means over 200 ms bins shifted by 25 ms steps.

To determine whether a particular classification accuracy was better than chance, we repeated the decoding procedure using a randomized assignment of cues to trials (thereby shuffling spike rates), also over 100 iterations. These two distributions of accuracy scores, one for the real (unshuffled) data and one for the randomized data, were then compared using the receiver operating characteristic (ROC) area-under-the-curve (AUC), combined with a bootstrap test for significance over 1000 iterations (in which the group membership of a given accuracy score was randomized to create the null distribution). Because the observed variance of the accuracy scores depends on details, such as training/testing ratio and number of Monte-Carlo cross-validations performed, the ROC AUC reflects the adequacy of those chosen parameters to appropriately estimate decoder classification accuracy, providing additional information regarding its robustness and significance.

To understand how multiple behavioral factors may have together influenced feedback-epoch neuronal activity, we constructed a linear model for each neuron incorporating several categorical behavioral variables. Specifically, predictor variables consisted of the Task (cue learning vs spatial learning), the outcome [reward prediction error (RPE): Unexpected Positive, Expected Positive, Unexpected Negative, Expected Negative], the identity of the chosen Cue or spatial Location, and whether the animal repeated that choice of cue or location on the next trial (Will Repeat Cue and Will Repeat Location). The RPE was determined as in our prior work (Asaad and Eskandar, 2011) by comparing the outcome of the current trial with that of the immediately preceding trial. In other words, if the current trial was correct and the prior trial was incorrect, the current trial was deemed to be unexpectedly correct and therefore reflected a positive RPE. Negative RPEs occurred when the current trial was incorrect whereas the prior trial was correct. Zero RPEs resulted when the current and prior trial had the same positive or negative outcome. This approach was feasible because, as we previously observed, the single preceding trial had the strongest influence on the animals' strategy. The linear models allowed for one-way interactions between these predictors. The response variable was a neuron's spike rate during the 500 ms feedback period.

To determine the similarity between a feedback-related representation of the cue and the earlier representation of that cue (when it was visually present in the cue array) during learning, we compared neuronal population activity in a single 200 ms bin starting 100 ms after feedback onset to that in a corresponding 200 ms bin starting 100 ms after cue array onset. These time bins were chosen because they captured the most common peak phasic response of neurons in this region, but the pattern and significance of results observed were insensitive to the precise time bins chosen. The similarity of population representations at those times was compared by computing the cosine between the population vectors as follows:

$$\frac{V_{cue} \cdot V_{feedback}}{\|V_{cue}\| \, \|V_{feedback}\|}$$

For this analysis, trials were sorted according to the trial number relative to achieving the learning criterion of four consecutive correct trials. These values were compared with the vector similarity recalculated for the condition in which spike counts were randomized across trials. This provides a baseline for this metric using the same distribution of neuronal activity within each session.

Results

We trained two nonhuman primates (M1 and M2) to perform a learning task that invoked both the structural and temporal credit-assignment problems. The task required the animals to learn (and subsequently relearn many times) which cue among four, presented earlier in a trial, signaled the correct spatial choice at the end of the trial (Fig. 1). We recorded 635 neurons (395 in M1; 240 in M2) from the dlPFC while animals performed this
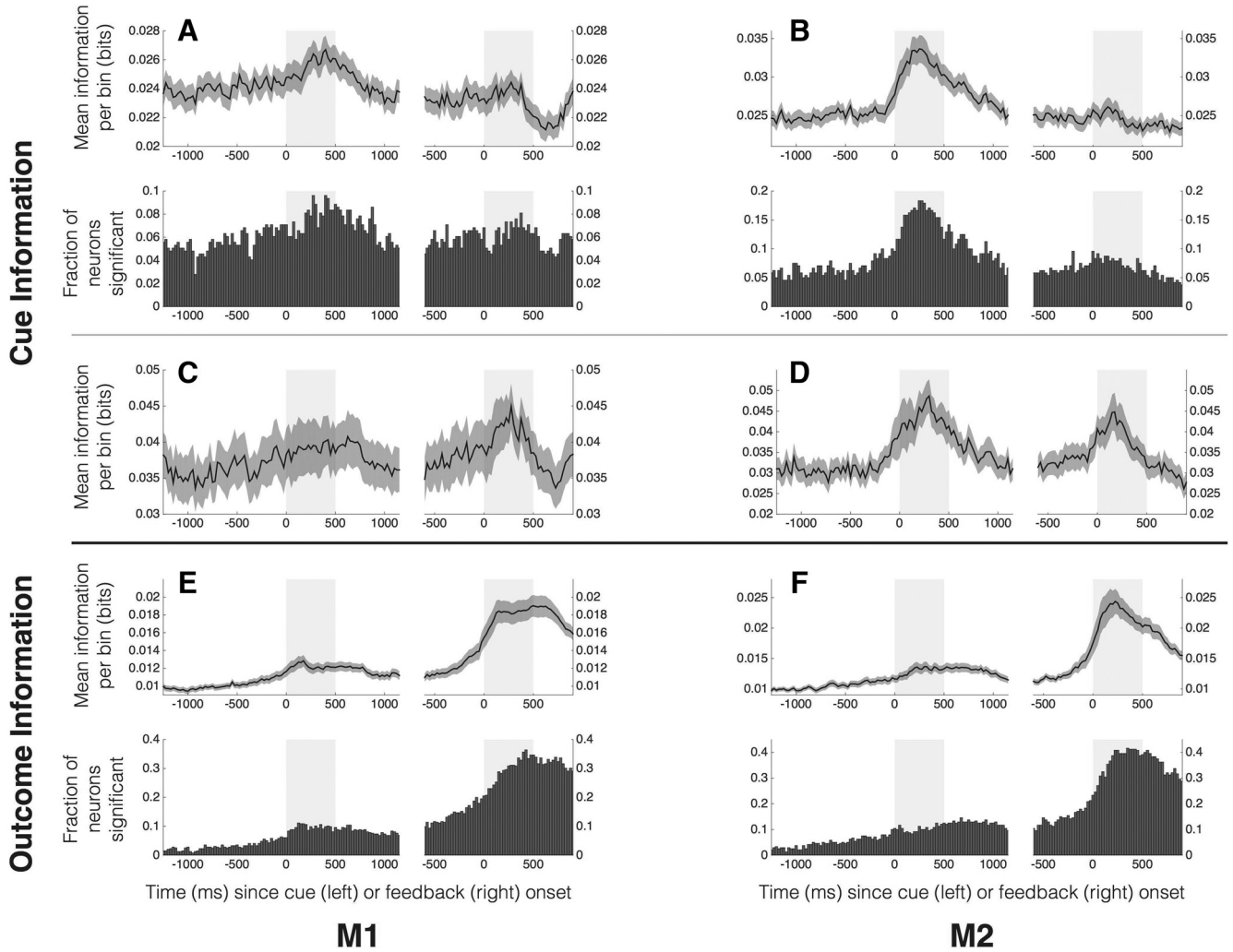
Figure 2. Cue-related and outcome-related information across neurons. *A–F*, The population-averaged information about cues (*A* and *B*) and outcomes (*E* and *F*) for each subject (M1: *A*, *C*, and *E*; M2: *B*, *D*, and *F*). In *C* and *D*, and in the top portion of *A*, *B*, *E*, and *F*, the information (in bits) is shown at each time point, averaged across all recorded neurons from each subject. The shaded region around each line indicated the mean ± SE. Note that information about the correct cue can be present throughout the trial because this is stable over an entire block. The bottom portions of *A*, *B*, *E*, and *F* show the number of neurons conveying significant information at each time point. *C* and *D* show the time course of information about the cues within just those neurons with significant cue-related information during the feedback period. Outcome-related information before the feedback reflects the prior trial's outcome, as previously described (Asaad and Eskandar, 2011). Note that the amount of information conveyed about cues or outcomes over the entire population is within the same order of magnitude (top portions of *A*, *B*, *E*, and *F*), but this information is distributed over many more neurons in the case of outcome representation (*C*, *D*, and bottom portions of *A*, *B*, *E*, and *F*).
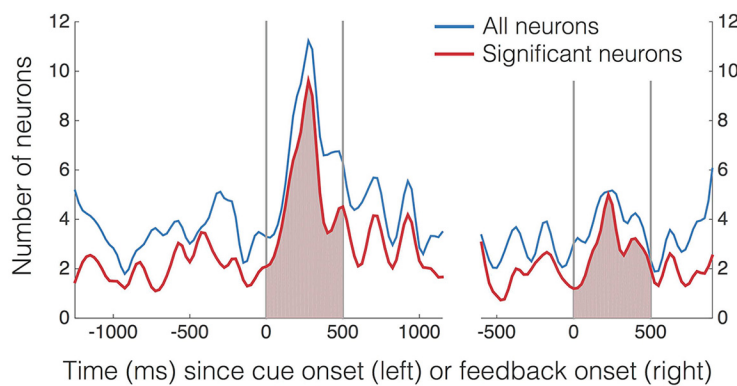

