A Comparison of Abstract Rules in the Prefrontal Cortex, Premotor Cortex, Inferior Temporal Cortex, and Striatum

Rahmat Muhammad[1], Jonathan D. Wallis[1,2], and Earl K. Miller[1]

Abstract

■ The ability to use abstract rules or principles allows behavior to generalize from specific circumstances. We have previously shown that such rules are encoded in the lateral prefrontal cortex (PFC) and premotor cortex (PMC). Here, we extend these investigations to two other areas directly connected with the PFC and the PMC, the inferior temporal cortex (ITC) and the dorsal striatum (STR). Monkeys were trained to use two abstract rules: "same" or "different". They had to either hold or release a lever, depending on whether two successively presented pictures were the same or different, and depending on which rule was in effect. The rules and the behavioral responses were reflected most strongly and, on average, tended to be earlier in the PMC followed by the PFC and then the STR; few neurons in the ITC reflected the rules or the actions. By contrast, perceptual information (the identity of the pictures used as sample and test stimuli) was encoded more strongly and earlier in the ITC, followed by the PFC; they had weak, if any, effects on neural activity in the PMC and STR. These findings are discussed in the context of the anatomy and posited functions of these areas. ■

INTRODUCTION

The ability to generalize principles or rules from experience is central to complex, goal-directed behavior. It endows cognitive flexibility by allowing us to deal with circumstances that have not been directly experienced. For example, we learn the "rules" for restaurant dining from specific experiences and can then apply them to new restaurants. The prefrontal cortex seems ideally situated for the abstraction of such behavior-guiding principles (Miller, Freedman, & Wallis, 2002). It is at the apex of the cortical processing hierarchy; the prefrontal cortex (PFC) is interconnected with all cortical sensory systems as well as premotor cortical areas involved in generating volitional movements (Barbas, 2000; Fuster, 2000). Indeed, PFC activity does reflect abstractions such as perceptual categories, small numbers, and general rules (Nieder, Freedman, & Miller, 2002; Freedman, Riesenhuber, Poggio, & Miller, 2001; Wallis, Anderson, & Miller, 2001). Furthermore, at least one type of abstraction, learned perceptual categories, are more strongly reflected in PFC activity than in the inferior temporal cortex (ITC), a higher order cortical visual area that provides the PFC with highly processed visual information (Freedman, Riesenhuber, Poggio,

& Miller, 2003). Still, the respective contributions of the PFC and other brain structures to rule abstraction are not well understood because direct neurophysiological comparisons between candidate brain areas are rare. Here, we provide a comparison of the roles of four different brain areas in the representation of abstract rules.

We have previously reported that abstract rules "same" and "different" were reflected in neural activity in the lateral PFC and portions of the lateral premotor cortex (PMC) (Wallis & Miller, 2003; Wallis et al., 2001). We targeted these structures because previous studies have implicated them in rule learning and following (Genovesio, Brasted, Mitz, & Wise, 2005; Brasted & Wise, 2004; Toni, Rushworth, & Passingham, 2001; Murray, Bussey, & Wise, 2000; White & Wise, 1999; Passingham, 1993). Much of the previous work has involved learning of specific associations between a sensory cue and a motor response. For our studies, we made the rules "abstract" by training monkeys until they could apply them to novel stimuli that had no preexisting stimulus–response association. However, they are likely to involve similar substrates as specific cue–response (rule) learning because the abstract rules are built on the same if-then type of logic as specific rules. Our studies revealed differences in neural properties that provide clues into their relative functional specializations. Among other things, we found that the abstract rules were reflected more strongly and earlier in PMC activity than PFC activity, suggesting that the PMC is "closer" to the storage of these highly familiar rules than

the PFC. Here, we extend this comparison by adding two more areas that are also directly connected with the PFC and PMC, the caudate nucleus of the striatum (STR) and the anterior ITC.

The ITC was of interest because our monkeys applied the same and different rules to complex visual pictures and the ITC seems to play a major role in the recognition of such stimuli (Tanaka, 1996; Desimone, Albright, Gross, & Bruce, 1984). In addition, it is directly interconnected with the PFC (Seltzer & Pandya, 1989; Barbas, 1988). Furthermore, interactions between the PFC and ITC are necessary for normal learning and retention of conditional visuomotor associations (Bussey, Wise, & Murray, 2002). The STR, specifically the dorsal STR, was of interest because it is directly interconnected with the PFC and PMC and seems to be part of a frontobasal ganglia network for learning goal-directed behaviors (Pasupathy & Miller, 2005; Brasted & Wise, 2004; Nixon, McDonald, Gouhg, Alexander, & Passingham, 2004; Hollerman, Tremblay, & Schultz, 2000; Toni & Passingham, 1999; Winocur & Eskes, 1998; Graybiel, Aosaki, Flaherty, & Kimura, 1994). Furthermore, a recent study (Pasupathy & Miller, 2005) showed that neural correlates for the learning of specific rules (conditional visuomotor associations between a specific cue and a specific response) were more strongly reflected in STR activity than PFC activity. Therefore, we wondered whether abstract rules would produce a similar or different pattern of results.

Here we report differences between the four areas, including (but not limited to), observations that abstract rule and motor-response activity was significantly more abundant and stronger in the frontal cortex (PFC or PMC) than in the STR or ITC, whereas selectivity for the pictures used to make the same and different judgments was strongest in the ITC.

METHODS

Subjects

Three adult rhesus monkeys, *Macaca mulatta* (Monkey A: female, 5 kg; Monkey B: male, 6 kg; Monkey C: male, 11 kg), were used in the experiments. Recordings from the PFC and PMC of Monkeys A and B are described in Wallis and Miller (2003).

Behavioral Task

Two pictures were briefly presented separated by a short memory delay. Depending on which rule was in effect (same or different), the monkeys had to respond (i.e., release a lever) if the pictures were a match or a nonmatch, respectively. A trial began when the monkeys grasped a lever and fixated a central fixation spot (Figure 1). They were required to maintain fixation within 1.5° of the fixation spot for the duration of the trial. After 800 msec of fixation, a sample picture (1.8° in width and height, 800 msec duration) appeared at the center of gaze along with a cue (100 msec duration). The cue signaled the monkey which rule to follow on that trial (see below). The sample picture was followed by a 1500-msec delay and then a test picture.

The test picture was either a nonmatch (different from the sample) or a match (the same as the sample picture). For the ''same'' rule, monkeys had to release the lever if the test picture was a match to receive a reward (a drop of apple juice); if the test picture was a nonmatch, the monkey had to continue holding the lever through a second delay (500 msec). For the ''different'' rule, they had to release the lever if the test picture was a nonmatch; if it was a match, they had to continue holding the lever through the second delay. The second delay was always followed by a picture that required a release response (and subsequent reward). The second delay was used to elicit a behavioral response for each trial and thus ensure that the monkey was paying attention on every trial. It was not used in any of the analyses because only the first test picture required a decision. A different set of four randomly selected sample pictures was used for each daily recording session. Using four pictures ensured that the identity of the nonmatching pictures could not be predicted. As a result, the monkeys had to remember both the current rule and the sample picture (Wallis et al., 2001).

The cues used to signal the rules were either visual, auditory, or gustatory. Two cues, one from each modality, were used for each rule and varied across monkeys. For Monkey A, the ''same'' rule was indicated by a drop of juice or a low tone, and the ''different'' rule was indicated by no juice or a high tone. For Monkey B, juice or blue border signified same, whereas no juice or green border indicated different. For Monkey C, juice or blue border indicated same, whereas no juice or pink border indicated different. The purpose of the multiple, disparate cues for each rule was to determine whether neural activity was reflecting the abstract rule signified by the cue or the physical properties of the cue (Wallis et al., 2001). For each recording session, trials were randomized across all cues, samples, rules, and responses.

Recording Sites

The recording sites are pictured in Figure 2. PFC recordings in Monkey A were from the left hemisphere, and in Monkeys B and C were from both hemispheres. Because there was very little difference in sulci position between hemispheres and monkeys, the recording sites are plotted on the same figures. PMC recordings were bilateral in both Monkeys A and B. The positions of the recording chambers were determined from magnetic resonance imaging (MRI) scans. The ITC and STR chambers were positioned above areas TEa and the head and body of the caudate nucleus, respectively. ITC and STR recordings were bilateral in Monkey B and in the left
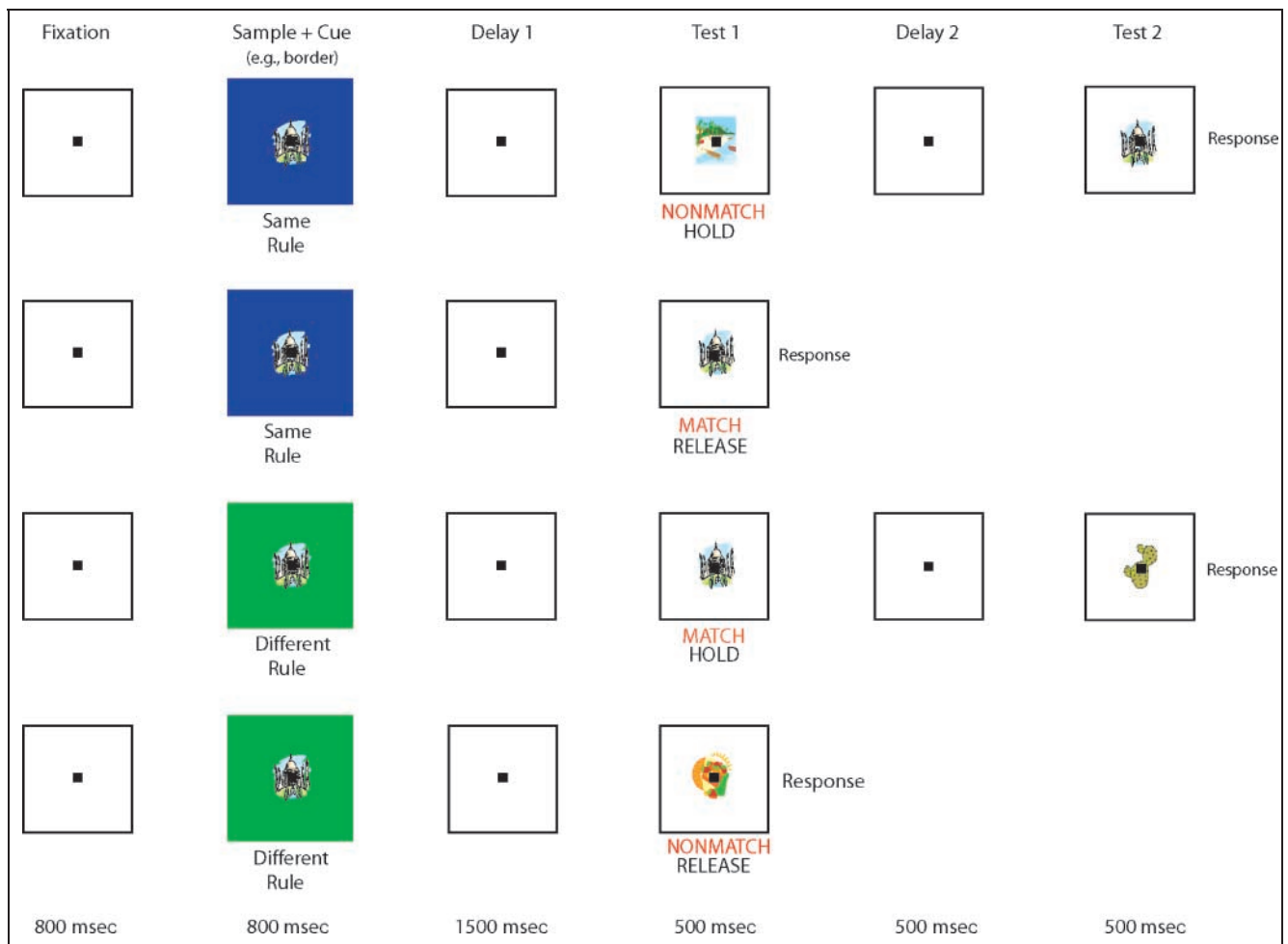
Figure 1. A schematic diagram of the behavioral task. After fixating, monkeys are simultaneously presented with a sample picture and rule-signaling cue (see Methods). The sample picture is followed by a delay and a test picture. Depending on the rule ("same" or "different"), monkeys had to either hold or release the lever when the first test object was a match or a nonmatch to the sample object. The second test object always required a release response.

