

Neuronal Correlates of Metacognition in Primate Frontal Cortex

Paul G. Middlebrooks^{1,2,3,4,*} and Marc A. Sommer^{1,2,3,5}

¹Department of Neuroscience

²The Center for the Neural Basis of Cognition

³The Center for Neuroscience at the University of Pittsburgh
University of Pittsburgh, Pittsburgh, PA 15213, USA

⁴Department of Psychology, Vanderbilt University, Nashville, TN 37240, USA

⁵Department of Biomedical Engineering, the Center for Cognitive Neuroscience, and the Duke Institute for Brain Sciences,
Duke University, Durham, NC 27708, USA

*Correspondence: paul.g.middlebrooks@vanderbilt.edu
<http://dx.doi.org/10.1016/j.neuron.2012.05.028>

SUMMARY

Humans are metacognitive: they monitor and control their cognition. Our hypothesis was that neuronal correlates of metacognition reside in the same brain areas responsible for cognition, including frontal cortex. Recent work demonstrated that nonhuman primates are capable of metacognition, so we recorded from single neurons in the frontal eye field, dorsolateral prefrontal cortex, and supplementary eye field of monkeys (*Macaca mulatta*) that performed a metacognitive visual-oculomotor task. The animals made a decision and reported it with a saccade, but received no immediate reward or feedback. Instead, they had to monitor their decision and bet whether it was correct. Activity was correlated with decisions and bets in all three brain areas, but putative metacognitive activity that linked decisions to appropriate bets occurred exclusively in the SEF. Our results offer a survey of neuronal correlates of metacognition and implicate the SEF in linking cognitive functions over short periods of time.

INTRODUCTION

We not only perform cognitive functions, we also evaluate and alter them. For example, after creating a lecture, we may reflect on how we organized its content. If the lecture is not ready yet, we may think about how to improve its logical structure. Monitoring and controlling cognitive processes is called metacognition (Flavell, 1976).

Researchers have incorporated metacognition into psychological frameworks (Nelson and Narens, 1990) and attempted to localize its neuronal basis in the human brain. Metacognitive skills are impaired in patients with lesions of medial and lateral frontal cortex (Pannu et al., 2005; Schnyer et al., 2004) and in subjects who experience transcranial magnetic stimulation over dorsolateral prefrontal cortex (Rounis et al., 2010). Functional magnetic resonance imaging has implicated multiple brain regions involved in metacognition, including dorsolateral

prefrontal cortex (Kikyo et al., 2002), medial prefrontal cortex (Chua et al., 2006), and cingulate cortices (Chua et al., 2006; Kikyo et al., 2002).

Little is known about how the brain encodes metacognitive processes at the single neuron level. An animal model would facilitate such research, and recent behavioral studies have provided evidence for some degree of metacognition in rats (Foote and Crystal, 2007), dolphins (Smith et al., 1995), rhesus monkeys (Hampton, 2001; Smith et al., 1998), and orangutans (Suda-King, 2008). When offered the chance to take a test or decline it, these animals may opt-out on relatively difficult trials, ensuring a small reward rather than risking no reward if they take the test and fail it. Gorillas, chimpanzees, bonobos, orangutans (Call, 2010), and rhesus monkeys (Hampton et al., 2004) seek information to improve future decisions, an example of metacognitive control, and rhesus monkeys can be trained to bet whether a past decision was correct or incorrect, an example of metacognitive monitoring (Kornell et al., 2007). We recently designed a streamlined version of such a betting task that involves visual stimuli and saccadic eye movement reports, and we reported evidence that monkeys can monitor their own decisions (Middlebrooks and Sommer, 2011).

Here, we recorded from single neurons in macaque frontal cortex during the betting task to search for neuronal activity related to metacognition, which we hypothesized may colocalize with neuronal activity related to cognition. Only two studies previously recorded single neuron activity related to possible metacognitive processing. Kiani and Shadlen (2009), using an opt-out task, reported that neuronal activity in monkey lateral intraparietal cortex correlated with choices to abort a task. Kepecs et al. (2008), using a delayed reward task, showed that activity in rat orbitofrontal cortex predicted whether an animal would opt out of waiting for reward after an incorrect decision. Our task is fundamentally different from the opt-out tasks used in both prior studies. A monkey had to make a decision and then place a bet on the correctness of that decision (Figure 1A). Appropriate wagers required retrospective monitoring, a metacognitive process. Every trial contained the same sequence of task events, and every trial required the monitoring of decisions, allowing us to directly compare activity between trials to identify neuronal correlates of decision-making, wagering, and monitoring.

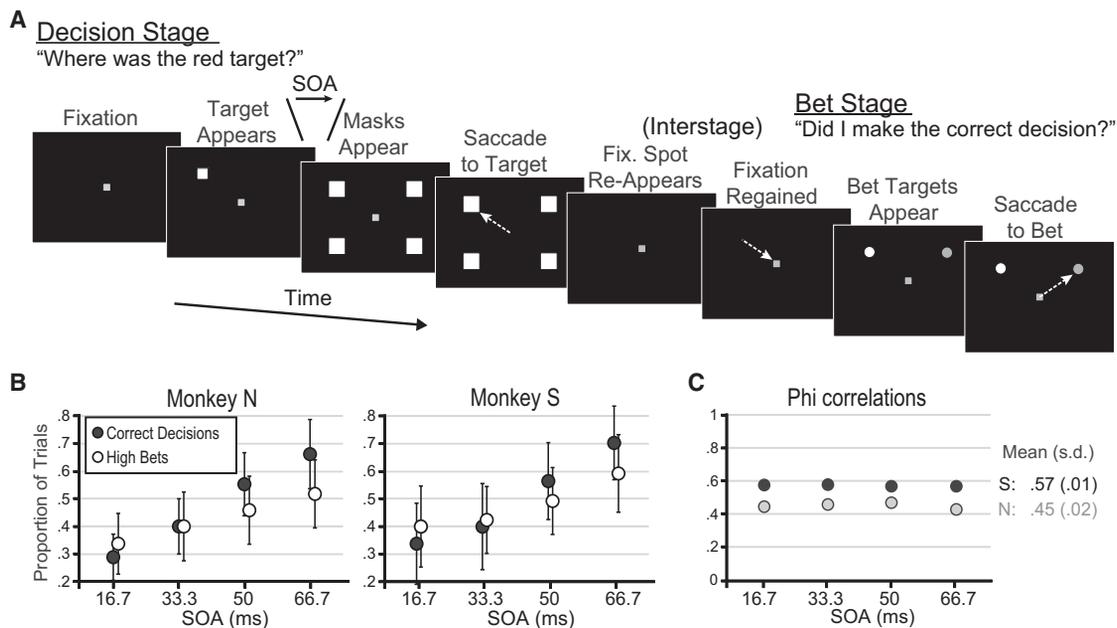


Figure 1. Task and Behavior

(A) Each trial consisted of a decision stage and a bet stage, separated by an interstage period. In the decision stage, monkeys foveated a fixation spot, a target appeared at one of four locations, and after a variable stimulus onset asynchrony (SOA), masks appeared at the four locations. A correct decision (shown) was made if a saccade (arrow) went to the target location. A saccade to any other location was an incorrect decision (not shown). During the interstage period, the fixation spot reappeared and monkeys foveated it to initiate the bet stage. Two bet targets appeared and monkeys placed their bet by making a saccade to one of them. They received the outcome of the bet (reward or timeout) to end the trial.

(B) Overall proportion of correct decisions (black circles) and high bets (white circles) made by each monkey as a function of SOA. Error bars represent standard deviations (SD).

(C) Overall phi correlations (Kornell et al., 2007) for Monkeys N (gray) and S (black) as a function of SOA. Mean and SD across SOAs are shown to the right. See also Figure S1.

We recorded from neurons in three frontal areas: the frontal eye field (FEF), dorsolateral prefrontal cortex (PFC), and supplementary eye field (SEF). Each area has neuronal activity related to vision, saccades, and reward (Boch and Goldberg, 1989; Bruce and Goldberg, 1985; Ding and Hikosaka, 2006; Funahashi et al., 1991; Kim et al., 2008; Mohler et al., 1973; Roesch and Olson, 2003; Russo and Bruce, 1996; Stuphorn et al., 2000; Watanabe, 1996). FEF and PFC contain neurons involved in decision making (Kim and Shadlen, 1999), target selection (Schall et al., 1995), attention (Iba and Sawaguchi, 2003; Thompson and Bichot, 2005), and maintaining information during a delay (Funahashi et al., 1989; Kim et al., 2008; Sommer and Wurtz, 2001). FEF neurons, in particular, predict upcoming decisions in a reverse-masking task (Thompson and Schall, 1999) that inspired the decision-making portion of our task. PFC neurons have been implicated in a range of high-level cognitive processes, including executive function (Miller and Cohen, 2001), abstract rule encoding (Wallis and Miller, 2003), and behavioral context (Johnston and Everling, 2006), suggesting that they collectively function to guide behavior for a desired outcome (Tanji and Hoshi, 2008). SEF neurons have been implicated in performance monitoring by signaling error, conflict, and reward (Nakamura et al., 2005; Stuphorn et al., 2000). Given these different characteristics, we predicted that FEF neurons would be more “low level” in encoding the decision alone,

whereas PFC and SEF would be more “high level” in linking the decision to the appropriate bet.

We analyzed neuronal activity from FEF, PFC, and SEF with respect to three main functions of the task: making decisions, placing bets, and linking decisions to appropriate bets. Activity in all three areas correlated with decisions and likewise with bets, but only activity in the SEF correlated with monitoring decisions to guide bets. Of the three areas, the SEF seems the most involved in metacognition.

RESULTS

Behavior

We previously provided a detailed analysis of the monkeys’ behaviors during sessions prior to neuronal recordings (Middlebrooks and Sommer, 2011). Here, we analyze behavioral data collected during the recording sessions of the present study (150 sessions for Monkey N, 182 for Monkey S). On average, each monkey made more correct decisions in the decision stage and placed more high bets in the bet stage as a function of longer SOA (Figure 1B, one-way ANOVAs, each $p < 0.001$). In principle, SOA alone could have provided information to guide betting; monkeys could have ignored their trial-by-trial decisions and just bet high more often if the masks appeared later or the task seemed easier. We analyzed the data from each SOA

separately to address this potential confound. Trial-by-trial analyses revealed that for each monkey, within each SOA, bets were correlated appropriately with decisions (χ^2 test, $p < 0.001$ for each SOA and each monkey; details in Middlebrooks and Sommer, 2011). We quantified performance across SOAs using two phi correlation methods (Kornell et al., 2007; Zar, 1999). Phi correlation values could range from zero (random betting) to one (perfect association between decisions and bets). Both monkeys' phi correlations, assessed with either method (Figure 1C; Figure S1 available online), were significant at each SOA and constant across SOAs (one-way ANOVA, $p > 0.05$).

Another potential confound is the use of motor-related cues. Monkeys could possibly detect their saccade latencies during the decision stage and use this information to help place bets. This explanation is feasible if latency distributions differ between correct-high versus correct-low trials and between incorrect-high and incorrect-low trials, but they did not (Table S1). All of these results replicate our prior findings (Middlebrooks and Sommer, 2011) and indicate that, within each trial during neuronal recordings, monkeys maintained information about their decision to guide their bet, a metacognitive strategy.

Single Neuron Recordings

We studied 87 neurons in the FEF (Monkey N: 35, Monkey S: 52), 112 in the PFC (N: 54, S: 58), and 133 in the SEF (N: 61, S: 72). As expected, neurons in all three areas were highly modulated during the task (Figure S2). The monkeys' betting behavior did not vary significantly between recording sessions in the three cortical areas (phi correlations for Monkey N: FEF, 0.51; PFC, 0.49; SEF, 0.47; for Monkey S: FEF, 0.59; PFC, 0.54; SEF, 0.54; no differences between areas by ANOVAs, $p > 0.05$, for both monkeys). Because the monkeys were well trained, the neuronal recording data included more correct-high and incorrect-low trials (the appropriate decision-bet pairings) than correct-low and incorrect-high trials (Table S2 shows the breakdown of trial outcomes).

Decision-Related Neuronal Activity

To test whether neurons encoded the decision, we compared all correct with all incorrect trials, regardless of subsequent bets (i.e., high and low bet trials pooled).

Sensory-Related Activity Comparison. First, we focused on neuronal activity related to the visual target. Using a similar masked target task, Thompson and Schall (1999) demonstrated that signals predictive of a monkey's decision occur in the early visual responses of FEF neurons, prior to the start of motor-related processes (reviewed by Schall and Thompson, 1999; Schall, 2001; see also Schall et al., 1995; Sato and Schall, 2003). We analyzed trials in which the target appeared in the hemifield contralateral to the neuron's location in the brain, because for FEF, visual receptive fields are typically contralateralized (Bruce and Goldberg, 1985). Contralateral biases are common in SEF and PFC too (Funahashi et al., 1989, 1990, 1991; Russo and Bruce, 1996), and we wanted to analyze data from all three areas in the same way for a fair comparison.

Single neuron examples are shown for FEF, PFC, and SEF (Figures 2A–2C). Each neuron was active during the early visual response (visual-1) and delay epochs (gray shadings), and each was more active on correct than incorrect trials in both epochs

(t test, $p < 0.05$). At the population level, all three frontal areas showed this effect (Figures 2D–2F; Table 1). We repeated these analyses using only those neurons that were significantly active within each epoch, and this yielded the same results (Table S3). These findings extend the results of Thompson and Schall (1999) to show that visual and delay activity correlate with decisions in a masked target task in the SEF and PFC as well as in FEF.

Motor-Related Activity Comparison. To analyze activity related to decision saccades, we compared the correct and incorrect trials for which a saccade was made into the contralateral field. We analyzed activity just before and after the saccade (presaccadic-1 and postsaccade epochs, respectively). Only the SEF population had activity in these epochs that differentiated correct from incorrect decisions (Table 1). Repeating this analysis on the subsets of neurons active within each epoch (i.e., only neurons with significant pre- or postsaccadic activity), SEF neurons were more active during correct than incorrect decisions within the postsaccade epoch (Table S3) but not the presaccadic-1 epoch. FEF and PFC showed no effect in either epoch.

Relationship to Bet-Related Activity. We expected bet-related activity to resemble decision-related activity, given the high trial-by-trial correlations between decisions and bets: correct decisions were mostly followed by high bets and incorrect decisions by low bets (Table S2). To analyze bet-related activity explicitly, we compared high bet with low bet trials regardless of preceding decisions (i.e., pooled correct and incorrect trials). The results, as expected, were similar to those from the decision-related activity analysis and are summarized in the Supplemental Information (Bet-related activity section of Supplemental Results; Tables S4 and S5).

Metacognition-Related Neuronal Activity

To test whether neuronal activity correlated with metacognitive monitoring, we compared trials when the monkey made the same decision but different bets. Our rationale was that metacognition is the process that links a decision to a bet, allowing for purposeful wagering instead of random wagering. Signals related to metacognition should differ between trials when a decision is followed by an appropriate versus inappropriate bet.

We first compared neuronal activity between correct-high (CH) and correct-low (CL) trials. This was a straightforward analysis because visual stimuli and saccade directions were equivalent in CH and CL trials throughout the decision stage. We included only those trials in which the targets were located in, and the saccades were directed into, the contralateral field. The critical time period was the interstage epoch: the time span after the decision was reported but before the bet targets appeared.

In the FEF, neuronal activity was no different in CH versus CL trials during the interstage epoch. A single neuron example (Figure 3A) was equally active for CH and CL trials during the interstage epoch (gray shading), and the same negative result was found for the FEF population (Figure 3D, left). FEF population activity profiles overlapped for CH and CL trials (Figure 3D, right). In the FEF, visual receptive fields and movement fields are often much smaller than a hemifield, so for a more careful test of FEF activity, we then limited our analyses to directions

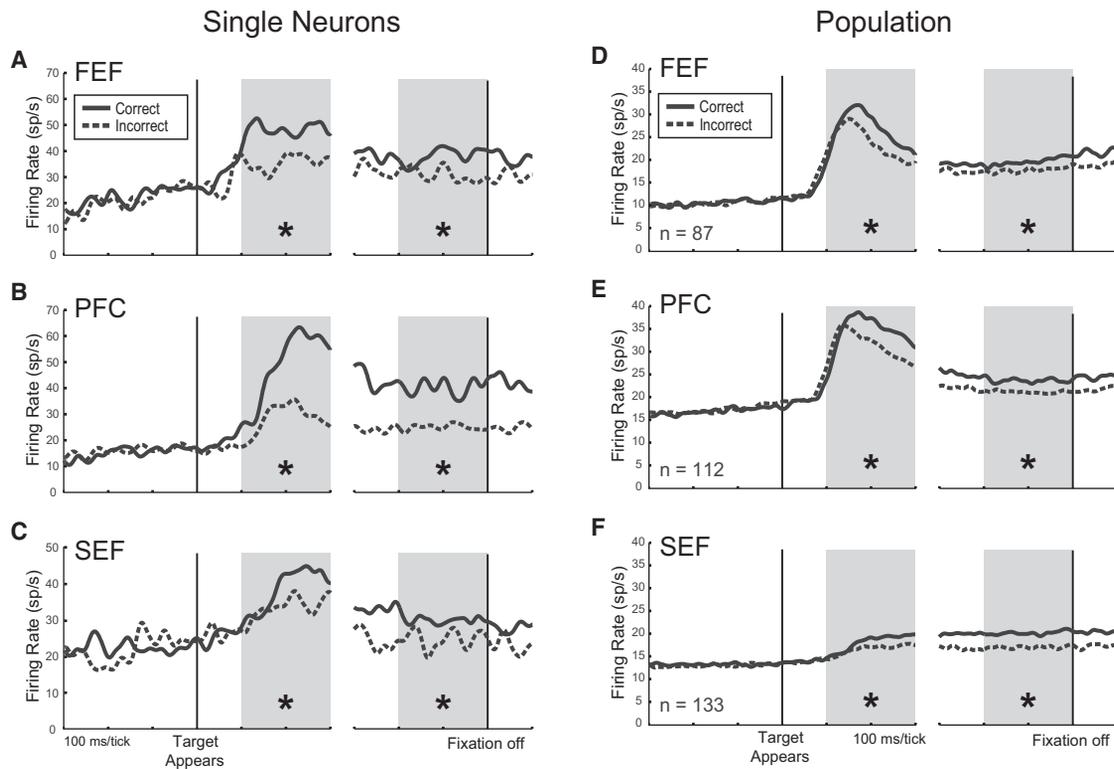


Figure 2. Decision-Related Neuronal Activity

Within each panel, neuronal firing rates for all correct trials (solid lines) and incorrect trials (dashed lines) are aligned to decision stage target onset (left) and fixation spot offset (right), and gray shadings indicate visual-1 (left) and delay epochs (right). Asterisks indicate a significant difference in activity ($p < 0.05$) within the epoch.

