

Incorporation of new information into prefrontal cortical activity after learning working memory tasks

Ethan M. Meyers^{a,1}, Xue-Lian Qi^b, and Christos Constantinidis^b

^aDepartment of Brain and Cognitive Sciences, McGovern Institute, Massachusetts Institute of Technology, Cambridge, MA 02139; and ^bDepartment of Neurobiology and Anatomy, Wake Forest University School of Medicine, Winston-Salem, NC 27157

Edited by Ranulfo Romo, Universidad Nacional Autonoma de Mexico, Mexico City, D.F., Mexico, and approved February 7, 2012 (received for review January 19, 2012)

The ability to learn new tasks requires that new information is integrated into neural systems that already support other behaviors. To study how new information is incorporated into neural representations, we analyzed single-unit recordings from the prefrontal cortex (PFC), a brain region important for task acquisition and working memory, before and after monkeys learned to perform two behavioral tasks. A population-decoding analysis revealed a large increase in task-relevant information, and smaller changes in stimulus-related information, after training. This new information was contained in dynamic patterns of neural activity, with many individual neurons containing the new task-relevant information for only relatively short periods of time in the midst of other large firing rate modulations. Additionally, we found that stimulus information could be decoded with high accuracy only from dorsal PFC, whereas task-relevant information was distributed throughout both dorsal and ventral PFC. These findings help resolve a controversy about whether PFC is innately specialized to process particular types of information or whether its responses are completely determined by task demands by showing there is both regional specialization within PFC that was present before training, as well as more widespread task-relevant information that is a direct result of learning. The results also show that information is incorporated into PFC through the emergence of a small population of highly selective neurons that overlay new signals on top of patterns of activity that contain information about previously encoded variables, which gives insight into how information is coded in neural activity.

neural coding | task learning | Macaque | vision | principal sulcus

The prefrontal cortex (PFC) is a brain region involved in planning, decision making, working memory, and learning new context dependent behaviors (1–3). Although many studies have found task-related activity in the PFC for a variety of different behaviors (4–6), it is often unclear whether this task-related information had always existed in PFC or whether it emerged as a result of learning. Furthermore, in studies in which it seems likely that new information arises as a result of training (7, 8), how this new information interacts with preexisting information is not well understood. Given that we must continually learn to perform new tasks, while simultaneously maintaining the ability to perform previously learned behaviors, understanding how new information is integrated into existing neural processing is fundamental to understanding how the brain enables complex and adaptive human behaviors.

To gain insight into how learning a new task affects processing in PFC, we analyzed single-unit activity from neurons before and after two monkeys were trained to perform two distinct delayed match-to-sample tasks (9, 10) (Fig. 1 and Fig. S1). Using a neural-population decoding analysis, we were able to directly assess what information was represented in the population before training and what new information arose because of learning a new task. Our analyses sought to examine several questions concerning the content and coding of information including: (i) Does learning a new task change the amount of information

about basic stimulus features or does it only change the amount of information about more complex task-related variables? (ii) Does the new information arise because of the emergence of a few highly selective neurons or is information evenly distributed across the population? (iii) Do neurons become specialized to process only one type of information, as suggested by some studies (11), or can the same neuron carry multiple types of information as other studies suggest (12)? (iv) Is the new information contained in a dynamic population code (13–15), or is there one stationary pattern of neural activity that contains the new information? (v) Are there differences in the information content between dorsal and ventral PFC, and does learning affect these two brain regions equally (16–20)? Thus, this work gives insight into how new information is incorporated in neural systems and helps clarify the key computations that are occurring in PFC (21, 22).

Results

Neural recordings were made from two monkeys while they passively viewed a sequence of two stimuli and after they were trained on two delayed-match-to-sample tasks. Before training, the monkeys fixated a central point and passively observed two stimuli that were separated by a delay (Fig. 1C). After training, the monkeys viewed the same sequence of stimuli and made a saccade that indicated whether the two stimuli were identical (Fig. 1D and Fig. S1A). In the “feature task,” the monkeys indicated whether the symbols were the same (Fig. 1D); in the “spatial task,” the monkeys indicated whether the square symbol appeared at the same location (Fig. S1). Neurons were sampled with an unbiased procedure, recording from all neurons that could be isolated, and a decoding procedure was used that jointly analyzed the activity of 750 neurons at a time recorded in separate sessions (*Methods*). Results from the feature task are shown in the main text of the article (*Results*), and the results from the spatial task are shown in the supplemental figures (*SI Text*). (Overall, the results were very similar between the two tasks.)