Figure 3. Timing of peak cue-related information. The time of maximal information about the cues is plotted for every neuron (blue) or only neurons carrying statistically significant information (red). For each neuron, the time of peak cue-related information was determined and added to this histogram to depict the number of neurons that conveyed their maximal cue-selective information at each time point. The shaded areas under the red line were integrated to obtain the numbers of neurons whose cue-related information was maximal in the cue or feedback epochs (see text). Note the local peak in the number of neurons whose maximal information about the cues was conveyed during the feedback period.

task to determine whether the activity of these neurons provided the necessary information for performing the credit assignment.

Behavior

A detailed analysis of the animals' behavior during this task was reported previously (Asaad and Eskandar, 2011). Three key observations are relevant here. First, animals tended to learn more from correct responses than from incorrect responses. In other words, they did not rely heavily on counterfactual reasoning (accumulated information about which choices are not correct from prior negative outcomes); rather, negative outcomes led to random or nearly random guessing. Second, their strategies were driven in large part by the immediately
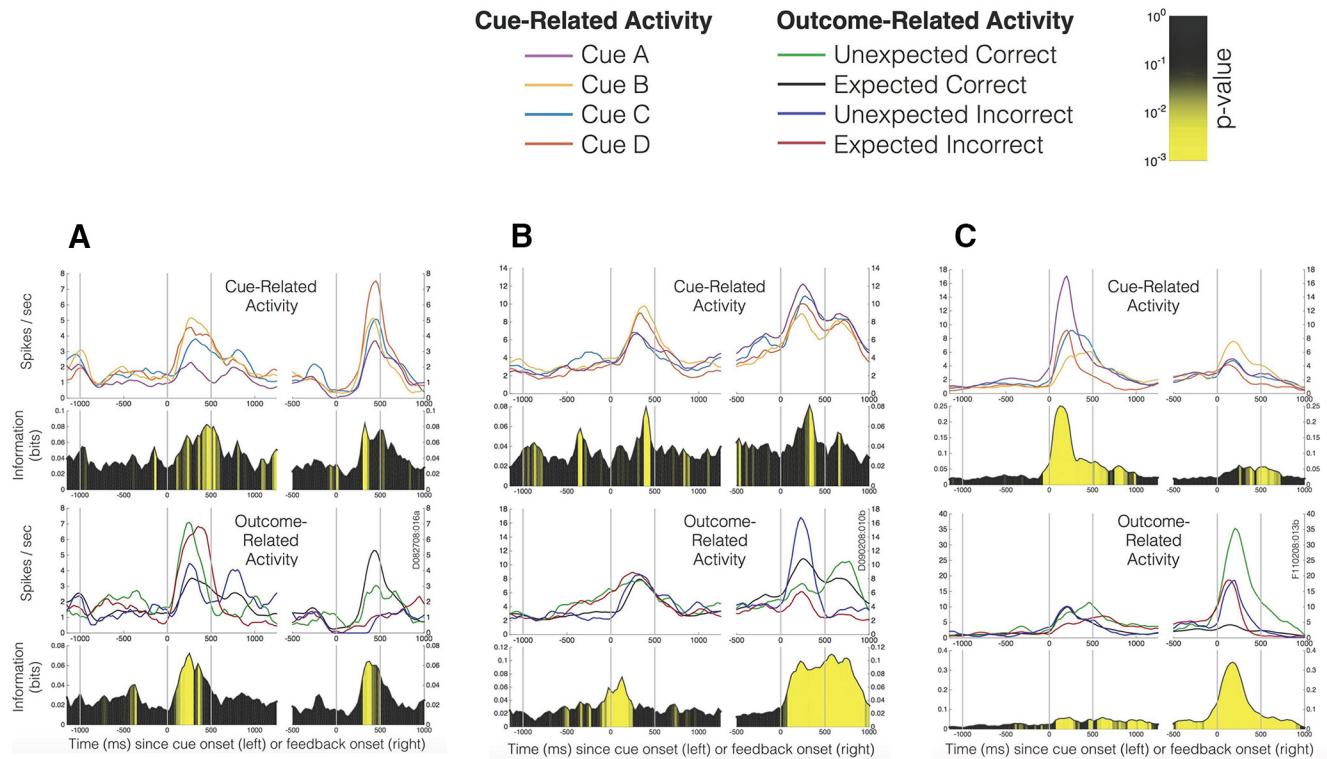
Figure 4. Examples of individual neurons with both cue and outcome selectivity. *A–C*, Three individual neurons. The top portion shows each neuron's activity sorted according to the cue picture that was designated correct, while the bottom portion shows those neurons' activities sorted according to the outcome. The top portion of each row shows the activity of these neurons in spikes per second, whereas the bottom portions show the information content in bits (assessed across the 4 cue or outcome exemplars; see Materials and Methods). The shading in the bottom portions reflects the significance of the information metric based upon a bootstrap reshuffling of the assignment of trials to conditions. Note that information about the correct cue picture could be present before the appearance of the cue array because a single cue was designated correct for an entire block. Note also that significant information about outcome can be present before the feedback epoch because the prior trial's outcome was reflected in the outcome categories used here (Asaad and Eskandar, 2011; Donahue and Lee, 2015).

preceding single trial's outcome. Third, superimposed on this, animals relied on a default spatial strategy; that is, they typically reselected a chosen spatial location on a subsequent trial in response to a positive outcome. They converted to a cue-based strategy (reselecting the location marked by a particular cue object) once the spatial strategy failed. This dynamic is shown in Figure 1C, which plots the probability of reselecting a particular cue or spatial location as a function of trial number within a block.