ITC and right STR in Monkey C. All of these recordings were conducted with multiple electrodes, from 8 to 24 electrodes implanted in one to three brain areas simultaneously. Simultaneous recordings from the PFC, ITC, and STR were conducted in Monkey C. This allowed a detailed comparison of neural properties while avoiding potential confounding factors (e.g., differences in the level of the monkey's experience with the task during recordings from different areas.

Data Analysis

We compared the activity of four different brain regions (PFC, PMC, ITC, and STR) during performance of the same/different abstract rule task. The PMC neuron population and a subset of the PFC neuron population are the same as previously reported in Wallis and Miller (2003). We have added PFC data from Monkey C and ITC and STR data from Monkeys B and C.

For some analyses, we divided the trials into three contiguous nonoverlapping epochs: sample, delay, and

test. The sample epoch was from sample onset to sample offset. The delay epoch began after sample offset and lasted until the end of the delay. The test epoch began with test picture onset and lasted until its offset (500 msec if no behavioral response, typically earlier with a behavioral response). Baseline activity was averaged over the 500 msec prior to sample onset. All analyses were conducted using data from correct trials.

We compared neural activity across the four brain regions using a sliding receiver operating characteristic (ROC) analysis and analyses of variance (ANOVAs). ROC was used to quantify how strongly different aspects of the task were encoded. Briefly, an ROC analysis measures the degree of overlap between two response distributions. For each selective neuron, the preferred and unpreferred conditions were compared, giving two distributions, P and U respectively, of neuronal activity. For example, for a rule-selective neuron these distributions would be the neuron's firing rates when the same rule was in effect and when the different rule was in effect. An ROC curve was then generated by taking each
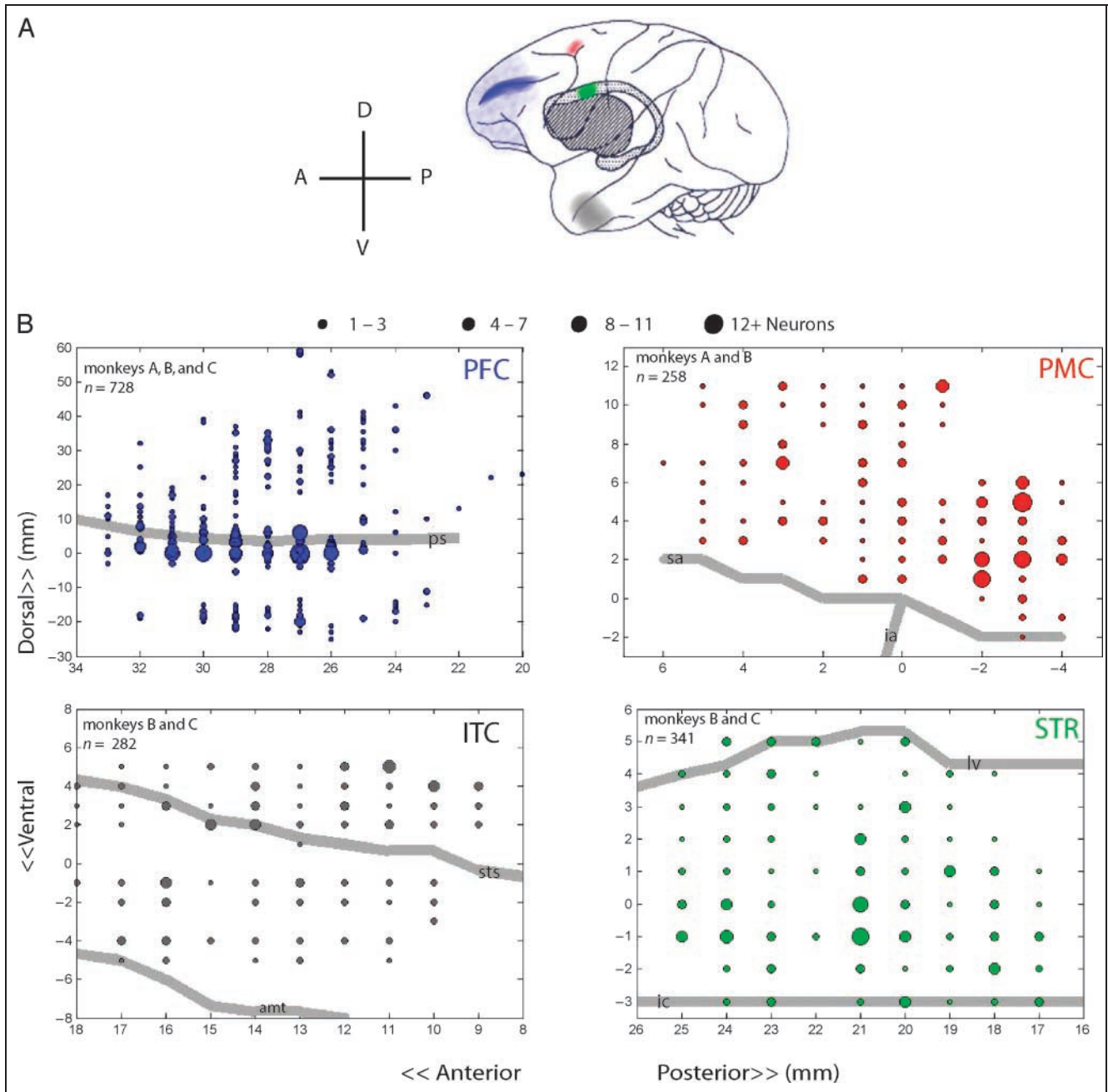
Figure 2. Recording locations in the PFC, PMC, ITC, and STR. The general areas of recording are differentially shaded on a lateral view of a rhesus macaque brain. The intensity of shading within each area is an indication of the number of recording sites in that region. Locations of the PMC (area 6/F2) recordings (shaded red) in monkeys A and B were dorsal to the superior arcuate (sa). Recordings from PFC (shaded blue) include portions of dorsolateral PFC (areas 9, 46, and 9/46), ventrolateral PFC (areas 47/12 and 45), and orbitofrontal PFC (areas 11, 13, and 14). ITC recording sites (TEa shaded gray) from monkeys B and C were between the anterior medial temporal sulcus (amt) and the superior temporal sulcus (sts). The location of dorsal STR recordings (shaded green) from monkeys B and C was confined to the more anterior part of the caudate. D = dorsal; V = ventral; A = anterior; P = posterior. Flattened representations of electrode penetration sites for each area are shown in the bottom of the figure. The size of the dots indicates the number of recordings performed at that site. The numbers of monkeys used and neurons recorded from each area are indicated. The reconstruction method for PFC and PMC recording sites is described in Wallis and Miller (2003). The same method was used to reconstruct ITC and STR recording sites. In all cases the anterior–posterior position was measured with respect to the interaural line. The dorsoventral position was measured with respect to the principal sulcus for PFC recordings, the genu of the arcuate sulcus for PMC recording, the superior temporal sulcus for ITC recordings, and the internal capsule for STR recordings. ps = principal sulcus; sa = superior arcuate sulcus; ia = inferior arcuate sulcus; sts = superior temporal sulcus; amt = anterior medial temporal sulcus; ic = internal capsule; lv = lateral ventricle.

observed firing rate of the neuron and plotting the proportion of P that exceeded the value of that observation against the proportion of U that exceeded the value of that observation. The area under the ROC curve was then calculated. A value of 0.5 would indicate that the two distributions completely overlap (because the proportion of P and U exceeding that value is equal), and as such that the neuron is not selective. A value of 1.0, on the other hand, would indicate that the two distributions are completely separate (i.e., every value drawn from U is exceeded by the entire of P, whereas none of the values of P are exceeded by any of the values in U) and so the neuron is very selective. This method of analysis has the advantage that it is independent of the neuron's firing rate and so can be used to compare neurons with different baseline firing rates and dynamic ranges. It is also nonparametric and so does not require the distributions to be gaussian. Furthermore, the ROC value can be thought of as the probability that an independent observer could identify the condition that had been presented using the neuron's firing rate.