Our analysis examined the amount of information that could be decoded from the neuronal population (evaluated as the performance of a linear classifier) about stimulus identity or location (i.e., which of eight stimuli were shown in the feature and spatial tasks) and about the match/nonmatch status of a trial. We observed little difference in the amount of information about the identity of the first stimulus before training (Fig. 2A, blue trace) compared with after training (Fig. 2A, red trace). In contrast, there was a massive increase in information about the match/nonmatch status of the trial in PFC after training. Before

Author contributions: X.-L.Q. and C.C. designed research; E.M.M., X.-L.Q., and C.C. performed research; E.M.M. contributed new reagents/analytic tools; E.M.M. analyzed data; and E.M.M. and C.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: emeyers@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1201022109/-DCSupplemental.

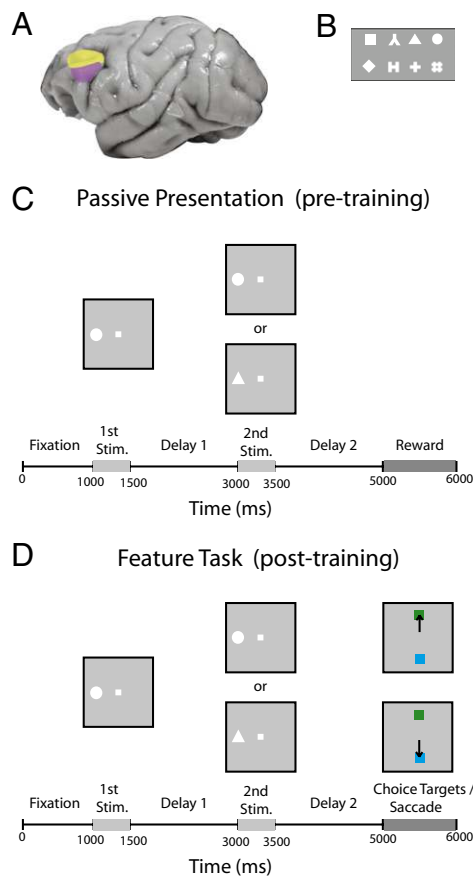


Fig. 1. Brain regions and the feature task. (A) Dorsal (yellow) and ventral (magenta) regions of lateral PFC where the recordings were made. (B) Stimuli used in the feature task. The stimuli extended 2° of visual angle. (C) Passive fixation task that was used before training. (D) Feature task. The monkeys viewed the same sequence of images as in the passive task; however, at the end of the experiment, monkeys needed to make a saccade to the green target if the stimuli matched or to the blue target if the stimuli did not match.

training, when the match/nonmatch status of a trial was irrelevant for the task, we could not decode information about whether the two stimuli matched in shape at accuracies that were above chance (Fig. 2B, blue trace). However, after training, when the match/nonmatch status was critical for correctly completing the task, we could decode this information at above-chance levels shortly after the onset of the second stimulus [permutation test, $P < 0.005$; see colored bars at the bottom of Fig. 2A and B], and it was possible to decode this information with close to 100% accuracy during most of the following delay period (Fig. 2B, red trace). This increase in match/nonmatch information across the population was attributable to a small subset of neurons that became highly selective for match/nonmatch information after training (points above horizontal line in Fig. 2C), as well as a larger number of neurons that showed a small increase in their match/nonmatch selectivity. In fact, these highly selective neurons were so informative that the top eight most selective neurons contained almost all of the information that was present in the entire population (Fig. S2A). However, the less strongly selective neurons still contained significant amounts of redundant information, as evidenced by the fact that when we excluded the top 128 most selective neurons, we still obtained above-chance decoding accuracies (Fig. S2B). Similar results were observed for spatial stimuli (Fig. S3): equivalent levels of stimulus information could be decoded during the first stimulus presentation before

and after training, and an increase in the match/nonmatch information was observed after training. A subtle difference between tasks was an increase in position information in anticipation of the second stimulus presentation after training in the spatial task.

