Title: The supplementary motor area encodes reward expectancy in eye movement tasks

Authors: M. Campos[1], B. Breznen[2], K. Bernheim[2], and R. A. Andersen[1,2]

Departments: Computation and Neural Systems[1] and Division of Biology[2]

Institution: California Institute of Technology

Abbreviated Title: Reward expectancy expressed in SMA neurons

Correspondence: Michael Campos
California Institute of Technology
Computation and Neural Systems
MC 216-76
Pasadena, CA 91125
Phone: 626-395-8337
Fax: 626-795-2397
E-mail: mcampos@caltech.edu

Date: April 13, 2005

FINAL ACCEPTED VERSION

Abstract: Neural activity signifying the expectation of reward has been found recently in many parts of the brain, including midbrain and cortical structures. These signals can facilitate goal-directed behavior or the learning of new skills based on reinforcements. Here we show that neurons in the supplementary motor area (SMA), an area concerned with movements of the body and limbs, also carry a reward expectancy signal in the post-saccadic period of oculomotor tasks. While the monkeys performed blocks of memory-guided and object-based saccades, the neurons discharged a burst after a ~200 ms delay following the target acquiring saccade in the memory task, but often fired concurrently with the target acquiring saccade in the object task. The hypothesis that this post-saccadic bursting activity reflects the expectation of a reward was tested with a series of manipulations to the memory-guided saccade task. It was found that, while the timing of the bursting activity corresponds to a visual feedback stimulus, the visual feedback is not required for the neurons to discharge a burst. Second, blocks of no-reward trials reveal an extinction of the bursting activity as the monkeys come to understand that they would not be rewarded for properly generated saccades. Finally, the delivery of unexpected rewards confirmed that, in many of the neurons, the activity is not related to a motor plan to acquire the reward (e.g. licking). Thus we conclude that reward expectancy is represented by the activity of SMA neurons, even in the context of an oculomotor task. These results suggest that the reward expectancy signal is broadcast over a large extent of motor cortex, and may facilitate the learning of new, coordinated behavior between different body parts.

Keywords: FRONTAL CORTEX, REWARD, SACCADE

Introduction:

There has been substantial progress in recent years on the identification and characterization of the network of brain areas that are involved in the processing of reward. Reward expectancy signals have been found in many cortical areas such as the medialfrontal (Matsumoto et al. 2003; Shidara and Richmond 2002), dorsolateral prefrontal (Barraclough et al. 2004; Kobayashi et al. 2002), orbitofrontal (Hikosaka and Watanabe 2000; Tremblay and Schultz 2000), and parietal cortices (Musallam et al. 2004; Platt and Glimcher 1999; Sugrue et al. 2004). Subcortical regions expressing reward expectancy include the caudate (Watanabe et al. 2003), striatum (Cromwell and Schultz 2003; Hassani et al. 2001; Tremblay et al. 1998), superior colliculus (Ikeda and Hikosaka 2003), and midbrain dopamine neurons (Satoh et al. 2003; Schultz et al. 1997). Reward related signals have also been found in the dorsomedial frontal cortex (DMFC), an anatomical region that includes our current area of interest, the supplementary motor area (SMA).

The DMFC has been shown to participate in volitional (Schlag and Schlag-Rey 1985), or goal-oriented motor acts (Mann et al. 1988). It contains at least three well-studied motor-representation areas that are thought to be involved in higher order control of behavior: the SEF (Schlag and Schlag-Rey 1987, 1985), the SMA (Luppino et al. 1991; Matsuzaka et al. 1992) and the presupplementary motor area (pre-SMA) (Fujii et al. 2002; Nakamura et al. 1998; Shima and Tanji 2000). These three motor areas of DMFC can be distinguished based on anatomical connectivity (Luppino et al. 1993; Parthasarathy et al. 1992), and physiological responsivity (Matsuzaka et al. 1992). The SMA and pre-SMA are located in the DMFC on and above the medial wall in the frontal lobe. An orofacial region occupies the rostral end of the SMA, and further rostral is the pre-SMA. Intracortical microstimulation evokes movements in both areas, though the movements evoked in the pre-SMA require longer trains of pulses that produce more complex movements (Fujii et al. 2002). The pre-SMA has been implicated in planned motor acts (Matsuzaka and Tanji 1996), and the acquisition (Nakamura et al. 1998), planning and regulating (Shima and Tanji 2000) of sequential procedures. Additionally, pre-SMA neurons respond more often to visual stimulation compared to SMA neurons. The SMA consists of a rostrocaudal progression of orofacial, forelimb, and hindlimb movement representations (Mitz and Wise 1987). Lateral to the SMA, microstimulation will evoke eye movements. This area is defined as SEF (Fujii et al. 2002). Several studies have shown that electrical microstimulation at low currents (<50 µA, and sometimes as low at 10 µA) will elicit saccades in SEF (Chen and Wise 1995; Fujii et al. 1995; Mann et al. 1988; Russo and Bruce 1993; Tehovnik and Sommer 1996).