(A–C) Single neuron examples. Each neuron was more active during correct than incorrect trials in both epochs.

(D–F) Population activity. In all three areas, activity was greater for correct than for incorrect trials in both epochs (Table 1 shows corresponding numerical data). Population spike density functions are the average of all individual neuron spike density functions from each area.

See also Figure S2.

associated with the visual receptive field and/or movement field for each neuron; however, the results were still negative (Figures S3A and S3D).

PFC neuron activity was marginally better at distinguishing CH from CL trials. An example neuron (Figure 3B) was more active for CH trials than CL trials during the interstage epoch. In the

Table 1. Decision-Related Activity: Population

| FEF | Baseline | Visual-1 | Delay | Presaccadic-1 | Postsaccade |
|-----------|------------|-------------------|-------------------|---------------|-------------------|
| Correct | 10.4 (0.9) | 26.8 (2.0) | 19.5 (1.6) | 29.3 (2.6) | 20.2 (2.4) |
| Incorrect | 10.2 (0.9) | 23.8 (1.7) | 17.6 (1.4) | 30.5 (2.7) | 20.4 (2.5) |
| p Value | 0.33 | <0.001* | 0.017* | 0.16 | 0.75 |
| PFC | | | | | |
| Correct | 16.0 (1.3) | 33.9 (2.6) | 23.5 (1.7) | 25.7 (2.2) | 25.9 (2.4) |
| Incorrect | 16.4 (1.3) | 30.7 (2.4) | 20.7 (1.5) | 25.0 (2.0) | 28.0 (2.3) |
| p Value | 0.47 | <0.001* | <0.001* | 0.39 | 0.86 |
| SEF | | | | | |
| Correct | 13.4 (1.1) | 18.5 (1.4) | 20.1 (1.5) | 22.2 (1.5) | 22.7 (1.5) |
| Incorrect | 13.1 (1.1) | 17.0 (1.3) | 17.02 (1.2) | 20.9 (1.4) | 20.5 (1.3) |
| p Value | 0.20 | <0.001* | <0.001* | 0.024* | <0.001* |

Population decision-related firing rates during decision stage epochs. For each cortical region, all correct versus all incorrect firing rates (spikes/s) are shown with standard errors in parentheses and paired t test p values underneath. Asterisks and bold fonts represent significant differences ($p < 0.05$) between correct and incorrect trials.

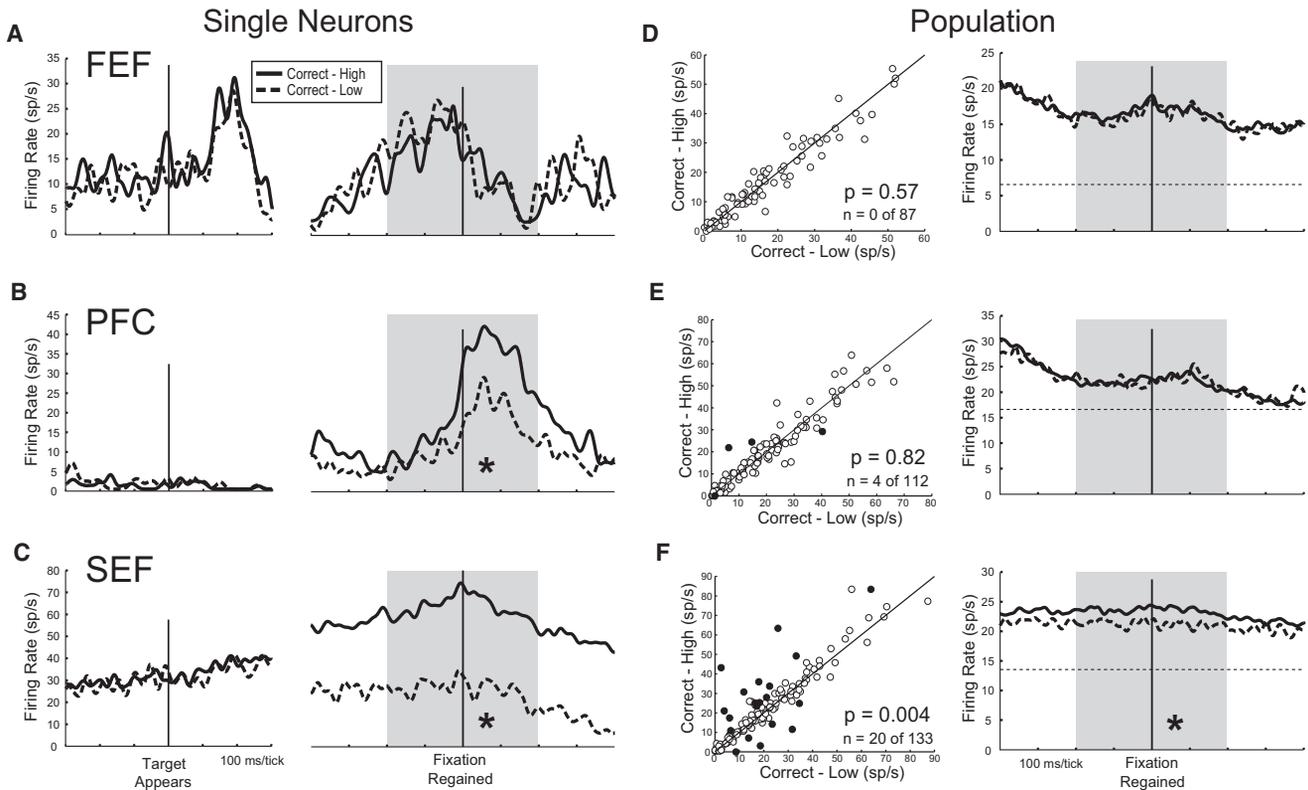


Figure 3. Metacognition-Related Activity

(A–C) For the single neuron examples, firing rates for all correct-high (CH) trials (solid lines) and correct-low (CL) trials (dashed lines) are aligned to decision stage target onset (left) and on regaining fixation to begin the bet stage (right). Gray shading indicates the interstage epoch. For the population data (D–F), scatterplots (left) show CH versus CL firing rates for each neuron during the interstage epoch, p values of t tests of population CH versus CL activity, and numbers (n) of individual neurons with significant CH versus CL activity (each denoted with a filled dot). Activity profiles (right) show population spike density functions, with baseline activity levels (dashed horizontal lines) provided for reference. Asterisks indicate significant differences within the interstage epoch ($p < 0.05$ for single neurons and < 0.025 for population data). The FEF neuron (A) was not significantly different between trial types (post-ANOVA t test, $p > 0.05$), but the PFC neuron (B) and SEF neuron (C) had significantly greater activity in CH than in CL trials (both $p < 0.001$).

(D) In the FEF, single neurons (left) and population profiles (right) showed no significant differences in activity between CH and CL trials.

(E) In the PFC, a few single neurons showed CH versus CL differences (left), but this was not significant at the population level, and population profiles overlapped (right).

(F) In the SEF, many individual neurons showed CH versus CL differences (left), this was significant at the population level, and population profiles were higher for CH than CL trials throughout the interstage epoch (right). Table 2, Interstage column, shows corresponding numerical data.

See also Figure S3.

PFC population, however (Figure 3E, left), there was no average activity difference between CH and CL trials, the incidence of individually significant neurons was not greater than expected by chance (4/112 neurons compared with 5/112 expected false positives at the $p < 0.05$ criterion for individual neurons; Fisher's exact test, $p = 0.999$), and average activity profiles for CH and CL trials overlapped (Figure 3E, right). Results were similarly negative for analyses restricted to visual and movement fields (Figures S3B and S3E).

The SEF seemed to be the major player in sustaining a metacognitive signal. The SEF neuron in Figure 3C, for example, was 2.5 times more active during the interstage epoch for CH than CL trials. Overall, 15% (20/133) of individual SEF neurons had significantly different activity in CH versus CL trials (Figure 3F, left, filled circles), a proportion significantly greater than expected by chance (a false positive rate of 6/133 neurons was

expected at $p < 0.05$; 20/133 neurons is significantly greater; Fisher's exact test, $p = 0.0063$). CH activity exceeded CL activity for 70% (14/20) of the individually significant neurons and at the population level (Figure 3F, right). The SEF results were similarly positive for analyses restricted to visual and movement fields (Figures S3C and S3F).

In the SEF, differential CH-CL activity could emerge long before the interstage epoch. Individual neurons showed a variety of time courses. Figures 4A and 4B show example CH $>$ CL neurons, and Figure 4C shows an example CH $<$ CL neuron. Pooling together the subset of 14 neurons for which CH activity significantly exceeded CL activity during the interstage epoch, it can be seen that the average effect started well before the interstage epoch and had a long time course (Figure 5A). CH and CL activity separated before the cue to make a saccade, that is, during the time when monkeys may have made their

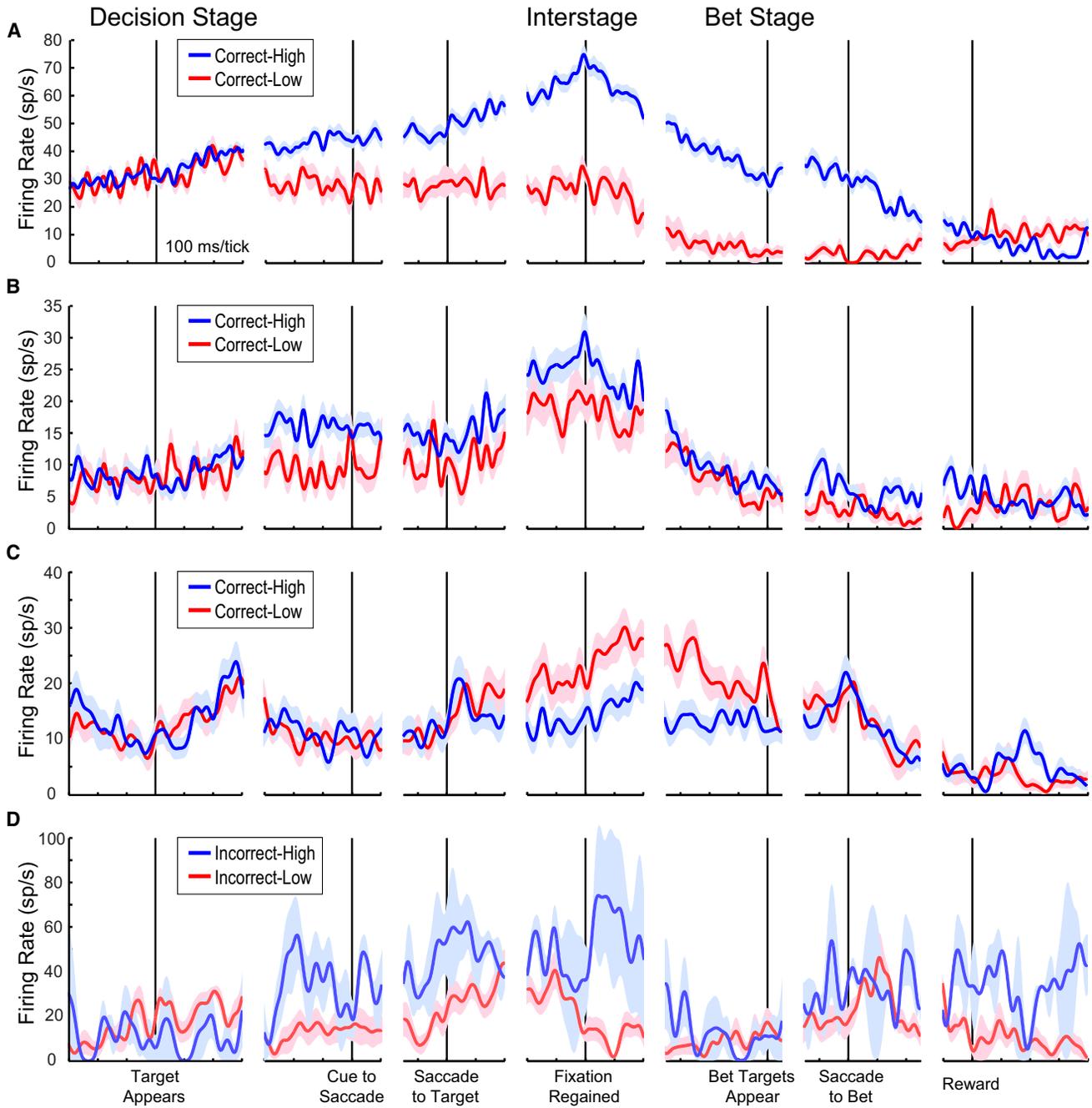


Figure 4. Time Courses of Example SEF Neurons

Activity profiles depict means (lines) and SEMs (shading) and are aligned to events indicated at bottom.

(A and B) Two example neurons for which correct-high (CH) activity significantly exceeded correct-low (CL) activity during the interstage epoch.

(C) Example neuron for which CL activity significantly exceeded CH activity during the interstage epoch.

(D) Example neuron for which incorrect-high (IH) activity significantly exceeded incorrect-low (IL) activity during the interstage epoch. Its noisy IH activity was typical because of the small number of trials (IH was the least likely trial outcome).

decision but before they reported it. The subset of six neurons with the reverse effect (CH < CL) had a more transient average time course (Figure 5B). Overall activity was higher for the CH > CL subset than for the CH < CL subset, including during the baseline period (first 300 ms of time courses), hinting that

the former subset may include more inhibitory interneurons than the latter subset (e.g., Connors and Gutnick, 1990; Constantinidis and Goldman-Rakic, 2002). We cannot provide further support for this possibility, however, because we did not store action potential waveforms (e.g., Mitchell et al.,

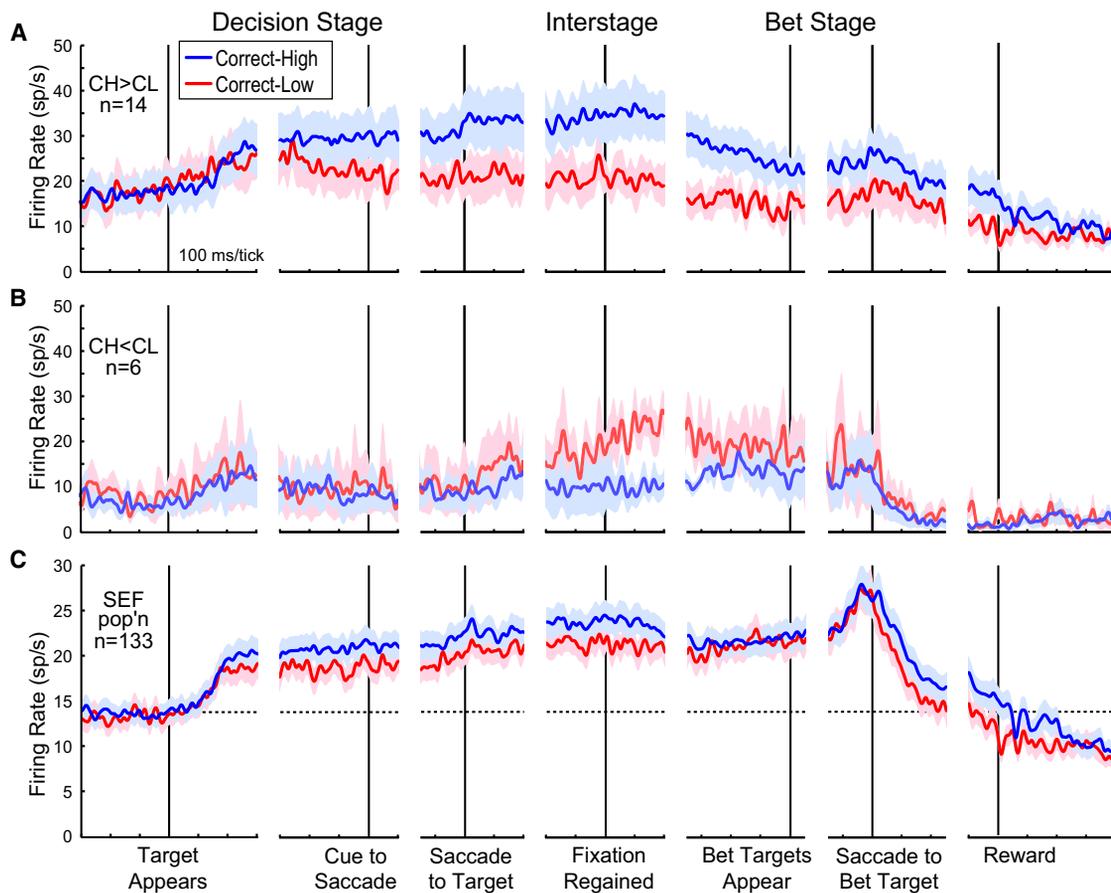


Figure 5. Time Courses of SEF Population Activity

Conventions as in Figure 4. Average activity profiles are shown for (A) the 14 SEF neurons for which CH activity significantly exceeded CL activity during the interstage epoch, (B) the six SEF neurons for which CL activity significantly exceeded CH activity during the interstage epoch, and (C) the entire population of SEF neurons. In the population, SEF activity distinguished CH from CL trials ~200 ms after target onset (“Target Appears”), and this differential activity was maintained through the interstage period (after “Fixation Regained” and before “Bet Targets Appear”). Table 2, bottom row, shows results of statistical analyses for this time range (Baseline through Interstage epochs). Population differential activity re-emerged after the saccade to the bet target. Only contralateral data are included here. Population baseline firing rate is shown with a horizontal dashed line. See also Figure S4.

