MIT Libraries | DSpace@MIT

MIT Open Access Articles

Learning substrates in the primate prefrontal cortex and striatum: = activity related to successful actions

The MIT Faculty has made this article openly available. ***Please share*** how this access benefits you. Your story matters.

Citation: Histed, Mark H., Anitha Pasupathy, and Earl K. Miller. "Learning Substrates in the Primate Prefrontal Cortex and Striatum: Sustained Activity Related to Successful Actions." Neuron 63.2 (2009): 244–253. Web. 5 Apr. 2012.

As Published: http://dx.doi.org/10.1016/j.neuron.2009.06.019

Publisher: Elsevier

Persistent URL: http://hdl.handle.net/1721.1/69950

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike 3.0

Massachusetts Institute of Technology

NIH Public Access
Author Manuscript
Neuron. Author manuscript; available in PMC 2010 July 30.

Published in final edited form as:
Neuron. 2009 July 30; 63(2): 244–253. doi:10.1016/j.neuron.2009.06.019.

Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions

Mark H. Histed[†,*], **Anitha Pasupathy**[*], and **Earl K. Miller**
The Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Summary

Learning from experience requires knowing whether a past action resulted in a desired outcome. The prefrontal cortex and basal ganglia are thought to play key roles in such learning of arbitrary stimulus-response associations. Previous studies have found neural activity in these areas, similar to dopaminergic neuron signals that transiently reflect whether a response is correct or incorrect. However, it is unclear how this transient activity, which fades in under a second, influences actions that occur much later. Here we report sustained outcome-related responses in single neurons of both areas, which last for several seconds until the next trial. Moreover, the outcome on a single trial influences the neural activity and behavior on the next trial: behavioral responses are more often correct and single neurons more accurately discriminate between the possible responses when the previous trial was correct. These long-lasting signals about trial outcome provide a way to link one action to the next, and may allow reward signals to be combined over time to implement successful learning.

Introduction

Both the lateral prefrontal cortex (PFC) and the caudate nucleus (Cd) of the basal ganglia have been implicated in learning abstract associations. Anatomically, these two regions are extensively interconnected with each other and the rest of the brain, including sensory, motor, and higher-level associational areas (Petrides and Pandya, 2006; Petrides and Pandya, 2007; Wise et al., 1996; Passingham, 1995; Fuster, 1997). They are thus well-positioned to control complex behavior. Frontal cortical areas and basal ganglia nuclei are interconnected in parallel “loops” (Houk and Wise, 1995; Alexander et al., 1986; Middleton and Strick, 2000; Alexander et al., 1990), suggesting close interaction during their function. Further, the deactivation or manipulation of neural function in these two areas affects learning behavior, showing both areas to be necessary for learning (Fellows and Farah, 2005; Petrides, 1985; Petrides, 1994; Gaffan and Harrison, 1989; Murray et al., 2000; Nakamura and Hikosaka, 2006b; Miyachi et al., 1997; Williams and Eskandar, 2006; Nakamura and Hikosaka, 2006a).

[†]To whom correspondence should be addressed. Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Goldenson, Rm 243, mark_histed@hms.harvard.edu, Phone: +1 617 432 1326, Fax: +1 (617) 734-7557.
[*]These authors contributed equally to this work.
Current address for AP: Department of Biological Structure, University of Washington, Seattle, WA, USA

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Neurophysiological studies in the PFC and Cd have also linked neuronal responses in both areas to flexible learning. These studies have demonstrated that information about the stimuli, the behavioral responses, and the association between the two are encoded by neurons in the PFC and Cd (Asaad et al., 1998; Pasupathy and Miller, 2005). During such learning, moreover, PFC and Cd neurons modify their activity to more strongly reflect this acquired knowledge about the learned association (Chen and Wise, 1995; Pasupathy and Miller, 2005; Asaad et al., 1998; Barnes et al., 2005; Murray et al., 2000). Finally, activity of many PFC and Cd neurons reflect task outcome or the delivery of reward — another important piece of information critical for guiding learning. Learning depends on using feedback from the environment about the outcome of actions, and in the laboratory this feedback is typically the delivery of a food reward for desired (correct) behavior. Neural signals related to reward are closely associated with the midbrain dopaminergic system, whose neurons fire transiently in relation to reward delivery (Ljungberg et al., 1992; Schultz et al., 1993a). These neurons project to many brain areas, but they strongly innervate the basal ganglia and the PFC (Anden et al., 1966; Berger et al., 1988; Williams and Goldman-Rakic, 1993). Unsurprisingly, then, PFC and BG neurons have been found to show activity after reward delivery and behavioral response feedback (Schultz et al., 1993b; Schmitzer-Torbert and Redish, 2004; Barraclough et al., 2004; Barnes et al., 2005; Apicella et al., 1991; Ichihara-Takeda and Funahashi, 2006; Watanabe, 1989; Lau and Glimcher, 2007; Seo et al., 2007).

Thus, in the PFC and BG, neural correlates of both outcome and learning have been documented, but it is still unclear how these interact and whether the outcome related signals are used to modify neural activity and behavior. This is because reward-related activity occurs at the end of the trial and has mainly been reported to be quite transient. Reward responses last just a few hundred milliseconds after the delivery (or withholding) of the reward (Lau and Glimcher, 2007), whereas the next opportunity for behavior (i.e., the next behavioral trial) and the associated task-related neural activity is typically seconds away. Thus, it has been unclear how such temporally disparate signals interact, though several ideas have been put forth, especially in the context of neuroeconomics (e.g. Rangel et al., 2008; Montague and Berns, 2002; Rushworth and Behrens, 2008; Doya, 2008). Specifically, in the frontal cortex, some studies show that past reward history can modulate task-related activity (Barraclough et al., 2004; Uchida et al., 2007; Seo et al., 2007; Seo and Lee, 2009). In the hippocampus, this has been seen in some outcome-modulated neurons (Wirth et al., 2009). However, little is known about the transient vs. sustained nature of outcome-related single neurons in learning tasks, and the role of the basal ganglia has not been explored.

How neurons encode trial outcome — transiently or by sustained firing — has important implications for the mechanism of learning. Prior work has suggested two ways that learning might occur. First, the outcome of previous trials could be stored in synaptic strengths, represented by a connection weight in a neural network model. The transient responses to reward would then be used to change synapses after each trial, affecting on the next trial only neurons' excitability or responses (Barraclough et al., 2004; Sugrue et al., 2005). This would be supported by transient reward responses. However, there is a second possibility: The outcome of each trial might be stored in the sustained firing patterns of the neurons. Then, the dynamic state of the network could store the learned association without any required change in synaptic strength (Maass et al., 2002; Ganguli et al., 2008). Outcome-related activity sustained until the next trial could then be combined with the learned representations to select the next action. This latter model predicts that sustained neural firing related to outcome should be observed between trials and the learning induced changes will be evident on the next trial. But until now, no such sustained reward-related firing has been observed in these areas.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Here, we report data that supports this second model, shedding light on the neural mechanisms linking environmental feedback to neural plasticity by showing that learning can indeed be implemented by changes in network state. As animals learned associations between visual stimuli and saccade responses, we studied the responses of neurons in the PFC and Cd. We found that the activity of many neurons in the PFC and Cd reflects the delivery or withholding of reward (i.e., whether a trial was correct or incorrect). This activity can be sustained, and we observed that it often lasts for several seconds, the entire period between trials. Finally, we found that the outcome of a single trial also did impact the neural representation of the learned association, as if information about outcome was being combined with task information to cause learning-related changes. Response selectivity was stronger on a given trial if the previous trial had been rewarded and weaker if the previous trial was an error. This was independent of whether the animal had just begun to learn the association or was already quite good at it. Together, these results describe how learning in PFC and Cd is shaped by behavioral outcome signals.