To compare ROC values between areas, we used a Wilcoxon's rank sum test assessed at $p < .01$. The ROC was also used to measure the time course of neuronal selectivity thus allowing estimation of each neuron's selectivity latency. The ROC was computed by averaging activity over a 200-msec window that was slid in 10-msec steps over the course of the trial. To measure latency, we used the point at which the sliding ROC curve equaled or exceeded 0.6 for three consecutive 10-msec bins. Latency was defined as the center of the first time bin. Although 0.6 is an arbitrary criterion, it was chosen because it yielded latency values that compared favorably with values that would be determined by visually examining the spike density histograms. Other measures yielded similar results, such as values reaching three standard deviations above baseline ROC values and when ROC values exceeded the 99th percentile of the baseline values. Power analysis was used to determine if a sufficient number of neurons had reached criterion to meaningfully compare the latency of selectivity between areas. A bootstrap analysis was used to determine if the ROC values were significantly different from chance (for details, see Wallis & Miller, 2003).

We used a three-way ANOVA to identify neurons whose average firing rate during the sample and delay epochs varied significantly with trial factors (evaluated at $p < .01$). The factors used were the modality of the cue (Monkey A: taste/auditory cue, Monkey B: taste/visual cue), the rule that the cue signified (same or different), and which of the four pictures was presented as the sample stimulus. We defined rule-selective neurons as those that showed a significant difference in their firing rates between the two different rules, regardless of either the cue that was used to instruct the monkey or the picture that was used as the sample stimulus. Thus, a rule-selective neuron was one that showed a main effect

of rule and no interaction with the other two factors. We also used this analysis to define picture-selective neurons (those that had a main effect of picture and no interaction with the other two factors). Differences between two areas were determined using a chi-square test at $p < .05$.

We tested for three different effects during the test epoch. Selectivity for the test picture was determined using a Kruskal–Wallis one-way ANOVA. Neurons whose activity varied with the match/nonmatch status of the test picture were assessed with a Wilcoxon's rank sum test. Selectivity for the behavioral response (go vs. nogo) was determined by a Wilcoxon's rank sum test to compare activity during the hold versus release trials. For all of these tests, a criterion level of $p < .01$ was used and differences were determined using chi-square tests.

RESULTS

Behavior

The performance of all three monkeys indicates that they were proficient at the task (Monkey A, 85% correct; Monkey B, 95% correct; Monkey C, 89% correct). Overall performance was slightly but significantly better during sessions where we recorded from the PMC and STR compared to PFC, but in all cases the monkeys' performance was at a high level (PMC: 93%, $n = 17$; PFC: 88%, $n = 82$; STR: 93%, $n = 50$; ITC: 90%, $n = 29$; Tukey HSD $[0.01] = 3.59$, $p < .01$).

We used Wilcoxon's matched-pairs signed rank test to compare the animals' performance on the same versus different trials, on the trials for which the cue was juice/ no juice versus visual or auditory cues, and on hold or release trials. Performance on the same and different trials was identical and remained so across the recording sessions for each area ($p > .3$). Likewise, there was no preference to hold or release the lever across these sessions ($p > .2$). As previously noted (Wallis et al., 2001), the monkeys performed better for the juice/ no-juice cues (91% correct) than for the visual or auditory cues (85% correct), probably because the juice/ no-juice cues were very salient, but this also remained constant across recording sessions for all four areas ($p < 5 \times 10^{-4}$).

Neuronal Properties

We recorded from a total of 1609 neurons from all three monkeys across all four areas (PFC: $n = 728$; PMC: $n = 258$; ITC: $n = 282$; STR: $n = 341$). A three-way ANOVA (see Methods) on each neuron's average activity across the sample or delay epochs was used to determine the percentage of neurons that showed significant selectivity for the rules (same and different), the sample picture (four different pictures), and the motor response (hold vs. release, evaluated at $p < .01$). We will

first consider the effects of rules and pictures; the motor response factor will be discussed later with other end-of-trial effects. To determine if the percentages of neurons showing effects were significantly different between areas, we used chi-square tests with a Bonferroni-corrected alpha level of .00167 for multiple comparisons. For all comparisons below, we use all recorded neurons in each area, regardless of whether or not they showed any task-related selectivity or even responsiveness. All the differences we report were apparent when we only considered neurons with significant selectivity. However, using all recorded neurons is superior because it gives us a comparison of the neuronal properties in each brain area that is unbiased by any statistical selection criterion.

Rule Selectivity

Figure 3 summarizes the proportion of neurons in each area that showed significant rule selectivity and/or picture selectivity during the sample and delay epochs. The PFC showed a significantly greater proportion of rule selective neurons in the sample versus the delay epochs (chi-square $p < .05$) and the PFC and ITC showed a significantly greater incidence of picture selectivity in the sample versus the delay epoch (chi-square $p < .05$). For simplicity in comparisons across areas, we will collapse across the sample and delay epochs; if a neuron showed selectivity in both the sample and delay epochs, it was only counted once. Note that rule and picture-selectivity are not mutually exclusive; a neuron could be picture selective in the sample period and rule selective in the delay period.

There was a significantly greater incidence of rule selectivity in the PMC (48% of all recorded neurons or 125/258) than the PFC (41% or 297/728), a greater incidence in the PFC than the STR (26% or 89/341), and a greater incidence in the STR than the ITC (12% or 34/282, chi-square, all comparisons $p < .01$). In all areas, about half of the rule neurons were more strongly activated by same and half were more strongly activated by different.

Figure 4 shows the activity of a ''rule-selective'' neuron in the PFC. This neuron had a higher firing rate throughout the sample and delay epochs during the trials in which the different rule was in effect than when the same rule was in effect. Note that the level of activity for each rule was similar regardless of which specific cue signaled the rule. Its rule-related activity was also similar regardless of which sample picture the monkey held in memory.

To quantify the strength of rule selectivity, we applied the sliding ROC analysis (see Methods) on the activity of each and every recorded neuron (we did not preselect neurons for any task-related selectivity or even responsiveness). The ROC values are proportional to the absolute difference in firing rate for each neuron between same trials and different trials relative to neural variance. The values range from 0.5 (i.e., no difference in firing rate and therefore no rule information) to almost 1 (i.e., perfect discrimination between rules). Furthermore, because the ROC analysis was ''slid'' in 10-msec steps, we could estimate the latency for each neuron to begin to show a rule effect (see Methods).

Figure 5 shows plots of the ROC values for each and every recorded neuron for all four brain areas. Each

Figure 3. Proportion of neurons selective for the rules and for the pictures during the sample and delay epochs. Each bar represents the percentage of neurons out of the total number recorded in each brain region, which showed significant selectivity for the rules or the picture (three-way ANOVA, $p < .01$). The black portions of each bar represent the fraction of neurons that preferred the rules. The gray portion represents the fraction of neurons that preferred pictures. Proportions are collapsed across the sample and delay epochs and if a neuron showed an effect in both epochs, it is only counted once. *Significant difference in incidence of rule selectivity; **significant difference in incidence of picture selectivity.
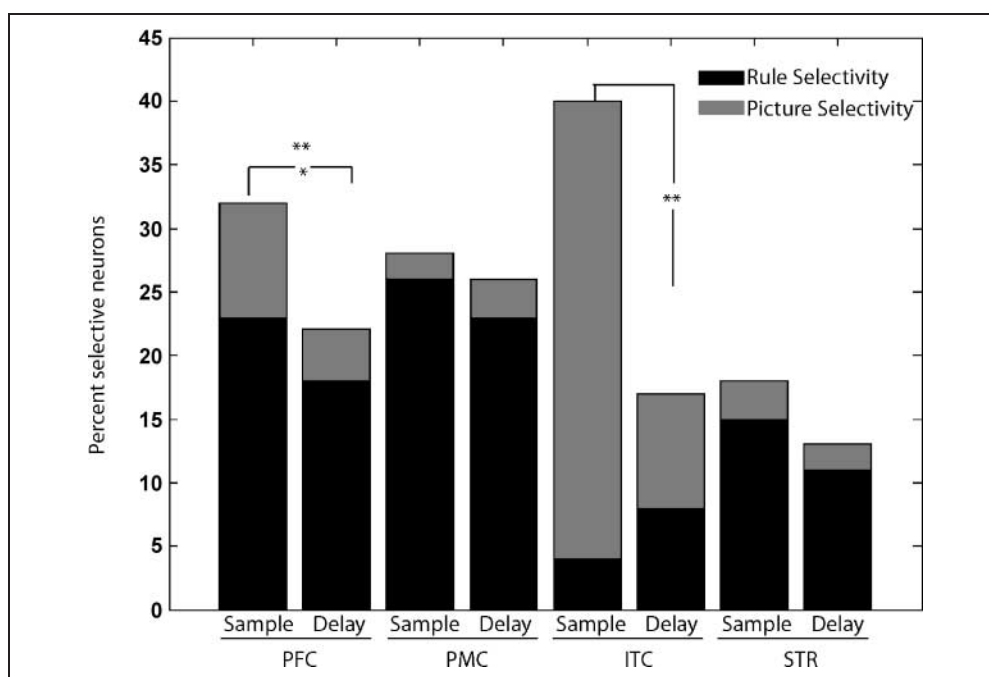
Figure 4. Average firing rate histograms and raster plots from a rule-selective neuron recorded from the PFC. Bin width: 50 msec bins. The simultaneous onset of sample and 100-msec cue stimulus is indicated by the gray bar. Sample offset is indicated by the line at 800 msec. Each horizontal raster line corresponds to neural activity on a single trial. Plots are color coded by task condition (see legend).
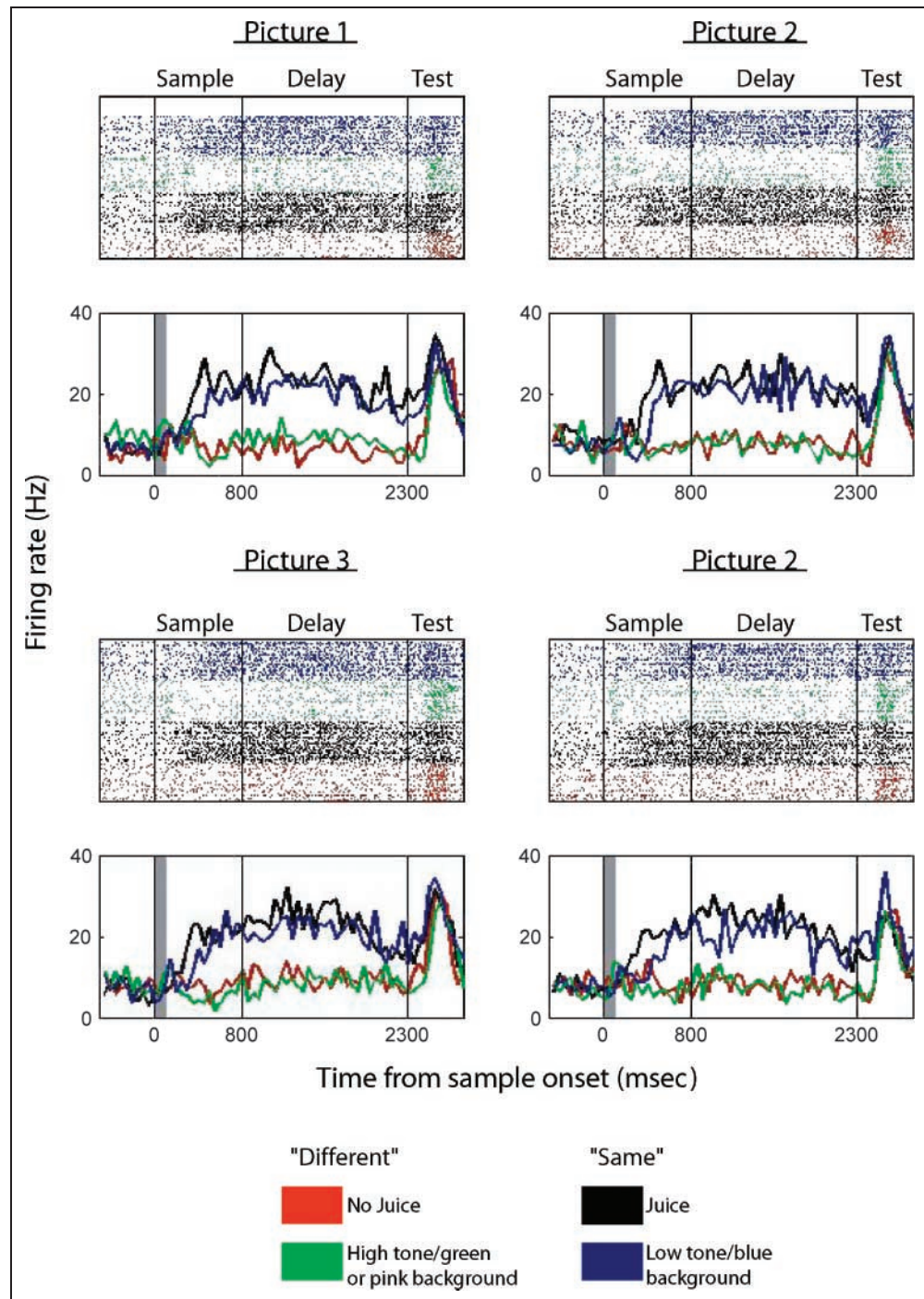