Last, because blocks were chosen in a pseudorandom fashion, we confirmed that animals did not use information from prior blocks to narrow their choices and accelerate learning in subsequent blocks. Specifically, examining the first cycle of four blocks in which animals learned to designate each of the four cue objects as correct, there was no significant difference in the average initial performance in each block (average percentage correct in the first five trials ranged from 28.0 to 30.9%, $p = 0.93$ by one-way ANOVA). So, although animals could have deduced by the fourth block the identity of the correct cue with 100% accuracy, they did not appear to use such a strategy, consistent with their more local lack of reliance on negative feedback to narrow choices within a block.

Neuronal representations at the time of feedback

Credit assignment first requires a representation of the relevant feature—here, the cue object that correctly signaled the rewarded saccade target—when the success or failure of a particular choice became evident. Critically, the generic feedback provided to the animals in this task to indicate a correct choice was identical across all conditions and so revealed nothing about the identity of the relevant cue; rather, information about the cue at the time of feedback must be internally generated by the animals to solve the task. Indeed, we observed that information about the correct cue was available throughout the trial (because the identity of the correct cue remained stable throughout a block), including during feedback (Fig. 2). More specifically, 161 of 635 individual prefrontal neurons (25.4% overall; 96 of 395 and 65 of 240 in each animal) conveyed significant information about the identity of the correct cue at some point during feedback about the outcome.

Because the calculation of information about the correct cue depended upon comparing spike rates across blocks of trials, slow changes in baseline neuronal activity (e.g., due to mechanical or physiological "drift") could potentially artifactually influence the number of neurons observed to carry information about this feature. Therefore, we repeated the information calculation after eliminating neurons observed to show significant information during a 500 ms "baseline" epoch preceding cue array onset (see Materials and Methods). This eliminated 35.1% of neurons (138 of 395 in M1; 85 of 240 in M2). Of the remaining neurons, 18.0% (74 of 412; 42 of 257 in M1; 32 of 155 in M2) continued to convey significant information about the cues during feedback (71% of the proportion observed across all neurons). Note that this is a conservative estimate, because the design of the task allowed that information about the cues could persist between trials, and so some of the excluded neurons could in principle have been conveying useful information even during the baseline epoch.

The importance of the cue representation during the feedback epoch was evident in examining the time course of information
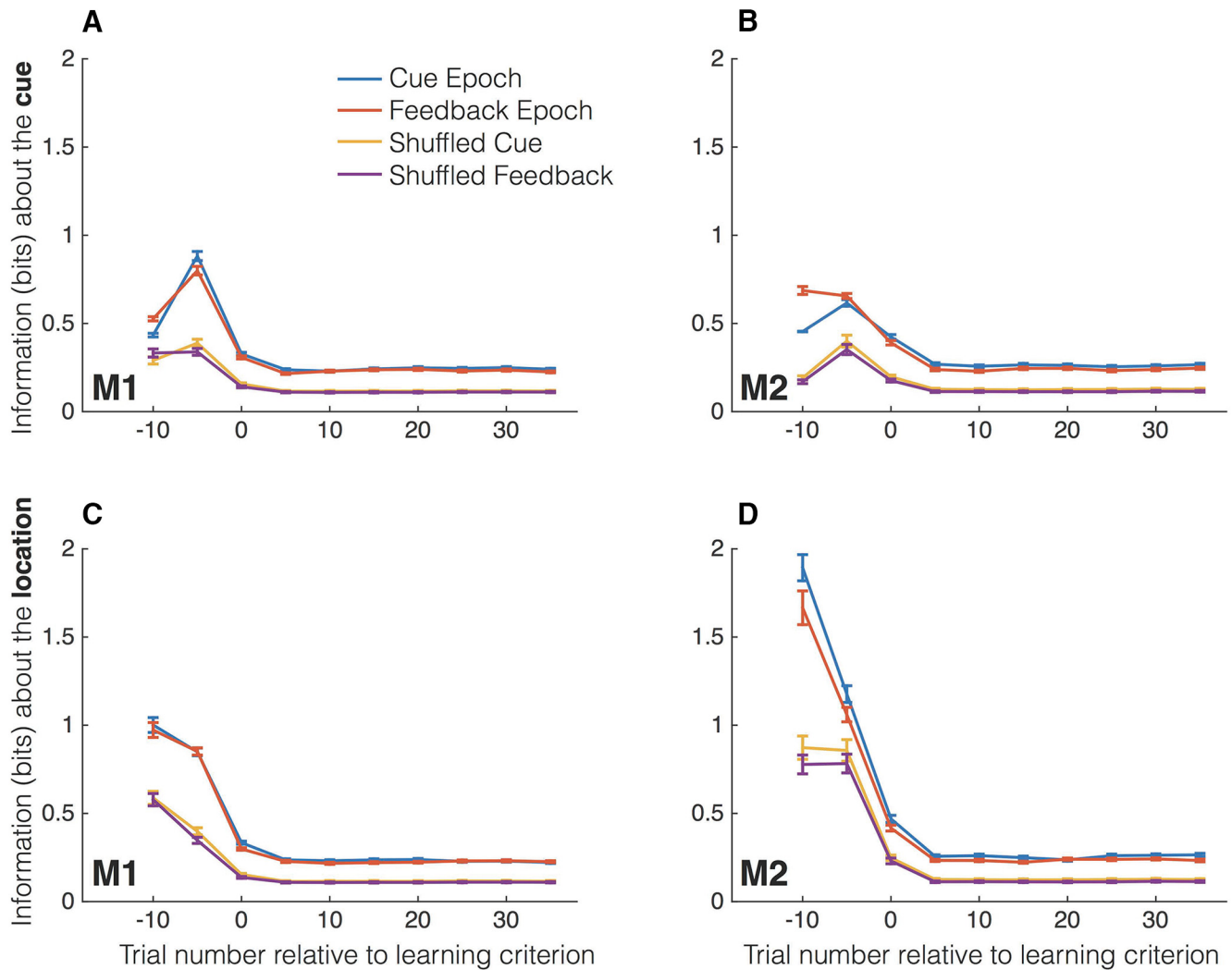
Figure 5. Cue and spatial information as a function of learning. ***A–D***, The information conveyed by neurons during the cue epoch (blue) or feedback epoch (red) about the selected cue (***A***, ***B***) or spatial location (***C***, ***D***) is plotted as a function of correct trial number relative to learning criterion (±SEs). The same metric calculated for data in which the assignments of trials to cue objects (***A***, ***B***) or spatial locations (***C***, ***D***) were shuffled are shown in yellow and purple, respectively. Data are shown separately for subjects M1 (***A***, ***C***) and M2 (***B***, ***D***). Note that the information measured here is an order of magnitude higher than observed in Figure 2, in large part due to differences in entropy in this calculation which considered fewer trials for each data point. Note the somewhat delayed peak in cue information relative to spatial information, which may reflect animals' initial reliance on a spatial strategy before switching to a cue-based strategy.

about the cues across a trial. When considering just those neurons with significant cue selectivity during the feedback period, a prominent second peak of information was evident during that epoch (Fig. 2C,D). Similarly, the feedback period exhibited a local maximum in the number of neurons that conveyed their maximal information about the cues during that time (Fig. 3). Specifically, the incidence of neurons that were maximally selective for the cues during feedback (52 of 635; 8.2% overall) was approximately half the incidence of neurons conveying maximal cue-related information during the cue interval itself (101 of 635; 15.9% overall). In other words, for a notable subgroup of neurons, the feedback period was the most potent task epoch to elicit cue-related information. This second peak in cue encoding at the time of feedback was consistent with amplification or reactivation of the required information when the contingent outcome became apparent.