Results

In order to assess how outcome signals could be used to guide learning, we trained animals to perform an associative learning task. Animals learned arbitrary associations between each of two picture cues, both new each day, and a leftward or rightward eye movement response (Fig. 1). The task and behavioral performance are described in detail in Pasupathy and Miller (2005). Animals learned the association by trial and error, and once they were performing well (>90% on each picture; see Methods), the associations were reversed without any explicit cue. By repeatedly reversing the associations, we could examine multiple instances of learning and relearning. Animals performed at least three reversals during each recording session.

We found that the activity of many neurons reflects the behavioral outcome (correct vs. error) in both the PFC and Cd. Single neurons in both the PFC (Fig. 2A) and Cd (Fig. 2B) show immediate changes in activity based on whether the behavioral response was correct or an error. Some neurons show an increase in activity after corrects (Fig 2, left column) while others show an increase in activity after errors (Fig 2, right column). In the PFC and Cd, both types of responses are roughly equal in number (greater for correct: 54%, 101/186 in Cd; 47%, 112/237 in PFC; amongst cells modulated by outcome in the first 500ms after response, $p<0.05$ via non-parametric ANOVA; see also Supp. Fig. 3).

PFC and Cd neurons maintain outcome information in sustained activity

Neurons in both the PFC and Cd are known to show transient responses to rewarding stimuli (e.g. Schultz et al. 1993b). However, it is not known how information about previous actions might be carried in the brain from one trial to the next, so that it can be used in learning. We found that many neurons in both PFC and Cd carry this sustained information. Single neurons in both areas convey strong, sustained outcome information across the entire 4–6 second inter-trial interval (Fig. 2C–D).

We used a tuning index, computed from the area under a receiver operating characteristic (ROC) curve, to quantify the outcome information carried by different neurons in the population. To measure the time course of the outcome-related selectivity, we computed this index in a sliding time window, 200 ms long (Fig. 2, bottom panels: A3, B3, C3, D3). If outcome-related selectivity is low, a neuron's firing will be identical after correct and after error, and the ROC area will be 0.5. In contrast, if a neuron perfectly encodes whether a response is correct or not, the ROC area will be 1. This analysis showed that sustained information about outcome is present in an average over the population in each brain region (Figure 2E). Further, we found that information about outcome peaks shortly after the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

reward and lasts until the next trial. In summary, we found that in both PFC and Cd, neurons carry information about the outcome of previous trials until the next trial, where this signal is available for guiding the animals' next response.

Single correct responses increase direction selectivity on the next trial

We also found that the outcome of one trial strongly impacts how much information neurons carry on the next trial about the learned association. We have previously shown that in this task, PFC and Cd responses carry association information through selectivity for the direction of the learned response (Pasupathy and Miller, 2005). Here, we describe how this direction selectivity on a given trial is altered by the outcome on a preceding trial. Specifically, we found that a correct trial increases direction selectivity on the next trial, while incorrect trials reduce it (Fig. 3). An example PFC neuron (Fig. $3A_1$) illustrates these effects. This neuron reflects both the outcome of the previous trial and the learned direction response. These effects are quantified, respectively, by an outcome ROC (Fig. $3A_3$) and a direction ROC (Fig. $3A_4$). The neuron fires at a higher rate when the previous trial is correct than if it was an error. Simultaneously, it encodes the learned association — it also fires more when the upcoming saccade is rightward than leftward. And, the strength of the association selectivity depends on the outcome of the previous trial, because after a correct response this selectivity is stronger. An illustrative example from the Cd (Fig. 3B) also shows stronger direction selectivity after a correct behavioral response. In these neurons, i) information about the outcome from the preceding trial is available on the subsequent trial and, ii) neural activity that reflects the upcoming learned behavior is modulated by the previous trial's outcome.

This increase in direction selectivity after a correct trial is seen across the population of recorded neurons (Fig. 4). For both areas, there was significantly greater direction selectivity when the previous trial had been correct than when it had been incorrect (Fig. 4AB). We also quantified this effect for each neuron by subtracting the mean ROC value for the cued saccade direction when the previous trial was correct from that when the previous trial was incorrect. Across the population in both areas, these differences are positive ($p<0.001$ in both cases; sign test for non-zero median), showing greater selectivity for the cued saccade direction if the previous trial had been correct (Fig. 4CD; see also Supp. Results). Increases in accuracy after correct trials were also reflected in the animals' behavior: performance on a given trial is more likely to be correct if the previous trial was correct than if it was incorrect (Fig. 4E).

We used the area under the direction ROC curve to quantify the neural changes that accompany learning. In our past work (Pasupathy and Miller, 2005), we separated neural selectivity into cue, response direction, and association components by partitioning the total variance, using a two-way model for cue and direction with an interaction term. (Thus, association selectivity is principally the degree to which neurons simultaneously encode cue and response direction; cf. Pasupathy et al., 2005, Asaad et al., 1998). Here, we used ROC area for direction because it captures, in a single measure, selectivity for the learned response and also the majority of association selectivity, both of which change over learning. We also repeated our past methods to examine direction and association selectivity separately, and found that both show the same effects as when they are combined in the direction ROC area. They each show stronger tuning after correct trials, and weaker tuning after errors (Supp. Fig. 5).

Thus, we found that single behavioral responses have strong effects on both animal behavior and neural activity — a correct trial strengthens both neural selectivity and the probability of a correct behavioral response, while after an error both are much nearer to chance performance.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Behavioral accuracy is improved after a correct response

The outcome of a single trial influences direction selectivity on the next trial. However, animals' behavioral performance improves slowly with learning. Thus, in theory it might be possible that the effect of a single trial on the next trial's response was due merely to these slow changes. This was unlikely due to the large changes caused by a single correct or error trial (Fig. 4), but we confirmed that it was not the case by comparing performance both early and late in learning (Fig. 5). The increase in selectivity when the previous trial is correct is seen during the first half of a block of learning trials, when many errors are made, as well as on the second half, when performance was better and fewer errors are made. On error trials, direction selectivity is smaller and thus closer to what would be expected by chance. The fact that reversals are not accompanied by an explicit cue probably encouraged the animals to favor a trial and error strategy, where error trials often resulted in guessing on the next trial (see Discussion). The behavioral impact of a single trial's outcome was also seen at the start and end of learning: in the first 10 trials after reversal, animals made 72% correct responses on trials after corrects and 53% following errors, while in the last 20 trials, animals perform at 92% following corrects and 57% following error trials.

A weaker signature of outcome can persist beyond single trials

We report above how each trial's outcome affects direction selectivity on the next trial. While this was the strongest effect, we also observed a weaker signature of changes arising from more than one trial in the past. For example, one might expect a cluster of correct trials to result in greater direction selectivity than one correct trial preceded by several error trials. This can be seen in the population data (Fig. 2E). There is an elevated baseline outcome ROC before the time of the response (red and blue dotted lines elevated above chance level, 0.5). When we repeated the analysis with random reassignment of the current trial's correct and error status, this effect was still present (data not shown). However, it fell to chance when we reshuffled trial numbers, which breaks the link between trials nearby each other in time. This implies that it is not merely due to the statistical structure of the spike trains we recorded, but was a true signature of neurons that reflect behavioral outcomes over more than one trial. Despite the presence of this weaker multi-trial effect, a single trial produces an increment in this long-term information (difference between solid and dotted lines in Fig. 2E). Furthermore, the magnitude of the single trial effect is at least as large as all the multi-trial effects summed together (difference between dotted lines and 0.5 level).