Three recent studies have explicitly connected the SEF to reward variables. Amador et al. discovered reward-predicting and reward-detecting neuronal activity in SEF (Amador et al. 2000). Schall and colleagues used the countermanding task to characterize three different types of neurons in the SEF – error, conflict, and reinforcement neurons, and suggested that these could serve a performance monitoring function (Stuphorn et al. 2000). Roesch and Olson found modulations of neural activity in response to both reward and punishment (Roesch and Olson 2003, 2004), and concluded that these modulations during the early stages of the trials correlate with the motivation and not reward expectation. In this study we present neural activity reflecting reward expectation during a later stage of the trials, specifically after the monkey

performs the instructed behavior. Our findings are similar to the reports of Amador and colleagues and Stuphorn and colleagues, however, the reward expectancy signal we describe is found in the SMA, while these other studies were recording from nearby SEF. Taken together, these results suggest that a reward expectancy signal may be present throughout the DMFC.

In this study we present evidence that a reward expectancy signal is expressed in the neural activity of the SMA during the performance of an oculomotor task. The signal is not related to the metric of the eye movement. Rather, it encodes expectation of the reward after the successful completion of the instructed behavior. The results presented here began as a discovery during a project that was originally intended to investigate the contribution of the SEF to saccades to objects. Some early recordings in the SMA uncovered a post-saccadic bursting activity that we hypothesized might be related to the expectation of reward, and experiments devised during the course of the project confirmed this hypothesis. While future studies of reward expectancy in SMA might use tasks that are specifically designed to investigate reward variables, with the two oculomotor tasks employed here we are able to establish two novel findings. First, we show that a reward expectancy signal is present in the SMA, an area that has been thought to be concerned only with movements of the body and limbs, during an oculomotor task. Second, we show the coupling of the signal's onset time with a secondary reinforcer. These findings suggest a general learning mechanism that would reinforce all motor representations in DMFC that are active just before the animal can expect to receive a reward. A preliminary account of this study has appeared previously (Campos et al. 2003).

Methods:

Studies were performed on two behaving, male rhesus monkeys (Macaca mulatta). Each was chronically fitted with a stainless steel head post for head immobilization, and a recording chamber over a craniotomy for electrode insertions. All procedures were approved by the Caltech Institutional Animal Care and Use Committee.

Stimuli and tasks. Monkeys were seated in a dimly lit room, 42 cm from a tangent screen. Stimuli were rear-projected with 800x600 resolution and a refresh rate of 72 Hz using a custom built software display client with OpenGL. Task logic was controlled by National Instruments real time LabView software.

Two eye movement tasks were used; a memory-guided saccade task and an object-based saccade task. In both tasks the monkey was instructed to perform a saccade from a central fixation point to one of 43 targets placed at regular intervals to cover the entire visual field out to 17 deg of visual angle in every direction from central fixation.

In the memory-guided saccade task (Figure 1a) monkeys were required to maintain central fixation while a peripheral target was briefly flashed, wait until the central fixation point extinguished, and then saccade to the remembered location. After successfully holding fixation at the target location, the target re-appeared to provide visual feedback of the correct eye position. The monkey then had to maintain fixation on the visible target for an additional interval of 250 before receiving a juice reward of about 0.2 ml.

In the object-based saccade task (Figure 1b), an object (isosceles triangle) was presented behind the central fixation point while the monkey fixated there. The object was cued for one of two possible locations on the object, and then, after a delay period, the object extinguished and reappeared at a peripheral location and new orientation. The monkeys were required to saccade to the previously cued part of the object in the new location and orientation. The cued locations of the object were chosen so that the correct saccade ended in the same screen location as the targets in the memory-guided task. After maintaining fixation on the cued part of the object for 250 ms, the monkeys were rewarded with a drop (about 0.2 ml) of juice.

The memory-guided and object-based saccade tasks were designed to investigate the neural computations supporting object-based saccades; however, the important difference between the tasks for the purposes of this study was actually what happened after the saccade was completed. In the object task the target was visible at the time of the saccade, and the monkey could acquire a visible target. In the memory task the target reappeared 250 ms after the saccade to the remembered location. Thus in the object task the animals received earlier feedback from a secondary reinforcement stimulus.

In a recording session a block of memory-guided saccades preceded a block of object based saccades. The memory-guided saccade block consisted of 3 correct saccades to each location. The object-based saccade block consisted of 12 correct saccades to each location. Control trials were performed during the memory-guided saccade block at the discretion of the experimenter.

Recording Procedure. Neurons were accessed on vertical penetrations with glass coated platinum-iridium electrodes (Fred Haer Co.). The electrodes were advanced with a Fred Haer or Narashige microdrive system, through a blunt stainless steel guide tube pressed against the dura. Neurons were generally found 1-3mm beneath the exterior of the dura.

Waveforms were amplified and isolated online with a commercial hardware and software package (Plexon Inc.). Cell activity was monitored with custom built online data visualization software written in Matlab.