2007), and we found no significant differences in spiking statistics between the subsets (see the Spike Burstiness section in Supplemental Results; Anderson et al., 2011).

In the entire population of SEF neurons (Figure 5C), population differential activity emerged early in the decision stage and then was maintained, steadily and significantly, through the interstage epoch. The numerical data corresponding to this sustained effect are listed in Table 2, bottom row. The SEF population results were the same when we extended the analysis beyond contralateral space to all directions (summarized in Figure S4A, top). When we considered only the subset of SEF neurons significantly active within each epoch, we found a similar pattern differentiating CH versus CL activity (Figure S4A, middle and bottom). Finally, the population-level CH > CL effect during the interstage epoch was significant for each monkey individually (Table S6).

The complementary approach to testing whether neuronal activity correlates with metacognitive behavior is to compare

incorrect-high (IH) versus incorrect-low (IL) trials. Analyses of IH and IL trials are complicated by two issues, however. First, the target location is not coincident with the saccade destination, by definition of an incorrect trial. Incorrect saccades may go to the other target location in the same hemifield or to either target location in the other hemifield. Thus, different directions had to be analyzed as a function of epoch (see IH versus IL section in Supplemental Results). Second, IH trials were the rarest outcome (only 10% of all trials; Table S2), resulting in few trials to analyze for many neurons. Nevertheless, we performed the IH versus IL analyses and found, as with the CH versus CL analyses, significant effects during the interstage epoch in the SEF population ($p = 0.005$) but not in the PFC or FEF populations (Figures S3G–S3I). For most of the individually significant SEF neurons (9/10), IH activity exceeded IL activity. These neurons showed a variety of time courses (one is shown in Figure 4D), but on average they distinguished between IH and IL trial outcomes continuously from the end of the decision stage to

Table 2. Metacognition-Related Activity: Population

| FEF | Baseline | Visual-1 | Delay | Presaccadic-1 | Postsaccade | Interstage |
|---------|------------|------------|-------------------|---------------|-------------------|-------------------|
| CH | 10.4 (0.9) | 27.2 (2.0) | 19.9 (1.7) | 27.8 (2.3) | 20.5 (2.2) | 16.4 (1.5) |
| CL | 10.6 (1.0) | 27.3 (2.1) | 19.3 (1.7) | 26.9 (2.2) | 21.0 (2.1) | 16.1 (1.5) |
| p Value | 0.64 | 0.82 | 0.40 | 0.37 | 0.59 | 0.57 |
| PFC | | | | | | |
| CH | 15.8 (1.3) | 33.9 (2.6) | 23.3 (1.7) | 25.4 (2.1) | 27.0 (2.4) | 21.9 (2.0) |
| CL | 15.9 (1.3) | 32.7 (2.4) | 23.1 (1.7) | 24.3 (1.9) | 26.3 (2.4) | 22.0 (2.0) |
| p Value | 0.79 | 0.23 | 0.82 | 0.23 | 0.35 | 0.82 |
| SEF | | | | | | |
| CH | 13.6 (1.1) | 18.5 (1.4) | 20.1 (1.6) | 22.2 (1.5) | 22.9 (1.5) | 23.6 (1.6) |
| CL | 13.1 (1.2) | 17.7 (1.5) | 18.6 (1.4) | 20.1 (1.4) | 21.0 (1.4) | 21.3 (1.5) |
| p Value | 0.25 | 0.10 | <0.001* | 0.017* | <0.001* | <0.001* |

Population metacognition-related firing rates linking decisions to bets. For each cortical region, all correct-high (CH) versus all correct-low (CL) firing rates (spikes/s) are shown with standard errors in parentheses and paired t test p values underneath. Asterisks and bold fonts represent significant differences between CH and CL trials ($p < 0.025$ criterion, see [Experimental Procedures](#)).

the end of the trial ([Figure S4B](#)). We directly compare IH to CH activity, and IL to CL activity, below in the [Reward Anticipation](#) section.

For the FEF and PFC populations, CH versus CL differences failed to reach significance not only during the interstage epoch, as described previously, but also in every epoch ([Table 2](#), top and middle rows). IH and IL activity differences were similarly insignificant across epochs (except for one epoch in the FEF; [Supplemental Results](#), IH versus IL section). Finally, no CH-CL or IH-IL differences were significant in any epoch for the subsets of FEF and PFC neurons that were significantly active in each epoch (data not shown).

Effects of SOA

We varied the SOA in the task to elicit large numbers of correct and incorrect trials (and their associated bets) to analyze. This raises two questions. Did varying SOAs contribute to differences in trial durations between trial outcomes (e.g., CH versus CL) that could have influenced our neuronal results? And, more to the point, did metacognition-related signals in SEF vary with SOA? Regarding the first question, we did not expect SOA distributions (and thus trial lengths) to vary appreciably between trial outcomes, given that metacognitive behavior was stable across SOAs (e.g., [Figure 1C](#)). Betting depended on trial-by-trial decisions, not SOAs. The only exception might be if a monkey “guessed” during the more difficult, shorter SOA trials; it might choose a target randomly and then bet low to be safe. If its choice was correct, the outcome would be a CL trial. Hence, short SOAs might be slightly more prevalent in CL trials than in CH trials. Consistent with this expectation, we found that average SOAs were 48.3 ms (SD 17.9 ms) in CH trials and 45.1 ms (SD 18.2 ms) in CL trials, a slight but significant difference (Mann-Whitney U test, $p < 0.001$). This 3.2 ms difference in mean trial duration was negligible compared to the overall trial duration of ~ 2 s, so it is unlikely to have influenced our neuronal data. Regarding the second question, we analyzed whether our main indicators of metacognitive signals, CH firing rates, CL firing rates, and differential CH-CL activity, varied across SOAs. We analyzed each of these three data sets for all six

epochs of [Table 2](#) (baseline through interstage), for contralateral directions and all directions. Firing rates did not vary significantly as a function of SOA for any of the 36 tests (ANOVAs, $p > 0.05$ for all). In sum, variations in SOA were critical for the task design but had no significant influence on the neuronal effects that we found, just as they had no influence on metacognitive behavior (e.g., [Figure 1C](#)).

Bet Stage-Related Activity

We also analyzed CH versus CL differences for time periods after the interstage epoch, through the bet stage of the task. Briefly, at the population levels, none of the three cortical areas had activity correlated with metacognition after the interstage epoch and before the bet. In the SEF, CH and CL signals clearly became different again after the bet, through delivery of the reward ([Figure 5C](#); [Supplemental Results](#), bet stage-related activity section; [Tables S7 and S8](#)). In the SEF population, this disappearance and resurgence of CH > CL activity might be explained by opposing dynamics of CH > CL and CH < CL neurons. The individually significant CH > CL neurons sustained their signal through the entire bet stage ([Figure 5A](#)), but the CH < CL neurons were transiently active in the late interstage and early bet stage ([Figure 5B](#)), so they may have effectively nullified the CH > CL signal during that time at the population level.

Reward Anticipation

Many neurons in the SEF encode reward anticipation ([Roesch and Olson, 2003](#); [So and Stuphorn, 2010](#)). In our experimental design, reward amounts were determined entirely by behavior: the decision and the bet. We could not know what reward amounts the monkeys expected on given trials, but it is likely that they placed high bets in anticipation of high reward and low bets in anticipation of low reward. If our SEF neurons represented reward anticipation, this might explain the higher firing rates in CH versus CL trials and IH versus IL trials. Quantitatively, the reward anticipation hypothesis predicts that activity should be equal for all trials in which the *same* bet was made after *different* decisions: firing rates should be indistinguishable between CH and IH trials and between CL and IL trials. We found that, to the contrary, SEF activity strongly differentiated between

CH and IH trials and between CL and IL trials through the decision stage and into the bet stage. As with our usual analyses, we considered trials for which targets were located within, and saccades were directed into, the contralateral field. During the decision stage (Table S9), the CH-IH difference in population activity began in the visual-1 epoch and lasted through the interstage epoch. In the subsets of neurons with significant activity in each epoch, the same pattern of results was observed with the exception of the presaccadic-1 epoch. SEF activity also was different in the decision stage between CL and IL trials. As a population, the difference was significant during the delay and interstage periods. For the subsets, CL-IL firing rates were different from the visual-1 epoch through the interstage epoch, except in the presaccadic-1 epoch. Thus, although we would not rule out effects of reward anticipation during the decision stage, we found little evidence for it.

During the bet stage (Table S10), SEF population activity became more similar between CH and IH trials and between CL and IL trials; differences in activity between these trial outcomes diminished and eventually ceased. This implies that neuronal correlates of reward anticipation may have contributed more to SEF population activity near the end of the trial. On a related note, SEF neurons are known to modulate with reward delivery (Stuphorn et al., 2000). We also observed reward modulation, in that SEF neurons had higher firing rates for the worst reward outcome (IH, resulting in timeout and no reward) than any of the other trial outcomes (IH: 16.1 ± 1.6 sp/s, CH: 12.0 ± 1.4 sp/s, $p < 0.001$; CL: 10.3 ± 1.2 sp/s, $p < 0.001$; IL: 11.0 ± 1.3 sp/s, $p < 0.001$; see also Table S10).

Influence of Past Trial Outcomes

Given the elevated activity in the SEF at the end of IH trials, we asked whether intertrial effects may have influenced our data. In strategic decision-making tasks, choices and neuronal activity can be affected by the outcomes of previous trials (Barracough et al., 2004; Seo and Lee, 2009), suggesting that carryover of neuronal activity from one trial to the next could guide choices (Sutton and Barto, 1998). First, using our behavioral data and considering all directions of target locations and saccades, we analyzed the rates at which monkeys switched their bets from trial to trial, that is, the rates of making low bets after CH or IH trials or high bets after CL or IL trials. If a bet is influenced by previous trial outcome, monkeys should switch bets with relatively low likelihood after CH and IL trials (“win-stay” strategy) but with high likelihood after IH and CL trials (“lose-switch” strategy). We found no such intertrial effects: the rates of placing low bets after high bet trials (i.e., switching after CH or IH trials) were no different from the average rate of placing low bets (Figures 6A and 6B, left data). The same was true for rates of placing high bets after low bet trials (Figures 6A and 6B, right data; t tests, all $p > 0.05$).

At the neuronal level, we found carryover of previous trial information that differed between brain areas. Data were pooled over all directions. In the SEF population, baseline firing rates were higher after IH trials than after other trial outcomes (Figure 6E; paired t tests, all $p < 0.05$). The effect was individually significant for 13% (17/133) of the SEF neurons. The effect disappeared as soon as the decision stage began (target appearance) and did not return throughout the course of the trial; no other epochs

in SEF distinguished between previous trial outcomes (paired t tests, all $p > 0.05$).

In contrast, neurons in both PFC and FEF carried information about previous IH trials into various decision stage epochs of the next trial. PFC carried substantial previous-trial information, as seen previously (Barracough et al., 2004). Like in the SEF, baseline firing rates in the PFC were higher after IH trials than after other trial outcomes (Figure 6D, paired t tests, all $p < 0.05$). The effect was individually significant for 11%, 12/112, of the PFC neurons. This IH-related signal was sustained through the next two (visual-1 and delay) epochs (data not shown). In the FEF, previous IH trials had no effect on baseline activity but led to significantly higher firing rates during the post-saccade period (paired t tests, all $p < 0.05$, data not shown).

In sum, IH trials seemed to affect neuronal activity in the next trial. In the SEF, this influence ended after the baseline period, matching the monkeys' behavior in that previous-trial information did not account for current trial performance. In PFC and FEF, previous-trial information persisted into the current trial to affect neuronal activity in various epochs.

DISCUSSION

We recorded from single neurons in the FEF, PFC, and SEF while monkeys performed a visual oculomotor task in which they monitored their own decisions. Neuronal activity correlated with decisions and bets was found in all three areas, but joint activity that linked decisions to appropriate bets was found exclusively in the SEF. This putative metacognitive activity began swiftly in the SEF during the decision stage and continued into the bet stage. Monkey behavior was independent of previous trial outcome, as was SEF activity (but not PFC or FEF activity).

We had predicted that both the SEF and PFC would participate in metacognitive monitoring, but our data supported a role only for the SEF. The putative metacognitive activity in SEF arose early in trials (Figures 5A and 5C), beginning soon after the start of the decision-related signal (Figure 2F) and before the monkey reported its decision with a saccade. The time course suggests that monitoring a decision occurs in near simultaneity with making the decision. This seems analogous to the time course of monitoring motor operations (“corollary discharge”); when motor areas finalize a movement command, upstream areas monitor it within milliseconds (Sommer and Wurtz, 2004). It should be noted that most (15/20) of our individual SEF neurons with a metacognitive signal also exhibited a decision-related signal. This close relationship between metacognitive and decision-related signals may be no coincidence: in the SEF, decision-related signals may evolve into metacognitive signals. A decision-related signal that outlasts the decisive act (the saccade to the target) provides information that could be monitored for later behavior (the bet). Although decision-related signals occurred in all three areas, our data suggest differences in how the signals are used. In SEF, the prolonged decision-related signal seems to be maintained for internal use (e.g., determining the bet to make). In PFC and FEF, the briefer signal may guide only immediate acts (e.g., planning the decision saccade).

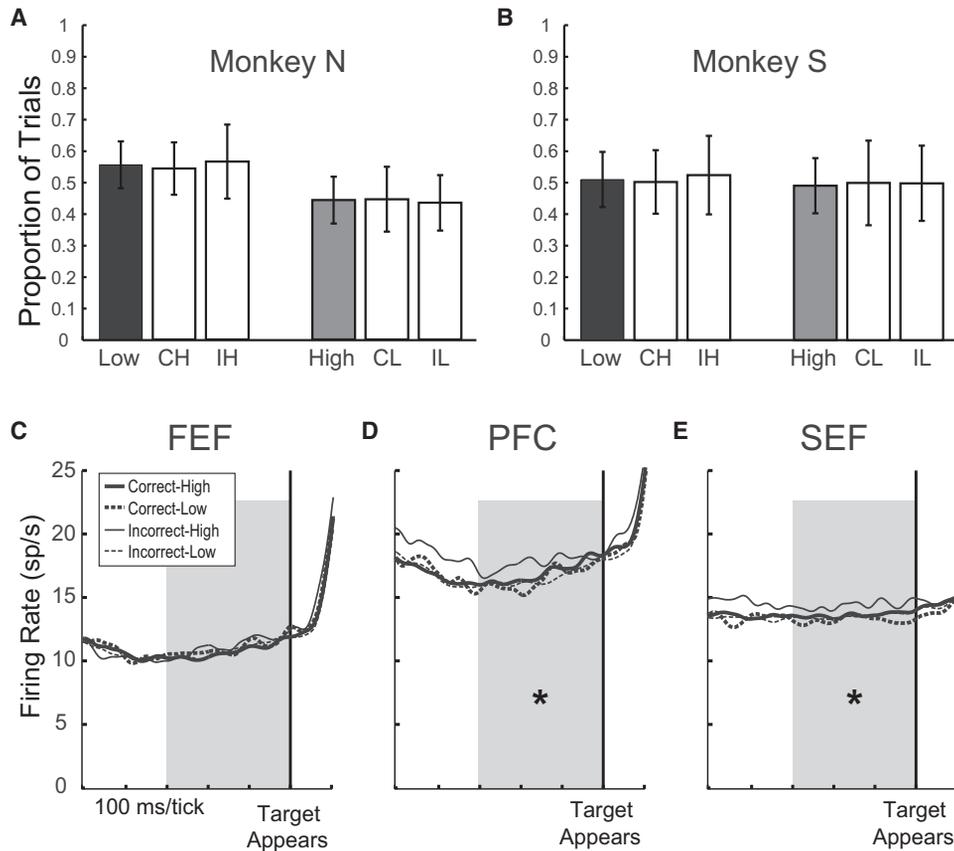


Figure 6. Intertrial Effects

(A and B) Rate of placing bets as a function of previous trial outcome for each monkey. In each graph, the average low bet rate (black bar) is plotted next to low rates after previous CH and IH trials (white bars), and the average high bet rate (gray bar) is plotted next to high rates after previous CL and IL trials (white bars). Error bars are standard deviations. None of the “switch” rates were different from the respective average bet rates (paired t tests, $p > 0.05$).