Recent work has shown that neural activity in PFC shows complex temporal dynamics, with individual neurons changing the way they code information over the course of a trial (13–15). To test whether the task-relevant match/nonmatch information that emerged after training was also encoded by dynamic population activity, we applied a decoding analysis in which we trained the classifier using data from one time period (as indicated by the y axis on Fig. 3A and Fig. S4A) and tested the classifier using data from a different time period (as indicated by the x axis on Fig. 3A and Fig. S4A). If the information is represented by a stationary pattern of activity (i.e., if patterns of neural activity that encode the match/nonmatch trial status are the same at all time points), then training the classifier using one time point with high information should lead to high classification accuracy at all other time points where the information is present. Conversely, if information is represented by dynamic patterns of neural activity, then training the classifier at any one time point with high information should lead to high classification accuracy only at that time point. Fig. 3A (Right) clearly shows that high decoding accuracy is only obtained when the classifier is trained and tested on data from the same time point relative to stimulus onset. Over the second delay period, the decoding accuracy dropped from 98% correct to 68% correct when the classifier was trained on data taken 500 ms before the time the classifier was tested. Thus, new task-relevant information that emerges as a result of training is contained by a dynamic pattern of neural activity. The dynamic nature of the information coding was also evident in the mean firing rates of individual neurons. Examining activity of highly selective match/nonmatch neurons (Fig. 3B and Fig. S4B) revealed that the task-relevant information in several neurons was present for only short periods of time relative to the duration of the second delay period stimulus when this information was present (e.g., the middle neuron in Fig. 3B and the rightmost neuron in Fig. S4B), giving rise to the dynamic coding of match/nonmatch information seen at the population level. Additionally, firing-rate modulations that occurred throughout the trial were attributable to individual neurons carrying information about different variables at different points, as can be seen by the fact that the highly match/nonmatch selective neurons also contained large amounts of stimulus identity information (Fig. S5). Thus, task-relevant information is incorporated into PFC by interleaving/overlapping new information into ongoing dynamic activity that is carrying information about other variables and consequently the absolute firing rate level of a single neuron at a particular time point is often highly ambiguous if the context of the larger population is not taken into account.

The results presented thus far combined data from dorsal PFC (areas 46 and 8a) and ventral PFC (areas 12 and 45) (Fig. 1A). Previous research has led to conflicting findings about whether there are distinct functional differences between dorsal and ventral PFC (16–20), despite the clear anatomical connectivity differences between these two regions (23). To assess whether these brain areas have similar amounts of stimulus identity and match/nonmatch information, we applied the same decoding analyses separately to data from each of these areas. The results show that both dorsal and ventral PFC contained above-chance match/nonmatch information after the training (Fig. 4A; permutation test, $P < 0.005$). In contrast, information about the identity of the stimuli was largely confined to dorsal PFC, at least for the limited stimulus set that we tested neurons with (Fig. 4B; permutation test, $P < 0.005$). Similar results were seen in the spatial task (Fig. S6) (although the onset of match/nonmatch information was a bit longer in ventral PFC in the spatial task).