Transient outcome effects are large

The transient outcome responses shown by prefrontal and caudate neurons have received relatively little emphasis (but see Lau and Glimcher, 2008; Fujii and Graybiel, 2003), though these responses are quite large. More precisely, the large ROC value for transient correct and error responses indicate that the neurons carry a large amount of information about correct vs. error. Because we use the same ROC analysis method to examine both outcome and direction information, we can compare the relative strength of these effects (Supp. Fig. 1). The transient outcome ROC often shows a value between 0.7 and 0.9 (Supp Fig 1B–C), similar to the direction ROC during the saccade (Supp Fig 1C), and larger than the other information these neurons represent (Supp Fig 1A–B). Thus, whether the transient outcome signal reflects mainly input or local processing, its strength implies it is an important signal in these two areas.

Discussion

Here we report two main results. First, in a learning task, neurons in the PFC and caudate nucleus show sustained activity that reflects a trial's correct or incorrect status, which lasts until the next trial. Second, the neural representation of the learned information in this task is

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

changed by a single trial's outcome: correct trials improve the strength of direction selectivity on the next trial.

Implication for learning models; relation to the dopamine system

There are at least two ways that the brain might store information for seconds or longer about behavioral outcome. At these timescales, information might reside in changes in the strength of synaptic connections, resulting in different sized responses to future stimuli. Or, it might be stored in the activity of the neurons, maintained by sustained neural firing rate. Demonstrating the feasibility of both methods, neural network models have been devised that use each method for information storage (Hopfield, 1982; Rumelhart et al., 1986; Maass et al., 2002; Drew and Abbott, 2006; Ganguli et al., 2008). Previously, mainly transient reward responses had been reported in the frontal cortex and basal ganglia (Schultz et al., 1993b; Schmitzer-Torbert and Redish, 2004; Barnes et al., 2005; Apicella et al., 1991; Ichihara-Takeda and Funahashi, 2006; Watanabe, 1989; Lau and Glimcher, 2007; Lau and Glimcher, 2008; though note that modulation of frontal lobe task responses can depend on reward history: Barraclough et al., 2004; Uchida et al., 2007; Seo et al., 2007; Seo and Lee, 2009), and similarly transient responses have been seen in the dopamine system of the basal forebrain, which sends strong connections to the areas we studied. Thus, because transient responses were seen in the frontal lobe, the basal ganglia, and in dopamine neurons, prior work suggested that the "synaptic strength" hypothesis might be the mechanism for storing information about past responses (Jackson et al., 2006). This was also supported by observation of task-related modulation by reward (Barraclough et al., 2004; Seo et al., 2007; Uchida et al., 2007). But learning seemed to be too fast to result from synaptic changes. While there are ways that synaptic strengths can vary transiently (through synaptic depression or facilitation, e.g. Thomson and Deuchars, 1994; Tsodyks and Markram, 1997), long-lasting synaptic changes require protein synthesis (Frey et al., 1996) and therefore take tens of minutes to occur. If synaptic changes did underlie this learning, it had not been explained how such fast yet long-lasting changes might occur. Thus, there has been an inconsistency in our understanding of the mechanism for learning in these areas: the types of changes thought to be required took much longer than the time available to make them. Consistent with models that have proposed how network state can store memories (Maass et al., 2002; Ganguli et al., 2008), our data demonstrate that this can be seen in the sustained activity of single neurons.

Note that while we found sustained firing rate changes, it is possible that learning also results in synaptic changes. In fact, the ability to remember associations over hours, days or more almost certainly requires a remodeling of connection strengths somewhere in the brain. However, given that frontal cortex is known to show sustained changes in activity in memory tasks (Fuster and Alexander, 1971; Funahashi et al., 1989) and other complex tasks (Fuster et al., 2000; Wallis et al., 2001 2001), it is consistent with our understanding of these areas that sustained rate changes also encode outcome information. Having both outcome and direction information available puts the frontal cortex and basal ganglia in an excellent position to combine them and thus perhaps guide synaptic strength adjustments, so that both types of changes may co-exist during learning. Because all information relevant to the task is present in both areas, they may be the principal place where such learning is instantiated.

Sustained outcome responses fill a gap in our knowledge of the neural responses necessary for learning. The transient and sustained responses are, however, likely to be intimately related. For example, the transient responses may trigger sustained responses. Two recent studies (Williams and Eskandar, 2006; Nakamura and Hikosaka, 2006a) support this idea. Microstimulation of the striatum led to improvements in learning, and moreover, these improvements were seen only when the microstimulation occurred at the time of the reward. It may be that these transient outcome responses reflect a large input from the dopamine

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

system, and microstimulation applied at the time of the outcome signal interferes with the transformation of this dopamine input into the sustained changes we observed. This kind of transformation has recently been reported in PFC *in vitro* by Sidiropoulou et al. (2009), who found that dopamine inputs can depolarize single neurons, leading to sustained firing rate changes. While we saw this type of sustained activity, we found roughly equal numbers of neurons that increase firing to correct (when dopamine neurons typically increase activity) as increase firing to error (when dopamine neurons typically decrease activity). But dopamine has also been reported to inhibit frontal neurons' firing (Otani et al., 1999), and depolarization by dopamine may also trigger recurrent network mechanisms, possibly in frontal-basal ganglia loops (Alexander et al., 1990), which inhibit some neurons and excite others.

Our results cannot be explained by drift/baseline changes

One potential concern might be that long-term changes in neuronal activity over many trials might affect our results, whether due to baseline activity changes or possible changes in position of the electrode relative to the neuron. To deal with this issue, we included neurons in the analyses only if the neuron's activity was stable while the animal performed at least four repetitions of learning – i.e., the animal first learned one pairing followed by three reversals of the pairing, each of which the animal learned to the behavioral criterion level. Thus, long-term changes in the neurons' activity would tend to affect all types of trials equally, ruling out spurious effects where neurons would appear to respond to one stimulus or direction due to drift. Also, we corrected for bias in the ROC area by shuffling trials randomly (e.g. in Fig. 2E; see Methods; Supp Fig. 2). Since this method intermixed trials at the beginning and end of the recording sessions, it also controls for any effect of long-term drifts in activity.

Transient outcome responses: pure reward responses?

Animals can and often do learn given only secondary reinforcement that is not in itself a reward (Pavlov, 1927). As an example, human students will study for an exam in order to much later earn a high salary. Here we study only how a trial's outcome yields future changes in behavior and in neural activity. Because the exact nature of the stimulus used to provide feedback is not important for the changes we study, we have not examined whether the transient end-of-trial responses are associated with the primary or secondary reinforcement stimuli (cf. Wirth et al., 2009). This is because in either case, the signal is likely to arise from the midbrain dopaminergic system, whose neurons have been shown to fire in response to both types of reinforcers (Schultz et al., 1993a). Specifically, dopamine neurons code for reward predictions, and they begin to fire in response to many sorts of secondary reinforcers when these reinforcement stimuli predict future rewards (Schultz, 1998). Thus, whether these neurons fire for reward alone or for trial outcome, they strongly encode information about a key element of learning: whether responses were correct or incorrect.