Data Analysis. Bursting activity was identified using a burst detection algorithm (Hanes et al. 1995; Thompson et al. 1996). Bursts were initially detected by a threshold crossing of a surprise index (SI), which is the negative of the log of a calculated significance level. The significance level describes the likelihood that the observed number of spikes would occur in a given interval, considering the average firing rate of the cell, based on the assumption that the inter-spike intervals follow a Poisson distribution. The significance level used to calculate the threshold for the SI was 0.01. The mean of the Poisson distribution was calculated as the number of spikes in the trial divided by the duration of the trial. Since the mean can change from trial to trial, the algorithm assumes stationarity only over the duration of a single trial, and the threshold will adapt to changes in the baseline firing rate of the neuron over time. After the initial threshold crossing, the beginning and end of the burst were precisely identified, and multiple bursts could be identified in a single spike train (Thompson et al. 1996).

For ANOVA of firing activity in task intervals, the intervals were defined as follows. The baseline period was the interval between the acquisition of the fixation point and the cue appearance. The cue period was the interval that the cue was visible, and the memory period was the interval between the cue disappearance and the fixation point disappearance (the signal for the monkey to make the saccade). The saccade period was the 200 ms interval preceding the acquisition of the target, and the post-saccadic period was the interval from the target acquisition until the delivery of reward. All intervals were defined by these same events in both the memory and object-based tasks. The duration of the post-saccadic interval was 500 ms in the memory task, and 250 ms in the object-based saccade task.

Electrical Stimulation. A BAK instruments stimulator was used to deliver biphasic currents at 330 Hz of typically less than 200 µA in 100-500 ms trains through the recording electrodes.

Electromyography. Electromyography (EMG) recordings were performed in one monkey with a World Precision Instruments (DAM 80) AC/DC amplifier, and paired hook-wire electrodes (44 ga x 100 mm) from Viasys healthcare.

MR Imaging. Magnetic resonance (MR) imaging was performed at the Caltech Brain Imaging Center on a 3 T Siemens Trio. Anatomical images were acquired sagittally with 0.7 mm slice thickness using an in plane field of view of 168 x 168 mm on a 256 x 256 base matrix, yielding a final native voxel resolution of 0.656 x 0.656 x 0.7 mm. These images were realigned via multi-planar reformat to recording chamber landmarks using Siemens Syngo software (version MR 2003T DHHS.) This rotated volume was resliced at 0.7 mm spacing along the z-axis of the chamber and visualized using the AFNI software package (Cox 1996).

Results:

With a series of single electrode penetrations 173 cells were recorded in both tasks from two monkeys (monkey S: 100; R: 73). According to ANOVA of baseline firing rates versus the post-saccadic interval, 50 (S: 34; R: 16) neurons demonstrated a significant modulation in the post-saccadic interval, with 17 (S: 9; R: 8) of these modulated in the post-saccadic interval exclusively. Many of the neurons were also active during task periods. According to ANOVA of baseline firing rates versus cue, memory, and saccade intervals 84 (S: 55; R: 29) of the recorded cells were significantly ($p < 10^{-5}$) modulated during at least one of these intervals in both tasks. A breakdown of neurons with significant modulations for the individual periods of the memory saccade task (cue: 23, memory: 51, saccade: 71), show that there was substantial activity present in all task intervals, however this activity was generally not spatially tuned (see Table 2 and discussion below). Summary cell count information is provided in Table 1, along with results of control experiments.

Anatomic localization of the recording sites. The sites of all of the electrode penetrations included in this study are superimposed on axial MRI scans in Figure 2 (a,b). While recordings were taken on the surface of cortex, MRI sections for anatomical localization were chosen at a depth appropriate to clearly show the locations of the penetrations relative to surrounding sulci.

The sites which yielded the 50 neurons with significant (ANOVA, $p<10^{-5}$, see above) post-saccadic modulations are shown in red, and the remaining sites are shown in blue. Not all neurons recorded at the sites marked in red were modulated in the post-saccadic period. The red marker only indicates that at least one of these 50 neurons of interest was recorded at that site.

While much of SMA is in F3 on the mesial surface, there is also a portion of F3 on the dorsal surface, within about 3 mm of the midline that is also considered SMA proper (Luppino et al. 1991; Matsuzaka et al. 1992). The neurons of interest in this report, indicated in red on the axial slices in Figure 2, mostly cluster within this distance to the left (monkey's right) of the midline for monkey S, and to the right (monkey's left) of the midline for monkey R. No recordings were performed in SEF. In both monkeys some of the recordings were in area F2, lateral to SMA-proper (Luppino et al. 1991). In monkey S the majority of the recordings were directly medial to the genu of the arcuate sulcus, while in monkey R the recordings were medial and somewhat posterior. The SEF is medial to the arcuate sulcus and somewhat anterior, though there is some variability in the precise location of SEF as described in previous studies. See Sommer and Tehovnik 1999 for review.