While the presence of information about the correct cue at the time of feedback suggests neurons in the dlPFC could contribute information necessary for credit assignment, implicating them more directly in the credit assignment process requires a concurrent representation of the outcome. We previously observed that dlPFC neurons were highly engaged by trial outcomes in this task, such as whether a trial was correct or incorrect, and whether that outcome was surprising. In particular, most PFC neurons responded to unexpected outcomes, with approximately equal numbers activated by positive or negative reward-prediction errors (Asaad and Eskandar, 2011). Extending this, here we find that some neurons (123 of 635; 19.4% overall; 71 of 395 and 52 of 240 in each animal) conveyed information about both the outcome and the correct cue during feedback (Fig. 4). In fact, the magnitude of information for cues and outcomes was correlated such that neurons that conveyed more information about the outcome tended also to convey greater information about the cue object (M1: $r = 0.55$, $p < 0.001$; M2: $r = 0.38$, $p < 0.001$). This combined representation of outcome and relevant antecedent information is a necessary substrate for linking cause and effect for credit assignment, so neurons that integrate both types of information can participate directly in that function.

To understand how information about the animal's behavioral choice evolved with learning, we calculated the information
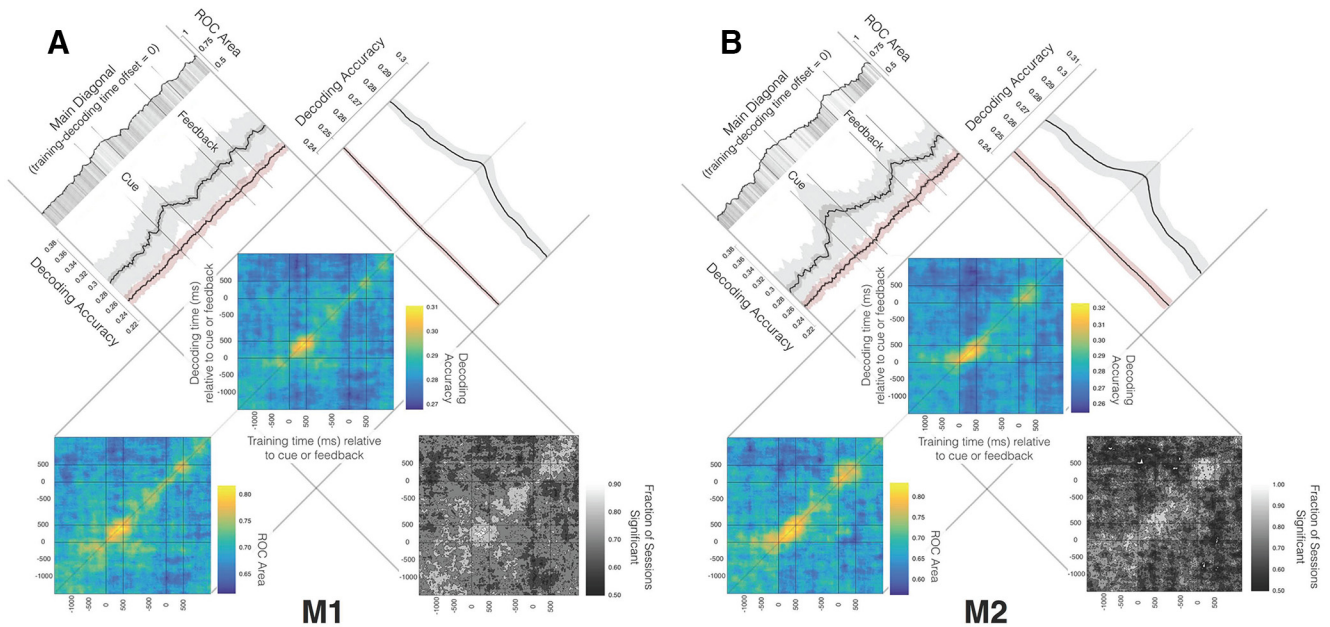
Figure 6. Cross-temporal decoding of cue-related neuronal activity. *A*, *B*, Population activity vectors from simultaneously recorded neurons were used to classify individual trials according to the correct cue (*A*, M1; *B*, M2). The accuracy of classification is depicted in the color scale of the central plot. The classifier was trained using a particular time bin (x-axis) and then tested against the same or different time bins (y-axis) from separate trials. Classification used a linear decoder that relied upon simply the minimum Euclidean distance between trained and tested vectors. The decoding accuracy when the same time bin was used for training and testing (across separate trials) resides in the main diagonal and is shown in the upper left (black line with gray areas representing SEs and SDs). A shuffled bootstrap procedure in which trials were randomly reassigned to cues was used to verify chance-level decoding (∼25% correct) in that circumstance (black line with red areas for SEs and SDs). The ROC results comparing actual versus shuffled decoding is shown at the bottom left, and the fraction of recording sessions with significant decoding according to the ROC shuffled bootstrap is shown at the bottom right. The far upper-left shows the ROC results along the main diagonal, with the shading corresponding to the fraction of significant sessions as in the bottom right. Cross-temporal decoding accuracy is depicted at the upper right, which is computed by taking the mean over each diagonal. The SDs and SEs are shown in light and dark gray, respectively (SEs may be imperceptible due to their small values). Note that while there is a peak in decoding accuracy when using nearby time bins (near the center of this plot), decoding accuracy does not return to chance even at large offsets between the training and decoding bins, necessitating some degree of stability in the neuronal representation across time. Exclusion of neurons with potentially unstable baseline activity (see Materials and Methods) did not significantly alter this result.

metric during the feedback period for short segments of trials organized according to stage of learning (see Materials and Methods). We found that information about the cue formerly at the chosen location and about the spatial location itself peaked very early during learning, and then rapidly declined once the correct cue object was learned (Fig. 5). In general, neurons conveyed more information about the spatial location of the choice during early learning, but the magnitude of information about the cue or spatial location stabilized at about the same low level once learning had occurred. Interestingly, the time course of information during early learning suggests that cue object information may peak somewhat later than spatial information, consistent with the animals' behavioral strategy, in which they tended to rely first on repeating a spatial choice but then switched to repeating a cue-based choice once that approach failed.

Figure 7. Decoding accuracy as a function of neuronal ensemble size. Decoding accuracies are plotted against the size of the corresponding neuronal ensembles for each session. The dots and lines represent the means across all training–decoding offsets ± SDs (SEs are too small to be visible). The lines depict the least-squares linear fit to each subject's data (M1: $r = 0.518$, $p = 0.002$; M2: $r = 0.673$, $p = 0.0008$). Note the y intercept for both animals is appropriately close to chance level (horizontal line, 0.25).