horizontal line shows ROC data for a single neuron across the course of the trial. The plots are sorted by each neuron's rule effect latency when the latency could be estimated; neurons that did not reach criterion for determining latency (see Methods) were left unsorted. As the ANOVAs suggested (see above), rule selectivity seemed overall strongest in the PMC, followed by the PFC, then by the STR, and weakest of all in the ITC.

For a direct comparison of the strength of rule selectivity across areas, we calculated the ROC values using the mean firing rate of each neuron across the sample and delay epochs. We compared the mean ROC values across the entire population of recorded neurons from each area by using a Wilcoxon's rank sum test. Table 1 summarizes the results. The average ROC values are relatively low because we averaged across all recorded neurons regardless of whether they showed a rule effect or even any task responsiveness. However, the pattern of significant differences was consistent with the above analyses. Rule selectivity was stronger in the
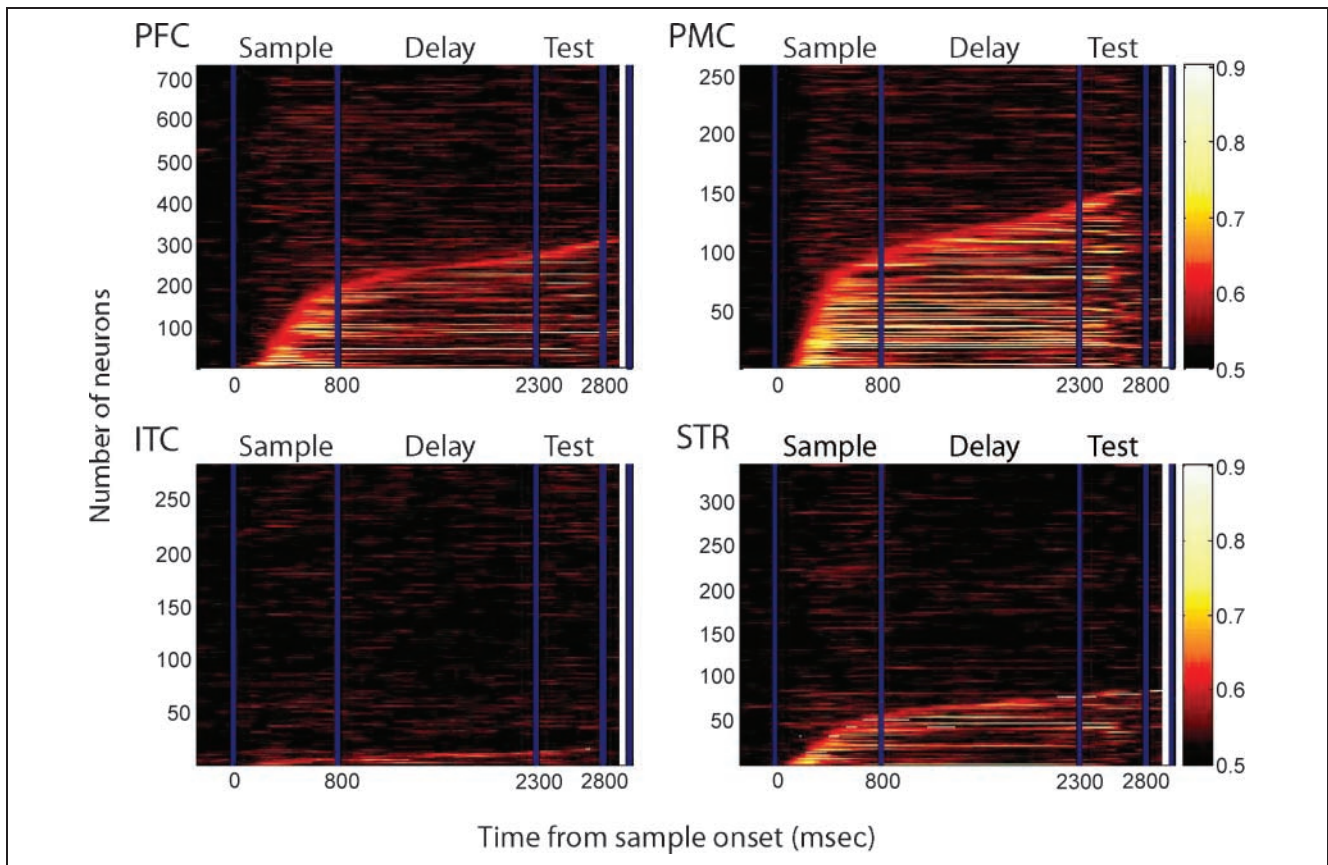
Figure 5. Sliding ROC analysis of rule selectivity. For each brain area, neurons with ROC values that equaled or exceeded 0.6 for three consecutive 10-msec time bins were sorted according to latency. Neurons that did not reach criterion were left unsorted. Each line corresponds to the ROC values for a single neuron over a 200-msec window "slid" in 10-msec steps over the length of the trial. All recorded neurons are included in this plot.

PMC followed by the PFC then the STR and finally the ITC (see Table 1).

Figure 6 shows the distribution of latencies for neurons that reached the criterion for determining latency (see Methods). This yielded 202 PFC, 98 PMC, 7 ITC, and 57 STR neurons (ITC neurons are not included in Figure 6 because so few neurons showed a rule effect). The latencies are highly variable, but there were significant differences between the populations. On average, rule selectivity appeared significantly earlier in the PMC (median = 280 msec) than in the PFC (median = 370 msec; Wilcoxon's rank sum test, $p < .05$). STR latencies (median = 350 msec) were not significantly different from those of the PFC or PMC. A power analysis (see Methods) indicated that the small number of ITC neurons with effects did not allow for a statistically meaningful comparison.

Picture Selectivity

In addition to remembering which rule was currently in effect, monkeys also had to identify and remember the sample image. Consequently, many neurons showed selectivity for the four images used as samples (and test stimuli) each day. Figure 7 shows an example of a single neuron from the ITC. It had a higher burst of activity after sample onset for one of the four pictures (Picture 3) regardless of the different rules or cues.

Figure 3 shows the incidence of picture-selective neurons in each area as determined by three-way ANOVA (described above and in Methods). In comparing the proportion of neurons with picture selectivity across areas, we again collapsed across the sample and delay epochs, and neurons showing effects in both epochs were only counted once. The pattern was quite different from that seen for rule selectivity. The proportion of picture-selective neurons was highest in the ITC (45% of all neurons or 126/282), followed by the PFC (13% or 94/728), and finally the PMC (5% or 12/258) and STR (4% or 15/341, chi-square test, all $p < .01$). The incidence of picture selectivity in the PMC and STR were not significantly different ($p = .96$).