Relation to previous work

Other laboratories have studied similar effects in other task contexts. Lee and colleagues (Barraclough et al., 2004; Seo et al., 2007; Seo and Lee, 2009) have demonstrated that past history of reward can modulate the task-related responses of neurons in a mixed-strategy game. They have found these effects in several frontal lobe areas, including the supplementary eye fields (also called the dorsomedial frontal cortex, or DMFC), the cingulate cortex, and the PFC. These studies, however, did not closely compare transient and sustained outcome-related activity (though they have found some signatures of this; e.g. Fig. 6, Seo and Lee, 2009). Wirth et al. (2009) studied the hippocampus and identified neurons that show outcome-related activity and also change their task-related responses based on

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

prior outcome (cf. Supp. Fig. 1). Narayanan and Laubach (2008) saw outcome-related effects in rat frontal but not motor cortex. Taken together with our work, these studies suggest that the effects we observed reflect general mechanisms for learning that are present in many learning-related brain areas. Future studies are needed to examine how information flows between these structures.

We have previously (Pasupathy and Miller, 2005) described how direction and cue selectivity evolves with learning. The present study explores a number of new phenomena. First, this report examines outcome-related signals. Second, we show here how single trials impact the strength of information about the task. Our prior work focused on the timecourse of selectivity, comparing latency of direction selectivity near the beginning (right after a reversal) and the end of learning (just before the next reversal). Here, we look at how the strength of direction selectivity on a single trial is affected by the trial that immediately precedes it, no matter if it is at the start or end of each block. In fact, we show that the single-trial effect is only weakly affected by the position in the block (Fig. 5). While apparently at odds, these two effects are complementary. Because learning results in more and more correct trials, there are fewer error trials at the end of learning than at the beginning. We find that a single error trial has a constant effect on the next trial no matter where it occurs, and that the accumulation of them at the beginning of a block produces the average effect we previously reported.

But why should neurons weight error trials at the start and end of learning similarly? We expected that the animal would obtain much more information from early than late error trials, as early error trials were key to re-learning the reversed association. We think that this is explained by the strategy the animals used. In this task, reversals occurred with no explicit cue. Because each error trial, especially at the end of learning, might have signaled a reversal, it makes sense to attempt a few guesses after an error, no matter where it occurred, to determine if a reversal had happened. Under the task constraints we imposed, this was a rational behavioral strategy (see also Fusi et al., 2007).

Both behavior and neural responses were more accurate after correct than error trials. This suggests that the animals learned more from correct trials than mistakes; in other words, a correct trial told the animal more about how to make future responses than an error trial. While this may be a strategy specific to this task, it may also be a more general strategy for animal learning that bears future investigation.

Conclusion

The results reported here show that these two areas, previously known to show learning-related changes, also have full information available to them to do all the neural computations necessary for learning. Here we have shown that cells show robust signals about the outcome of behavioral responses, and that these persist between trials. Furthermore, after a correct trial, cells increase their selectivity for the association to be learned, and likewise decrease it after an incorrect trial. This may represent a single-trial snapshot of the learning process — how single cells change their responses in real time as a result of information about what is the right action and what is the wrong one.

Methods

Behavioral task

Animals began each trial by fixating a central spot for 800 ms, followed by the appearance of the picture cue for 500 ms and then a 1000 ms memory delay period. The end of the delay was signaled by the disappearance of the fixation spot and the appearance of two identical saccade target spots, one on the left and one on the right. Animals made a saccadic eye

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

movement to one of the two possible saccade targets. Animals had to learn, by trial and error, an arbitrary association between the two cues and the two possible responses. After animals could perform the association well (after at least 30 correct responses and 90% correct trials over the previous 10 trials for each cue) the association was reversed with no signal, and they had to relearn the new association. Perhaps because there was no explicit signal, so that any error might signal a reversal, the animals relearned the association slowly after reversal (Fig. 1). Each recording session consisted of 3–8 reversals (4–9 trial blocks). By requiring animals to repeatedly relearn the associations, we could dissociate learning-related effects from artifactual effects that resulted from slow shifts over the course of a session, related i.e., to motivational changes or changes in the position of the electrode relative to a neuron.

The cues were complex color images and were new for each recording session so that animals had no prior response associated with a cue. Two other sets of cues, both with non-reversing cue-response associations, were intermixed with the two reversing cues (total 6 cues), a set of highly-familiar cues which were unchanged from day to day, and a set which were new each session. Data presented here comes from the first set of novel cues with reversing associations only.

When animals made the correct response, they received drops of juice paired with a tone for each drop. The first tone and drop began 100–130 ms after a correct saccade was completed. The next trial began in 5.5 seconds. If a correct response was made, the saccade targets were left in place for 500 ms to provide a fixation target and reduce post-reward saccades. If an incorrect response was made, a visual error stimulus (a large red square) was displayed during an additional 1 second delay before the start of the next trial. A black screen occupied the remaining interval (final 5–5.5 s) between trials. The time of outcome feedback was defined as the time of the beep for correct and the time of red square onset for error trials. To ensure that the slight difference in interval between trials (correct: 5.5 s, error, 6.5 s) did not affect our results, we computed outcome measures both forward from the end of each trial and backward from the start of the next (Fig. 2E). This did not change the effects.

Data analysis

The recording methods used here are described in Pasupathy and Miller (2005); the data set described there is the same set used here. All recording and animal procedures were in accordance with US National Institute of Health (NIH) guidelines and were conducted under the guidance of MIT veterinary staff and with the approval of the MIT Institutional Animal Care and Use Committee.

We recorded all neurons with sufficiently large signals without pre-selecting neurons for task-related responses like saccadic or visual responses (722 PFC neurons, 597 Cd neurons). For population direction analyses (Fig. 4, Fig. 5), we used the neurons which showed a significant effect of cue or of saccade via ANOVA, as well as being stably recorded for at least 3 reversals — 4 instances of learning (N=350, PFC; 249, Cd). All directional analyses used the actual saccade direction, not the cued direction on that block (identical on correct trials but different if the response was an error). For the population outcome ROC plot (Fig 2E), to compare effect magnitude across the time of reward and the inter-trial interval we used all neurons which showed a significant effect of reward in both time periods, via non-parametric ANOVA ($p<0.05$; N=94, PFC; N=85, Cd. Number of cells significant at $p<0.05$ in reward period by itself: N=237, PFC; N=186, Cd; in inter-trial interval: N=125, PFC; N=110, Cd). The effect in each interval was qualitatively similar and remained significant if we used all cells that were significantly modulated by reward in that interval.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The histograms (Figs. 2, 3) were calculated by convolving the spike train with a 140 ms square window. All ROCs were computed over a 200 ms sliding window, which gave slightly more statistical power than a 140 ms window. To compute the area under the ROC curve, for each neuron we divided the set of trials into two groups, i.e correct vs. error for the outcome ROC. Then, we constructed the ROC curve: the fraction of correct decisions ("hits") that an ideal observer would make vs. incorrect decisions ("false alarms") as the threshold is varied (Green and Swets). The area under this curve is the probability that an ideal observer successfully chooses the correct trial condition given the firing rate on that trial, and thus gives a measure of overlap of the two firing rate distributions (e.g. Dayan and Abbott, 2001). The ROC values we computed were similar between the two animals and so for the population figures we pooled each animal's data together. Because there is no *a priori* preferred case for e.g. reward vs. error (Supp. Fig. 3), we rectified ROC values around 0.5. To correct for biases in ROC values, we used a shuffle-corrector: for each cell and time point, we randomly shuffled trials between the two groups and repeatedly recomputed the ROC. Then, we subtracted the difference between the shuffled, random value and 0.5 from the measured ROC. (See Supp. Fig. 2 for further details.) Supporting the validity of these procedures, we found that post-correction, the mean fixation-period direction ROC was 0.5 (Fig 4A–B); further, these ROC results agreed with results found using a linear model (Supp. Fig. 5).