The stability of neuronal representations over time

If a representation is to be directly useful for credit assignment, it must be similar to the representation of the relevant feature itself, and not simply an independent, even if informative, representa-
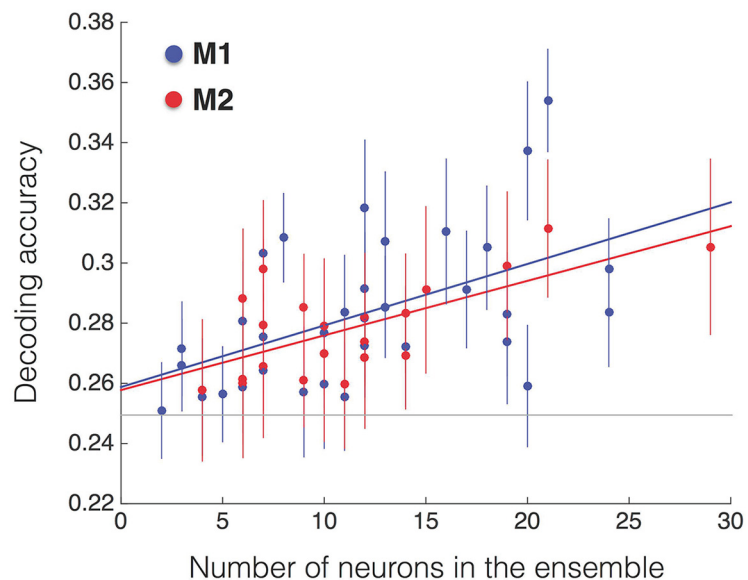
tion. In other words, neuronal activity might discriminate between the visual cues at the time of feedback, but do so in a manner independent of the representation that was evident at the time of cue presentation. Prior studies have indeed suggested that
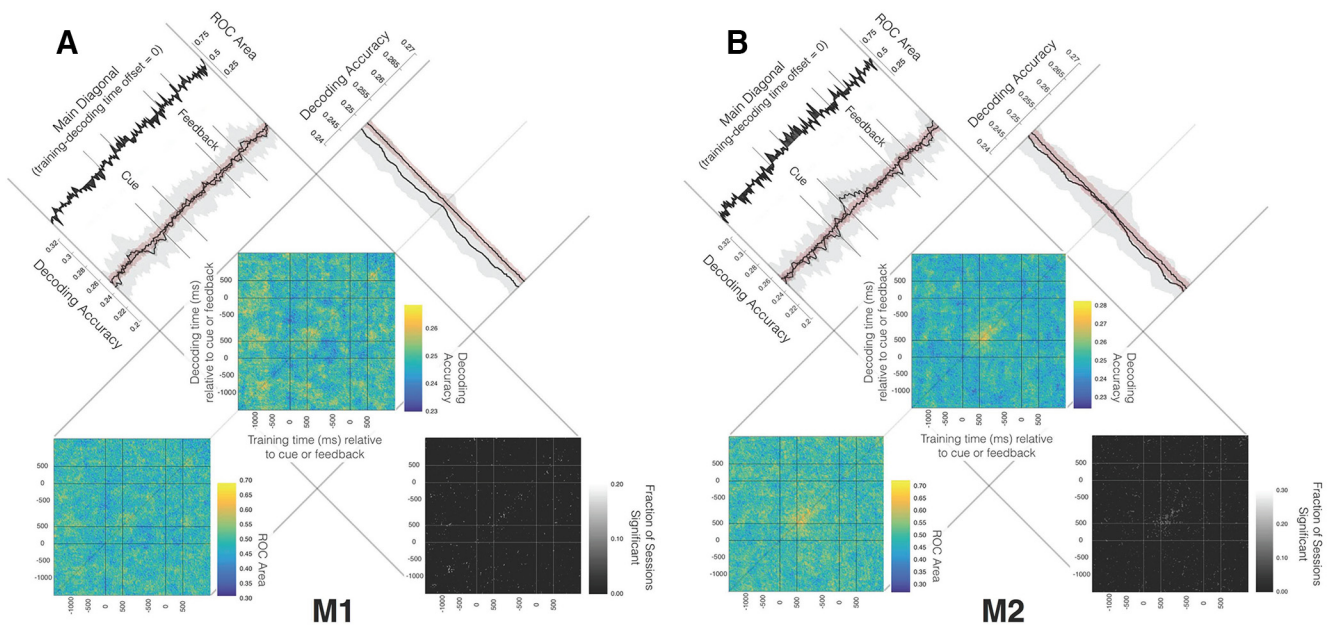
Figure 8. Cross-temporal decoding of cue-related activity in the spatial task. Conventions and methods are the same as in Figure 5. Here, the population decoding was applied to assess the amount of information conveyed by simultaneously recorded neurons about the cue during the spatial task, where the identity of the cue was irrelevant to learning. **A**, Data for M1. **B**, Data for M2.

the representation of stimuli over time is not stable in the PFC, such that the neuronal representation at one time is not related to the neuronal representation of the same stimulus at a different time (Meyers et al., 2008; Sigala et al., 2008; Barak et al., 2010; Stokes et al., 2013). However, in our task, a stable representation of the cue over time was necessary so that reinforcement mechanisms acting at the time of feedback could ultimately drive attention to that cue when it next appeared in the subsequent trial to properly guide an animal's choice.

The activity of single neurons is typically noisy, and so examining their individual responses to the cue during cue array presentation versus during feedback may yield little apparent similarity (Fig. 4). However, applying a simple, linear population decoding method to simultaneously recorded neurons revealed that there was in fact significant consistency across these representations (Fig. 6). Specifically, a decoder trained at the time of feedback could be applied to classify, above chance performance, the identity of a cue using neural activity at the time of cue array presentation and, conversely, a decoder trained during cue presentation could decode the identity of a cue according to its neural representation at the time of feedback.

Decoder performance depended on the number of neurons within each ensemble (Fig. 7). While classification accuracy was only modestly above chance, on average and in absolute terms, larger ensembles yielded more successful decoding. This suggests the observed results did not reflect ceiling performance, and improved accuracy would likely be achieved with larger populations of neurons. Note that the information being decoded here is dependent upon inferred covert attention to a particular object presented simultaneously with other objects in a cue array, rather than simply upon the identity of a single object that is the subject of direct fixation, and the data comprising this figure include the cross-temporal (off-diagonal) decoding bins. For comparison, if one considers only the bins on-diagonal, maximal decoding accuracy for a single session reaches ~38%; further limiting the decoding to only the time bins during cue array presentation yields a maximum decoding accuracy of ~46%. These

results approximately illustrate the maximum decoding accuracy achievable with this simple linear decoder applied to our dataset.

The context specificity of neuronal representations for credit assignment

Next, credit assignment requires a representation activated only when necessary to avoid misattribution of credit for successful or unsuccessful outcomes to available but irrelevant features. Therefore, we examined whether the feedback-related representation of the cue was actively engaged only when necessary, or if it reflected simply an automatic process triggered by the visual stimulus at the intended choice location. To do this, we compared neuronal activity between the credit assignment task and a "spatial" task in which the cues themselves were irrelevant for learning the correct response. Both tasks used the same visual stimuli and required the same motor responses. However, in the spatial task, the correct choice was determined solely by its spatial location rather than by the cue that had appeared earlier at that location. We found that neurons no longer conveyed the identity of the cue object when it was not necessary for credit assignment: Only 10 of 395 (2.5%) and 6 of 240 (2.5%) neurons in each animal reflected the correct cue at the time of feedback in the spatial task.

Likewise, application of the same population analysis, used above, to decode the identity of the cue at the chosen location in the spatial task yielded very different results compared with its application in the main credit-assignment task. Decoding accuracy was barely distinguishable from chance performance, and the ROC analysis revealed few significant time bins and few sessions with significant decoding performance (Fig. 8). These findings demonstrate that feedback-epoch representations of the correct cue were actively engaged only as needed, and are broadly consistent with prior results demonstrating active selection of relevant information for representation in the dlPFC (Rainer et al., 1998a; Everling et al., 2002; Lebedev et al., 2004).