A similar pattern of results was obtained with a sliding ROC analysis conducted on each and every recorded neuron (Figure 8). These ROC values were calculated using the difference in activity between the most and

Table 1. Strength of Selectivity for Task Factors Averaged Across the Sample, Delay, and Test Epochs as Determined by ROC Analysis

| | Rule | | Picture | | | | |
	Sample	Delay	Sample	Delay	Test	Match/Nonmatch Test	Response Test
Median ROC values							
PFC	0.5397	0.546	0.5378	0.5346	0.5419	0.5217	0.5379
PMC	0.5489	0.5379	0.5315	0.5328	0.5319	0.523	0.5723
ITC	0.5204	0.5213	0.5844	0.5466	0.5798	0.5237	0.5178
STR	0.5269	0.5282	0.5284	0.5274	0.5303	0.5151	0.5305
Wilcoxon's rank sum test p values							
PFC vs. PMC	0.005	<0.001	<0.001	0.1	<0.001	0.9	<0.001
PFC vs. ITC	<0.001	<0.001	<0.001	<0.001	<0.001	0.1	<0.001
PFC vs. STR	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.1
PMC vs. ITC	<0.001	<0.001	<0.001	<0.001	<0.001	0.5	<0.001
PMC vs. STR	<0.001	<0.001	0.1	0.01	0.5	<0.001	<0.001
ITC vs. STR	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

The values in this table are the median ROC values across all (randomly selected) recorded neurons, regardless of their responsiveness or selectivity. They are accompanied by p values from multiple Wilcoxon's rank sum tests comparing the areas.

least preferred pictures (see Methods). Once again, each line corresponds to one neuron and we sorted the traces by their picture-selectivity latency or they were left unsorted. Picture selectivity was strongest in the ITC followed by the PFC and it was weak in both the PMC and STR.

This was confirmed by comparing average ROC values to the most and least preferred pictures using activity averaged across the sample and delay epochs. The results are summarized in Table 1. Again, the average ROC values are relatively low because they are averaged across every recorded neuron with no preselection based on significant effects or responsiveness. Picture selectivity was strongest in the ITC, followed by the PFC and finally the PMC and STR.

We used the sliding ROC analysis to determine latencies for picture selectivity following sample onset; 140 PFC, 15 PMC, 149 ITC, and 31 STR neurons reached latency criterion (see Methods). Again, each population showed variability in individual latencies and both areas show a relatively high proportion of neurons with short latencies. However, there were differences in the population medians. It was significantly shorter in the ITC (median = 160 msec) than the PFC (median = 220 msec, $p < .01$). Although picture selectivity median latency occurred later in the STR (median = 330) and the PMC (median = 280), not enough neurons reached criterion in these areas to allow meaningful statistical comparisons (power test, see Methods). Figure 9 shows
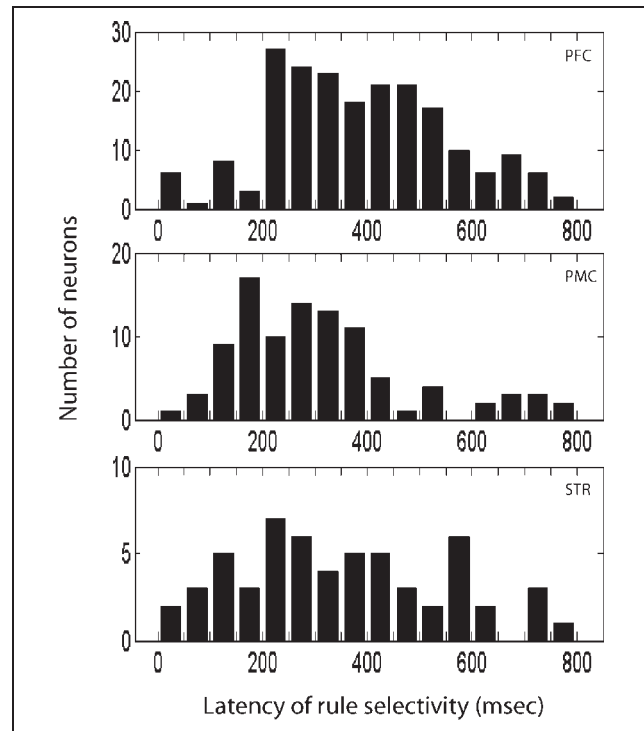
Figure 6. Latency of rule selectivity for single neurons. Distribution of latencies of the onset of rule selectivity following cue stimulus for all neurons for which latency could be determined. Latency was defined as the point at which the values of the sliding ROC analysis equaled or exceeded 0.6 for three consecutive 10-msec time bins (see Methods).

Figure 7. Average firing rate histograms and raster plots from a picture-selective neuron recorded from the ITC. Bin width: 50-msec bins. See Figure 4 for conventions.
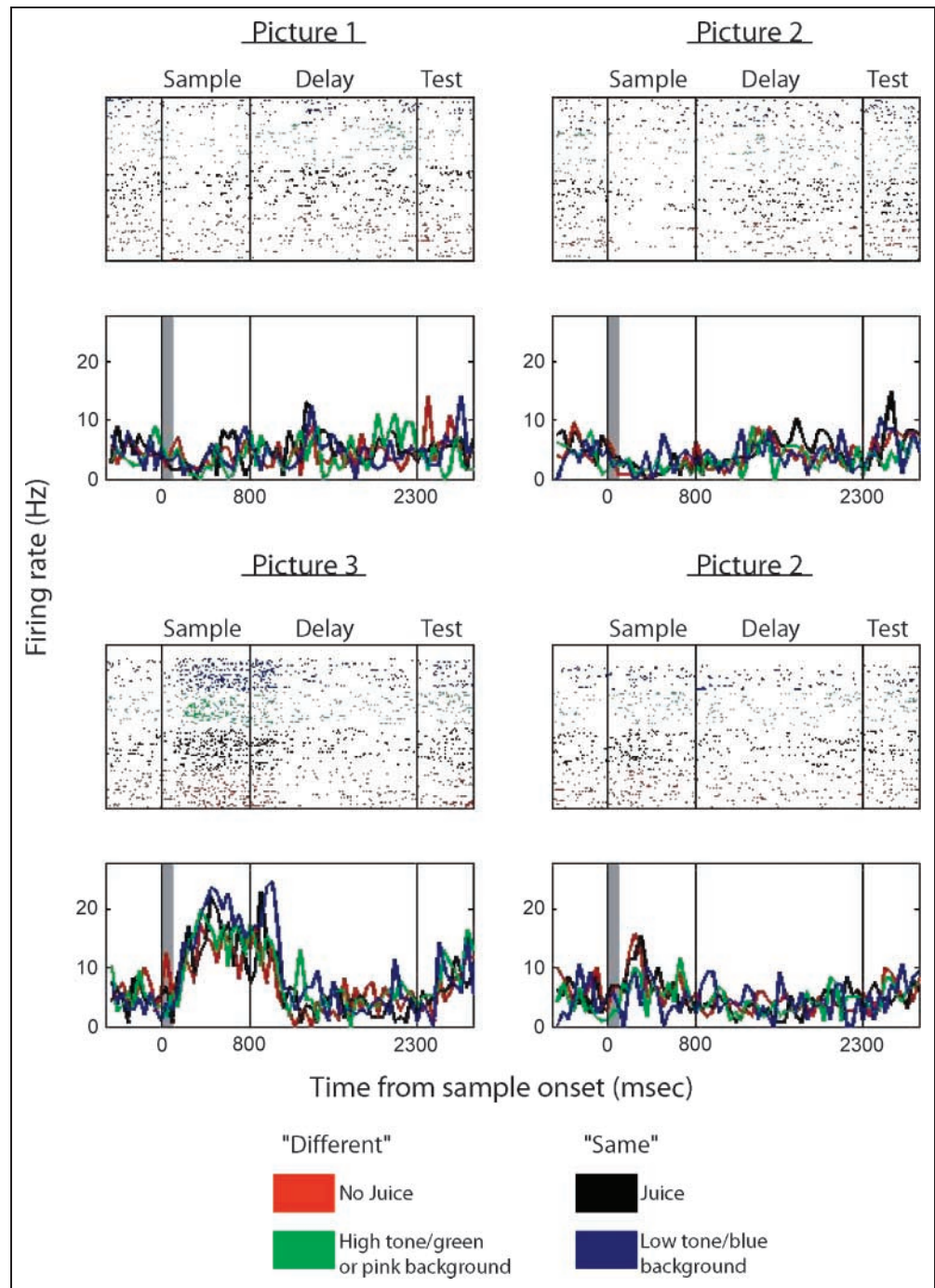

the latency distributions for the PFC and ITC, the only two areas with a large proportion of picture-selective neurons.

Activity during the Test Epoch

During the test epoch, the monkeys saw the second (test) picture and determined if it was a match or nonmatch to the sample picture. They then responded by either releasing or continuing to hold the lever depending on the current rule and the match/nonmatch

status of the test picture. We compared neural correlates of match/nonmatch selectivity and selectivity related to the behavioral response (hold vs. release).