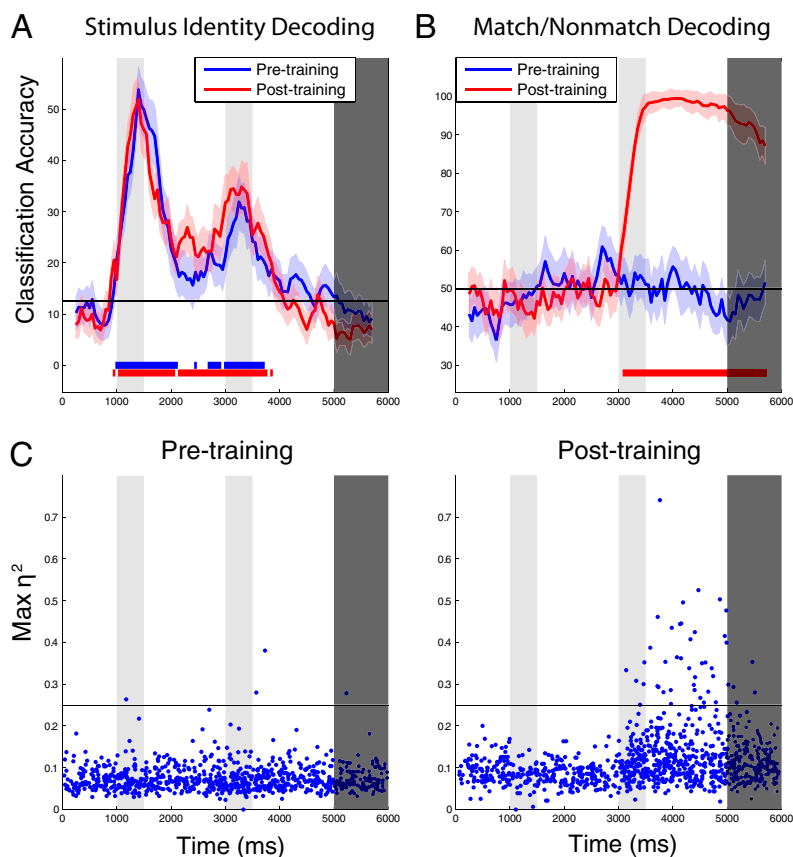


Fig. 2. Information in PFC pre- and posttraining in the feature task. (A) Comparison of information about the identity of the first stimulus pretraining (blue) and posttraining (red). The gray shaded regions indicate the times when the first, second, and decision stimuli were shown, the black horizontal line indicates the level of decoding expected by chance, the color shaded regions indicate 1 SE in the decoding accuracy if different neurons were used, and the red and blue bars at the bottom of the figure indicate times when the decoding accuracy was above chance (permutation test, $P < 0.005$). As can be seen, training had little effect on stimulus identity information. (B) Comparison of information about the match/nonmatch trial status pretraining (blue) and posttraining (red). As can be seen, there is a large increase in match/nonmatch status information after training. (C, Match/nonmatch selectivity of individual neurons before training (Left) and after training (Right). (The black horizontal line is for visualization purposes to make the pre- and posttraining differences easier to compare.) The η^2 statistic measures the proportion of the trial-by-trial variance in firing rates explained by whether a trial is a match or a nonmatch trial. Each point corresponds to the η^2 value of a single neuron at the latency when the neuron had its maximal selectivity (see *Methods*).

Thus, there are significant differences in how basic stimulus information is processed by these brain regions, whereas the newly learned task-relevant information was more distributed. Evidence also exists for variations in anatomical structure and function within the dorsal and ventral prefrontal cortex (23, 24). In our data set, comparison of area 46 with 8a revealed no qualitative differences within the dorsal PFC (Fig. S7). The match/nonmatch information was not restricted to the prefrontal cortex: recordings made from the posterior parietal cortex (areas LIP and 7a) of one of the monkeys after training also revealed a similar level of match/nonmatch information in that brain region (Fig. S8), indicating that the task-relevant information could potentially be wide-spread across several cortical areas.

Discussion

The results presented above give insight into how new task-relevant information is incorporated into existing processing, in different regions of PFC. Our results show that whereas there was little change in the amount of basic stimulus information before training (Fig. 2A), more complex information about whether the stimuli matched became present throughout PFC only after this information became relevant to the monkeys' behavior (Fig. 2B). Additionally, our analyses revealed that the majority of basic stimulus information was restricted to dorsal PFC (Fig. 4B), whereas the new task-relevant information was

much more widely distributed (Fig. 4A and Fig. S8). The fact that all regions of PFC contained the new task-relevant information is consistent with the adaptive coding model, which proposes that PFC can adapt to encode information about any property that is relevant for behavior (3). Additionally, the finding that basic visual information was restricted to dorsal PFC is consistent with domain-specific theories of PFC that claim that there are differences between different regions of lateral PFC (16, 17, 21) and raises questions about the validity of strict integrative theories, which claim there are no regional differences (18, 19). Anatomical studies have found that the cortical areas that project to dorsal PFC are different from those that project to ventral PFC and that there are extensive intra-area connections within PFC (23, 25). Based on our results and these anatomical findings, we hypothesize that the long-range anatomical connections between PFC and other cortical regions constrain the types of information that PFC can encode about basic stimulus properties and that, through learning, the task-relevant information becomes distributed more broadly via the intraarea connections within PFC.