A

M1 n = 395	Task	RPE	Cue	Location	Will Repeat Cue	Will Repeat Location		Total
Task	9 (2.3%)							59 (14.9%)
RPE	12 (3.0%)	19 (4.8%)						130 (32.9%)
Cue	20 (5.1%)	54 (13.7%)	31 (7.8%)					132 (33.4%)
Location	20 (5.1%)	54 (13.7%)	42 (10.6%)	37 (9.4%)				135 (34.2%)
Will Repeat Cue	6 (1.5%)	16 (4.1%)	16 (4.1%)	11 (2.8%)	2 (0.5%)			40 (10.1%)
Will Repeat Location	3 (0.8%)	14 (3.5%)	13 (3.3%)	11 (2.8%)	7 (1.8%)	5 (1.3%)		45 (11.4%)

B

M2 n = 240	Task	RPE	Cue	Location	Will Repeat Cue	Will Repeat Location		Total
Task	8 (3.3%)							51 (21.3%)
RPE	10 (4.2%)	22 (9.2%)						83 (34.6%)
Cue	18 (7.5%)	33 (13.8%)	19 (7.9%)					71 (29.6%)
Location	19 (7.9%)	42 (17.5%)	16 (6.7%)	36 (15.0%)				89 (37.1%)
Will Repeat Cue	2 (0.8%)	11 (4.6%)	10 (4.2%)	10 (4.2%)	5 (2.1%)			30 (12.5%)
Will Repeat Location	6 (2.5%)	10 (4.2%)	8 (3.3%)	9 (3.8%)	3 (1.3%)	1 (0.4%)		32 (13.3%)

Figure 9. Relationships between behavioral variables and feedback epoch activity. A linear model was fit to the feedback epoch spike rates of individual neurons to assess the influence of the animals' current and upcoming choices on neuronal activity (spike rate in the 500 ms feedback epoch). *A*, Results for M1. *B*, Results for M2. Predictor variables consisted of the Task (cue learning vs spatial learning), the outcome (RPE; see Materials and Methods), the identity of the chosen Cue or spatial Location, and whether the animal would repeat that choice of cue or location on the next trial (Will Repeat Cue and Will Repeat Location). The "Total" column on the right shows the number of neurons whose activity was found to be significantly ($p < 0.01$) dependent upon the factor listed at the left, either singularly or in interaction with another factor; numbers here may not sum to the simple totals taken from the left because neurons were counted only once even if they depended on a particular factor in more than one way (such as in 2 different interactions). Repeating this analysis while excluding neurons with potentially unstable baseline activity (see Materials and Methods) did not significantly alter any of these results (all values within ±4%).

The relationships between feedback-related neuronal activity and learning

To explore the relationship between feedback period activity and behavioral variables, including subsequent choices, we fit a linear model to the spike rates of each neuron during the 500 ms feedback epoch. We assessed the following predictors: Task (Object vs Spatial); RPE (outcome and expectedness); Cue object formerly at the chosen location (A, B, C, or D); Chosen spatial location (upper left, upper right, lower right, or lower left); and whether the animal will repeat the choice of that object or that spatial location on the subsequent trial. We allowed for one-way interactions between these potential predictors, and counted the number of neurons whose feedback-epoch activity significantly depended on these factors. We found that reward-prediction error, the object formerly at the chosen location and the chosen location itself were the factors that most often significantly influenced a neuron's feedback epoch response (Fig. 9). More than a third of neurons that conveyed information about the reward-prediction error did so in conjunction with a representation of the object (M1: 54 of 130, 41.5%; M2: 33 of 83, 39.8%). Interestingly, ~10–13% of neurons had feedback-related activity that predicted an animal
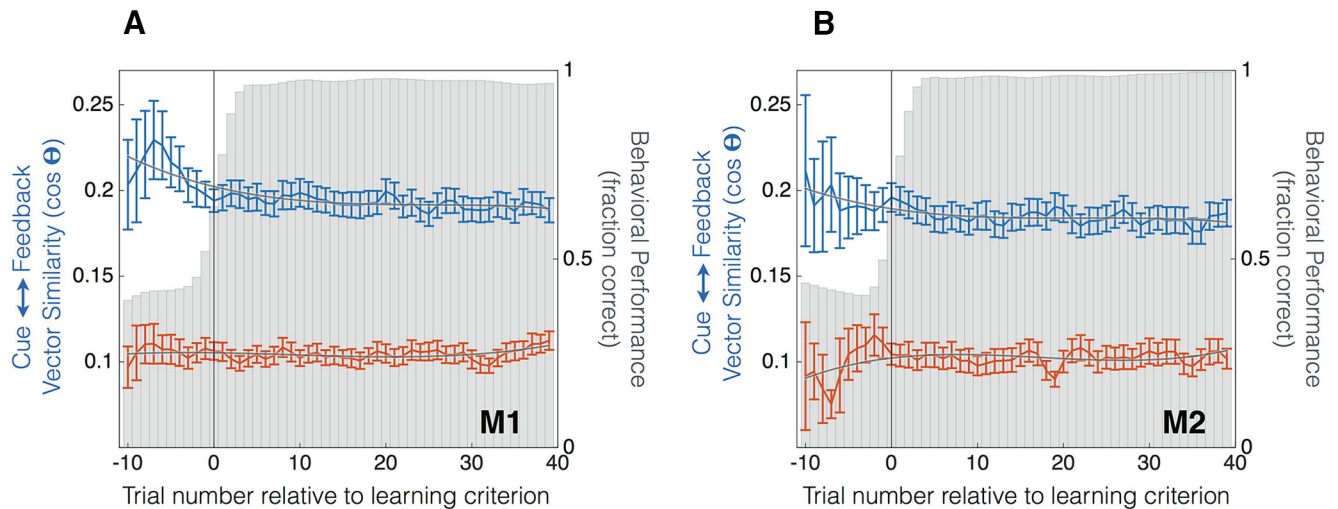
Figure 10. Cross-temporal fidelity of cue representations across the cue and feedback epochs during learning. *A*, *B*, The similarity of neuronal representations of the cue across time for subjects M1 (*A*) and M2 (*B*) was assessed by taking the cosine between population vectors derived from the cue and feedback epochs of correct trials and plotting this according to trial number relative to learning criterion (first of 4 consecutive correct trials). Shown is the mean vector similarity for each trial (blue) ± SE (left axis). A third-order polynomial fit is overlaid to depict the trend. The shuffled (control) vector similarity values are plotted in red. Concurrent behavioral performance is plotted as a bar graph in the background (right axis). Data are smoothed using a three-trial sliding average.

would subsequently reselect that cue object or spatial location on the next trial.

Finally, we examined whether the fidelity of the feedback representation was related to learning behavior. Specifically, we hypothesized that neural representations of a particular credit-deserving feature at different times within a trial—specifically, when that feature was actually present versus when the outcome was revealed—should be most similar during early learning, when credit assignment was needed to forge the link between these events. Therefore, we quantified the degree to which population activity at the time of feedback matched activity earlier in the trial, at the time of cue presentation. Applying a simple cosine vector similarity measure (see Materials and Methods) revealed that similarity was indeed greatest in earlier trials and decremented gradually as learning progressed (Fig. 10; for a simple linear fit for subject M1: $r = -0.65$, $p < 0.0001$; M2: $r = -0.44$, $p = 0.0012$; for both subjects combined: $r = -0.6972$, $p < 0.0001$), consistent with the particular importance of credit assignment during early learning. No relationship between similarity and learning was observed in the control analysis that recalculated the similarity measure using shuffled spike rates (M1: $r = 0.05$, $p = 0.19$; M2: $r = 0.19$, $p = 0.72$).