Figure 10 shows examples of two single neurons whose activity varied with the match/nonmatch status of the test picture. One showed stronger activity to nonmatches (Figure 10A), the other stronger activity to matches, regardless of the rule or behavioral response required (Figure 10B). We identified such neurons by using a t test (assessed at $p < .01$) that compared average test epoch activity on match versus nonmatch
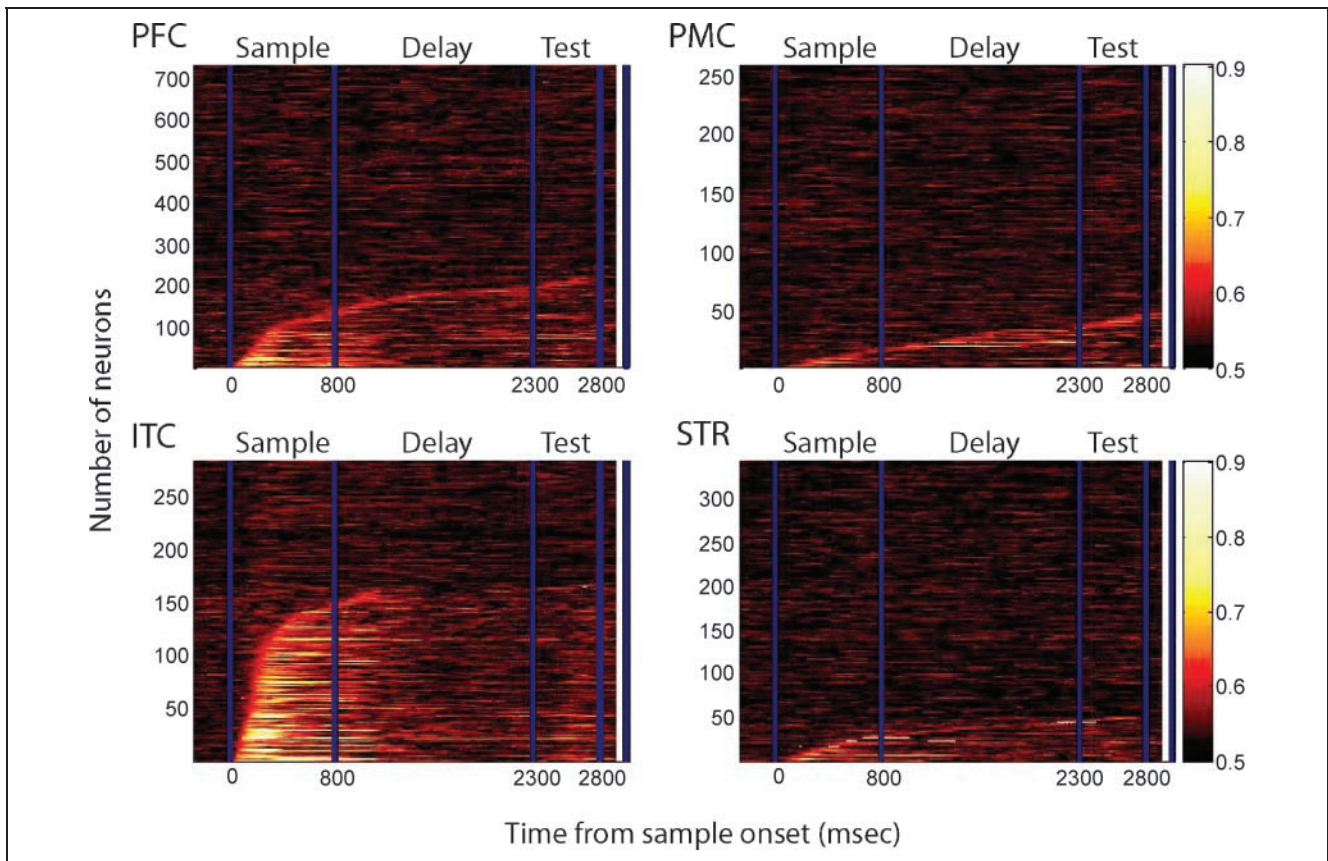
Figure 8. Sliding ROC analysis of picture selectivity. See Figure 5 for conventions.

trials. We found that a similar proportion of neurons in the PFC (21% or 152/728), PMC (24% or 61/258), and ITC (23% or 64/282) had a small but significant match/nonmatch effect. There was a significantly smaller proportion of these neurons in the STR (13% or 45/341, chi-square, $p < .01$).

Sliding ROC analyses for match versus nonmatch are shown in Figure 11. ROC values based on activity averaged across the test epoch indicated that the match/nonmatch effect was significantly weaker in the STR when compared to the PFC, PMC, and ITC (Table 1); it did not differ significantly among the latter three areas. The small numbers of PMC ($n = 12$), ITC ($n = 24$), and STR ($n = 17$) selective for match/nonmatch activity did not allow for a statistically meaningful comparison of match/nonmatch onset latency.

Neurons whose activity reflected the behavioral response (hold or release) are shown in Figure 12A (PMC) and B (PFC). Both had a higher firing rate on release versus hold trials regardless of the rule or match/nonmatch status of the test picture. We identified neurons that showed an effect of the behavioral response using a t test (assessed at $p < .01$), on average test epoch activity. There were significant differences in the proportion of selective neurons for all comparisons between areas (chi-square, all $p < .01$). Behavioral

response selectivity was significantly most prevalent in the PMC (72% or 187/258) followed by the PFC (47% or 339/728), the STR (41% or 140/341), and finally the ITC (18% or 51/282). This was also illustrated by the sliding ROC analysis (Figure 13). Furthermore, the ROC values from the average activity during the test period (see Methods) indicated strongest average effects of the behavioral response in the PMC, followed by the PFC and STR, which were not different from each other, and finally by the ITC (Table 1). The latencies of behavioral response activity could be determined for 268 PFC, 160 PMC, 30 ITC, and 123 STR neurons. Overall, it was significantly earlier in the PMC (median = 280 msec) relative to the PFC (median = 340 msec; Wilcoxon's rank sum test, $p < 1 \times 10^{-7}$) and also earlier relative to the STR (median = 340 msec; Wilcoxon's rank sum test, $p < 1 \times 10^{-3}$; see Figure 14). The number of ITC neurons that reached criterion was too small for a meaningful statistical comparison (power analysis, see Methods).

DISCUSSION

This study compared and contrasted neural correlates of rule-guided actions in four brain areas: the prefrontal, premotor, and inferior temporal cortices and the dorsal
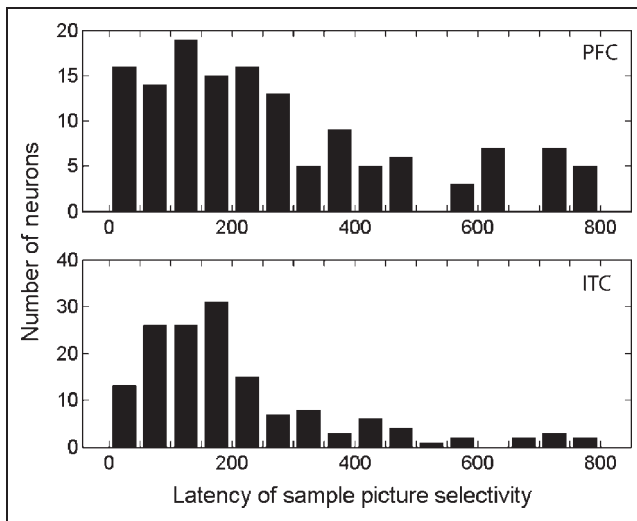
Figure 9. Latency of sample picture selectivity. See Figure 6 for conventions.

striatum. As in our previous study comparing the PFC and PMC (Wallis & Miller, 2003), we found some overlap: Two or more task variables (the rules, the pictures, the match/nonmatch status of the test picture and the behavioral responses) were reflected in the activity of every area tested and the PFC reflected them all. However, there were differences. The rules and the behavioral responses were reflected most strongly and, at least, on average, tended to be earlier in the PMC

followed by the PFC and then the STR; few neurons in the ITC reflected the rules or the actions. By contrast, perceptual information (the identity of the pictures used as sample and test stimuli) was encoded more strongly and, on average, earlier in the ITC, followed by the PFC; they had weak, if any, effects on neural activity in the PMC and STR. The match/nonmatch status of the test picture had the weakest effect, but it tended, on average, to appear in the PFC first.

The Perception–Action Arc

It seems that the PFC was more of a ''crossroad'' for this task that than the other three areas; it was the one area where all the major task variables were represented. This, of course, makes sense because the PFC is at an anatomical crossroad. It is the only brain area in this study that is directly interconnected with the other three and, in general, it is one of the most well-connected brain areas, directly connected with most of the cerebral cortex (including the PMC and ITC) and many subcortical structures (such as the dorsal STR).

The relatively strong representation of perceptual information in the ITC, rule representation/response information in the PMC, and both in the PFC fits well with their conceptualization as cortical components of a perception–action arc (Fuster, 1995). Perceptual information (identity of the pictures, match/nonmatch status) was strongest and tended to appear earliest in the