The trial time periods were defined as follows: Transient selectivity after the outcome feedback ("reward period", Fig. 6) was computed from the time of the correct or error feedback to 500 ms afterwards. Outcome selectivity lasting from one trial to the next ("inter-trial period") was computed 2–4000 ms after the outcome feedback, and computing it from 2500 ms to 500 ms before the start of the next trial produced nearly identical results. The cue period was from the onset of the cue till its offset 500 ms later, and the delay period is 1000 ms long, from cue offset. The saccade period was chosen to cover pre- and post-saccadic peaks (Bruce and Goldberg, 1985) and was defined as from the offset of the fixation point signaling the beginning of the response period to 500 ms later.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank M.R. Cohen for valuable suggestions and comments on the manuscript, and K. MacCully for expert technical assistance. This work was supported by the NINDS and the Tourette's Syndrome Association (A.P.).

References

Alexander GE, Crutcher MD, DeLong MR. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. Prog Brain Res 1990;85:119–146. [PubMed: 2094891]

Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 1986;9:357–381. [PubMed: 3085570]

Anden NE, Hfuxe K, Hamberger B, Hokfelt T. A quantitative study on the nigro-neostriatal dopamine neuron system in the rat. Acta Physiol Scand 1966;67:306–312. [PubMed: 5967596]

Apicella P, Ljungberg T, Scarnati E, Schultz W. Responses to reward in monkey dorsal and ventral striatum. Exp Brain Res 1991;85:491–500. [PubMed: 1915708]

Asaad WF, Rainer G, Miller EK. Neural activity in the primate prefrontal cortex during associative learning. Neuron 1998;21:1399–1407. [PubMed: 9883732]

Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature 2005;437:1158–1161. [PubMed: 16237445]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Barraclough DJ, Conroy ML, Lee D. Prefrontal cortex and decision making in a mixed-strategy game. Nat Neurosci 2004;7:404–410. [PubMed: 15004564]

Berger B, Trottier S, Verney C, Gaspar P, Alvarez C. Regional and laminar distribution of the dopamine and serotonin innervation in the macaque cerebral cortex: a radioautographic study. J Comp Neurol 1988;273:99–119. [PubMed: 3209731]

Bruce CJ, Goldberg ME. Primate frontal eye fields. I. Single neurons discharging before saccades. Journal of Neurophysiology 1985;53:603–635. [PubMed: 3981231]

Chen LL, Wise SP. Supplementary eye field contrasted with the frontal eye field during acquisition of conditional oculomotor associations. J Neurophysiol 1995;73:1122–1134. [PubMed: 7608759]

Dayan, P.; Abbott, LF. Theoretical neuroscience: Computational and mathematical modeling of neural systems. Cambridge (Massachusetts): MIT Press; 2001.

Doya K. Modulators of decision making. Nat Neurosci 2008;11:410–416. [PubMed: 18368048]

Drew PJ, Abbott LF. Extending the effects of spike-timing-dependent plasticity to behavioral timescales. Proc Natl Acad Sci USA 2006;103:8876–8881. [PubMed: 16731625]

Fellows LK, Farah MJ. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. Cereb Cortex 2005;15:58–63. [PubMed: 15217900]

Frey U, Frey S, Schollmeier F, Krug M. Influence of actinomycin D, a RNA synthesis inhibitor, on long-term potentiation in rat hippocampal neurons in vivo and in vitro. J Physiol (Lond) 1996;490 (Pt 3):703–711. [PubMed: 8683469]

Fujii N, Graybiel AM. Representation of action sequence boundaries by macaque prefrontal cortical neurons. Science 2003;301:1246–1249. [PubMed: 12947203]

Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. Journal of Neurophysiology 1989;61:331–349. [PubMed: 2918358]

Fusi S, Asaad WF, Miller EK, Wang XJ. A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. Neuron 2007;54:319–333. [PubMed: 17442251]

Fuster, JM. The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe. 2. Lippincott: Williams, and Wilkens; 1997.

Fuster JM, Alexander GE. Neuron activity related to short-term memory. Science 1971;173:652–654. [PubMed: 4998337]

Fuster JM, Bodner M, Kroger JK. Cross-modal and cross-temporal association in neurons of frontal cortex. Nature 2000;405:347–351. [PubMed: 10830963]

Gaffan D, Harrison S. A comparison of the effects of fornix transection and sulcus principalis ablation upon spatial learning by monkeys. Behav Brain Res 1989;31:207–220. [PubMed: 2914072]

Ganguli, S.; Huh, D.; Sompolinsky, H. Proc Natl Acad Sci USA. 2008. Memory traces in dynamical systems.

Green, DM.; Swets, JA. Signal Detection Theory and Psychophysics. New York: Wiley; 1966.

Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci USA 1982;79:2554–2558. [PubMed: 6953413]

Houk JC, Wise SP. Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. Cereb Cortex 1995;5:95–110. [PubMed: 7620294]

Ichihara-Takeda S, Funahashi S. Reward-period activity in primate dorsolateral prefrontal and orbitofrontal neurons is affected by reward schedules. J Cogn Neurosci 2006;18:212–226. [PubMed: 16494682]

Jackson A, Mavoori J, Fetz EE. Long-term motor cortex plasticity induced by an electronic neural implant. Nature 2006;444:56–60. [PubMed: 17057705]

Lau B, Glimcher PW. Action and outcome encoding in the primate caudate nucleus. J Neurosci 2007;27:14502–14514. [PubMed: 18160658]

Lau B, Glimcher PW. Value representations in the primate striatum during matching behavior. Neuron 2008;58:451–463. [PubMed: 18466754]

Ljungberg T, Apicella P, Schultz W. Responses of monkey dopamine neurons during learning of behavioral reactions. J Neurophysiol 1992;67:145–163. [PubMed: 1552316]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Maass W, Natschläger T, Markram H. Real-time computing without stable states: a new framework for neural computation based on perturbations. Neural computation 2002;14:2531–2560. [PubMed: 12433288]

Middleton FA, Strick PL. Basal ganglia and cerebellar loops: motor and cognitive circuits. Brain Res Brain Res Rev 2000;31:236–250. [PubMed: 10719151]

Miyachi S, Hikosaka O, Miyashita K, Karadi Z, Rand MK. Differential roles of monkey striatum in learning of sequential hand movement. Exp Brain Res 1997;115:1–5. [PubMed: 9224828]

Montague PR, Berns GS. Neural economics and the biological substrates of valuation. Neuron 2002;36:265–284. [PubMed: 12383781]

Murray EA, Bussey TJ, Wise SP. Role of prefrontal cortex in a network for arbitrary visuomotor mapping. Exp Brain Res 2000;133:114–129. [PubMed: 10933216]

Nakamura K, Hikosaka O. Facilitation of saccadic eye movements by postsaccadic electrical stimulation in the primate caudate. J Neurosci 2006a;26:12885–12895. [PubMed: 17167079]

Nakamura K, Hikosaka O. Role of dopamine in the primate caudate nucleus in reward modulation of saccades. J Neurosci 2006b;26:5360–5369. [PubMed: 16707788]

Narayanan NS, Laubach M. Neuronal correlates of post-error slowing in the rat dorsomedial prefrontal cortex. Journal of Neurophysiology 2008;100:520–525. [PubMed: 18480374]

Otani S, Auclair N, Desce JM, Roisin MP, Crépel F. Dopamine receptors and groups I and II mGluRs cooperate for long-term depression induction in rat prefrontal cortex through converging postsynaptic activation of MAP kinases. J Neurosci 1999;19:9788–9802. [PubMed: 10559388]

Passingham, RE. The Frontal Lobes and Voluntary Action. Oxford: Oxford University Press; 1995.