Discussion

We found that neuronal activity in the dlPFC fulfilled the necessary criteria for enabling a solution to the credit-assignment problem.

First, many neurons encoded the identity of a relevant cue at the time of feedback, even though it was no longer present. Information about the cues (and chosen locations) peaked during early learning, and rapidly settled to a lower level once learning was complete. Importantly, many of these neurons simultaneously encoded the outcome of a choice. This concurrent representation is necessary to link the outcome with the causal cue. The representation of information in other PFC areas may not share this selectivity. For example, neurons in the orbitofrontal cortex have been reported to reflect prior actions at the time of feedback, but do so regardless of the outcome (Tsujimoto et al., 2009), which is inconsistent with a central role in credit assignment; however, that study and another (Seo et al., 2007) did

observe encoding of prior actions with respect to prior or current reward in the dlPFC, analogous to our results showing comodulation of neuronal responses by preceding sensory stimuli and outcomes. Notably, the types of outcome modulation we observed here were consistent with a reward-prediction error, demonstrating that this critical learning signal can interact directly with the neuronal representations of to-be-learned features in the dlPFC.

The observation that neuronal activity represented the cues at the time of feedback is distinct from the "chosen-value" representations previously observed throughout the PFC (Padoa-Schioppa, 2009, 2013; Sul et al., 2010; Kennerley et al., 2011; Donahue and Lee, 2015). Specifically, all the cues in our experiment were associated with identical reward, so neural activity could differentiate the cues solely by their visual features. The ability to link outcomes with potentially causative stimuli based upon their identity (more than simply reflecting their differential value) is a key requirement for credit assignment.

The application of a multivariate linear model to the feedback period activity of dlPFC neurons revealed that not only could feedback-related activity reflect the choice outcome and task features, such as the relevant cue object or chosen spatial location, but that in some neurons this activity predicted whether animals would reselect those features on the subsequent trial, consistent with the notion that this activity contributes to learning behavior.

Second, the neuronal representation of the cues at the time of feedback were sufficiently similar to the representation of the same cues at the time of their actual presentation, earlier in the trial, such that the identity of the correct cue could be determined from ensemble activity using a decoder trained at a different time within the trial. A stable representation over time can facilitate the linking of a behavioral outcome with an earlier causal feature. While statistically significant and stable population decoding was found throughout the trial and across broad temporal offsets, the magnitude of this decoding was modest, on the order of 3–5% improvements (in absolute terms) over chance performance, on average. Some of this was likely due to the relatively small ensembles of simultaneously recorded neurons considered, and some of this likely results from the use of a very simple linear decoder.

Therefore, decoding performance with larger ensembles and using more sophisticated, biologically plausible classifiers is likely to be significantly better (Pouget et al., 2000), though this remains to be tested directly in the context of credit assignment.

Third, when the same neurons were recorded in a task with identical visual and motor elements, but with a different rule that rendered the identity of the cue object irrelevant for learning, the cues were no longer strongly represented in neuronal activity at the time of feedback, neither within individual neurons nor across simultaneously recorded populations of neurons, demonstrating that the cue representation was actively engaged when necessary for learning.

Last, the fidelity of the feedback representation, compared with the earlier representation of the cues, was greatest when credit assignment was most critical, in the earlier trials of each block when learning occurred. These data suggest that stability of the neural code may be dynamically modulated by the needs of the task and state of learning.

Together, these results are consistent with the notion that neurons in the dlPFC provide the necessary selective and stable representation of relevant features at the time of feedback to enable credit assignment.

Previous work found that little or no similarity exists in the dlPFC neural representation of particular features across time (Meyers et al., 2008; Sigala et al., 2008; Barak et al., 2010; Stokes et al., 2013), in contrast to our current results. One possible reason for this difference may lie in the nature of the behavioral task. Our task required "on-line" learning and relearning, and explicitly required difficult credit assignment to achieve this learning; meanwhile, prior work used a well learned task in which animals had extensive experience with the particular cues, rules, and discriminations. As we previously observed, the magnitude of dlPFC neuronal activity related to learning decreases over extended experience (Asaad et al., 1998), and here we observed that the cross-temporal representational similarity also decreases quite rapidly over at least a few tens of trials; more dissimilarity may develop gradually over time, perhaps even to the point where no apparent similarity remains. Therefore, the examination of representational similarity during early, on-line learning and the use of a task that required representational stability for credit assignment may have been critical to observe stability in the neural code. Importantly, our results do not argue for a lack of representational "drift" as information is passed among neuronal ensembles, but simply demonstrate that significant information can indeed persist in a stable form.

The persistence of some aspect of neuronal activity related to the credit-deserving feature into the time of delayed feedback is known more generally as an eligibility trace (Sutton and Barto, 1998), and reinforcement interacting with this eligibility trace is what confers the selectivity necessary for proper credit assignment. While spiking activity, as we observed here, may provide an overt eligibility trace, additional nonspiking aspects of neuronal function may nevertheless contribute to this process (Fiete and Seung, 2006; Izhikevich, 2007; Urbanczik and Senn, 2009). Our results show, however, that stable spiking activity is indeed one viable mechanism for solving the temporal credit-assignment problem.

Credit assignment is undoubtedly a complex process to which a variety of brain regions contribute key components. For example, previous work has implicated other areas of the PFC as well as the parietal cortex. Specifically, the fMRI BOLD signal in the lateral orbitofrontal cortex of monkeys was observed to correlate with win–stay/lose–shift (contrasted with win–shift/lose–stay)

behavior, reflecting successful choices that depended upon proper credit assignment (Chau et al., 2015), and lesions of the lateral orbitofrontal cortex led to a "spread of effect" (Ogden, 1933), whereby credit was misattributed to preceding events (Noonan et al., 2010). In humans, BOLD signals correlated with attribution of credit to attended versus nonattended cues were observed in the medial and orbitofrontal cortices (Akaishi et al., 2016); our study did not directly assess potential differences in the assignment of credit to attended versus nonattended options because there was no direct measure of the locus of attention, and these animals generally learned much more from positive than from negative feedback in this task (Asaad and Eskandar, 2011), limiting our ability to examine substrates of counterfactual reasoning. Meanwhile, neurons in the lateral intraparietal sulcus responded more vigorously after a choice that should have been assigned credit for a delayed reward, regardless of the timing of that event within a sequence of choices (Gersch et al., 2014), suggesting these events may be "tagged" for later reference (perhaps through the augmentation of an eligibility trace that determines the magnitude of an event's contribution in subsequent credit assignment).

Importantly, credit assignment may not be a unitary process. There may be both implicit and explicit learning mechanisms that operate in parallel to enable solutions to the credit-assignment problem (Fu and Anderson, 2008), and so these dlPFC neuronal representations may ultimately contribute to one or multiple processes.

Our results show that information necessary to perform credit assignment resides in the dlPFC. To what extent other cortical or subcortical areas contribute inputs to this process, are recipients of information computed here, or interact more dynamically to enable credit assignment across a broader circuit is not yet clear, and should be the subject of future work.