Our results also give insight into how new task-relevant information is coded at the population and individual neuron level. At the population level, we observed that the new information was contained in a dynamic population code (13–15), with different neurons carrying information at different latencies relative to the start of the trial (Fig. 3A). These time-dependent representations

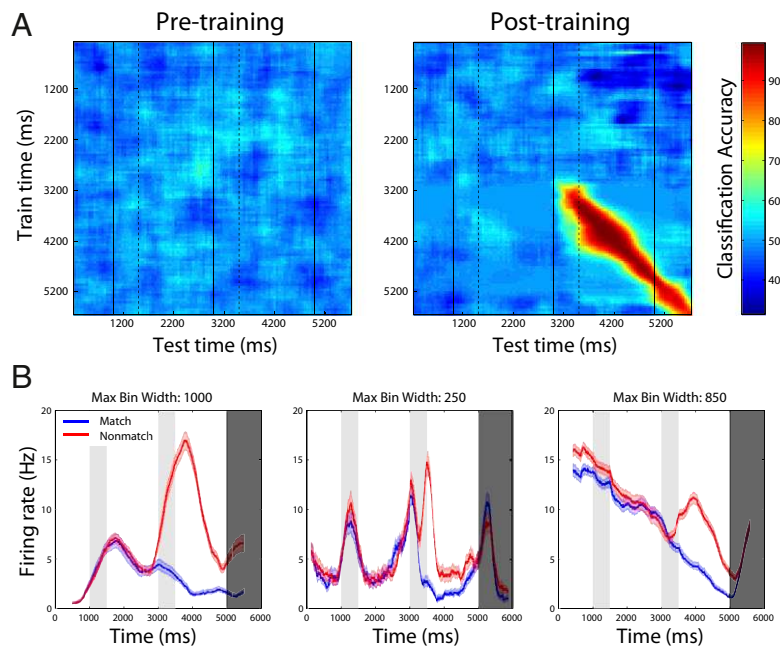


Fig. 3. Dynamic coding of task relevant information after training in the feature task. (A) Results from training a classifier at one time period (*y* axis) and testing the classifier at a second time period (*x* axis) for decoding the match/nonmatch trial status, either pretraining (*Left*) or posttraining (*Right*). The black solid vertical lines indicate the times when the first, second, and match/nonmatch stimuli appeared, and the black dashed lines indicate the offset times of the first and second stimuli. After the monkey was trained, high classification accuracies are seen only when the classifier is built and tested using data from around the same time periods, which shows that different patterns of neuron activity contain the task relevant information at different time points in the experiment. (B) Firing rates for the match trials (blue) and nonmatch trials (red) for the three most selective neurons in feature task. As can be seen, differences in firing rates between match and nonmatch trials appear to be added on top of other firing rate changes that are occurring over the course of a trial (and are carrying information about other variables). Additionally, some neurons (e.g., *Middle*) only contain large firing rate differences between match and nonmatch trials for short periods of time, which give rises to the dynamic coding of information at the population level. (These neurons are typical examples of the population of neurons that have large amounts of match/nonmatch information). Error bars indicate 1 SEM.

could enable the PFC to keep track of when particular events occurred and thus might be involved in the neural representation of time. Our population analyses also revealed that a small subset of highly match/nonmatch selective neurons emerged after training and these neurons contained almost all of the task-relevant information that was present in the larger population (Fig. 2C and Fig. S24). These results raise the possibility that only a small percentage of the population might be critical for processing particular types of information at any one point in time and that the redundant information seen in the rest of the population (Fig. S2B) could help make the circuit more robust in the face of damage to these highly selective neurons. This finding also has implications for the way in which neural data are analyzed; methods that rely exclusively on average selectivity of neurons over the whole population may miss the importance of such highly selective neurons.

Finally, at the single-neuron level, we observed that the new task-relevant information was often contained in relatively short time windows that were present in the midst of other large firing-rate modulations that occurred throughout the trial (Fig. 3B and Fig. S4B). These findings suggest that, unlike the results reported in other brain regions (27), information in PFC is not coded solely by the maximum firing rate of a neuron (e.g., see the rightmost neuron in Fig. 3B), but, rather, that the firing rates of neurons need to be evaluated in relation to the activity of other neurons in the population. Additionally, we observe that these other firing-rate modulations carry information about other variables (Fig. S5), which shows individual neurons in PFC are multiplexing different types of information in a single spiking sequence. Such multiplexing of information could be an efficient strategy that allows the large number of time-dependent representations in PFC to be encoded by a much smaller number of neurons. Overall, these findings shed light on how novel information is incorporated into

PFC activity and how neural activity codes information, which should lead to richer theories of how PFC controls behavior and how information is coded in neural activity more generally.