Pasupathy A, Miller EK. Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature 2005;433:873–876. [PubMed: 15729344]

Pavlov, IP. Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. London: Oxford University Press; 1927.

Petrides M. Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. Neuropsychologia 1985;23:601–614. [PubMed: 4058706]

Petrides M. Frontal lobes and behaviour. Curr Opin Neurobiol 1994;4:207–211. [PubMed: 8038578]

Petrides M, Pandya DN. Efferent association pathways originating in the caudal prefrontal cortex in the macaque monkey. J Comp Neurol 2006;498:227–251. [PubMed: 16856142]

Petrides M, Pandya DN. Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. J Neurosci 2007;27:11573–11586. [PubMed: 17959800]

Rangel A, Camerer C, Montague PR. A framework for studying the neurobiology of value-based decision making. Nat Rev Neurosci 2008;9:545–556. [PubMed: 18545266]

Rumelhart, DE.; Hinton, GE.; Williams, RJ. Learning internal representations by error propagation. In: Rumelhart, DE.; McClelland, JL., editors. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Cambridge, MA: The MIT Press; 1986. p. 318-362.

Rushworth MF, Behrens TE. Choice, uncertainty and value in prefrontal and cingulate cortex. Nat Neurosci 2008;11:389–397. [PubMed: 18368045]

Schmitzer-Torbert N, Redish AD. Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. J Neurophysiol 2004;91:2259–2272. [PubMed: 14736863]

Schultz W. Predictive reward signal of dopamine neurons. J Neurophysiol 1998;80:1–27. [PubMed: 9658025]

Schultz W, Apicella P, Ljungberg T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. J Neurosci 1993a; 13:900–913. [PubMed: 8441015]

Schultz W, Apicella P, Ljungberg T, Romo R, Scarnati E. Reward-related activity in the monkey striatum and substantia nigra. Prog Brain Res 1993b;99:227–235. [PubMed: 8108550]

Seo H, Barraclough DJ, Lee D. Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. Cereb Cortex 2007;17(Suppl 1):i110–117. [PubMed: 17548802]

Seo H, Lee D. Behavioral and neural changes after gains and losses of conditioned reinforcers. J Neurosci 2009;29:3627–3641. [PubMed: 19295166]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Sidiropoulou K, Lu FM, Fowler MA, Xiao R, Phillips C, Ozkan ED, Zhu MX, White FJ, Cooper DC. Dopamine modulates an mGluR5-mediated depolarization underlying prefrontal persistent activity. Nat Neurosci. 2009

Sugrue LP, Corrado GS, Newsome WT. Choosing the greater of two goods: neural currencies for valuation and decision making. Nat Rev Neurosci 2005;6:363–375. [PubMed: 15832198]

Thomson AM, Deuchars J. Temporal and spatial properties of local circuits in neocortex. Trends in Neurosciences 1994;17:119–126. [PubMed: 7515528]

Tsodyks MV, Markram H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc Natl Acad Sci USA 1997;94:719–723. [PubMed: 9012851]

Uchida Y, Lu X, Ohmae S, Takahashi T, Kitazawa S. Neuronal activity related to reward size and rewarded target position in primate supplementary eye field. J Neurosci 2007;27:13750–13755. [PubMed: 18077686]

Wallis JD, Anderson KC, Miller EK. Single neurons in prefrontal cortex encode abstract rules. Nature 2001;411:953–956. [PubMed: 11418860]

Watanabe M. The appropriateness of behavioral responses coded in post-trial activity of primate prefrontal units. Neurosci Lett 1989;101:113–117. [PubMed: 2505197]

Williams SM, Goldman-Rakic PS. Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. Cereb Cortex 1993;3:199–222. [PubMed: 8100725]

Williams ZM, Eskandar EN. Selective enhancement of associative learning by microstimulation of the anterior caudate. Nat Neurosci 2006;9:562–568. [PubMed: 16501567]

Wirth S, Avsar E, Chiu CC, Sharma V, Smith AC, Brown E, Suzuki WA. Trial outcome and associative learning signals in the monkey hippocampus. Neuron 2009;61:930–940. [PubMed: 19324001]

Wise SP, Murray EA, Gerfen CR. The frontal cortex-basal ganglia system in primates. Crit Rev Neurobiol 1996;10:317–356. [PubMed: 8978985]

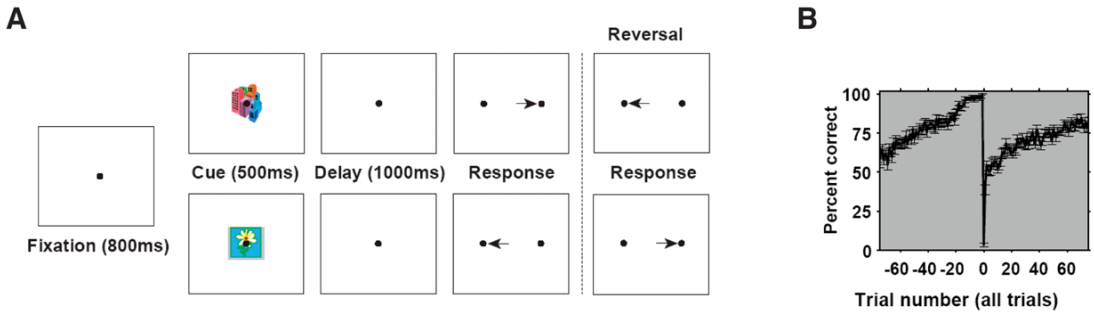

Figure 1. Behavioral task

A, schematic of the associative learning task. Animals were required to learn, by trial-and-error, an arbitrary association between a picture cue and a directional eye movement response. On each trial, they held their eye position on a central fixation point for 800 ms, and then the cue was turned on for 500 ms. After a 1000 ms memory delay period, the fixation point was extinguished and the animals made their response; the correctness of the response was signaled immediately after the saccade (see Methods). After animals had learned this association, we reversed the pairing with no explicit signal and animals relearned the reversed association. B, Average learning curve, showing performance before and after reversal. X-axis: trial number; at trial 0, the association was reversed with no signal, almost always causing an error (trial 1). Within a few trials, performance reverted to near 50% and then gradually increased as animals learned the new pairing. Error bars: S.E.M.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