References

Akaishi R, Kolling N, Brown JW, Rushworth M (2016) Neural mechanisms of credit assignment in a multicue environment. J Neurosci 36:1096–1112. CrossRef Medline

Asaad WF, Eskandar EN (2008) A flexible software tool for temporally precise behavioral control in Matlab. J Neurosci Methods 174:245–258. CrossRef Medline

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

Asaad WF, Rainer G, Miller EK (1998) Neural activity in the primate prefrontal cortex during associative learning. Neuron 21:1399–1407. CrossRef Medline

Asaad WF, Santhanam N, McClellan S, Freedman DJ (2013) High-performance execution of psychophysical tasks with complex visual stimuli in MATLAB. J Neurophysiol 109:249–260. CrossRef Medline

Barak O, Tsodyks M, Romo R (2010) Neuronal population coding of parametric working memory. J Neurosci 30:9424–9430. CrossRef Medline

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple tests. J R Stat Soc B 85:289–300.

Chau BK, Sallet J, Papageorgiou GK, Noonan MP, Bell AH, Walton ME, Rushworth MF (2015) Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. Neuron 87:1106–1118. CrossRef Medline

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173. CrossRef Medline

Dias R, Robbins TW, Roberts AC (1997) Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. J Neurosci 17:9285–9297. 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

Everling S, Tinsley CJ, Gaffan D, Duncan J (2002) Filtering of neural signals by focused attention in the monkey prefrontal cortex. Nat Neurosci 5:671–676. CrossRef Medline

Fiete IR, Seung HS (2006) Gradient learning in spiking neural networks by dynamic perturbation of conductances. Phys Rev Lett 97:048104. CrossRef Medline

Fu WT, Anderson JR (2008) Solving the credit assignment problem: explicit and implicit learning of action sequences with probabilistic outcomes. Psychol Res 72:321–330. CrossRef Medline

Funahashi S, Bruce CJ, Goldman-Rakic PS (1993) Dorsolateral prefrontal lesions and oculomotor delayed-response performance: evidence for mnemonic "scotomas". J Neurosci 13:1479–1497. Medline

Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. Science 173:652–654. CrossRef Medline

Gersch TM, Foley NC, Eisenberg I, Gottlieb J (2014) Neural correlates of temporal credit assignment in the parietal lobe. PLoS One 9:e88725. CrossRef Medline

Gochin PM, Colombo M, Dorfman GA, Gerstein GL, Gross CG (1994) Neural ensemble coding in inferior temporal cortex. J Neurophysiol 71:2325–2337. Medline

Izhikevich EM (2007) Solving the distal reward problem through linkage of STDP and dopamine signaling. Cereb Cortex 17:2443–2452. CrossRef Medline

Kennerley SW, Behrens TE, Wallis JD (2011) Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat Neurosci 14:1581–1589. CrossRef Medline

Kovach CK, Daw ND, Rudrauf D, Tranel D, O'Doherty JP, Adolphs R (2012) Anterior prefrontal cortex contributes to action selection through tracking of recent reward trends. J Neurosci 32:8434–8442. CrossRef Medline

Lara AH, Wallis JD (2014) Executive control processes underlying multi-item working memory. Nat Neurosci 17:876–883. CrossRef Medline

Lara AH, Wallis JD (2015) The role of prefrontal cortex in working memory: a mini review. Front Syst Neurosci 9:173. CrossRef Medline

Lebedev MA, Messinger A, Kralik JD, Wise SP (2004) Representation of attended versus remembered locations in prefrontal cortex. PLoS Biol 2:e365. CrossRef Medline

Mansouri FA, Buckley MJ, Tanaka K (2007) Mnemonic function of the dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. Science 318:987–990. CrossRef Medline

Markowitz DA, Curtis CE, Pesaran B (2015) Multiple component networks support working memory in prefrontal cortex. Proc Natl Acad Sci U S A 112:11084–11089. CrossRef Medline

Meyers EM, Freedman DJ, Kreiman G, Miller EK, Poggio T (2008) Dynamic population coding of category information in inferior temporal and prefrontal cortex. J Neurophysiol 100:1407–1419. CrossRef Medline

Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci 16:5154–5167. Medline

Noonan MP, Walton ME, Behrens TE, Sallet J, Buckley MJ, Rushworth MF (2010) Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc Natl Acad Sci U S A 107:20547–20552. CrossRef Medline

Ogden RM (1933) Thorndike's proof of the law of effect. Science 77:240. CrossRef Medline

Padoa-Schioppa C (2009) Range-adapting representation of economic value in the orbitofrontal cortex. J Neurosci 29:14004–14014. CrossRef Medline

Padoa-Schioppa C (2013) Neuronal origins of choice variability in economic decisions. Neuron 80:1322–1336. CrossRef Medline

Parker A, Gaffan D (1998) Memory after frontal/temporal disconnection in monkeys: conditional and nonconditional tasks, unilateral and bilateral frontal lesions. Neuropsychologia 36:259–271. CrossRef Medline

Pasternak T, Lui LL, Spinelli PM (2015) Unilateral prefrontal lesions impair memory-guided comparisons of contralateral visual motion. J Neurosci 35:7095–7105. CrossRef Medline

Pouget A, Dayan P, Zemel R (2000) Information processing with population codes. Nat Rev Neurosci 1:125–132. CrossRef Medline

Rainer G, Asaad WF, Miller EK (1998a) Selective representation of relevant information by neurons in the primate prefrontal cortex. Nature 393:577–579. CrossRef Medline

Rainer G, Asaad WF, Miller EK (1998b) Memory fields of neurons in the primate prefrontal cortex. Proc Natl Acad Sci U S A 95:15008–15013. CrossRef Medline

Reveley C, Gruslys A, Ye FQ, Glen D, Samaha J, E Russ B, Saad Z, K Seth A, Leopold DA, Saleem KS (2016) Three-dimensional digital template atlas of the macaque brain. Cereb Cortex. Advance online publication. Retrieved June 21, 2017. CrossRef Medline

Rossi AF, Pessoa L, Desimone R, Ungerleider LG (2009) The prefrontal cortex and the executive control of attention. Exp Brain Res 192:489–497. CrossRef Medline

Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE (2011) Frontal cortex and reward-guided learning and decision-making. Neuron 70:1054–1069. CrossRef Medline

Saad ZS, Reynolds RC, Argall B, Japee S, Cox RW (2004) SUMA: an interface for surface-based intra- and inter-subject analysis with AFNI. Paper presented at 2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano, Arlington, VA, April.

Samengo I (2002) Information loss in an optimal maximum likelihood decoding. Neural Comput 14:771–779. CrossRef Medline

Seo H, Barraclough DJ, Lee D (2007) Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. Cereb Cortex 17 [Suppl 1]:i110–i117. CrossRef Medline

Sigala N, Kusunoki M, Nimmo-Smith I, Gaffan D, Duncan J (2008) Hierarchical coding for sequential task events in the monkey prefrontal cortex. Proc Natl Acad Sci U S A 105:11969–11974. CrossRef Medline

Simmons JM, Minamimoto T, Murray EA, Richmond BJ (2010) Selective ablations reveal that orbital and lateral prefrontal cortex play different roles in estimating predicted reward value. J Neurosci 30:15878–15887. CrossRef Medline

Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J (2013) Dynamic coding for cognitive control in prefrontal cortex. Neuron 78:364–375. 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

Sutton R, Barto A (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT.

Tsujimoto S, Genovesio A, Wise SP (2009) Monkey orbitofrontal cortex encodes response choices near feedback time. J Neurosci 29:2569–2574. CrossRef Medline

Urbanczik R, Senn W (2009) Reinforcement learning in populations of spiking neurons. Nat Neurosci 12:250–252. CrossRef Medline

Victor JD, Purpura KP (1996) Nature and precision of temporal coding in visual cortex: a metric-space analysis. J Neurophysiol 76:1310–1326. Medline