Methods

Recording Methods and Task. Before training, 726 and 111 neurons were recorded from the 2 rhesus monkeys (*Macaca mulatta*) while they passively viewed the stimuli in the feature task, and 810 and 210 neurons were recorded after training (for the spatial task, 595 and 113 neurons were recorded before training; 814 and 214 were recorded after training). A grid system was used for the recordings, and a map of penetrations was generated by aligning the placement of the electrodes within the grid to a magnetic resonance image of the cortical surface. Dorsal PFC in this study was defined as the area containing the two banks of the principal sulcus (≤ 2 mm from the principal sulcus) and extending posterior to the arcuate sulcus, which incorporates the posterior aspect of area 46 and parts of area 8a. Ventral PFC was defined as the area in the convexity of the PFC lateral to the principal sulcus (>2 mm from the center of the principal sulcus), thus incorporating parts of areas 12 and 45. Neurons were not prescreened for stimulus selectivity; however, if a recorded neuron had a restricted receptive field, the stimuli in the feature task were typically presented in its center. The average signal-to-noise ratio of spike waveforms (in neurons that were significantly modulated by task events) was 8.0 in the pretraining population and 7.2 in the posttraining population (10). To complete the task, the monkeys needed to maintain fixation within 2° of the center of the screen. After training, the monkeys additionally needed to saccade to a green target stimulus on match trials and a blue stimulus on nonmatch trials (the location of the green and blue stimuli were randomly counter-balanced across trials, so that the decoding match/nonmatch information was not merely reflecting a planned movement direction). The stimuli were paired in each experimental session, so that, on nonmatch trials, the first stimulus was always shown with the same second nonmatching stimulus. The surgical procedures, recording methods, task details, and anatomical localization methods have been described previously (9, 10).

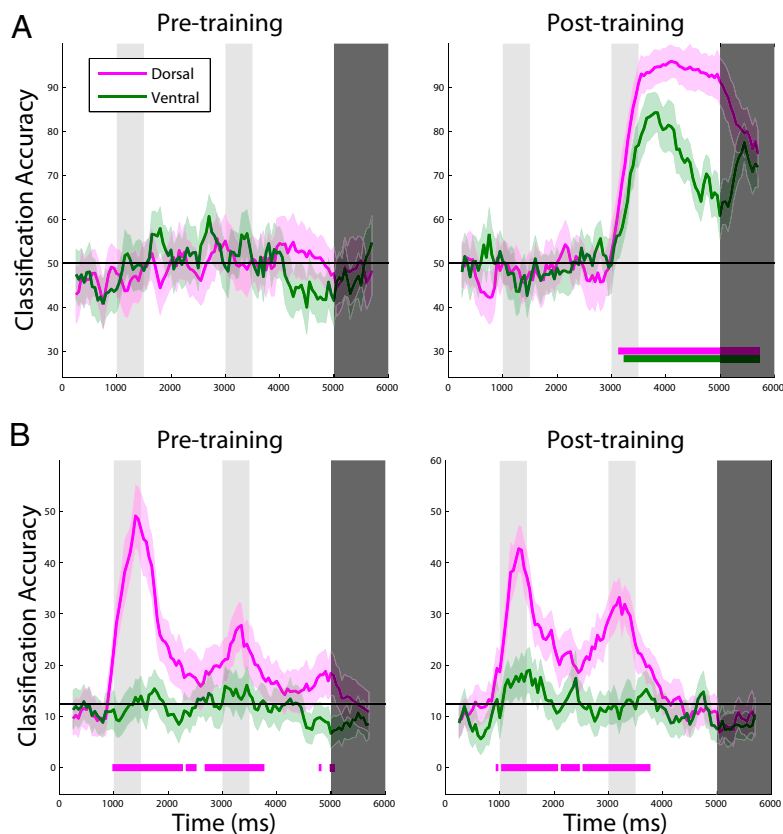


Fig. 4. Comparing information in dorsal PFC (magenta) vs. ventral PFC (green). (A) Match/nonmatch information pretraining (*Left*) and posttraining (*Right*) reveals that, after training, there is task relevant match/nonmatch information in both dorsal and ventral PFC. (B) In contrast, information about which stimulus was shown (stimulus identity information) was seen only in dorsal PFC in both the pretraining and posttraining data (left and right plots, respectively).