(C–E) Neuronal activity during baseline period (shaded) as a function of previous trial outcome for the populations of FEF, PFC, and SEF neurons. Asterisks indicate whether IH activity was greater than all three of the other trial outcomes by paired t tests ($p < 0.05$).

Metacognition-related activity in SEF had not been reported previously. No fMRI studies reported human SEF signals during metacognition tasks, although many fMRI results have implicated regions interconnected with SEF, such as anterior cingulate and medial prefrontal regions (Chua et al., 2006; Kikyo et al., 2002). Our recording strategy was to study every neuron encountered, so our population data may be considered a representative sample of SEF neurons (leaving aside issues of sampling biases related to neuron size, e.g., Sommer and Wurtz, 2000). The signals that we found in individually significant neurons were prominent, but the gross signal in the entire SEF population was small (Figure 5C; Table 2), suggesting that it may not be distinguishable with fMRI. In future work we will concentrate our recording efforts on only those SEF neurons that show metacognition-related activity (differential CH versus CL and IH versus IL signals) to investigate them in more detail.

Prior recording studies of monkey SEF reported neurons signaling reward, errors, conflict, and/or inhibition of planned saccades, collectively referred to as performance monitoring (Nakamura et al., 2005; Stuphorn et al., 2000). We found two lines of evidence for reward signals in the SEF: elevated firing

rates during the reward epoch of CH versus CL trials and information about worst-outcome, IH trials, in the reward period that carried over to the next trial (a “lack of reward” signal). Neither signal can explain our putative metacognitive activity in SEF because both start after the bet on one trial and end before the next trial’s decision. Regarding error signals (Stuphorn et al., 2000), an “error” in our task is not straightforward. An error could be a trial that earned no reward (IH), but we did not observe increased or decreased firing rates on IH trials until around the time of reward, as mentioned. A subtler interpretation is that an error occurred when less reward was earned than potentially available (CL trials). Yet, we did not see SEF activity greater on CL than CH trials in any epoch or transient decreases in activity on CL trials. Finally, a transient error signal might occur after any incorrect decision (e.g., during the postsaccade and/or interstage epochs), since incorrect decisions were always less advantageous than correct decisions. We did not observe SEF neurons with that sort of signal either. In short, we saw little or no evidence of error signals in our SEF data.

We found, as well, that reward anticipation (Roesch and Olson, 2003; So and Stuphorn, 2010) was not a plausible explanation for

the metacognitive signals. Our experiment did not explicitly vary reward anticipation, but it could be argued that “bet anticipation” is the same thing, as long as the animals expected all high bets to yield high reward and all low bets to yield low reward. We found little evidence for bet or reward anticipation. The activity of our SEF neurons differentiated between trials that culminated in identical bet selection (CH versus IH and CL versus IL trials). This differential activity occurred throughout the decision stage and interstage periods, when putative metacognitive signals dominated. Signals related to identical bet selection became less distinguishable in the bet stage, suggesting that reward anticipation signals “took over” in the betting phase of the task. Our results cannot resolve the extent to which metacognition and reward anticipation signals are conveyed by separate SEF neurons or multiplexed in single neurons. In a recent study (So and Stuphorn, 2012), monkeys performed a gambling task in which a delay was imposed between when the monkey made a choice and when reward was delivered. SEF neurons recorded during the task carried multiple signals; some activity patterns varied with expected reward, some with experienced reward, and others with the difference between expected and experienced reward. Similar signals in SEF were reported during a token-based gambling task (Seo and Lee, 2009), in which reward was delivered after earning a sufficient number of tokens across trials. These reports complement our conclusion that SEF signals correlate with metacognitive monitoring only within a trial, not across trials. This comparison between studies highlights a key difference between our task and most other gambling tasks. Our monkeys gained no advantage by adjusting their bets based on previous trial outcomes; the reward yielded by a bet depended only on the decision made by the monkey earlier in the same trial. Our reward probabilities depended critically on the ability to monitor decisions (details in Middlebrooks and Sommer, 2011). In probabilistic gambling tasks, on each trial the reward probabilities are set by computer according to some distribution, and thus monkeys learn to keep track of those expected probabilities in addition to, or instead of, their own behavior. A salient goal of future work would be to design experiments that manipulate both reward expectation and metacognitive monitoring in systematic ways, to reconcile the extent that both signals may be carried by SEF neurons.

It was also possible that the neurons may have been coding the actual (as opposed to expected) upcoming reward. We found, however, that neuronal firing rates across trial outcomes did not parallel relative reward values, so actual rewards were not predicted by firing rates. Lastly, riskiness (McCoy and Platt, 2005) could be proposed as an alternative account of our data. If the neurons were signaling levels of risk, we would expect high firing rates for all high bets and low firing rates for low bets, but we did not observe this pattern (for more on these issues, see Supplemental Discussion).

Neither the FEF nor the PFC showed much evidence of metacognition-related activity. Instead, activity in both areas was correlated with the initial stage of the task: making the decision. This supported our initial prediction about the FEF, which was based on similar results from Thompson and Schall (1999). As discussed in that prior study and related work from the Schall laboratory, differences in FEF visual responses correlate with

making decisions but are not trivially explained by other factors (e.g., saccade preparation; see Supplemental Discussion). In the PFC, we expected to find prominent metacognitive signals because it has been implicated previously in human metacognition (Rounis et al., 2010). The PFC is a large, functionally heterogeneous region (e.g., Romanski, 2004), and our posterior sampling of it (Figure S2A) may have missed metacognition-related areas. However, the neurons we recorded featured all of the familiar hallmarks of dorsolateral PFC (Funahashi et al., 1989, 1990, 1991): visual responses, strong delay activity, and postsaccadic activity (Figure S2C). In the context of visual-saccadic tasks, the neurons seemed typical. It could be that metacognitive processing in PFC (and/or FEF) occurs in specific, yet rare, neurons. FEF and PFC activity also may be more dependent on spatial parameters of the task than SEF. FEF neurons can have quite spatially restricted visual receptive and movement fields (Bruce and Goldberg, 1985), but even when we analyzed target locations confined to those fields, we found no metacognition-related effects.

Our results complement a recent report that LIP activity correlated with monkeys' tendency to opt-out of making a decision (Kiani and Shadlen, 2009), suggesting that the activity signals confidence. Both the fundamental task design and the visual stimuli used in the LIP study differed from those used here. Moving-dots stimuli (Kiani and Shadlen, 2009) require evidence accumulation over time, but the decision stage of our task requires detection of a single brief stimulus. A possible advantage of our task is that its brief stimulus presentation demands a more immediate monitoring of the decision to guide the eventual metacognitive judgment. Given the short latency at which the metacognitive signals separated and the long duration of the separation, SEF neuronal activity seems to transcend general confidence and correspond more to monitoring of the monkeys' percept. Another possible advantage of our task is that we were able to establish that the metacognition-related signals in SEF represented processes beyond reward anticipation, which was less clear in LIP using the opt-out task (Kiani and Shadlen, 2009) or in OFC using a delayed reward task (Kepecs et al., 2008).

Studies of metacognition naturally lead to questions about broader implications. One interpretation is that metacognition is associated with conscious awareness (Nelson, 1996), but we favor a more conservative view that self-monitoring does not presuppose self-awareness (Reder and Schunn, 1996). As we argued previously (Middlebrooks and Sommer, 2011), metacognition may be to cognition as corollary discharge is to action; both describe the ability of the brain to internally monitor its operations. Just as it appears that all animals that move have internal circuits for monitoring their movements (Crapse and Sommer, 2008), all animals with even rudimentary cognitive abilities may monitor those abilities. This monitoring ability, however, does not necessarily imply states of self-awareness anywhere near the levels experienced by humans.

EXPERIMENTAL PROCEDURES

Surgery

Two male rhesus monkeys (labeled N, 6.6 kg, and S, 6.0 kg) were surgically prepared for neuronal recordings and eye position measurements. Using

aseptic procedures, ceramic screws and an acrylic implant were affixed to the skull. Recording chambers and a head-restraint socket (Crist Instruments, Hagerstown, MD, USA) were embedded in the implant. Chambers were positioned over FEF/PFC (one chamber with access to both regions) and SEF using stereotaxic coordinates (FEF/PFC: A25, L20; SEF: A25, midline). In the same surgery, we implanted scleral search coils. Animals recovered for 1–2 weeks before training resumed. Procedures were approved by and conducted under the auspices of the University of Pittsburgh Institutional Animal Care and Use Committee and were in compliance with the guidelines set forth in the United States Public Health Service Guide for the Care and Use of Laboratory Animals.

Tasks

Receptive Field Mapping Tasks

To determine appropriate target locations for the metacognition task (described below), we initially characterized the receptive field of each neuron using simple visual oculomotor tasks (see Sommer and Wurtz, 2004). First, the monkey made visually guided saccades to targets in eight directions (cardinal directions and diagonals). After the neuron's preferred direction was established, the monkey performed visually guided saccades of varying amplitudes in that direction. If necessary, directions and amplitudes were adjusted, and the tasks were repeated to refine the assessment of the field. Once the receptive field center was located, we had the monkey make memory-guided saccades to that location, to distinguish visual-, delay-, and saccade-related activity (Mays and Sparks, 1980). We accepted neurons with any combination of these signals.

Metacognition Task

The task was described previously in detail (Middlebrooks and Sommer, 2011). Each trial consisted of a Decision Stage and a Bet Stage, separated by an interstage period (Figure 1A). In the decision stage the animal was required to detect and report the location of a masked visual target (Thompson and Schall, 1999), and in the bet stage was required to report, via a wager, whether a correct or incorrect decision had been made in the decision stage (Shields et al., 2005). Appropriate betting, thus optimal reward delivery, required the animal to maintain a representation of its decision. It is the maintenance of that decision signal, and its use for betting, that we refer to as metacognition. To obtain reward on any trial, completion of both the decision and bet stages was required.

Decision Stage. The monkey fixated a spot for 500–800 ms (randomized; Figure 1A, left). Then, a dim target appeared in one of four possible locations (also randomized). The locations were constant in a session but could vary between sessions; eccentricities ranged from 5–25 degrees and directions, relative to horizontal, ranged from 0–60 degrees. For each neuron, these parameters were chosen so that, when possible, at least one target location was in the receptive field center. The locations were mirror symmetric across the vertical meridian. After the target appeared, identical mask stimuli (white squares) appeared at all four locations. The interval between target appearance and masks appearance, the stimulus onset asynchrony (SOA), was randomized by trial (16.7, 33, 50, or 66.7 ms) to vary task difficulty. After the masks appeared, a random delay of 500–1,000 ms ensued, during which the monkey maintained fixation, while the masks remained visible. Then, the fixation spot was extinguished, cueing the monkey to report its decision by making a saccade to the perceived target location within 1,000 ms. The monkey received no performance feedback until after the bet stage, but the computer tracked whether the decision was correct (saccade landed in an electronic window around the target location) or incorrect (saccade landed anywhere else). If at any time during the decision stage the monkey broke fixation, made a saccade before cued to go, or failed to make a saccade, the trial was aborted (and repeated later) and the next trial immediately began.

Bet Stage. A new fixation spot appeared 350 ms after the decision saccade that concluded the decision stage (Figure 1A, right). The monkey foveated the spot and, 500–800 ms later, two bet targets appeared: a red "high-bet" target and a green "low-bet" target (for Monkey N; color assignments reversed for Monkey S). In a session the two locations were constant, but the appearance of high-bet or low-bet targets varied randomly between the locations. One location was in the center of the receptive field and the other was at the mirror symmetric location in the other hemifield. A monkey reported its bet by making

a saccade to one of the targets, then received reward or timeout as described below, and the trial ended. A monkey optimized its reward if it bet high after a correct decision and low after an incorrect decision. If, during the bet stage, the monkey broke fixation or made a saccade to a non-bet-target location, the trial was aborted and a brief timeout ensued before a new trial began.

Reward. The amount of reward delivered after each trial was based on how appropriate the bets were relative to the decisions. If the monkey made a correct decision and bet high, it earned maximum reward: five drops of water. If the monkey made an incorrect decision and bet high, it received no reward and a 5 s timeout. Betting low earned a sure but minimal reward: three drops after a correct decision and two after an incorrect decision. The reward schedule was based on previous studies (e.g., Kornell et al., 2007) and was fine-tuned to elicit best performance.

Neuronal Recordings

A single tungsten electrode (0.3–1 M Ω impedance at 1 kHz; FHC, Bowdoinham, ME, USA) was lowered through a 23 g guide tube using a custom micro-drive system (<http://lsr-ftp.nei.nih.gov/lsr/StepperDrive/>). A plastic grid with 1 × 1 mm hole spacing (Crist Instruments, Hagerstown, MD, USA) was attached inside the recording chamber. The FEF was confirmed with microstimulation by evoking saccades at low current threshold (<50 μ A; Bruce and Goldberg, 1985). The PFC was recorded from the same chamber as FEF. PFC recordings included locations a few millimeters anterior to identified FEF, in areas ventral, dorsal, and within the principal sulcus (identified by the isolation of neurons at lower depths than locations dorsal or ventral). The SEF was identified by moderate-current microstimulation (typically 50–100 μ A) that evoked or delayed saccades (Russo and Bruce, 1996; Schlag and Schlag-Rey, 1987). Standard extracellular recording techniques were used to isolate action potentials of single neurons (Sommer and Wurtz, 2004). All data were collected using the REX real-time system (Hays et al., 1982) and analyzed using MATLAB (R20010a, The MathWorks, Inc.).

Analyses

We defined multiple epochs throughout the metacognition task and measured and analyzed the average firing rates within these epochs. **Baseline** was 300 ms before decision stage target onset. During the decision stage, we analyzed a *visual-1* epoch 100–300 ms after target onset. The *visual-1* epoch was selected to start after the masks appeared in every trial, to end before the onset of our epoch for delay activity, and to capture a broad duration of visual-related activity. Also in the decision stage, we analyzed a *delay* epoch 200 ms before fixation offset, a *presaccadic-1* epoch 50 ms before saccade onset, and a *postsaccadic* epoch 100–300 ms after saccade onset. After the decision stage, we defined an *interstage* epoch as the 400 ms surrounding the time the animal regained fixation to initiate the bet stage, from 200 ms before until 200 ms after that time. In the bet stage, we analyzed a *visual-2* epoch 50–150 ms after bet target onset. The start of this epoch was sooner than that of the *visual-1* epoch because there were no masks and we could simply capture the visual response starting at the earliest latencies in the areas under study (generally ~50 ms in the FEF; Pouget et al., 2005). We truncated this epoch at 150 ms after bet target onset to minimize inadvertent measurement of saccade-related activity, given that there was no imposed delay before the bet saccade. Also in the bet stage, we analyzed a *presaccadic-2* epoch 50 ms before bet saccade onset, a *reward anticipation* epoch 250 ms before reward delivery, and a *reward* epoch 50–250 ms after reward.

We performed two types of population analyses. First, we included the entire population of recorded neurons. Then, we focused on only the subsets of neurons that were significantly modulated within particular epochs. A neuron was deemed significantly active in a given epoch if its average firing rate in the epoch on all correct trials (high and low bets pooled) was above its baseline firing rate as determined by paired t tests ($p < 0.05$ criterion). Modulations below baseline were rare, and such neurons were excluded from the second analysis.

To analyze decision-related activity, the average firing rate in each epoch was compared between correct trials and incorrect trials (regardless of bets). For single neuron analyses, comparisons were made using two-sample t tests ($p < 0.05$ criterion). For population analyses, comparisons were made using paired t tests ($p < 0.05$ criterion). Analysis of bet-related activity was

analogous, except we compared average firing rates between all high-bet and all low-bet trials (regardless of decisions).