Figure 10. Average firing rate histograms and raster plots from PFC neurons with match/nonmatch effects. See Figure 4 for conventions. Onset and offset of the test stimulus are indicated by the thin, vertical lines. Gray bar: mean reaction time of monkeys on ''release trials'' ± 1 standard deviation.
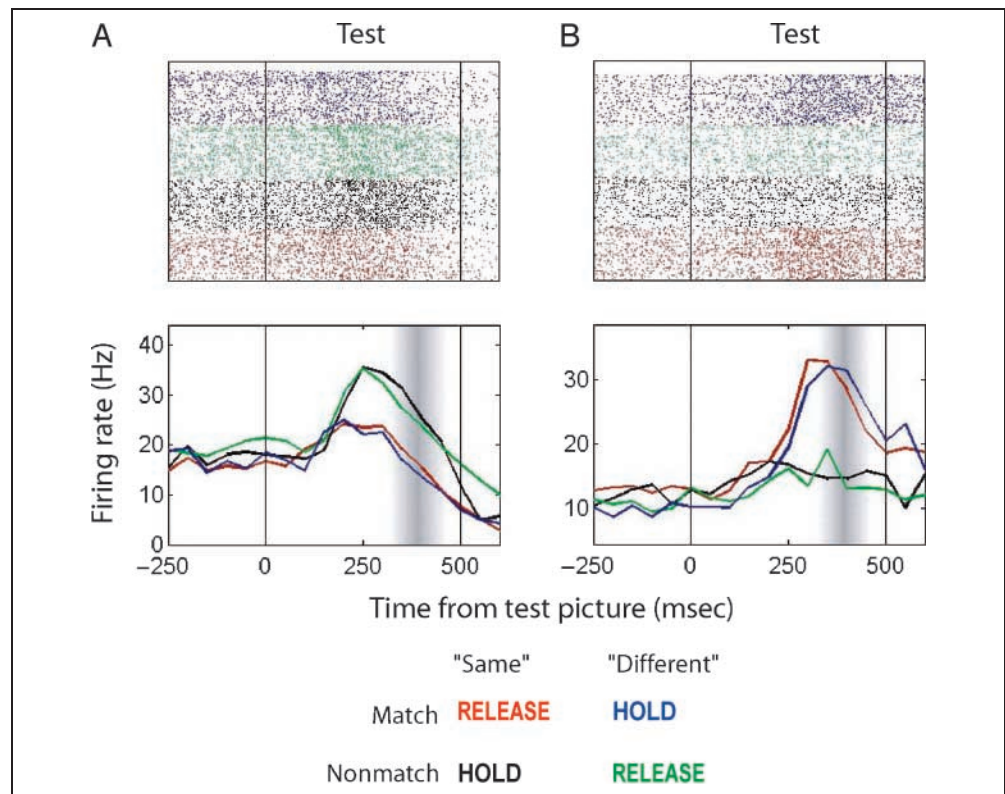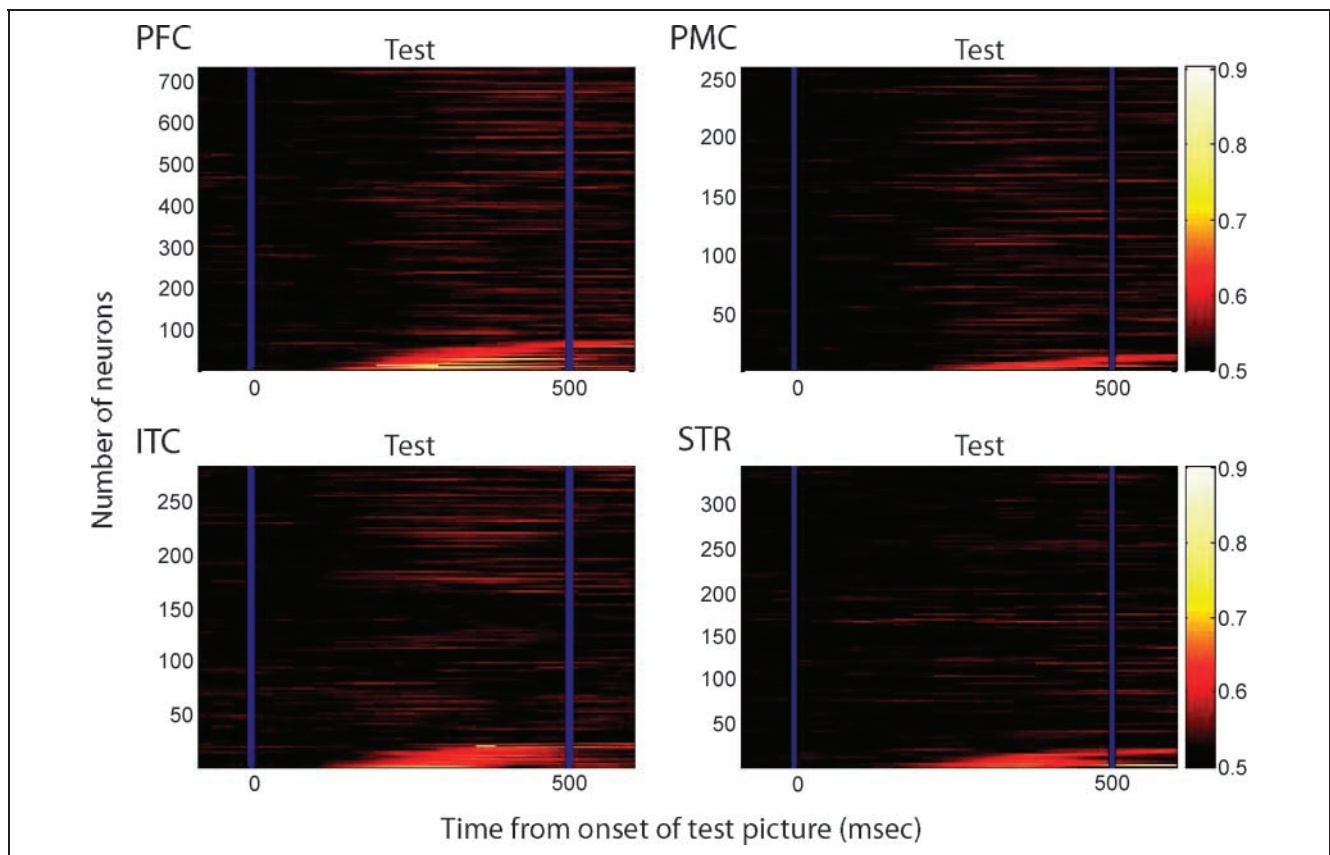
Figure 11. Sliding ROC analysis of match/nonmatch selectivity. See Figure 5 for conventions.

ITC, a temporal sensory cortical area long thought to play a central role in object recognition, and then in the PFC, which receives direct projections from the ITC. The ITC does not project directly to the PMC (Passingham, 1993) and dorsal STR, and perceptual information was weakest in the PMC and STR. By contrast, more action-related information (the rules and behavioral response) were strongest and earliest in the frontal cortex (PFC and PMC) and the STR, which receives direct projections from them. They have long been associated with generating volitional movement.

Rules and Responses

Abstracted, generalized rules are advantageous because they are highly efficient. They can be applied to many circumstances and thus forgo the necessity to learn anew about, and memorize details of, every specific experience. Deficits in switching between different abstract rules are a cardinal feature of PFC damage (Stuss et al., 2000; Owen, Roberts, Polkey, Sahakian, & Robbins, 1991; Nelson, 1976; Milner, 1963) and we have previously reported an abundance of PFC and PMC neurons that encoded the rules used here as well as the stronger (and tendency for earlier) effects in the PMC than the PFC (Wallis & Miller, 2003).

The stronger PMC rule effects may be because the rules were highly familiar to the animals; they had performed this task for over a year. Evidence suggests that the PFC is more critical for new learning than for familiar routines. PFC damage preferentially affects new learning; animals and humans can still engage in complex behaviors as long as they were well learned before the damage (Dias, Robbins, & Roberts, 1997; Knight, 1984; Shallice, 1982; Shallice & Evans, 1978), and PFC neurons are more strongly activated during new learning than during the performance of familiar cue–response associations (Asaad, Rainer, & Miller, 1998). Human imaging studies report a decrease in blood flow to the PFC as a task becomes more familiar (Raichle et al., 1994) and greater blood flow to the dorsal PMC than the PFC when subjects are performing familiar versus novel tasks (Boettiger & D'Esposito, 2005). Also, with increasing task familiarity, there is a relative shift in blood flow from areas associated with focal attention, such as the PFC, to motor regions (Della-Maggiore & McIntosh, 2005). Therefore, it may be that the PFC is primarily involved in new learning, but with familiarity, rules become more strongly established in motor system structures. Although both the PMC and STR receive inputs from the PFC, our study suggests that the PMC has primacy; its effects of rule (and behavioral response)

Figure 12. Average firing rate histograms and raster plots for a PMC neuron (A) and a PFC neuron (B) with behavioral response-related activity. See Figure 10 for conventions.

	"Same"	"Different"
Match	RELEASE	HOLD
Nonmatch	HOLD	RELEASE
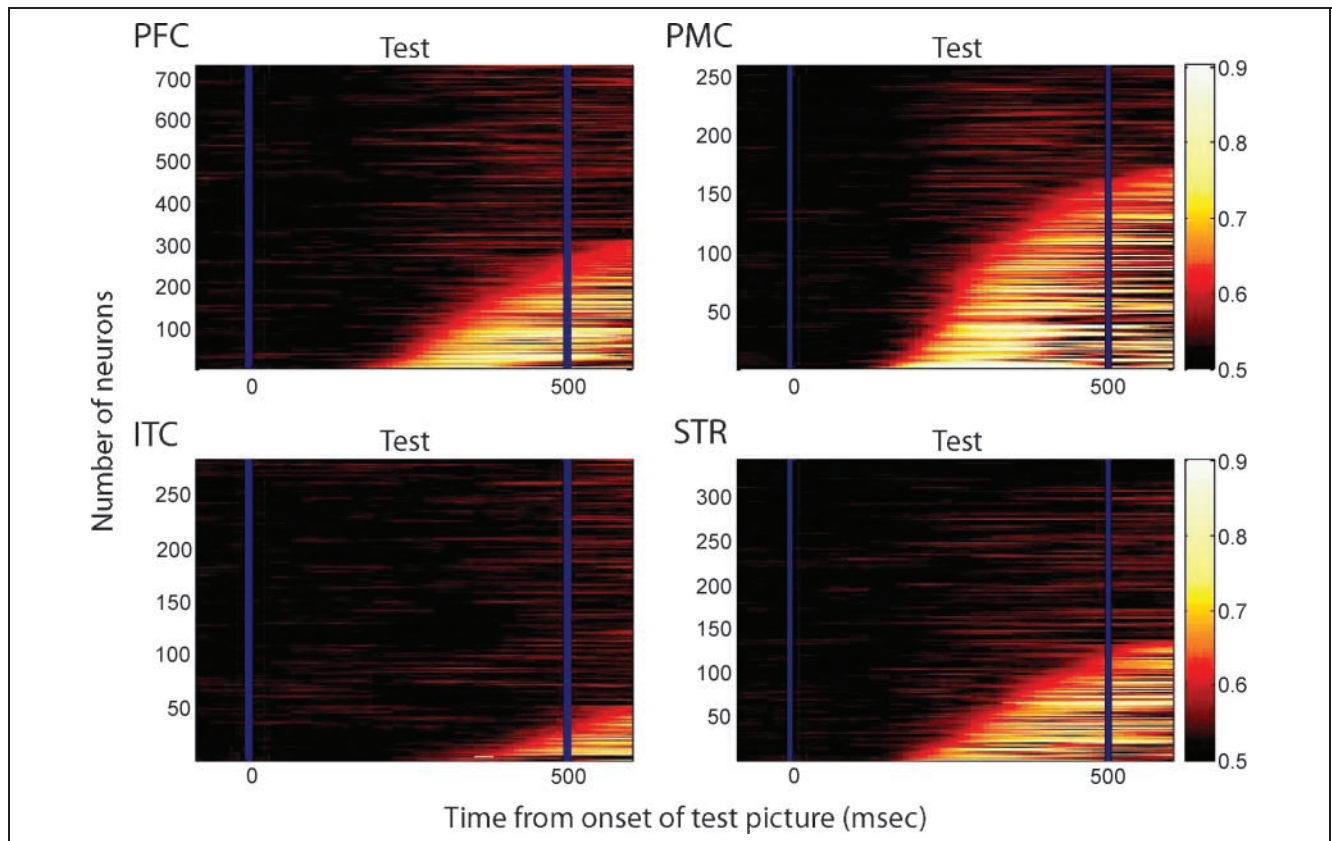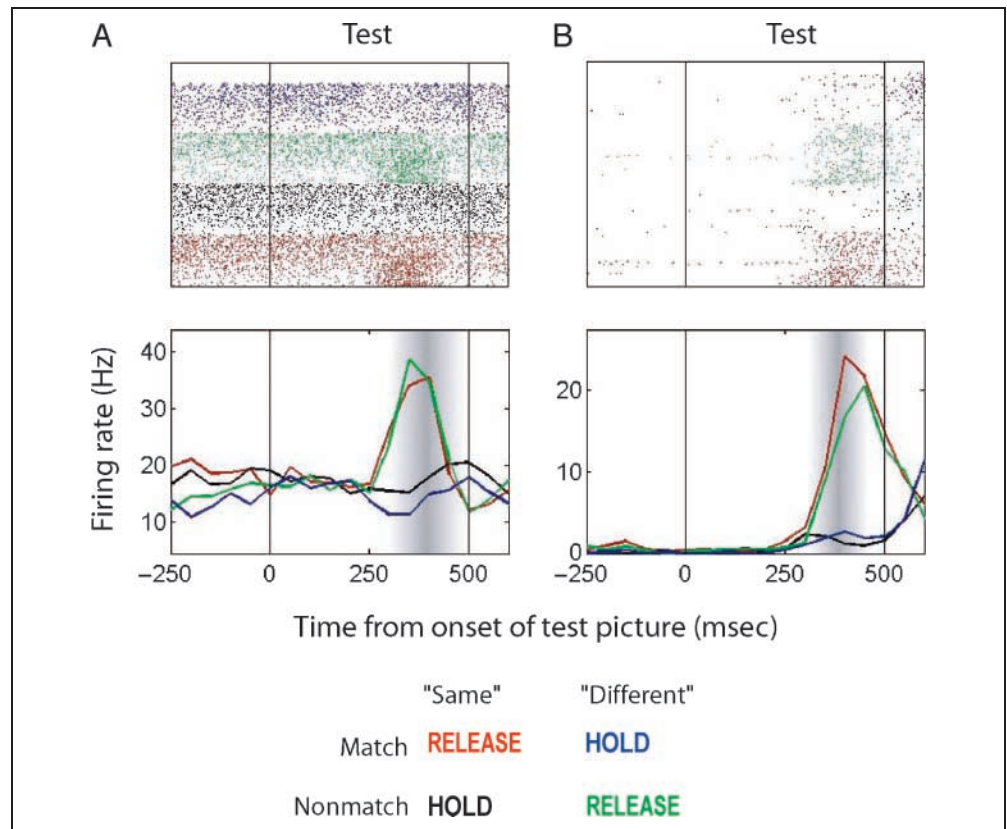
Figure 13. Sliding ROC analysis of behavioral response selectivity. See Figure 5 for conventions.
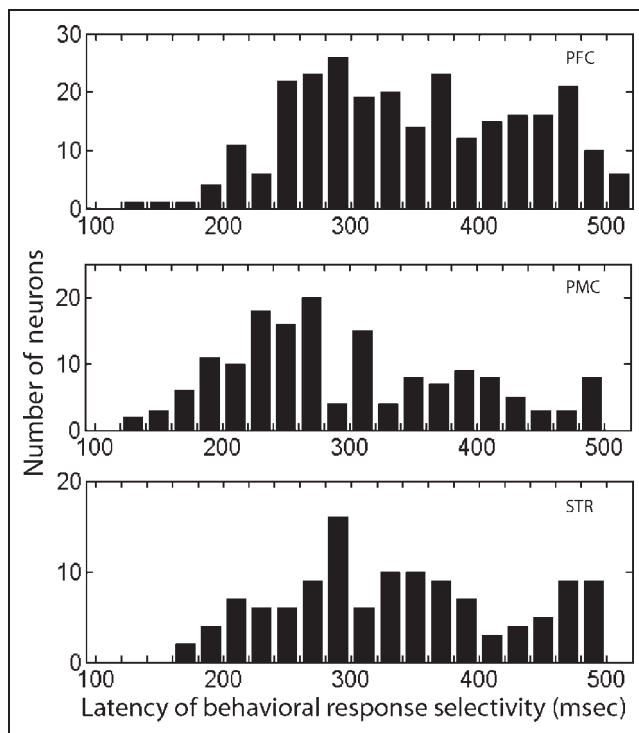
Figure 14. Distribution of estimated latencies (see Methods) for neurons showing behavioral response-related activity. See Figure 6 for conventions.