Data Analysis. The decoding analysis methods have been described previously (14, 28, 29). Briefly, a resample cross-validation procedure was used in which a maximum correlation coefficient classifier was trained on firing rates of pseudopopulations of neurons (i.e., populations of neurons that were recorded independently but treated as if they were recorded simultaneously). The firing rates were the average spiking activity in 500-ms bins sampled at 50-ms sliding intervals. This classifier was then used to decode the stimulus identity on the feature task and the stimulus position on the spatial task (Figs. 2A and 4B and Figs. S3A and S6B). For these tasks, 11 trials of each stimulus were used for training the classifier, and testing was done using one trial for each stimulus. Eight different stimuli were used in the feature task, and eight different locations were used in the spatial task, so chance decoding on these tasks was $1/8 = 12.5\%$. It should be noted that because for a given neuron the first stimulus was always paired with the same nonmatch stimulus on nonmatching trials, information about the first stimulus after the time when the second stimulus was shown (e.g., Fig. 2A and Figs. S3A and S7B) could be attributable to the second stimulus (30). When decoding the match/nonmatch trial status on the feature task (Fig. 2B, 3A, and 4A), 44 match and nonmatch trials were used for training the classifier and 4 trials from each condition were used for testing. When decoding the match/nonmatch trial status on the spatial task (Figs. S3B, S4A, and S6A), 48 match and nonmatch trials were used for training and 4 trials from each condition were used for testing. The chance decoding accuracy for match/nonmatch information was $1/2 = 50\%$. All neurons that had recordings from 12 repetitions of each stimulus on were included in the identity decoding analysis, and all neurons that have 48 (52) repetitions on the match/nonmatch feature (spatial) task were used. This led to at least 84% of the recordings being used for all analyses. To have maximum power in our results, data from both monkeys were combined in all analyses. To make a fair comparison in the decoding analyses, 750 neurons were randomly selected from the pre- and posttraining datasets and passed to the classifier (on the spatial task, only 600 neurons were used). When comparing the decoding accuracies for dorsal vs. ventral neurons, 250 neurons were

used in the feature task and 200 neurons were used in the spatial task. This procedure was repeated 50 times using different neurons, creating different pseudopopulations and using different training/ test splits each time, and the results were averaged over these 50 runs. The error bars were estimates of 1 SEM decoding accuracy that would occur if different neurons had been selected and were created by sampling the 750 (or 600) neurons with replacement, and taking the SD of the decoding results over 50 bootstrap runs. To evaluate whether the decoding results were above chance, the labels from the trials were randomly shuffled and the decoding procedure was run using 10 bootstrap iterations. This procedure was repeated 200 times to get a null distribution of the decoding accuracies that would occur by chance, and significant time periods were defined as those in which the real decoding accuracy exceeded all of the values in the null distribution [i.e., $P < 1/200$ ($P < 0.005$)].

To find the most selective neurons in Figs. S2, S3D, and S3E, an ANOVA was applied to all of the training data for each neuron, and the neurons with the smallest P value were classified as the most selective neurons. The effect size for the selectivity of individual neurons (Fig. 2C and Fig. S3C) was calculated using the η^2 statistic, which measures the proportion of the variance explained by the match/nonmatch labels (i.e., η^2 is the between-class sum of squares divided by the total sum of squares). The neurons selected in Fig. 3B were chosen by calculating the ANOVA P values using firing rates in bin sizes from 50–1,000 ms, sampled every 5 ms, and choosing the neurons with the smallest P values. Because neurons have different windows of selectivity, the smoothing bin size for these neurons was based on the bin size that led to the smallest P value. (The bin sized used for smoothing is shown above the firing rate plots for each neuron.) To calculate the decrease in decoding accuracy when training and testing the classifier at different times, the delay period was defined as the time period 250 ms after the offset of the second stimulus to 250 ms before the onset of the decision stimuli, which corresponds to 4,250–4,750 ms after the start of the trial.

ACKNOWLEDGMENTS. We thank Tomaso Poggio for his guidance and Gabriel Kreiman, Beata Jarosiewicz, and Ram Ramachandran for their comments on the paper. This research was sponsored by the following grants: National Institutes of Health Grant EY017077, National Science Foundation Grants 0640097 and 0827427, the Defense Advanced Research Projects Agency

Defense Sciences Office, and Air Force Office of Scientific Research Grants FA8650-50-C-7262 and FA9550-09-1-0606. Additional support was provided by The Tab Williams Family Endowment Fund, Adobe, the Honda Research Institute, a King Abdullah University Science and Technology grant (to B. DeVore), NEC, Sony, and especially by The Eugene McDermott Foundation.