Microstimulation. Electrical stimulation experiments show the progression of body movement responses typical of the SMA (Mitz and Wise 1987). Since eye movements were not observed to be elicited in either of the monkeys by stimulation of 50 µA, which is the upper limit of the low threshold criterion for eliciting eye movements in the SEF (Russo and Bruce 1993), or even currents as high as 200 µA, the recordings were not in the oculomotor area SEF.

Population characteristics. The average spike activity recorded during memory saccades from all sites for each monkey is summarized in Figure 2 (c,d). The average firing rate aligned to the target acquire event (end of saccade) is shown. The activity from monkey R (Figure 2d) is exclusively post-saccadic. In monkey S saccadic and memory period activity was also observed (Figure 2c), which could be due to the more anterior placement of the chamber. In monkey S the dominant peak of activity still occurs after a delay following the target acquire event.

While there were many neurons with modulated activity during different epochs of the task (see Table 2), very few were spatially tuned. A 2-way ANOVA between baseline firing rates, and 1) the firing rates from different task intervals, and 2) the spatial locations of the targets, was used to confirm this observation. A very small number of neurons passed the significance test ($p < 10^{-3}$) for dependence of firing rate on task interval and target location (cue: 1; memory: 7; saccade: 6).

A shift in burst onset times. The post-saccadic burst in both trial types (a,b – memory, c,d – object) for one of these neurons in Figure 3 is illustrated with raster plots of spike traces aligned to the target acquire event (a,c) and the reward delivery (b,d). Bursts of activity identified with the burst detection algorithm (methods) are shown as horizontal blue lines beneath the spike trains. The bursts in the object task (c) begin at a time that could be related to saccade generation. However, the bursts of activity in the memory task (a) come substantially later, revealing that these bursts do not participate in the generation of a saccade, or at least not in the context of the memory saccade task. For this neuron the post-saccadic firing terminates with reward delivery. Other neurons (see Figure 5 for example) were also observed to terminate just before or soon after reward delivery.

In Figure 4a histograms for the time to burst, relative to the target acquire event in each trial type are shown for the recording presented in Figure 3. There is a clear separation of these two groups (ANOVA, $p << 10^{-5}$). The mean bursting times relative to the target acquire event is 105 ms in the object task, and 537 ms in the memory task. The bursts in both tasks terminate with the delivery of the reward after successful completion of the task.

The difference of the mean time to burst in the memory task and the object task for the population of neurons that discharged a burst in at least 30% of the trials in both tasks (N = 30) is plotted as a histogram in Figure 4b. In general, the bursting activity came later, relative to the target acquire event, in the memory task compared to the object-based task. Neurons in this category showed a mean shift in the onset time of the burst of 202 ms. This number is comparable to, though slightly less than, the amount of the time the animal was required to fixate the remembered target location in the memory task before the reappearance of the target (250 msec).

The onset time of the burst corresponded to the appearance of the visual feedback, which was immediate in the object-based task, but delayed in the memory-guided saccade task. In both cases the visual feedback could serve as a predictor of a reward. The hypothesis that bursting activity reflects an expectation of reward was then tested in a series of control experiments outlined below.

Bursting does not accompany non-rewarded target acquisitions. It could be argued that the neurons simply signal the acquisition of any target, regardless of the expectation of

reward. We tested this possibility by comparing the activity of the neurons after the initial acquisition of the fixation point with the activity after the reappearance of the target in the memory-guided saccade task. We used ANOVA on two intervals: the first interval was between the fixation acquire event and the appearance of the cue (250 ms), and the second interval was between the target onset and the reward delivery (250 ms). This analysis reveals that of the 50 neurons with a significant post-saccadic modulation, none were significantly active during the initial acquisition of the fixation point.

Bursting is not a visual response. Control trials of the memory-guided saccade task in which the visual feedback was withheld were run to test whether or not the bursting activity is related to the visual feedback signal. As shown in Figure 3 and again in Figure 5, removal of the visual feedback (indicated in the figures with the green bar composed of green stars) does not eliminate the onset of the bursting activity, though it may reduce the intensity, or vary the onset time. Figure 5a shows an example in which the post-saccadic bursting activity was slightly extended by this control, though otherwise unchanged.

The bursting signal is therefore not indicating the reappearance of the target, though the visual reinforcement serves to sharpen and intensify the neural discharge. This control was run on 34 neurons, and 12 of them showed no significant difference in the mean firing rate from the time the target appeared (or should have appeared) until the end of the trial (ANOVA, $p < 10^{-5}$) in control vs. normal trials. Of the remaining neurons, many exhibited a temporal shift in their active periods, or a decrease in firing, but only 1 showed an extinction of the bursting activity. This control shows that visual feedback could be dissociated from the reward delivery, and the neural response remained.