To analyze metacognition-related activity, the aim was to compare trials in which decisions were identical, but bets (our observables of the monkey's internal state) were different. We compared average firing rates in each epoch between correct-high trials (correct decisions followed by high bets) and correct-low trials (correct decisions-low bets), or between incorrect-high trials (incorrect decisions-high bets) and incorrect-low trials (incorrect decisions-low bets). For single neuron analyses, one-way ANOVAs were first calculated between all four trial conditions. If significant at $p < 0.05$, multiple comparisons (Tukey-Kramer tests) were performed between individual conditions ($p < 0.05$ criterion). For population analyses, paired t tests were calculated between trial outcomes at $p < 0.025$, Bonferroni corrected from 0.05 because we used the same data to analyze reward expectation as well (see Results). Finally, to focus on activity related to targets in a neuron's visual receptive field, or to saccades made into its movement field, we analyzed memory-guided saccade data to ascertain the direction that yielded the strongest visual and presaccadic discharges. We used an epoch 50–150 ms after target onset for the visual response and an epoch 50 ms before saccade onset for the presaccadic activity. The receptive field and/or movement field was defined as the direction that elicited the maximum firing rate within the relevant epoch. In addition, the firing rate was required to be greater than the neuron's baseline firing rate (200 ms before target onset), assessed by t test. We used that direction for our analyses of metacognition task activity that were restricted to the best visual target direction and best saccade direction.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, ten tables, Supplemental Results, and Supplemental Discussion and can be found with this article online at <http://dx.doi.org/10.1016/j.neuron.2012.05.028>.

ACKNOWLEDGMENTS

This research was supported by the National Institute of Mental Health (Kirschstein NRSA F31 MH087094 to P.G.M.) and the National Eye Institute (EY017592 to M.A.S.).

Accepted: May 22, 2012

Published: August 8, 2012

REFERENCES

- Anderson, E.B., Mitchell, J.F., and Reynolds, J.H. (2011). Attentional modulation of firing rate varies with burstiness across putative pyramidal neurons in macaque visual area V4. *J. Neurosci.* *31*, 10983–10992.
- Barracough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* *7*, 404–410.
- Boch, R.A., and Goldberg, M.E. (1989). Participation of prefrontal neurons in the preparation of visually guided eye movements in the rhesus monkey. *J. Neurophysiol.* *61*, 1064–1084.
- Bruce, C.J., and Goldberg, M.E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. *J. Neurophysiol.* *53*, 603–635.
- Call, J. (2010). Do apes know that they could be wrong? *Anim. Cogn.* *13*, 689–700.
- Chua, E.F., Schacter, D.L., Rand-Giovannetti, E., and Sperling, R.A. (2006). Understanding metamemory: neural correlates of the cognitive process and subjective level of confidence in recognition memory. *Neuroimage* *29*, 1150–1160.
- Connors, B.W., and Gutnick, M.J. (1990). Intrinsic firing patterns of diverse neocortical neurons. *Trends Neurosci.* *13*, 99–104.
- Constantinidis, C., and Goldman-Rakic, P.S. (2002). Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. *J. Neurophysiol.* *88*, 3487–3497.
- Crapse, T.B., and Sommer, M.A. (2008). Corollary discharge across the animal kingdom. *Nat. Rev. Neurosci.* *9*, 587–600.
- Ding, L., and Hikosaka, O. (2006). Comparison of reward modulation in the frontal eye field and caudate of the macaque. *J. Neurosci.* *26*, 6695–6703.
- Flavell, J.H. (1976). Metacognitive aspects of problem solving. In *The Nature of Intelligence*, L.B. Resnick, ed. (Hillsdale, NJ: Erlbaum), pp. 231–236.
- Foote, A.L., and Crystal, J.D. (2007). Metacognition in the rat. *Curr. Biol.* *17*, 551–555.
- Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* *61*, 331–349.
- Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1990). Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms. *J. Neurophysiol.* *63*, 814–831.
- Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1991). Neuronal activity related to saccadic eye movements in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* *65*, 1464–1483.
- Hampton, R.R. (2001). Rhesus monkeys know when they remember. *Proc. Natl. Acad. Sci. USA* *98*, 5359–5362.
- Hampton, R.R., Zivin, A., and Murray, E.A. (2004). Rhesus monkeys (*Macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Anim. Cogn.* *7*, 239–246.
- Hays, A.V., Richmond, B.J., and Optican, L. (1982). A UNIX-based multiple process system for real-time data acquisition and control. *WESCON Conference Proceedings*, *2*, 1–10.
- Iba, M., and Sawaguchi, T. (2003). Involvement of the dorsolateral prefrontal cortex of monkeys in visuospatial target selection. *J. Neurophysiol.* *89*, 587–599.
- Johnston, K., and Everling, S. (2006). Neural activity in monkey prefrontal cortex is modulated by task context and behavioral instruction during delayed-match-to-sample and conditional prosaccade-antisaccade tasks. *J. Cogn. Neurosci.* *18*, 749–765.
- Kepecs, A., Uchida, N., Zariwala, H.A., and Mainen, Z.F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature* *455*, 227–231.
- Kiani, R., and Shadlen, M.N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* *324*, 759–764.
- Kikyo, H., Ohki, K., and Miyashita, Y. (2002). Neural correlates for feeling-of-knowing: an fMRI parametric analysis. *Neuron* *36*, 177–186.
- Kim, J.N., and Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* *2*, 176–185.
- Kim, S., Hwang, J., and Lee, D. (2008). Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron* *59*, 161–172.
- Kornell, N., Son, L.K., and Terrace, H.S. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychol. Sci.* *18*, 64–71.
- Mays, L.E., and Sparks, D.L. (1980). Saccades are spatially, not retinocentrically, coded. *Science* *208*, 1163–1165.
- McCoy, A.N., and Platt, M.L. (2005). Expectations and outcomes: decision-making in the primate brain. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* *191*, 201–211.
- Middlebrooks, P.G., and Sommer, M.A. (2011). Metacognition in monkeys during an oculomotor task. *J. Exp. Psychol. Learn. Mem. Cogn.* *37*, 325–337.
- Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* *24*, 167–202.
- Mitchell, J.F., Sundberg, K.A., and Reynolds, J.H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* *55*, 131–141.
- Mohler, C.W., Goldberg, M.E., and Wurtz, R.H. (1973). Visual receptive fields of frontal eye field neurons. *Brain Res.* *61*, 385–389.
- Nakamura, K., Roesch, M.R., and Olson, C.R. (2005). Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. *J. Neurophysiol.* *93*, 884–908.

- Nelson, T.O. (1996). Consciousness and metacognition. *Am. Psychol.* *51*, 102–116.
- Nelson, T.O., and Narens, L. (1990). Metamemory: a theoretical framework and new findings. *Psychol. Learn. Motiv.* *26*, 125–173.
- Pannu, J.K., Kaszniak, A.W., and Rapcsak, S.Z. (2005). Metamemory for faces following frontal lobe damage. *J. Int. Neuropsychol. Soc.* *11*, 668–676.
- Pouget, P., Emeric, E.E., Stuphorn, V., Reis, K., and Schall, J.D. (2005). Chronometry of visual responses in frontal eye field, supplementary eye field, and anterior cingulate cortex. *J. Neurophysiol.* *94*, 2086–2092.
- Reder, L.M., and Schunn, C.D. (1996). Metacognition does not imply awareness: strategy choice is governed by implicit learning and memory. In *Implicit Memory and Metacognition*, L.M. Reder, ed. (Mahwah, NJ: Lawrence Erlbaum Associates), pp. 45–77.
- Roesch, M.R., and Olson, C.R. (2003). Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J. Neurophysiol.* *90*, 1766–1789.
- Romanski, L.M. (2004). Domain specificity in the primate prefrontal cortex. *Cogn. Affect. Behav. Neurosci.* *4*, 421–429.
- Rounis, E., Maniscalco, B., Rothwell, J.C., Passingham, R.E., and Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cogn. Neurosci.* *1*, 165–175.
- Russo, G.S., and Bruce, C.J. (1996). Neurons in the supplementary eye field of rhesus monkeys code visual targets and saccadic eye movements in an oculocentric coordinate system. *J. Neurophysiol.* *76*, 825–848.
- Sato, T.R., and Schall, J.D. (2003). Effects of stimulus-response compatibility on neural selection in frontal eye field. *Neuron* *38*, 637–648.
- Schall, J.D. (2001). Neural basis of deciding, choosing and acting. *Nat. Rev. Neurosci.* *2*, 33–42.
- Schall, J.D., and Thompson, K.G. (1999). Neural selection and control of visually guided eye movements. *Annu. Rev. Neurosci.* *22*, 241–259.
- Schall, J.D., Hanes, D.P., Thompson, K.G., and King, D.J. (1995). Saccade target selection in frontal eye field of macaque. I. Visual and premovement activation. *J. Neurosci.* *15*, 6905–6918.
- Schlag, J., and Schlag-Rey, M. (1987). Evidence for a supplementary eye field. *J. Neurophysiol.* *57*, 179–200.
- Schnyer, D.M., Verfaellie, M., Alexander, M.P., LaFleche, G., Nicholls, L., and Kaszniak, A.W. (2004). A role for right medial prefrontal cortex in accurate feeling-of-knowing judgements: evidence from patients with lesions to frontal cortex. *Neuropsychologia* *42*, 957–966.
- Seo, H., and Lee, D. (2009). Behavioral and neural changes after gains and losses of conditioned reinforcers. *J. Neurosci.* *29*, 3627–3641.
- Shields, W.E., Smith, J.D., Guttmanova, K., and Washburn, D.A. (2005). Confidence judgments by humans and rhesus monkeys. *J. Gen. Psychol.* *132*, 165–186.
- Smith, J.D., Schull, J., Strote, J., McGee, K., Egnor, R., and Erb, L. (1995). The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*). *J. Exp. Psychol. Gen.* *124*, 391–408.
- Smith, J.D., Shields, W.E., Allendoerfer, K.R., and Washburn, D.A. (1998). Memory monitoring by animals and humans. *J. Exp. Psychol. Gen.* *127*, 227–250.
- So, N.Y., and Stuphorn, V. (2010). Supplementary eye field encodes option and action value for saccades with variable reward. *J. Neurophysiol.* *104*, 2634–2653.
- So, N.Y., and Stuphorn, V. (2012). Supplementary eye field encodes reward prediction error. *J. Neurosci.* *32*, 2950–2963.
- Sommer, M.A., and Wurtz, R.H. (2000). Composition and topographic organization of signals sent from the frontal eye field to the superior colliculus. *J. Neurophysiol.* *83*, 1979–2001.
- Sommer, M.A., and Wurtz, R.H. (2001). Frontal eye field sends delay activity related to movement, memory, and vision to the superior colliculus. *J. Neurophysiol.* *85*, 1673–1685.
- Sommer, M.A., and Wurtz, R.H. (2004). What the brain stem tells the frontal cortex. I. Oculomotor signals sent from superior colliculus to frontal eye field via mediodorsal thalamus. *J. Neurophysiol.* *91*, 1381–1402.
- Stuphorn, V., Taylor, T.L., and Schall, J.D. (2000). Performance monitoring by the supplementary eye field. *Nature* *408*, 857–860.
- Suda-King, C. (2008). Do orangutans (*Pongo pygmaeus*) know when they do not remember? *Anim. Cogn.* *11*, 21–42.
- Sutton, R.S., and Barto, A.G. (1998). *Reinforcement Learning, Volume 9* (Cambridge, MA: MIT Press).
- Tanji, J., and Hoshi, E. (2008). Role of the lateral prefrontal cortex in executive behavioral control. *Physiol. Rev.* *88*, 37–57.
- Thompson, K.G., and Schall, J.D. (1999). The detection of visual signals by macaque frontal eye field during masking. *Nat. Neurosci.* *2*, 283–288.
- Thompson, K.G., and Bichot, N.P. (2005). A visual salience map in the primate frontal eye field. *Prog. Brain Res.* *147*, 251–262.
- Wallis, J.D., and Miller, E.K. (2003). From rule to response: neuronal processes in the premotor and prefrontal cortex. *J. Neurophysiol.* *90*, 1790–1806.
- Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* *382*, 629–632.
- Zar, J.H. (1999). *Biostatistical Analysis* (New Jersey: Prentice Hall).

Neuron, Volume 75

Supplemental Information

Neuronal Correlates of Metacognition

in Primate Frontal Cortex

Paul G. Middlebrooks and Marc A. Sommer

INVENTORY OF SUPPLEMENTAL INFORMATION

Four Supplemental Figures:

- Figure S1 is related to Figure 1 of the main text and shows the results of an alternate test of phi correlation.
- Figure S2 is related to Figure 2 of the main text and shows where the neurons were recorded (Figure S2a) and the overall, average activity profiles recorded from those sites (Figure S2b-d).
- Figure S3 is related to Figure 3 of the main text and shows further SEF population analyses, for correct decisions when targets were in the visual receptive field or movement field (Figure S3a-f) and for incorrect decisions (Figure S3g-i).
- Figure S4 is related to Figure 5 of the main text and shows more time course data, for correct decisions sorted by direction and by neurons active in the epoch (Figure S4a) and for incorrect decisions (Figure S4b).

Ten Supplemental Tables:

- Table S1, Saccade Latencies
- Table S2, Proportions of Trial Outcomes
- Table S3, Decision-related Activity: Population Subsets
- Table S4, Bet-related Activity: Population
- Table S5, Bet-related Activity: Population Subsets

- Table S6, Metacognition-related Activity: Comparisons between Monkeys
- Table S7, Bet Stage-related Activity: Population
- Table S8, Bet Stage-related Activity: Population Subsets
- Table S9, SEF Activity Preceding High Bets (CH vs. IH) or Low Bets (CL vs. IL) in Decision Stage and Interstage
- Table S10, SEF Activity During High Bets (CH vs. IH) or Low Bets (CL vs. IL) in Bet Stage and Reward Period

Supplemental Results

Supplemental Discussion

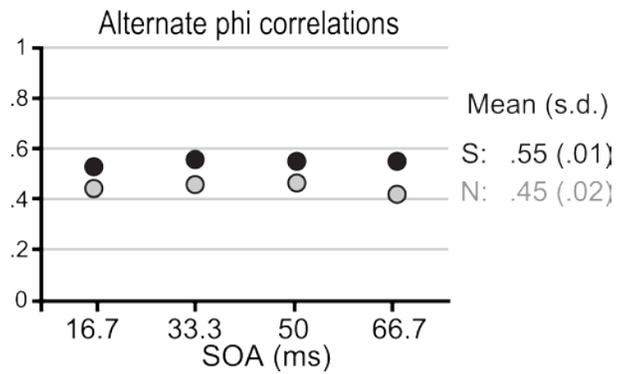
SUPPLEMENTAL INFORMATION**Supplemental Figures**

Figure S1. Results of alternate calculation for phi correlations (Related to Figure 1). This alternate method (Zar, 1999), takes into account whether the correlations are positive or negative. Conventions as in Figure 1c.

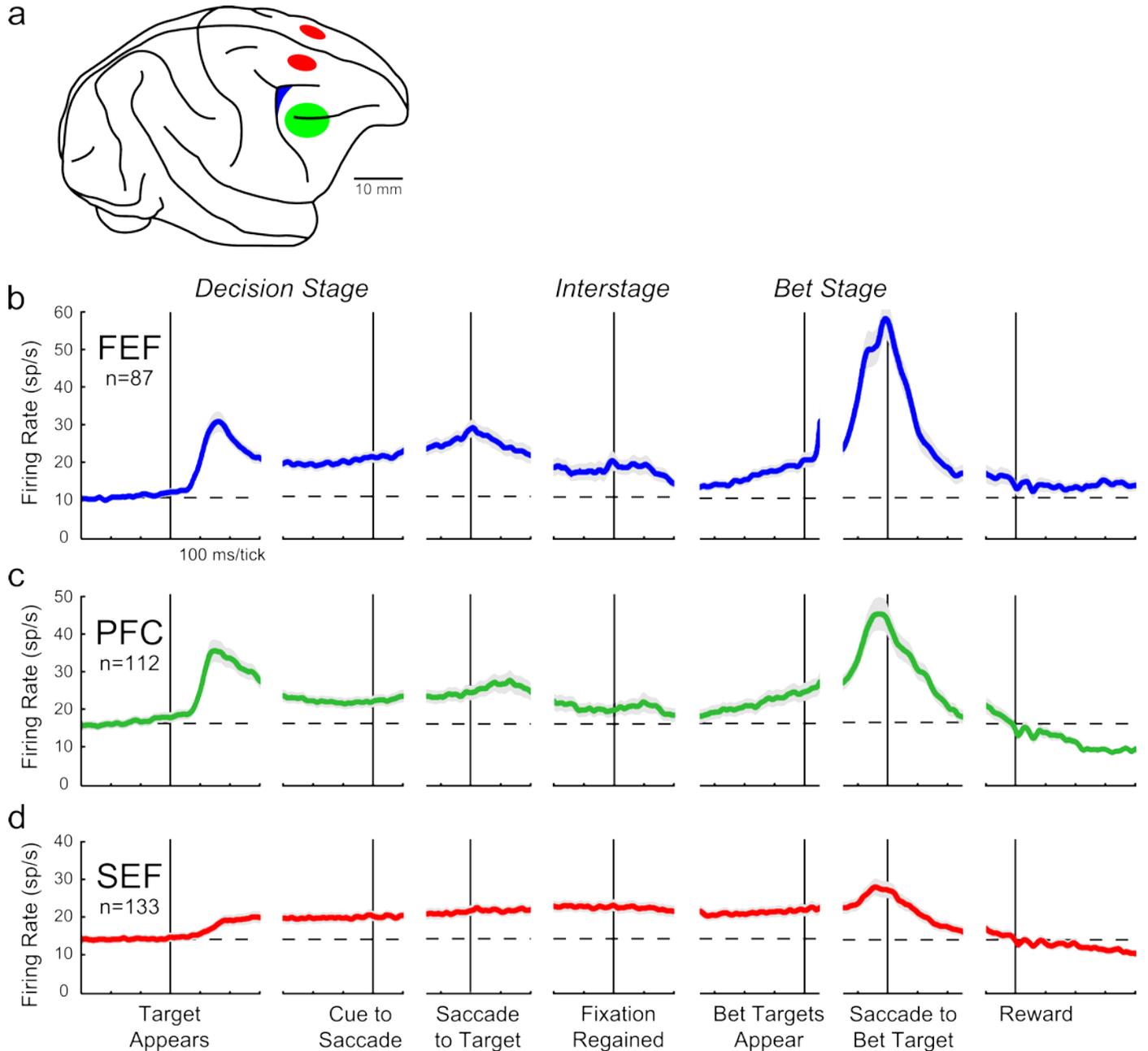


Figure S2. Populations of neurons recorded from FEF, PFC, and SEF (Related to Figure 2). (a) Summary of recording locations: right FEF (blue; in anterior bank of arcuate sulcus), right PFC (green; posterior third of entire PFC), and both SEFs (red). All of the sites were tested in both monkeys. (b-d) Overall activity profiles for our FEF, PFC, and SEF neuronal populations during the metacognition task. All trial outcomes and directions pooled. Traces show mean firing rates (lines) and SEM boundaries (shading) aligned to each task event.