were stronger and earlier than the STR making it more likely that this information flows from the PMC to the STR rather than the other way around. We do not mean to imply, however, that the PMC is not involved in new learning. Many studies have shown that it is (Passingham, 1993). Rather, it just may be that the PFC is more involved in new arbitrary learning than the PMC.

In principle, the weaker representation of rules in the STR could be due to their familiarity; a recent fMRI study found significantly greater blood flow to the anterior striatum when the subjects were learning rules than when they were using familiar rules (Boettiger & D'Esposito, 2005). However, it also possible that the rules were weaker in the STR because of their abstract nature (i.e., they were not tied to specific cue–response associations). The STR does seem central to the learning and following of specific stimulus–response associations (Packard & Knowlton, 2002), and a recent study indicates a more robust representation of familiar, specific cue–response associations in the STR than the PFC (Pasupathy & Miller, 2005).

Conclusions

In sum, our four-way comparison between different brain areas yielded a pattern of results consistent with known anatomical connections among them as well as their posited general roles in brain function, visual

recognition in the temporal lobe, and action control (rules and responses) in the frontal lobe. It also indicated that familiar abstract rules were stronger in the premotor than prefrontal cortex and weaker still in the dorsal striatum. Although this provides some insight into their respective contributions to this particular behavior, further insight can be gained by determining whether similar or different patterns of neural representation occur during different conditions, such as the learning of new rules and/or following specific stimulus–response associations.

REFERENCES

Asaad, W. F., Rainer, G., & Miller, E. K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron, 21,* 1399–1407.

Barbas, H. (1988). Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *Journal of Comparative Neurology, 276,* 313–342.

Barbas, H. (2000). Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. *Brain Research Bulletin, 52,* 319–330.

Boettiger, C. A., & D'Esposito, M. (2005). Frontal networks for learning and executing arbitrary stimulus–response associations. *Journal of Neuroscience, 25,* 2723–2732.

Brasted, P. J., & Wise, S. P. (2004). Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *European Journal of Neuroscience, 19,* 721–740.

Bussey, T. J., Wise, S. P., & Murray, E. A. (2002). Interaction of ventral and orbital prefrontal cortex with inferotemporal cortex in conditional visuomotor learning. *Behavioral Neuroscience, 116,* 703–715.

Della-Maggiore, V., & McIntosh, A. R. (2005). Time course of changes in brain activity and functional connectivity associated with long-term adaptation to a rotational transformation. *Journal of Neurophysiology, 93,* 2254–2262.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4,* 2051–2062.

Dias, R., Robbins, T. W., & Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: Restriction to novel situations and independence from "on-line" processing. *Journal of Neuroscience, 17,* 9285–9297.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science, 291,* 312–316.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A comparison of primate prefrontal and inferior

temporal cortices during visual categorization. *Journal of Neuroscience, 23,* 5235–5246.

Fuster, J. M. (1995). *Memory in the cerebral cortex.* Cambridge: MIT Press.

Fuster, J. M. (2000). Executive frontal functions. *Experimental Brain Research, 133,* 66–70.

Genovesio, A., Brasted, P. J., Mitz, A. R., & Wise, S. P. (2005). Prefrontal cortex activity related to abstract response strategies. *Neuron, 47,* 307–320.

Graybiel, A. M., Aosaki, T., Flaherty, A. W., & Kimura, M. (1994). The basal ganglia and adaptive motor control. *Science, 265,* 1826–1831.

Hollerman, J. R., Tremblay, L., & Schultz, W. (2000). Involvement of basal ganglia and orbitofrontal cortex in goal-directed behavior. *Progress in Brain Research, 126,* 193–215.

Knight, R. T. (1984). Decreased response to novel stimuli after prefrontal lesions in man. *Electroencephalography and Clinical Neurophysiology, 59,* 9–20.

Miller, E. K., Freedman, D. J., & Wallis, J. D. (2002). The prefrontal cortex: Categories, concepts and cognition. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 357,* 1123–1136.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology, 9,* 100–110.

Murray, E. A., Bussey, T. J., & Wise, S. P. (2000). Role of prefrontal cortex in a network for arbitrary visuomotor mapping. *Experimental Brain Research, 133,* 114–129.

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex, 12,* 313–324.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science, 297,* 1708–1711.

Nixon, P. D., McDonald, K. R., Gough, P. M., Alexander, I. H., & Passingham, R. E. (2004). Cortico-basal ganglia pathways are essential for the recall of well-established visuomotor associations. *European Journal of Neuroscience, 20,* 3165–3178.

Owen, A. M., Roberts, A. C., Polkey, C. E., Sahakian, B. J., & Robbins, T. W. (1991). Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia, 29,* 993–1006.

Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience, 25,* 563–593.

Passingham, R. (1993). *The frontal lobes and voluntary action.* Oxford, UK: Oxford University Press.

Pasupathy, A., & Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature, 433,* 873–876.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex, 4,* 8–26.

Seltzer, B., & Pandya, D. N. (1989). Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *Journal of Comparative Neurology, 281,* 97–113.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 298,* 199–209.

Shallice, T., & Evans, M. E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex, 14,* 294–303.

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., Hamer, L., Murphy, K. J., & Izukawa, D. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: Effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia, 38,* 388–402.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19,* 109–139.

Toni, I., & Passingham, R. E. (1999). Prefrontal–basal ganglia pathways are involved in the learning of arbitrary visuomotor associations: A PET study. *Experimental Brain Research, 127,* 19–32.

Toni, I., Rushworth, M. F., & Passingham, R. E. (2001). Neural correlates of visuomotor associations. Spatial rules compared with arbitrary rules. *Experimental Brain Research, 141,* 359–369.

Wallis, J. D., Anderson, K. C., & Miller, E. K. (2001). Single neurons in the prefrontal cortex encode abstract rules. *Nature, 411,* 953–956.

Wallis, J. D., & Miller, E. K. (2003). From rule to response: Neuronal processes in the premotor and prefrontal cortex. *Journal of Neurophysiology, 90,* 1790–1806.

White, I. M., & Wise, S. P. (1999). Rule-dependent neuronal activity in the prefrontal cortex. *Experimental Brain Research, 126,* 315–335.

Winocur, G., & Eskes, G. (1998). Prefrontal cortex and caudate nucleus in conditional associative learning: Dissociated effects of selective brain lesions in rats. *Behavioral Neuroscience, 112,* 89–101.