Bursting properties in the absence of reward. To see if, all else being equal, the absence of reward would have an effect on the neural activity, we occasionally withheld the reward for a block of trials during the memory-guided saccade task, even for correctly performed trials. In the no reward blocks, the monkeys generally continued to correctly perform the task for about thirty trials before stopping, and this comprehension of changing task conditions was reflected in the recorded neural activity. After a few trials, bursting activity would stop altogether. This control was run while recording 11 neurons, and all of these showed a significant difference in the mean firing rate from the target acquire event until the end of the trial (ANOVA, $p < 10^{-5}$) in control vs. normal trials. The vast majority (10) ceased firing during the pre-reward interval in the no-reward blocks, and the remaining neuron (of the 11 that were modulated) increased its firing rate after the reward period. An example neuron is shown in Figure 5b. The firing activity is gradually extinguished in the no reward block (black bar). In contrast, the activity during the no-visual feedback trials (green bar) has a less precise onset time, but does not extinguish. While the no-visual feedback trials show the effect of removing a predictor of reward, the no-reward blocks reveal the dynamic effects of the monkeys coming to understand that they should no longer expect a reward.

Unexpected reward trials. By removing the reward, the possibility that the bursting activity encoded an orofacial (e.g. licking) motor response was not eliminated. Every time the monkey expected a reward, he would presumably also plan a licking movement

to acquire it. The monkey would often stop licking the juice tube in the blocks of trials in which the reward was turned off, and this would correspond to the termination of the bursting activity. Of course, if the monkey no longer expected to be rewarded for the eye movements, he also had no reason to lick the juice tube.

To dissociate licking movements from the expectation of reward, control trials were run in which an unexpected reward was delivered. A bonus reward would be delivered with a 5% probability at the end of the fixation interval, just before the cue presentation. To quantify a response, an ANOVA ($p < 10^{-5}$) compared firing rates in the 200 msec interval during the bonus reward delivery, with the corresponding 200 msec period at the end of the fixation interval in normal trials. This first interval is the actual interval that the valve regulating the flow of reward was open. While running this control 25 neurons with reward related activity were recorded, and 23 of these demonstrated no correlated activity in the unexpected juice delivery period. This control shows that the majority of the recorded neurons are not responsive to rewards when they are not expected, ruling out the possibility that the neural activity is attributable to motor commands required to obtain the reward, such as licking and swallowing.

To address the possibility that the neural activity reflected monkey's postural responses or attempted postural responses, or preparation for either, we recorded muscle activity in three muscles active during postural adjustments. We recorded EMG (see Methods) from left and right *Latissimus dorsi* and right *Semitendinosus* of Monkey S during the performance of both tasks. We observed that these muscles were active during trunk movements and leg movements. While there was activity recorded from these muscle groups during the execution of the task, we found that it was not temporally locked to reward expectation. These negative EMG results rule out the possibility that the monkey is consistently making postural adjustments in anticipation of the reward delivery.

Discussion:
Bursting activity related to reward expectation was found in a cortical area that is not directly responsible for the generation of the behavior (saccade) that achieves the reward. This report separates itself from other reports by (1) investigating the reward expectation activity in the motor area SMA, and (2) showing a shift of activity in time course based on a secondary reinforcer, the visual feedback that usually predicts the upcoming reward. Below we outline and justify our findings, compare our results with the results of other studies, and suggest a functional role for the representation of reward expectancy in SMA during eye movement tasks.

Onset of reward-expectancy signal corresponds to a secondary reinforcer. The activity generally started with the secondary reinforcer and stopped with the delivery of the reward. The secondary reinforcer in this context was visual feedback that occurred before reward delivery. In the memory task the target reappears after 250 ms of fixation on the remembered target location. This visual feedback helped ensure accuracy in the initial learning of the task, but also became a *predictor* of the upcoming reward. In the object task the saccade target is visible, and so the monkey can be sure that he made a saccade to the target because he can see it. The onset time of the reward expectancy signal corresponds to the onset of the visual feedback in the tasks, either 250 ms after the correct saccade in the memory task, or immediately during the correct saccade in the object task.

As shown in control experiments, while the secondary reinforcer helped to synchronize the timing of the bursting activity, it was not necessary for the neurons to burst. This and other controls discussed below confirm that the bursting discharge carries a reward expectancy signal.

Control experiments establish the reward expectancy interpretation. In a series of control experiments the argument was built that this activity reflects an expectation of reward. First, the bursting activity was dissociated from visual feedback, with the demonstration that visual feedback is not required for the neurons to discharge, though it regularizes the timing of the onset. Second, when the reward was removed for a block of trials, the reward expectancy activity gradually disappeared, showing that this activity represents a dynamic variable corresponding to the comprehension of a changed task condition. Finally, the possibility that the neural trace signified a licking plan or a detection of reward was ruled out since there was generally no response to unexpected reward delivery.

Reward related activity in the supplementary eye fields. Reward related neural signals have already been described in the SEF (Amador et al. 2000; Roesch and Olson 2003; Stuphorn et al. 2000). We found a reward expectancy signal in the SMA that appears very similar to types of activity found by Amador and colleagues and Stuphorn and colleagues.