1. Miller EK (2000) The prefrontal cortex and cognitive control. *Nat Rev Neurosci* 1: 59–65.
2. Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.
3. Duncan J (2001) An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci* 2:820–829.
4. Nieder A, Freedman DJ, Miller EK (2002) Representation of the quantity of visual items in the primate prefrontal cortex. *Science* 297:1708–1711.
5. Romo R, Brody CD, Hernández A, Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399:470–473.
6. Romanski LM, Goldman-Rakic PS (2002) An auditory domain in primate prefrontal cortex. *Nat Neurosci* 5:15–16.
7. Shima K, Isoda M, Mushiake H, Tanji J (2007) Categorization of behavioural sequences in the prefrontal cortex. *Nature* 445:315–318.
8. Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291:312–316.
9. Meyer T, Qi X-L, Stanford TR, Constantinidis C (2011) Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *J Neurosci* 31: 6266–6276.
10. Qi X-L, Meyer T, Stanford TR, Constantinidis C (2011) Changes in prefrontal neuronal activity after learning to perform a spatial working memory task. *Cereb Cortex* 21: 2722–2732.
11. Roy JE, Riesenhuber M, Poggio T, Miller EK (2010) Prefrontal cortex activity during flexible categorization. *J Neurosci* 30:8519–8528.
12. Cromer JA, Roy JE, Miller EK (2010) Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron* 66:796–807.
13. Machens CK, Romo R, Brody CD (2010) Functional, but not anatomical, separation of “what” and “when” in prefrontal cortex. *J Neurosci* 30:350–360.
14. 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.
15. Warden MR, Miller EK (2007) The representation of multiple objects in prefrontal neuronal delay activity. *Cereb Cortex* 17(Suppl 1):i41–i50.
16. Goldman-Rakic PS (1996) Regional and cellular fractionation of working memory. *Proc Natl Acad Sci USA* 93:13473–13480.
17. Wilson FA, Scalaidhe SP, Goldman-Rakic PS (1993) Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* 260:1955–1958.
18. Rao SC, Rainer G, Miller EK (1997) Integration of what and where in the primate prefrontal cortex. *Science* 276:821–824.
19. Rainer G, Asaad WF, Miller EK (1998) Memory fields of neurons in the primate prefrontal cortex. *Proc Natl Acad Sci USA* 95:15008–15013.
20. Buckley MJ, et al. (2009) Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science* 325:52–58.
21. O'Reilly RC (2010) The what and how of prefrontal cortical organization. *Trends Neurosci* 33:355–361.
22. Wilson CRE, Gaffan D, Browning PGF, Baxter MG (2010) Functional localization within the prefrontal cortex: Missing the forest for the trees? *Trends Neurosci* 33:533–540.
23. Preuss TM, Goldman-Rakic PS (1991) Myelo- and cytoarchitecture of the granular frontal cortex and surrounding regions in the strepsirrhine primate Galago and the anthropoid primate Macaca. *J Comp Neurol* 310:429–474.
24. Badre D, D'Esposito M (2009) Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat Rev Neurosci* 10:659–669.
25. Pandya DN, Yeterian EH (1990) Prefrontal cortex in relation to other cortical areas in rhesus monkey: Architecture and connections. *Prog Brain Res* 85:63–94.
26. Tanji J, Hoshi E (2008) Role of the lateral prefrontal cortex in executive behavioral control. *Physiol Rev* 88:37–57.
27. Gold JJ, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
28. Zhang Y, et al. (2011) Object decoding with attention in inferior temporal cortex. *Proc Natl Acad Sci USA* 108:8850–8855.
29. Meyers EM, Kreiman G (2001) Tutorial on pattern classification in cell recording. *Visual Population Codes*, eds Kriegeskorte N, Kreiman G (MIT Press, Cambridge, MA), pp 517–538.
30. Meyers et al.; The incorporation of new information into prefrontal cortical activity after learning working memory tasks. Available at <http://cbcl.mit.edu/people/emeyers/pnas2012/>.