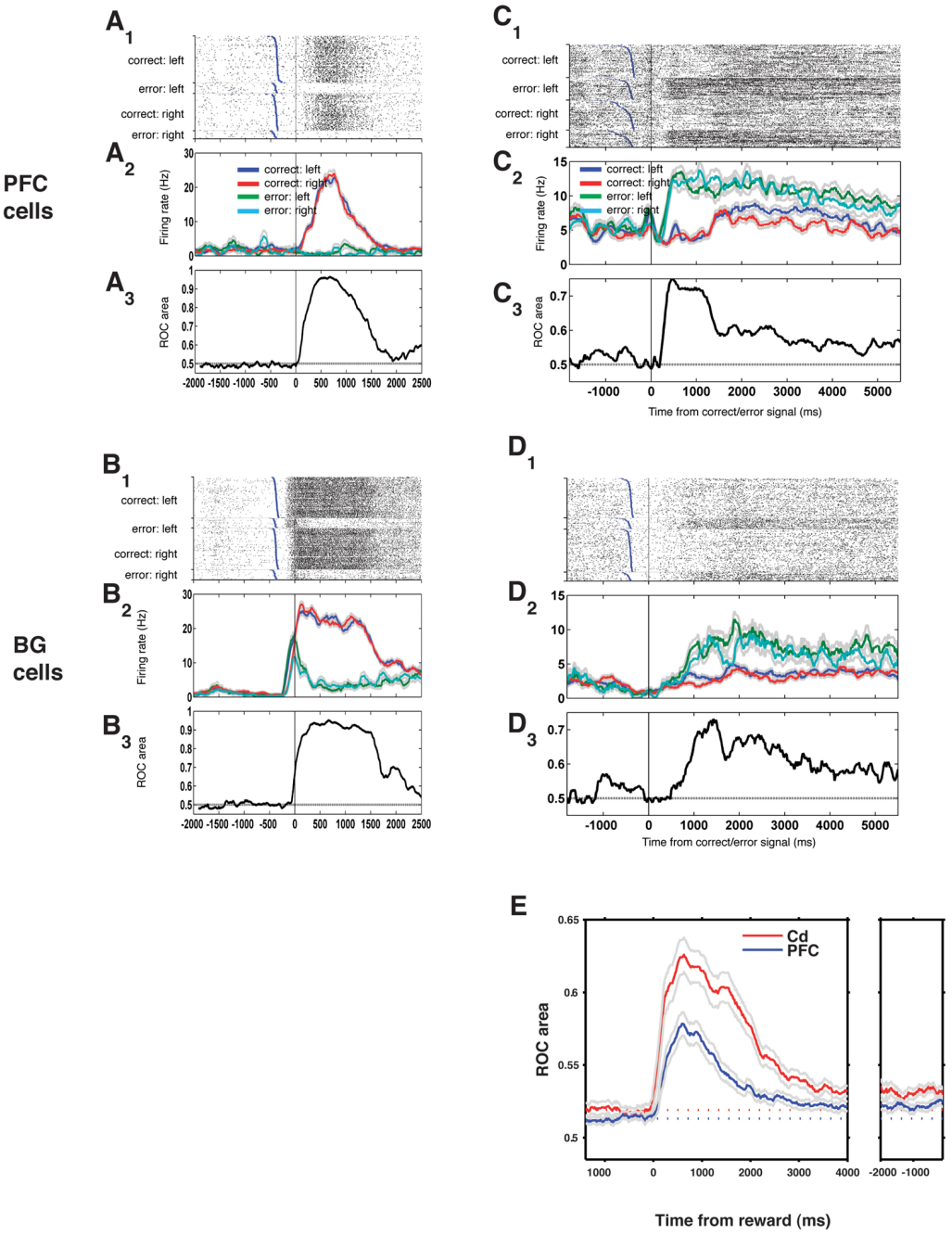

Figure 2. Cells signal correct or error outcome

A1–3: Single cell recorded from the PFC showing an increase in firing rate after correct outcome was signaled. All three panels show data from the same set of trials. X-axes: time from correct/error feedback signal. Top panel (A1): trial raster; each tick corresponds to a spike. Each row is a different trial; blue ticks, response times (end of saccadic eye movement); trials are sorted by response time within each of the four trial groups. Middle panel (A2): histogram of the same trials. Firing rates (colored lines) were computed by convolving the spike trains in A1 with a 140 ms square kernel. Gray lines: 1 S.E.M. Bottom panel (A3): information that this cell gives about correct vs. error at each time point, measured as area under ROC curve (Y-axis). B1-B3: a 2nd cell from Cd which exhibits a

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

similarly strong increase in firing rate on correct trials. C1-C3 and D1-D3: single PFC and Cd cells showing sustained responses about reward vs. error that lasted for several seconds into the next trial. Conventions as in A and B. E, population summary. Y-axis: mean reward information (reward ROC area) over the population of cells from each area. Blue, PFC mean (N=85; see Methods); red, caudate (N=94). Gray lines: 1 S.E.M. X-axis: time from correct/error feedback signal. Dotted lines indicate baseline information maintained from previous trial (see Discussion); elevation above this level shows additional information gained by neurons because of a single trial's reward. Left panel: data aligned on reward onset; right panel: aligned on the next trial's fixation onset (note inter-trial period length for errors: 6.5 s; for corrects: 5.5 s). The population of recorded cells from both areas signals whether single trials are correct or incorrect, and this information is maintained until the next trial.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

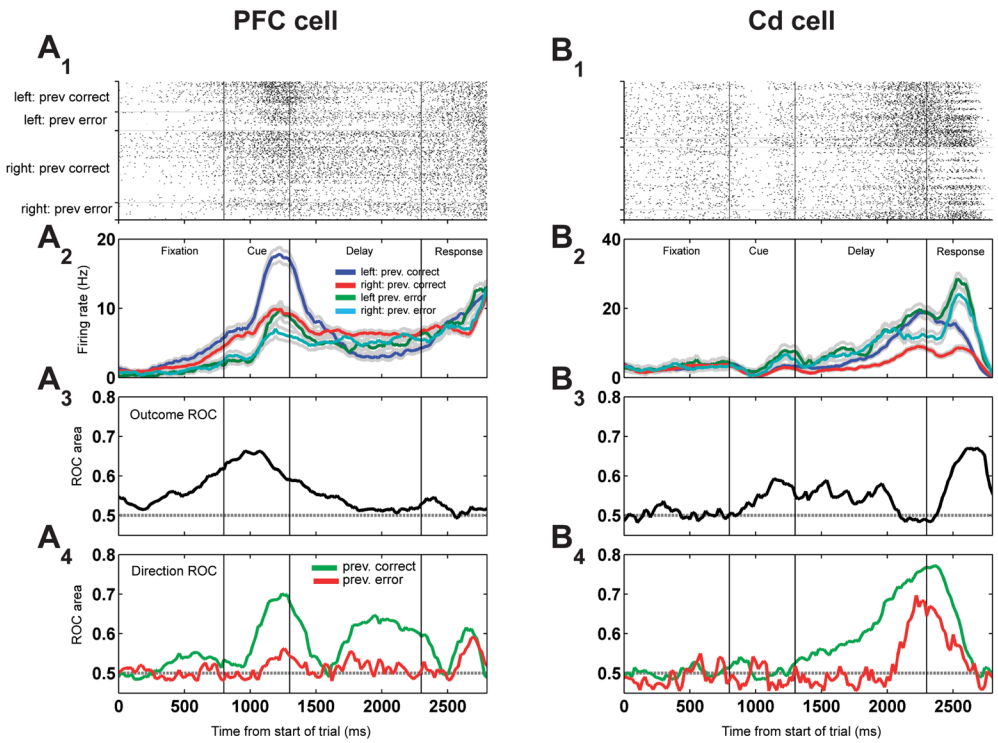

Figure 3. Direction signal is stronger when previous trial is correct: single cells