1) Reward expectancy and reward prediction. Reward prediction (RP) neurons have been described in SEF along with a set of complementary reward detecting (RD) neurons

(Amador et al. 2000). The neural activity we are describing as reflecting reward expectation is similar to the RP neural activity. We did not find evidence of RD activity.

The firing rates of RP neurons increase before the occurrence of a reward, then abruptly cease firing at reward delivery, just as we found in the bursting activity of many of the neurons in this study. Our results, combined with the results of Amador and colleagues, are therefore evidence that reward expectation can be found in both SMA and SEF, and likely throughout the DMFC. Our study adds to the findings of RP neurons by (1) recording neural responses during the unexpected delivery of reward and (2) submitting the monkey to short blocks of no-reward trials.

We choose to use the term reward expectancy, since "expectancy" captures the way the neural activity continues until reward delivery. Furthermore, this designation separates itself from reward prediction nomenclature found in the dopamine neuron literature. To predict is to foretell on the basis of experience, while to expect is to await or look forward to the coming or occurrence. The reward prediction signal found in midbrain dopamine neurons and the reward expectancy signal in the DMFC likely play different roles in learning and goal-oriented behavior (see below).

2) Reward expectancy and reinforcement. Reinforcement signals have been found in SEF using a countermanding saccade task (Stuphorn et al. 2000). The reinforcement neurons were shown to increase activation while awaiting reward. The term reward expectancy describes the function of this activity. Again, the results of this study are evidence that the reward expectation signal found in the SMA is also present in the SEF.

3) Reward expectancy vs. enhanced motivation. The post-saccadic burst cannot be a correlate of motivation (Roesch and Olson 2003, 2004) simply because it comes after the behavior it would presumably motivate. Our use of the burst detection algorithm (Hanes et al. 1995; Thompson et al. 1996) establishes that this bursting activity comes in the post-saccadic interval.

In the Roesch and Olson study, the preferred direction of a neuron was first identified, and then a memory-guided saccade task was run to and away from the preferred direction of the cell. Since we rarely found neurons to be spatially tuned, we may have been recording from different types of neurons in the SMA.

Interestingly, the authors noted that in areas in which reward effects were common, such as the SMAr, neurons "fired more strongly than reward-insensitive neurons during the period extending from the completion of the saccade to delivery of the ingested reward." The authors did not think their paradigm capable of distinguishing between various interpretations of the significance of this effect, such as preparation and execution of liking movements, or increased intensity of reward anticipation. In the present study, our control experiments show that the post-saccadic activity in SMA reflects reward expectation.

Reward expectancy vs. attention. The expected value of a reward has been shown to modulate activity during the performance of a task (Ikeda and Hikosaka 2003; Platt and Glimcher 1999; Shidara and Richmond 2002; Watanabe 1996). It has been argued that so far sufficient controls for these studies have not been performed to determine whether the cognitive state being manipulated is expected value or attention, since these two states

likely occur together and can be easily confounded. Likewise, studies examining attention may have recorded the effects of expected value (Maunsell 2004).

In the current study it is unlikely that the reward expectancy signal is actually an attention signal. The signal occurs after the task and is not spatially tuned and thus cannot reflect attention to the saccade location. It also cannot reflect attention to the reward since there was no activity when the reward was presented unexpectedly – novelty is a powerful attractor of attention.

Reward signals and reinforcement learning algorithms. A reinforcement learning algorithm (Sutton and Barto 1988) has been proposed to account for different reward-related signals that have been found in the brain, such as the error of reward prediction in midbrain dopamine neurons (Schultz et al. 1997). The prediction error signal is widely recognized as evidence for the implementation of a reinforcement learning algorithm in the brain. For example, the prediction error can serve to update action value estimates, so that the animal can have accurate estimates of the reward that can be expected for an action. In this formalism, the expected value, V, is updated after every trial according to the experienced reward by the equation:

$$V^{t+1} = V^t + \alpha(R^t - V^t)$$

Where R^t is the amount of reward obtained at time t, V^t is the amount of reward expected at time t, α is the learning rate, and V^{t+1} is the updated estimate of expected value at time t+1. In this formulation the time steps are individual trials, and the signal that corresponds to the error of reward prediction found in dopamine neurons (Schultz et al. 1997) is the term in the parentheses, $R^t - V^t$. The action value signal, V, that we are describing would not be used instead of a prediction error signal, R-V. Rather, both signals are supposed components of a larger reinforcement learning mechanism.

The dynamics of the reward expectancy signal in SMA corresponds to the dynamics of the expected value of the action, V. Specifically, this algorithm captures the way the post-saccadic firing activity in the SMA gradually dissipates in no-reward blocks. When the reward, R, is zero for a series of trials, the equation above will diminish the expected value of the action, V, until it reaches the new value of the reward, 0. The learning rate parameter, α, determines how quickly the estimate of the expected value approaches the new value. This equation also describes how the neural activity will return to normal firing when the reward is again delivered as usual.