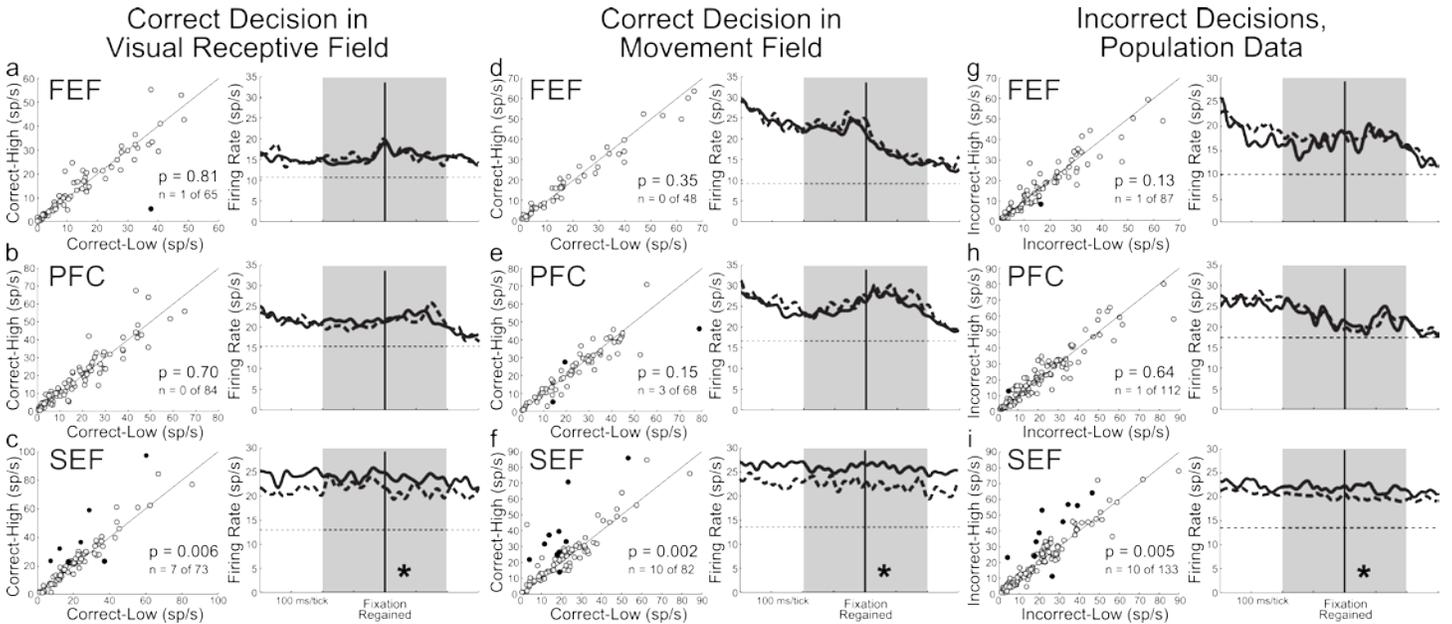


Figure S3. Further analyses of metacognition-related activity in SEF (Related to Figure 3). (a-c)

Scatterplots (left) show CH vs. CL firing rates for each neuron in the FEF, PFC, and SEF, restricted to those neurons with an identifiable visual receptive field, and to trials involving the single targets placed most optimally in that receptive field. Other details as in Figure 3d-f. (d-f) Same as (a-c), but restricted to neurons with an identifiable movement field and trials involving the single targets placed most optimally in that movement field (hence correct saccades were made into the movement field). (g-i) Comparison of IH vs. IL activity during the interstage epoch (same conventions as Figure 3d-f; this is the analog of the CH vs. CL analyses in the main text).

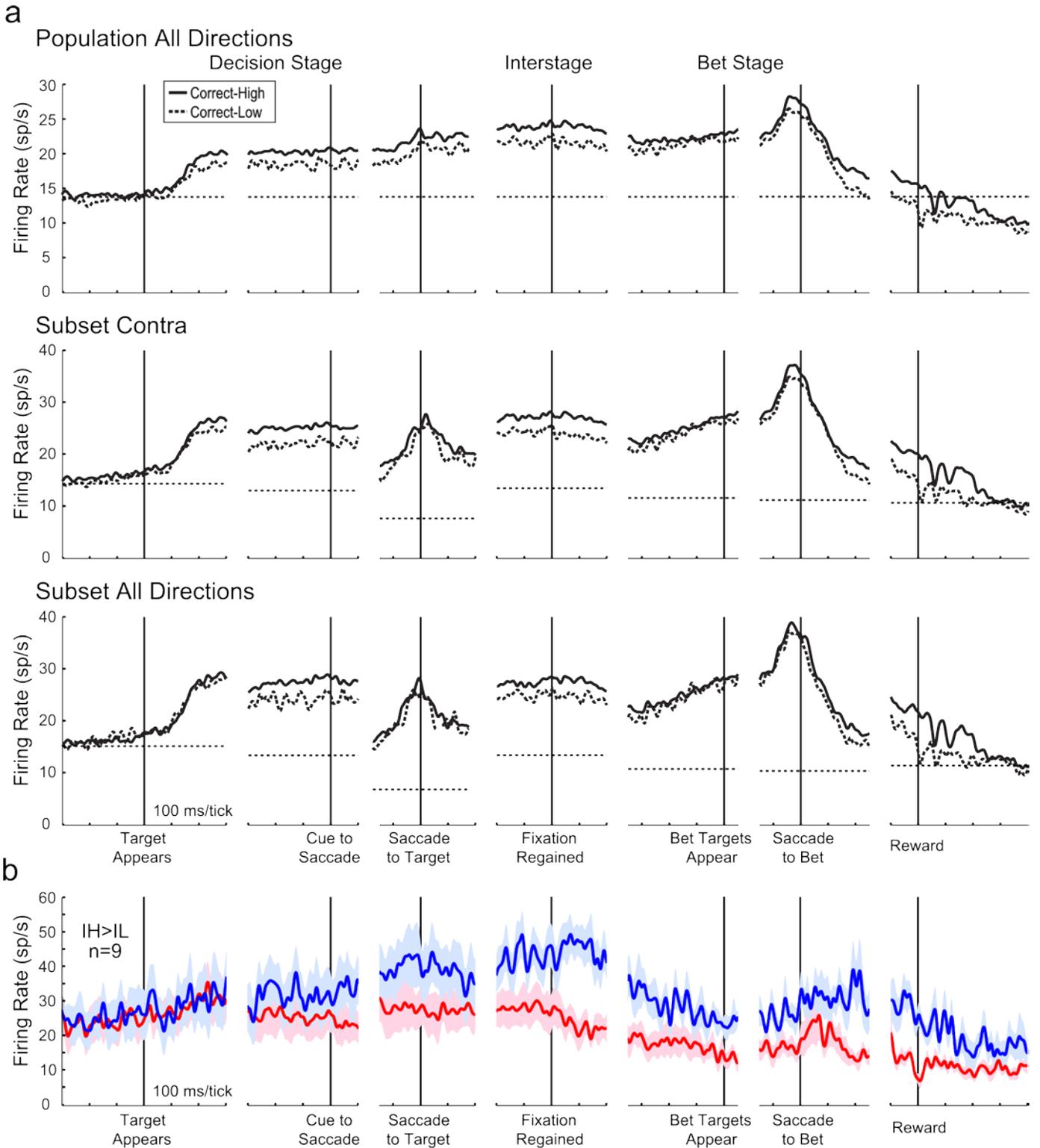


Figure S4. Further analyses of SEF population activity (Related to Figure 5). (a) *Top*: Population data pooled over all directions. *Middle*: Subset of neurons significantly active in each epoch, analyzed by contralateral hemifield. *Bottom*: Same, but analyzed for all directions. (b) The 9 SEF neurons for which IH activity was greater than IL activity during the interstage epoch.

Supplemental Tables

Saccade Latencies

| Outcome | Correct-High | Correct-Low | Incorrect-High | Incorrect-Low |
|----------|--------------|-------------|----------------|---------------|
| Monkey N | 222.3 | 221.7 | 237.8 | 231.9 |
| Monkey S | 167.4 | 163.6 | 174.2 | 177.5 |

Table S1. Latencies to saccade onset (in ms) during the Decision Stage for each monkey, across each trial outcome. Latencies did not differ for either animal between Correct-High vs. Correct-Low trials or between Incorrect-High vs. Incorrect-Low trials, $p > .05$ for all, Holm-Sidak multiple comparison tests.

Proportions of Trial Outcomes

| SOA | Correct-High | Correct-Low | Incorrect-High | Incorrect-Low |
|--------------|--------------|-------------|----------------|---------------|
| Monkey N | | | | |
| 16.7 | 0.68 (0.20) | 0.32 (0.09) | 0.19 (0.14) | 0.81 (0.57) |
| 33.3 | 0.68 (0.27) | 0.32 (0.12) | 0.21 (0.13) | 0.79 (0.48) |
| 50.0 | 0.68 (0.37) | 0.32 (0.18) | 0.19 (0.09) | 0.81 (0.36) |
| 66.7 | 0.67 (0.44) | 0.33 (0.22) | 0.20 (0.07) | 0.80 (0.27) |
| Monkey S | | | | |
| 16.7 | 0.67 (0.21) | 0.33 (0.11) | 0.26 (0.18) | 0.74 (0.50) |
| 33.3 | 0.67 (0.26) | 0.33 (0.13) | 0.25 (0.15) | 0.75 (0.46) |
| 50.0 | 0.70 (0.39) | 0.30 (0.16) | 0.23 (0.10) | 0.77 (0.34) |
| 66.7 | 0.75 (0.53) | 0.25 (0.17) | 0.22 (0.07) | 0.78 (0.23) |
| Both Monkeys | | | | |
| All SOAs | 0.71 (0.35) | 0.29 (0.14) | 0.20 (0.10) | 0.80 (0.41) |

Table S2. Proportions of trial outcomes for each monkey and SOA and, in bottom row, for the overall pooled data. The first number in each table entry is the proportion of Correct-High or Correct-Low trials vs. all Correct trials, or Incorrect-High or Incorrect-Low trials vs. all Incorrect trials. The parenthetical number in each entry shows the proportion of CH, CL, IH, or IL trials relative to all trials. In the bottom row, each number is not necessarily the average of the numbers above it, due to rounding errors and unequal data from each monkey.

Decision-related Activity: Population Subsets

| FEF | Baseline | Visual-1 | Delay | Presaccadic-1 | Postsaccade |
|-----------|------------|-------------------|-------------------|---------------|-------------------|
| n | | 63 | 52 | 41 | 19 |
| Correct | 10.4 (0.9) | 33.0 (2.2) | 25.9 (2.1) | 44.4 (3.5) | 39.4 (6.4) |
| Incorrect | 10.2 (0.9) | 28.5 (1.9) | 21.7 (2.0) | 44.7 (4.0) | 37.4 (6.4) |
| p | .33 | < .001* | < .001* | .81 | .44 |
| PFC | | | | | |
| n | | 80 | 54 | 25 | 20 |
| Correct | 16.0 (1.3) | 42.4 (3.2) | 32.8 (2.6) | 47.7 (6.0) | 40.8 (8.0) |
| Incorrect | 16.4 (1.3) | 37.4 (3.1) | 26.0 (2.4) | 43.0 (5.2) | 39.9 (8.4) |
| p | .47 | < .001* | < .001* | .10 | .64 |
| SEF | | | | | |
| n | | 68 | 76 | 27 | 46 |
| Correct | 13.4 (1.1) | 23.5 (2.3) | 26.5 (2.0) | 26.1 (3.0) | 28.8 (2.4) |
| Incorrect | 13.1 (1.1) | 21.0 (2.1) | 21.2 (1.6) | 25.3 (2.9) | 24.1 (2.2) |
| p | .20 | < .001* | < .001* | .28 | < .001* |

Table S3. Same as Table 1, but for the subsets of neurons within each epoch that had increased activity relative to baseline. For each cortical region, the number (n) of included neurons in each epoch is listed, and all correct vs. all incorrect firing rates (spikes/s) are shown with standard errors in parentheses and paired t-test p-values underneath. Asterisks and bold fonts represent significant differences ($p < .05$) between correct and incorrect trials.

Bet-related Activity: Population

| FEF | Baseline | Visual-1 | Delay | Presaccadic-1 | Postsaccade |
|------|------------|--------------|------------|---------------|-------------|
| High | 10.3 (0.9) | 25.9 (1.9) | 18.7 (1.5) | 29.3 (2.6) | 19.6 (2.3) |
| Low | 10.3 (0.9) | 24.4 (1.7) | 18.1 (1.5) | 30.5 (2.8) | 20.7 (2.4) |
| p | .83 | .006* | .31 | .42 | .16 |

| PFC | Baseline | Visual-1 | Delay | Presaccadic-1 | Postsaccade |
|------|------------|--------------|--------------|---------------|-------------|
| High | 16.0 (1.3) | 32.7 (2.5) | 22.6 (1.6) | 25.6 (2.2) | 25.8 (2.4) |
| Low | 16.5 (1.3) | 31.4 (2.4) | 21.4 (1.5) | 25.4 (1.9) | 26.0 (2.3) |
| p | .20 | .003* | .020* | .82 | .62 |

| SEF | Baseline | Visual-1 | Delay | Presaccadic-1 | Postsaccade |
|------|------------|--------------|-------------------|---------------|--------------|
| High | 13.3 (1.1) | 18.2 (1.4) | 19.9 (1.5) | 22.3 (1.5) | 22.5 (1.5) |
| Low | 13.2 (1.1) | 17.5 (1.3) | 17.5 (1.3) | 20.8 (1.4) | 20.6 (1.3) |
| p | .40 | .036* | < .001* | .013* | .001* |

Table S4. Population bet-related firing rates during Decision Stage epochs. For each cortical region, all high bet vs. all low bet firing rates (spikes/s) are shown with standard errors in parentheses and paired t-test p-values underneath. Asterisks and bold fonts represent significant differences ($p < .05$) between high and low bet trials.

Bet-related Activity: Population Subsets

| FEF | Baseline | Visual-1 | Delay | Presaccadic-1 | Postsaccade |
|------|------------|-------------------|-------------------|---------------|--------------|
| n | | 63 | 52 | 41 | 19 |
| High | 10.3 (0.9) | 31.7 (2.1) | 24.8 (2.0) | 44.4 (3.6) | 38.4 (6.3) |
| Low | 10.3 (0.9) | 29.3 (2.0) | 22.4 (2.1) | 43.8 (3.8) | 38.3 (6.4) |
| p | .83 | .001* | .001* | .69 | .97 |
| PFC | | | | | |
| n | | 80 | 54 | 25 | 20 |
| High | 16.0 (1.3) | 40.6 (3.2) | 30.8 (2.6) | 49.0 (6.3) | 40.9 (8.0) |
| Low | 16.5 (1.3) | 38.4 (3.1) | 27.3 (2.4) | 42.2 (4.6) | 39.4 (8.1) |
| p | .20 | < .001* | < .001* | .019* | .32 |
| SEF | | | | | |
| n | | 68 | 76 | 27 | 46 |
| High | 13.3 (1.1) | 22.8 (2.2) | 25.8 (2.0) | 26.2 (3.0) | 27.7 (2.4) |
| Low | 13.2 (1.1) | 21.9 (2.1) | 22.2 (1.8) | 25.2 (3.0) | 24.5 (2.1) |
| p | .40 | .11 | < .001* | .26 | .001* |

Table S5. Same as Table S4, but for the subsets of neurons within each epoch that had increased activity relative to baseline. For each cortical region, the number (n) of included neurons in each epoch is listed, and all high bet vs. all low bet firing rates (spikes/s) are shown with standard errors in parentheses and paired t-test p-values underneath. Asterisks and bold fonts represent significant differences ($p < .05$) between high and low bet trials.