Left panels (A1-A4): single PFC cell showing increased direction selectivity after previous trial was correct versus when the previous trial was an error. A1: trial raster; conventions as in Fig. 2A1. Trials are arranged by the response direction the animal chose on a given trial and the correct/error status of the previous trial. A2: Histogram of firing rates, conventions as in Fig. 2A2. A3: Information carried by this cell (measured by ROC area) about the correct vs. error outcome of the previous trial, averaged over response direction of the current trial. A4: Information (ROC area) about the response direction of the current trial, plotted in green when the previous trial was correct and red when the previous trial was an error. Right panels (B1-B4): a single cell recorded from the caudate nucleus; conventions as in A1-A4. Both cells give more information about the animal's intended response (i.e. ROC area is larger) when the previous trial was correct.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

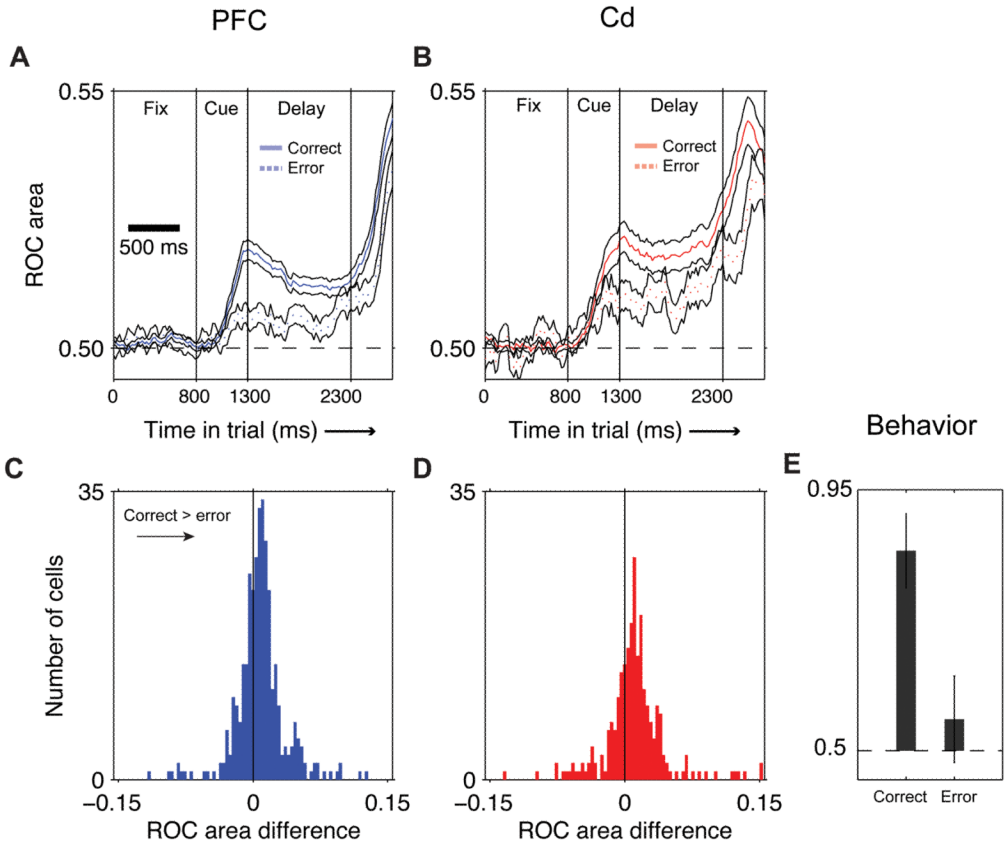

Figure 4. Direction signal is stronger when previous trial is correct: population summary

A, Averaged direction ROC values for all PFC cells when the previous trial was correct (solid blue line) vs. previous error (dotted blue line). Black lines: 1 S.E.M. X-axis, time in trial, Y-axis: average ROC value. B, Averaged direction ROC values for all Cd cells; conventions as in A. For both areas, information about direction is stronger after a correct trial than after an error trial. C-D, Distribution, over all cells, of the difference in ROC value after correct and after error. For each cell, we subtracted the delay period direction ROC value after correct trials from that after error trials. C, blue: PFC cells; D, red: Cd. The distributions are significantly shifted to the right (PFC: $p<10^{-7}$, Cd: $p<10^{-8}$, Wilcoxon test), showing stronger direction tuning after correct trials. E, Behavioral performance on the next trial after a correct or error trial. Error bars: std. dev. over 63 experimental sessions. Performance was much higher when the previous trial was correct than when the previous trial was an error.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

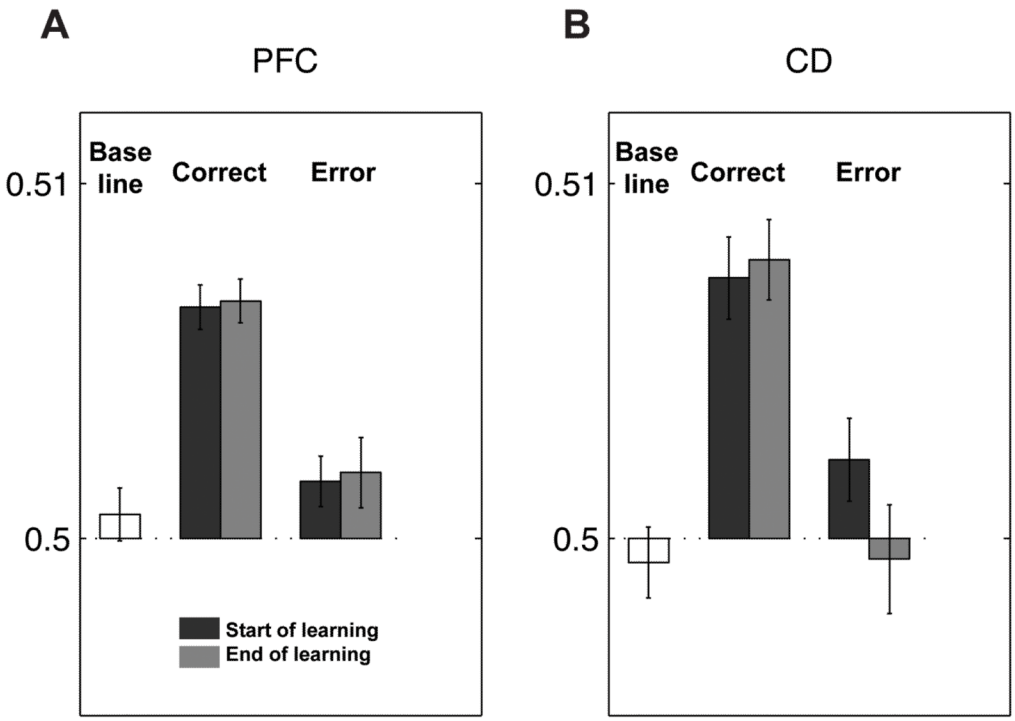

Figure 5. Increases in direction selectivity after correct trials occur both at the start and end of learning

Y-axis, the delay period direction selectivity (area under ROC curve) of all cells in the population after correct trials, middle bars, and after error trials, right bars. Each repetition of learning, from one reversal to the next, was divided into two sets of trials; the first half are shown as dark gray bars ("start of learning"), and light gray bars show the second half ("end of learning"). The ROC area from the fixation (baseline) period is shown at left. A, PFC neurons; B, Cd. These data show that the increases in direction selectivity after a correct trial exist both early and late in learning.