Functional significance of reward expectancy in DMFC -- a signal to guide learning. As in neural network models of reinforcement learning (Mazzoni et al. 1991; Suri and Schultz 1999), the reward signal found in DMFC could be used to train other parts of cortex to perform visuospatial tasks requiring arbitrary sensorimotor transformations. The reward expectancy signal found in the SEF (Amador et al. 2000; Stuphorn et al. 2000) is in position to shape future oculomotor behavior through its connections with the frontal eye fields (FEF) (Schall et al. 1993) and the superior colliculus (SC) (Fries 1984).

A reward expectancy signal is better than the detection of the reward itself for training purposes for two reasons. First, the reward expectancy signal implies that there is an internal model with an expected sensory outcome for a behavior, in this case the

secondary reinforcer. This model can be matched with a reward signal, and refined as often as rewards are delivered or unexpectedly withheld. Second, the reward expectancy signal comes at a time that is more proximal to the behaviors which earned the reward, and thus may be able to reinforce the high level motor signals in DMFC related to those behaviors.

Usefulness of reward expectancy in SMA during an eye movement task. A reward expectancy signal present in DMFC could maintain and enhance the high-level representations of behaviors that earn a reward. But why would this activity be present in the SMA, which is concerned with movements of the body and limbs, during an eye movement task? It is possible that the reward expectancy signal is maintained throughout DMFC so that it can enhance *any* volitional motor acts that precede reward. For instance, in hand eye coordination tasks the reward signal can reinforce activity in the limb area of SMA and SEF together. Other areas of SMA which do not have a convergence of activation of the motor map activation and reward signal would not be reinforced and would not produce learning (Sutton and Barto 1988). Thus, the expectation signal may be more widely broadcast than the motor activations of a particular behavior. This broader signal may serve to learn new coordinations of different body parts for particular tasks.

Figure Legends:
Figure 1. Time course of oculomotor tasks. Progression of tasks are shown in successive panels from the top left to bottom right. In the memory-guided saccade task **(a)**, the monkey is required to acquire a central fixation point at the start of the trial. After a variable delay, a cue is briefly flashed at one of 43 targets in the periphery. The possible targets cover the entire visual field out to 17 degrees. Following a hold interval the fixation point is extinguished and the monkey is required to saccade to the remembered target location and fixate there. After 250 ms the target reappears, and then following an additional 250 ms fixation, the animal is rewarded with a drop of juice. In the object-based saccade task **(b)**, the monkey begins the trial by acquiring a central fixation point. An object appears over the fixation point, and after a delay one side of the object is briefly cued. Following a hold period the object is extinguished and immediately reappears in a new location. The monkey is then required to saccade to the cued portion of the object and fixate there for 250 ms, before receiving a juice reward.
Figure 2. Sites of neural recording. Projections of chamber walls are indicated with a blue circle superimposed on axial MRI scans of Monkey S **(a)** and R **(b)**. Anatomical landmarks are of the arcuate sulci (AS), principle sulci (PS) and central sulcus (CS). Recording sites that yielded reward interval activity are shown as red dots, and the remaining recording sites are blue. Averaged output of all recorded neurons **(c, d)** shows the average firing rate for all recorded neurons for each monkey aligned on the target acquire event of memory saccade trials.
Figure 3. Shift in burst onset times. Peri-event time histograms in memory **(a,b)** and object **(c,d)** saccades tasks. Spikes are represented in red, aligned to the target acquire event **(a,c)**, and the reward delivery event **(c,d)**. The smoothed average firing rate for normal trials is plotted as a blue curve, for no-visual feedback trials in green. Horizontal blue lines indicate periods of burst activity for the spike trains above. Green stars forming a bar on the right edge of the panels indicate trials in which visual feedback was withheld. Cyan markers indicate the reappearance of the target.
Figure 4. Histograms of burst onset times. **a)** Distribution of time to burst for each trial of the memory (filled) and object (open) saccade tasks for the cell shown in Figure 3. Mean time to burst is: memory task: 537 ms, object task 105 ms. Memory task data clusters to the right of the object task data. **b)** Differences in time to burst onset after the target acquire event in memory vs. object tasks. Only cells with bursts in at least 30% of the trials in both tasks are shown (N=30). Mean difference: 202 ms. The population of average shift times is significantly different from 0 (t-test, $p < 10^{-5}$).
Figure 5. Example response to control trials. **(a)** Withheld visual feedback control trials. All trials shown are from the memory-saccade task. Green stars forming a bar on the edge of the panel indicate trials in which visual feedback was withheld. Smoothed average firing rates for normal trials are drawn in blue, and can be compared with the average firing rates during the withheld feedback trials drawn in green. **(b)** Withheld reward block of trials. All trials shown are from the memory-saccade task. Black stars forming a bar slightly inset from the right edge of the panel indicate the successfully completed trials in which the reward was not delivered. Green stars forming a bar on the

right edge of the panel indicate trials in which visual feedback was withheld. Only the spike trains from successfully completed trials are shown. Trials are arranged chronologically from top to bottom. Smoothed average firing rates for normal trials are drawn in blue, and can be compared with the average firing rates during the withheld reward trials shown in black, and the withheld feedback trials drawn in green. Cyan markers indicate the reappearance of the target.