Metacognition-related Activity: Comparisons between Monkeys

| Monkey | Neuronal Sample | Individual Neurons Significant | CH Firing Rate | CL Firing Rate | CH-CL p value |
|--------|-----------------|--------------------------------|----------------|----------------|---------------|
| S | All | 31% (22/72) | 26.6 (2.2) | 23.1 (1.7) | .003* |
| | Active | 40% (19/47) | 29.2 (2.4) | 24.5 (1.8) | .001* |
| N | All | 10% (6/61) | 20.5 (2.3) | 19.5 (2.3) | .002* |
| | Active | 14% (6/42) | 23.8 (2.7) | 22.6 (2.7) | .001* |

Table S6. Each monkey's metacognition-related SEF activity data during the interstage epoch, analyzed over all directions. Within each row corresponding to data from one monkey, the results of analyzing all neurons are presented above the results of analyzing only the neurons significantly active during the interstage epoch. Average firing rates (spikes/s) are shown with standard errors in parentheses. Asterisks and bold fonts represent significant differences of paired t-tests ($p < .025$).

Bet Stage-related Activity: Population

| FEF | Visual-2 | Presaccadic-2 | Reward Anticipation | Reward |
|-----|------------|---------------|---------------------|--------------|
| CH | 49.9 (4.6) | 50.8 (4.7) | 16.9 (2.0) | 11.5 (1.4) |
| CL | 51.3 (4.5) | 53.6 (4.8) | 18.9 (2.0) | 12.8 (1.3) |
| p | .35 | .13 | .005* | .16 |
| PFC | | | | |
| CH | 43.5 (3.8) | 44.3 (4.0) | 23.8 (2.1) | 11.7 (1.2) |
| CL | 44.1 (4.0) | 44.3 (4.0) | 26.1 (2.3) | 12.4 (1.1) |
| p | .71 | .99 | .04 | .28 |
| SEF | | | | |
| CH | 27.2 (2.2) | 27.8 (2.2) | 19.1 (1.6) | 12.0 (1.4) |
| CL | 26.1 (2.1) | 27.6 (2.4) | 16.4 (1.6) | 10.3 (1.2) |
| p | .10 | .80 | < .001* | .002* |

Table S7. Population Bet Stage-related firing rates. For each cortical region, all Correct-High (CH) vs. all Correct-Low (CL) firing rates (spikes/s) are shown with standard errors in parentheses and paired t-test p-values underneath. Asterisks and bold fonts represent significant differences ($p < .025$) between CH and CL trials.

Bet Stage-related Activity: Population Subsets

| FEF | Visual-2 | Presaccadic-2 | Reward Anticipation | Reward |
|-----|------------|---------------|---------------------|--------------|
| n | 64 | 63 | 33 | 25 |
| CH | 59.2 (5.0) | 61.1 (5.1) | 29.8 (3.5) | 20.6 (3.0) |
| CL | 61.2 (4.7) | 64.2 (5.2) | 31.9 (3.4) | 18.9 (2.5) |
| p | .29 | .17 | .08 | .39 |
| PFC | | | | |
| n | 77 | 72 | 65 | 17 |
| CH | 57.2 (4.6) | 62.0 (5.0) | 32.5 (2.9) | 26.7 (4.3) |
| CL | 58.2 (4.8) | 62.2 (4.9) | 34.3 (3.2) | 22.5 (3.5) |
| p | .66 | .93 | .30 | .016* |
| SEF | | | | |
| n | 72 | 69 | 63 | 28 |
| CH | 35.1 (2.9) | 37.8 (3.0) | 25.5 (2.5) | 22.6 (4.2) |
| CL | 33.9 (3.0) | 37.5 (3.5) | 21.6 (2.6) | 18.9 (3.9) |
| p | .23 | .85 | .002* | .018* |

Table S8. Same as Table S7, but for the subsets of neurons within each epoch that had increased activity relative to baseline. For each cortical region, the number (n) of included neurons in each epoch is listed, and all Correct-High (CH) vs. all Correct-Low (CL) firing rates (spikes/s) are shown with standard errors in parentheses and paired t-test p-values underneath. Asterisks and bold fonts represent significant differences ($p < .025$) between CH and CL trials.

SEF Activity Preceding High Bets (CH vs. IH) or Low Bets (CL vs. IL) in Decision Stage and Interstage

| | Population | Baseline | Visual-1 | Delay | Presaccadic-1 | Postsaccade | Interstage |
|------------------|------------|------------|--------------|-------------------|---------------|--------------|-------------------|
| High Bets | CH | 13.6 (1.1) | 18.5 (1.4) | 20.1 (1.6) | 22.2 (1.5) | 22.9 (1.5) | 23.6 (1.6) |
| | IH | 13.6 (1.1) | 16.9 (1.4) | 17.3 (1.3) | 20.8 (1.6) | 21.5 (1.5) | 21.4 (1.6) |
| | p | .49 | .002* | < .001* | .04 | .004* | .01* |
| Subset | | | | | | | |
| | n | | 62 | 72 | 30 | 42 | 90 |
| | CH | 13.6 (1.1) | 24.5 (2.4) | 27.9 (2.2) | 24.0 (2.7) | 27.8 (2.7) | 27.8 (2.0) |
| | IH | 13.6 (1.1) | 21.9 (2.4) | 21.4 (1.9) | 25.4 (3.2) | 24.9 (2.4) | 23.1 (1.0) |
| | p | .49 | .003* | < .001* | .30 | .008* | < .001* |
| Population | | | | | | | |
| Low Bets | CL | 13.1 (1.2) | 17.7 (1.5) | 18.6 (1.4) | 20.1 (1.4) | 21.0 (1.4) | 21.3 (1.5) |
| | IL | 13.0 (1.1) | 16.8 (1.3) | 17.0 (1.2) | 20.5 (1.4) | 20.4 (1.4) | 19.6 (1.4) |
| | p | .86 | .23 | .046* | .69 | .08 | .01* |
| Subset | | | | | | | |
| | n | | 62 | 72 | 30 | 42 | 90 |
| | CL | 13.1 (1.1) | 24.3 (2.5) | 24.2 (2.1) | 22.7 (2.8) | 27.8 (2.7) | 25.2 (1.8) |
| | IL | 13.0 (1.1) | 21.6 (2.2) | 21.4 (1.7) | 20.8 (2.3) | 23.1 (2.3) | 21.2 (1.7) |
| | p | .86 | .004* | .024* | .63 | .003* | < .001* |

Table S9. SEF firing rates compared between correct and incorrect high bet trials (CH vs. IH), *top section*, and correct and incorrect low bet trials (CL vs. IL), *bottom section*, during the Decision Stage and the Interstage epoch. Within each High or Low Bets section, results from the total population (upper row) and the subset of neurons with significant activity in each epoch (lower row) are shown separately. For each epoch (columns), firing rates for correct and incorrect trials are shown (spikes/s) with standard errors in parentheses and compared using paired t-tests (p-values underneath). Asterisks and bold fonts represent significant differences ($p < .025$) between data for which the same eventual reward was anticipated, but after different decisions.

SEF Activity During High Bets (CH vs. IH) or Low Bets (CL vs. IL) in Bet Stage and Reward Period

| | | | | | |
|----------------------|------------|-------------------|-------------------|---------------------|------------------|
| High Bets | Population | Visual-2 | Presaccadic-2 | Reward Anticipation | Reward |
| | CH | 27.2 (2.2) | 27.9 (2.2) | 19.1 (1.6) | 12.0 (1.4) |
| | IH | 25.3 (2.0) | 26.2 (2.2) | 18.6 (1.6) | 16.1 (1.6) |
| | p | .007* | .05 | .62 | <.001* |
| Subset | | | | | |
| | n | 72 | 69 | 63 | 28 |
| | CH | 35.1 (2.9) | 37.8 (3.0) | 25.5 (2.5) | 22.6 (4.2) |
| | IH | 31.6 (2.7) | 34.1 (3.2) | 21.9 (2.2) | 24.7 (4.5) |
| | p | < .001* | < .004* | .018* | .63 |
| <hr/> | | | | | |
| Low Bets | Population | Visual-2 | Presaccadic-2 | Reward Anticipation | Reward |
| | CL | 26.1 (2.1) | 27.6 (2.4) | 16.4 (1.6) | 10.3 (1.2) |
| | IL | 24.8 (2.1) | 26.1 (2.4) | 16.8 (1.6) | 11.0 (1.3) |
| | p | .10 | .11 | .43 | .04 |
| Subset | | | | | |
| | n | 72 | 69 | 63 | 28 |
| | CL | 33.9 (3.0) | 37.5 (3.5) | 21.6 (2.6) | 18.9 (3.9) |
| | IL | 31.0 (3.0) | 35.3 (3.5) | 20.9 (2.6) | 18.5 (3.9) |
| | p | .019* | .16 | .27 | .61 |

Table S10. Same as Table S9, but for Bet Stage and Reward periods.

Supplemental Results

Bet-related activity

We analyzed bet-related activity by comparing all high bet versus all low bet activity during the Decision Stage. Considering that behavioral analyses showed that most correct (or incorrect) decisions were followed by high (or low) bets, we expected that bet-related activity would parallel decision-related activity. This prediction was confirmed. For each cortical region, population activity in both the visual-1 and delay epochs predicted the eventual bets (except for the FEF delay period; Table S4). During saccade-related periods, SEF activity predicted bets but neither FEF nor PFC did. For the subsets of neurons specifically active in the epochs (Table S5), activity in all three areas predicted bets in both the visual-1 and delay epochs (except the SEF visual-1 period) and for saccade-related epochs, SEF postsaccadic and PFC presaccadic-1 activity predicted bets.

Spike Burstiness

The $CH > CL$ and $CH < CL$ subsets could consist of differing neuron phenotypes, as suggested by their differing overall activity levels. Inhibitory interneurons (“fast spiking” neurons) tend to have higher spontaneous and task-related firing rates than pyramidal (“regular spiking”) neurons (e.g. Connors and Gutnick, 1990; Mitchell et al., 2007). We could not perform the standard test for distinguishing between fast spiking and regular spiking neurons (comparisons of action potential widths; e.g. Mitchell et al., 2007), so we conducted a test on spiking statistics instead: burstiness analysis. Anderson et al. (2011) demonstrated that in primate cerebral cortex, putative inhibitory neurons (as identified with action potential width analysis) show a lower incidence of burstiness and lower median burstiness than putative pyramidal neurons. For the 20 neurons in our $CH > CL$ and $CH < CL$ subsets, we calculated each neuron’s burstiness/refractory index (BRI), which summarizes how often a neuron fires spikes in quick succession (Anderson et al., 2011). Briefly, a shift predictor (mean cross-correlation of all trials for the neuron) was subtracted from the neuron’s auto-correlation, and then that number was divided by the standard deviation of the shift predictor. The result, the BRI, is

analogous to a z-score; its units are in standard deviations, and a neuron with $BRI > 2$ was considered significantly bursty. We calculated BRI for spike trains during two epochs: the baseline epoch, when activity was as “spontaneous” (non-task-related) as possible, and the interstage epoch, when the main task-related function of interest, metacognition, occurred (analogous to the attentional, “sustained” period tested by Anderson et al., 2011). Data were collapsed over all trial outcomes.

We found no evidence from burstiness analysis to support the hypothesis that the $CH > CL$ and $CH < CL$ subsets consisted of different neuronal phenotypes. The *incidence of bursty neurons* in the baseline epoch was 64% (9/14 neurons) for the $CH > CL$ subset and 83% (5/6 neurons) for the $CH < CL$ subset. In the interstage epoch, the incidence was 78% (11/14 neurons) and 50% (3/6 neurons) respectively. In neither epoch was the incidence of bursty neurons significantly different between the two subsets (Fisher Exact Tests, $p > .3$ for both). The *median BRI* in the baseline epoch was 4.22 for the $CH > CL$ subset and 9.40 for the $CH < CL$ subset. In the interstage epoch, the medians were 8.72 and 4.80 respectively. In neither epoch was BRI significantly different between the two subsets (Mann-Whitney U tests, $p > .5$ for both). These negative results should be considered with caution, because small numbers of neurons were analyzed and burstiness alone (without action potential width data) is a weak predictor of neuronal phenotype (Anderson et al., 2011). Larger samples of metacognition-related SEF neurons that include action potential waveform data are needed to fully test the hypothesis that the $CH > CL$ and $CH < CL$ subsets consist of different neuronal phenotypes.

IH versus IL

The complementary approach to testing whether neuronal activity correlates with metacognitive behavior is to compare Incorrect-High (IH) and Incorrect-Low (IL) trials. As noted in the main text, because the target location was by definition not coincident with the saccade destination in incorrect trials, and because IH trials were rare, analysis of IH and IL trials was not as straightforward as for CH and CL trials. We separated the IH vs. IL analysis into a sensory-related activity comparison (visual-1 and delay epochs) and a motor-related

activity comparison (presaccadic-1 and postsaccadic epochs) as in our analysis of decision related activity. We used the motor-related activity comparison method to analyze the interstage epoch, because this epoch was during a time when the monkey's gaze returned to the center of the screen after making its (erroneous) decision saccade.

For the critical interstage epoch, our analysis of IH vs. IL trials yielded similar results as our analysis of CH vs. CL trials. The SEF population exhibited differential activity during the interstage epoch for IH vs. IL trials (IH: 21.4 ± 1.6 sp/s, IL: 19.6 ± 1.4 sp/s, $p = .005$), but the PFC and FEF populations did not. None of the cortical regions showed differential activity between IH and IL trials during the visual-1, delay, and presaccadic-1 periods. In the postsaccadic epoch neither the SEF nor PFC showed an effect, and the FEF showed a reverse effect; as a population, it was less active for IH trials than IL trials (IH: 19.0 ± 2.6 sp/s, IL: 21.97 ± 2.9 sp/s, $p = .008$). Finally, we repeated these analyses using only the subsets of neurons with significant activity in the given epochs. The SEF was again differentially active during the interstage epoch (IH: 23.1 ± 1.0 sp/s, IL: 21.18 ± 1.7 sp/s, $p = .02$), but the PFC and FEF were not. None of the remaining comparisons were significant.

Bet Stage-related activity

We analyzed the neuronal activity after the interstage epoch, through the Bet Stage, to the end of the trial. It is clear from inspection that, as the monkeys awaited the appearance of the bet targets, differential CH-CL activity in the SEF population waned (Figure 5c, range before "Bet Targets Appear"). Quantitatively, we found that none of the three cortical regions differentiated CH-CL or IH-IL in their visual responses to the bet targets (visual-2 epoch) or in their activities associated with saccades to the bet targets (presaccadic-2 epoch; paired t-tests, all $p > .025$). This held true for total population analyses (CH vs. CL data are shown in Table S7), and for the subsets of neurons that had increased activity relative to baseline (CH vs. CL data are shown in Table S8).

The SEF population activity became differentially modulated again after the monkeys placed their bets, in the reward anticipation and reward epochs (Tables S7 and S8; also apparent in Figure 5c, “Saccade to Bet” and “Reward” alignments). Comparable effects in the FEF and PFC were rare (significant only for the FEF population during the reward anticipation period, Table S7, and the PFC subset during the reward period, Table S8). For IH-IL comparisons, the SEF and PFC total populations had greater firing rates for IH trials during the reward epoch (SEF: IH: 16.1 ± 1.6 sp/s, IL: 11.0 ± 1.3 sp/s, $p < .001$; PFC: IH: 15.91 ± 1.5 sp/s, IL: 11.0 ± 1.0 sp/s, $p < .001$), but no other effects were significant (not shown).

Supplemental Discussion

Riskiness

Differences in neuronal activity between trial outcomes could have been related to riskiness of high vs. low bets. Were this the case, one would predict CH and IH trials to both have greater firing rates than CL and IL trials. Our analyses in the main text demonstrated that CH trials had greater firing rates than CL trials, and that IH trials had greater firing rates than IL trials, both of which are consistent with riskiness as an explanation. IH and CL trials, however, were not significantly different in any epoch throughout the trial (paired t-tests, all $p > .025$) except in the reward epoch (after the trial ended, too late to influence behavior). SEF activity was therefore inconsistent with a representation of riskiness.

Reward

Another alternative account is that the neuronal activity may be representing upcoming reward amounts for each trial outcome. Neuronal activity levels would map onto actual, rather than expected, reward values for each trial outcome. The prediction would be relatively high firing rates for CH trials and low firing rates for IH trials. As noted above, however, population IH trial activity was greater than IL trial activity and no different than CL trial activity, so reward is not a satisfactory explanation of the neuronal data.

Decision-related signals

Our finding of decision-related signals in the SEF and PFC, in addition to in the FEF (Thompson and Schall, 1999), was interesting but not central to our main result of metacognition-related activity in the SEF. It is important nonetheless to address some potential issues surrounding the interpretation of signals as “decision-related”. The masked target task has spatial features that could have introduced confounds. In our analysis of visual-1 and delay activity, we considered only trials in which the target was in the contralateral hemifield. This meant that on Correct trials, the saccade was made into the contralateral hemifield, but on Incorrect trials, the saccade was made into the ipsilateral hemifield. If signals related to preparing the saccade occurred early enough in a trial, they could have mimicked the signals that we interpreted as being correlated with Correct vs. Incorrect decisions. It is well established that saccade preparation/target selection signals are not, in fact, confounded with early visual activity (reviewed by Schall and Thompson, 1999, and Schall, 2001). Early visual responses represent visual stimuli whether they are targets or distractors, and saccadic target selection signals appear when the initial visual response wanes. The differential activity we (and Thompson and Schall, 1999) found that distinguished Correct from Incorrect trials started extremely early in many individual neurons (Figure 2a-c). Even in the averaged population profiles, where latencies of individual visual responses are “smeared” together, it can be seen that the differential responses occurred early. In the FEF and PFC data, in which there was a clear peak to the visual response, Correct-Incorrect differences started prior to that peak (Figure 2a,b). Motor-related signals would not start until well after that peak (Schall and Thompson, 1999; Schall, 2001). Motor-related signals could have contributed to activity in the delay epoch, but even if they did, our main result would be unchanged based on the visual-1 epoch results.

When we analyzed saccade-related activity, we re-sorted trials into those that included only saccades into the contralateral hemifield. This was to explicitly rule out contralateral-ipsilateral differences in activity related to preparation or execution of saccades. In the SEF, we still saw Correct-Incorrect differences, as a

continuation of the signal that started in the visual-1 epoch. The only trivial explanation for those differences would be if SEF neurons had persistent visual responses to briefly flashed, then masked, targets. In that scenario, the higher activity on Correct trials would be attributed to the contralateral target location, and the lower activity on Incorrect trials would be due to the ipsilateral target location. Especially given that SEF visual responses are not prominent to begin with, relative to FEF and PFC visual responses (Figure 2d-f and Figure S2b-d), the idea that the locations of flashed, masked stimuli are represented up to and even after the saccade seems highly unlikely. The most parsimonious interpretation was that the differential activity in the SEF during the presaccadic-1 and postsaccade epochs was correlated simply with making Correct vs. Incorrect decisions.

To Bet, or Not to Bet: That Is the Question of SEF Spikes

Kentaro Miyamoto,^{1,2} Toshiyuki Hirabayashi,^{1,2} and Yasushi Miyashita^{1,*}

¹Department of Physiology, The University of Tokyo School of Medicine, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

²These authors contributed equally to this work

*Correspondence: yasushi_miyashita@m.u-tokyo.ac.jp

<http://dx.doi.org/10.1016/j.neuron.2012.07.012>

The ability of animals to monitor their own cognitive processes is called metacognition. In this issue of *Neuron*, Middlebrooks and Sommer (2012) show that single-unit activity of SEF neurons exhibit a metacognitive signal while monkeys perform a postdecision wagering task.

When you look into a convex mirror, you will see yourself looking into the mirror (Figure 1A). You might pick up the mirror and move it around your face. Then you will see yourself reflected at various angles under the control of your hand's movement. Like the mirror reflecting us, we are endowed with the ability to monitor our own thoughts and cognition from various aspects. This ability is termed metacognition (Flavell, 1979). For instance, if you are cramming for an upcoming history exam, you may decide to focus on the material that you feel you understand the least. Or when you are reading a difficult book, you may reread a paragraph if you feel you did not initially grasp its meaning, and in some cases you may look up background information in an encyclopedia. Metacognition is the process by which you make a judgment on the basis of introspection of your own cognitive state. In this way, metacognition allows you to assess and regulate the current state of your cognitive activity so that you can determine how to act in a given situation (Dunlosky and Metcalfe, 2009).

Localization of metacognitive functioning in the human brain was attempted in a neuropsychological study of specific frontal lesions (Schnyer et al., 2004) and in an fMRI study of healthy subjects (Kikyo et al., 2002; Maril et al., 2003). Some frontal areas were found to be recruited when participants experienced a "feeling of knowing" what was to be recalled (Kikyo et al., 2002). Metacognitive ability had been thought to be unique to humans; however, recent studies show that rhesus monkeys also exhibit metacognitive behavior when performing cognitive

tasks (Hampton, 2001; Kiani and Shadlen, 2009; Kornell et al., 2007). Monkeys are capable of making reasonable "bets" on whether they were correct or incorrect in a perceptual or mnemonic test they had just taken. In this issue of *Neuron*, Middlebrooks and Sommer (2012) recorded the spiking activity of single neurons in the macaque frontal cortex during a metacognitive task (Figure 1B). This study is novel in its use of electrophysiology with high temporal and spatial resolution to capture a metacognitive process in macaque frontal cortex, a neural substrate that is shared by humans and monkeys.

The authors investigated the neuronal correlates of metacognition in this study using a postdecision wagering task (Middlebrooks and Sommer, 2011). This task comprised two stages (Figure 1B). In the first stage, monkeys performed an oculomotor delayed response to a presented cue stimulus (decision stage). Task difficulty was manipulated by randomly changing the time interval between the cue stimulus and the subsequent mask (stimulus onset asynchrony, SOA). After the decision (i.e., oculomotor response), and following a subsequent delay period, the monkeys chose one of two options by making another saccade (bet stage). One of the options ("high-bet") offered a larger reward only if the monkey made a correct saccade at the preceding decision stage, whereas the other option ("low-bet") guaranteed a smaller, but certain, reward regardless of whether the monkey made a correct decision. To earn the largest reward, the animals had to monitor their own decision in each trial and choose an appropriate option on the basis of a confidence in the decision, and this process is

metacognitive. The authors conducted single-unit recordings while the animals performed this task, which enabled them to examine the metacognitive signal at the single neuron level. They recorded the neuronal activity from three different areas in the frontal cortex (frontal eye field [FEF], dorsolateral prefrontal cortex [PFC], and supplementary eye field [SEF]) and examined which of these areas is most involved in metacognition.

Behavioral analysis first revealed that the monkeys performed this task as expected: the animals indeed made a correct decision more frequently when they chose the high-bet compared to when they chose the low-bet. This was true for each SOA, indicating that the monkeys placed their bets on the basis of trial-by-trial monitoring of their own decision, and not just on the basis of task difficulty.

Single-unit activity during this task was then analyzed for the FEF, PFC, and SEF in the frontal cortex. First, the authors compared neuronal activity for correct and incorrect decisions at the decision stage and found that all three areas exhibited significant increases in activity when the decision was correct. They next focused on activity during the time period between the decision and bet stages (interstage period). The authors hypothesized that the neuronal activity during this period would probably link the animal's decision and the subsequent bet, and thus encode the metacognitive signal. If neuronal activity encodes the animal's metacognition, there should be differences in activity between high- and low-bet conditions even for the same preceding decision. During the interstage

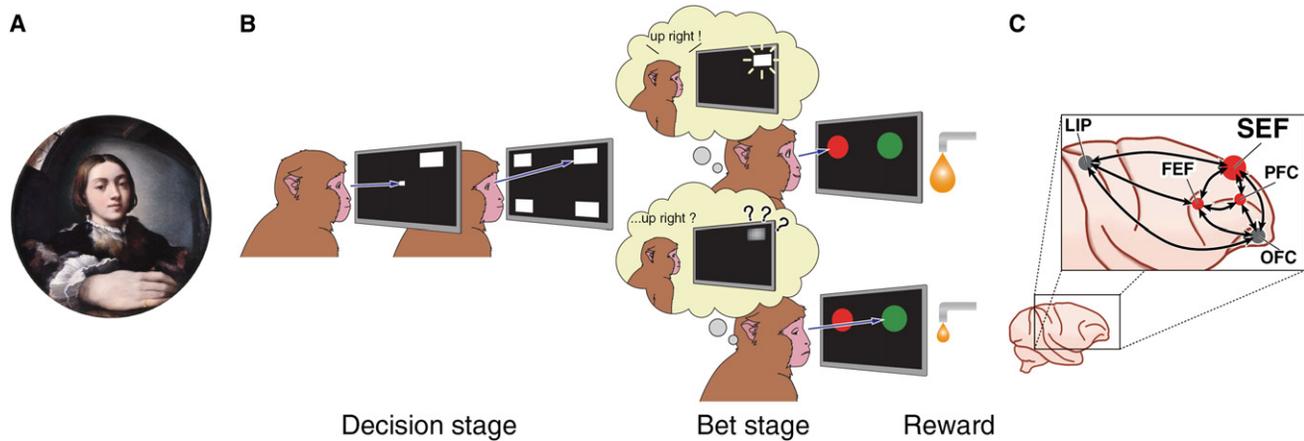


Figure 1. Metacognition in Macaque Monkeys: Frontoparietal Network for Decision Monitoring

(A) Metacognition is the ability to think about one's own thinking. In this picture, a man views himself reflected by a convex mirror (from Parmigianino's "Self-portrait in a Convex Mirror").
 (B) In Middlebrooks and Sommer (2012), monkeys first detected and reported the location of a peripheral target (decision stage) and then made a bet based on their decision (bet stage). When the monkeys chose the "high bet" (red circle), they earned the maximum reward for a correct decision but faced a penalty of a timeout without a reward for an incorrect decision. When the monkeys chose the "low bet" (green circle), they earned a minimal reward irrespective of the correctness of the decision.
 (C) Metacognition-related areas in which single-unit activity has been investigated so far. The red circles indicate the areas in which activity was recorded in the present study. The gray circles correspond to the areas targeted in monkeys by Kiani and Shadlen (2009) or in rats by Kepecs et al. (2008). The black arrows indicate anatomical connections (Cavada et al., 2000; Lynch and Tian, 2006).

period, the neuronal activity in FEF and PFC was indistinguishable when different bets were made following the same correct decision. However, SEF neurons exhibited significant differences in activity when high- and low-bets were made following the same correct decision. The activity was on average stronger for the high-bet compared to the low-bet. These results suggest that the activity of SEF neurons, but not that of PFC or FEF neurons, reflected the monkey's decision monitoring for the subsequent wagering.

The activity of SEF neurons has been shown to encode the animal's anticipation of a reward (Roesch and Olson, 2003; So and Stuphorn, 2010). Therefore, an important issue regarding the observed metacognitive signal is the involvement of reward anticipation. To address this, the authors examined differences in activity when the same bet was preceded by different (correct or incorrect) decisions. They hypothesized that SEF activity would be indistinguishable in these conditions if it encodes reward anticipation. They found that the activity of SEF neurons during the interstage period showed a significant difference between the conditions of correct and incorrect decisions followed by the same bet, suggesting that reward anticipation

of itself does not explain the activity of SEF neurons. This is a good control in their paradigm; however, the relationships between reward anticipation and the two-alternative forced choice of bets might be more complicated than the authors assumed. The relationships between metacognitive signal and reward anticipation should be examined more closely from various points of view in future studies.

Metacognition-related neuronal activity has been shown at the single-neuron level in a few previous studies. In particular, Kiani and Shadlen (2009) examined the neuronal signal encoding choice certainty in monkeys using an opt-out task paradigm. First, the monkeys were presented with moving dot stimuli with a given level of coherence. Monkeys were then given two forced choices, one of which indicated the correct direction of the dot motion and offered a reward. In half of the trials, a third opt-out choice was also presented in which the monkeys could receive a smaller, but certain, reward without choosing a direction. The authors recorded single-unit activity in the lateral intraparietal area (LIP) during this task and found that when the animal chose the opt-out option, the activity of LIP neurons was intermediate (i.e., between

the levels recorded when the correct target was located in and outside of the response field). The intermediate level indicates that the activity did not encode the saccadic target, suggesting that the activity of LIP neurons reflected monkey's certainty regarding the perceived direction. In this paradigm, the animal's decision and its monitoring could not be temporally segregated. In the present study, the decision stage and bet stage were temporally segregated with the linkage by the interstage period, so that the authors could extract the neuronal correlates of decision monitoring as a metacognitive process. The authors indeed found that the majority of SEF neurons that encoded decision monitoring during the interstage period also coded for the decision itself at the decision stage (i.e., different activity between correct and incorrect decisions) and discussed that the observed metacognitive signal of SEF neurons might have evolved from the decision signal. Both studies in monkeys, however, opened an important possibility that neuronal mechanisms underlying metacognitive functions can be tapped in the primate frontal and parietal cortices at the single-neuron level by devising an adequate behavioral paradigm. Furthermore, in a pioneering work

by [Kepecs et al. \(2008\)](#), they demonstrated that the activity of neurons in the rat orbitofrontal cortex (OFC) matched the model of the rat's uncertainty regarding their own past decision. Metacognitive signals in the corresponding area in monkeys should thus be examined in future studies, which will facilitate our understanding of the relationships between the metacognitive signals in different brain areas ([Figure 1C](#)).

The strength of the metacognitive signal observed in [Middlebrooks and Sommer \(2012\)](#) was several spikes per second on average, which is not a large proportion of all the spikes fired by these neurons. Therefore, readout mechanisms and the behavioral impact of the observed metacognitive signals should be considered carefully. This is related to the issue of across-areal neuronal circuitry for metacognition, which would include the SEF, LIP, and presumably OFC, among which

anatomical connections have been identified ([Figure 1C](#)) ([Cavada et al., 2000](#); [Lynch and Tian, 2006](#)). Clarifying the hierarchical relationships between these areas and differentiating their roles in metacognition should be the next step in understanding the neuronal circuitry that implements this cognitive process, which we humans profoundly exploit to lead our daily lives.

REFERENCES

- Cavada, C., Compañy, T., Tejedor, J., Cruz-Rizzolo, R.J., and Reinoso-Suárez, F. (2000). *Cereb. Cortex* 10, 220–242.
- Dunlosky, J., and Metcalfe, J. (2009). *Metacognition* (Thousand Oaks: Sage).
- Flavell, J.H. (1979). *Am. Psychol.* 34, 906–911.
- Hampton, R.R. (2001). *Proc. Natl. Acad. Sci. USA* 98, 5359–5362.
- Kepecs, A., Uchida, N., Zariwala, H.A., and Mainen, Z.F. (2008). *Nature* 455, 227–231.
- Kiani, R., and Shadlen, M.N. (2009). *Science* 324, 759–764.
- Kikyo, H., Ohki, K., and Miyashita, Y. (2002). *Neuron* 36, 177–186.
- Kornell, N., Son, L.K., and Terrace, H.S. (2007). *Psychol. Sci.* 18, 64–71.
- Lynch, J.C., and Tian, J.R. (2006). *Prog. Brain Res.* 151, 461–501.
- Maril, A., Simons, J.S., Mitchell, J.P., Schwartz, B.L., and Schacter, D.L. (2003). *Neuroimage* 18, 827–836.
- Middlebrooks, P.G., and Sommer, M.A. (2011). *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 325–337.
- Middlebrooks, P.G., and Sommer, M.A. (2012). *Neuron* 75, this issue, 517–530.
- Roesch, M.R., and Olson, C.R. (2003). *J. Neurophysiol.* 90, 1766–1789.
- Schnyer, D.M., Verfaellie, M., Alexander, M.P., LaFleche, G., Nicholls, L., and Kaszniak, A.W. (2004). *Neuropsychologia* 42, 957–966.
- So, N.Y., and Stuphorn, V. (2010). *J. Neurophysiol.* 104, 2634–2653.

Losing the Lust for Life: A New Role for an Old Feeding Peptide?

Benjamin B. Land¹ and Ralph J. DiLeone^{1,*}

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

*Correspondence: ralph.dileone@yale.edu

<http://dx.doi.org/10.1016/j.neuron.2012.07.018>

A recent paper in *Nature* ([Lim et al., 2012](#)) describes the effects of melanocortin receptors in the nucleus accumbens. The studies connect a hypothalamic peptide system with brain reward centers and show effects on specific neuronal populations and behavioral components of mood.

Food and Mood

It is hard to imagine something more integrated with our mood state than eating. The influences go in both directions, with intake affecting mood and mood states modulating eating. For example, depression can lead to either increases or decreases in intake. As with all complex neuropsychiatric conditions, elucidation of basic neurobiological mechanisms is a critical first step toward clarifying just how the brain integrates eating with emotions. A recent study from Robert Malenka and colleagues published in

Nature identifies molecules, circuits, and neuronal pathways by which hypothalamic derived peptides can influence hedonic states ([Lim et al., 2012](#)). Specifically, the study establishes mechanisms by which stress can lead to reduced intake and anhedonia.

Melanocortins and Their Receptors—Taking a Hint from Metabolism

The melanocortin agonist, alpha-MSH, is derived from the precursor peptide POMC. The POMC neurons of the arcuate

nucleus form the “stop” side of the hypothalamic feeding equation whereby activation of this population reduces intake. The paraventricular nucleus of the hypothalamus has been best studied as a site where the melanocortin MC4 receptor (MC4R) mediates these effects. However, the MC4R is broadly expressed in the brain, including the nucleus accumbens and dorsal striatum. Early work showed regulation of MC4R by opiates and a role for striatal MC4R signaling in cocaine reward ([Alvaro et al., 2003](#); [Hsu et al., 2005](#)), and more recent studies have