Table 1. Total number of recorded neurons for both monkeys. All statistical tests are ANOVA ($p < 10^{-5}$). Task related modulation indicates a modulation in one or more of cue, memory, saccade, or reward intervals relative to baseline. Reward period activity compares the baseline to reward interval; exclusive reward activity excludes neurons with significant cue or memory period activity. Control categories compare a time interval of interest (see text) associated with the control, in normal and control trials. Numbers in parentheses are the number of cells recorded while each control was run, and the numbers beside them are the number of cells with a significant modulation in firing rates according to ANOVA analyses described in the text.

Table 2. Cell counts for spatial tuning properties of recorded neurons. Active above baseline refers to a significant modulation according to an ANOVA test ($p < 10^{-5}$) comparing baseline firing rates to firing rates in cue, memory, or saccade intervals. Significant spatial and interval dependency indicates significant dependence of firing rates (2-way ANOVA, $p < 10^{-3}$) on the task interval, the spatial location of the targets, and the interaction between these two factors. All tests refer to data collected in memory saccade trials.

References:

1. **Amador N, Schlag-Rey M and Schlag J**. Reward-Predicting and Reward-Detecting Neuronal Activity in the Primate Supplementary Eye Field. *J Neurophysiol* 84: 2166-2170, 2000.

2. **Barraclough DJ, Conroy ML and Lee D**. Prefrontal cortex and decision making in a mixed-strategy game. 7: 404-410, 2004.

3. **Campos M, Breznen B and Andersen RA**. Reward expectancy in dorsomedial frontal cortex of the macaque monkey. *Soc. Neurosci. Abstr.*: 187.183, 2003.

4. **Chen L and Wise S**. Neuronal activity in the supplementary eye field during acquisition of conditional oculomotor associations. *J Neurophysiol.* 73: 1101-1121, 1995.

5. **Cox RW**. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research* **29**: **162-173**, 1996.

6. **Cromwell HC and Schultz W**. Effects of Expectations for Different Reward Magnitudes on Neuronal Activity in Primate Striatum. *J Neurophysiol* 89: 2823-2838, 2003.

7. **Fries W**. Cortical projections to the superior colliculus in the macaque monkey: A retrograde study using horseradish peroxidase. *J. Comp. Neurol.* 230: 55-76, 1984.

8. **Fujii N, Mushiake H, Tamai M and J T**. Microstimulation of the supplementary eye field during saccade preparation. *Neuroreport.* 6: 2565-2568, 1995.

9. **Fujii N, Mushiake H and Tanji J**. Distribution of Eye- and Arm-Movement-Related Neuronal Activity in the SEF and in the SMA and Pre-SMA of Monkeys. *J Neurophysiol* 87: 2158-2166, 2002.

10. **Hanes D, Thompson K and Schall J**. Relationship of presaccadic activity in frontal eye field and supplementary eye field to saccade initiation in macaque: Poisson spike train analysis. *Exp Brain Res.* 103: 85-96, 1995.

11. **Hassani OK, Cromwell HC and Schultz W**. Influence of Expectation of Different Rewards on Behavior-Related Neuronal Activity in the Striatum. *J Neurophysiol* 85: 2477-2489, 2001.

12. **Hikosaka K and Watanabe M**. Delay Activity of Orbital and Lateral Prefrontal Neurons of the Monkey Varying with Different Rewards. *Cereb. Cortex* 10: 263-271, 2000.

13. **Ikeda T and Hikosaka O**. Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron* 39: 693-700, 2003.

14. **Kobayashi S, Lauwereyns J, Koizumi M, Sakagami M and Hikosaka O**. Influence of Reward Expectation on Visuospatial Processing in Macaque Lateral Prefrontal Cortex. *J Neurophysiol* 87: 1488-1498, 2002.

15. **Luppino G, Matelli M, Camarda R and G R**. Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *J Comp Neurol.* 338: 114-140, 1993.

16. **Luppino G, Matelli M, Camarda R, Gallese V and Rizzolatti G**. Multiple representations of body movements in mesial area 6 and the adjacent cingulate cortex: an intracortical microstimulation study in the macaque monkey. *J Comp Neurol.* 311: 463-482, 1991.

17. **Mann S, Thau R and Schiller P**. Conditional task-related responses in monkey dorsomedial frontal cortex. *Exp Brain Res.* 69: 460-468, 1988.

18. **Matsumoto K, Suzuki W and Tanaka K**. Neuronal Correlates of Goal-Based Motor Selection in the Prefrontal Cortex. *Science* 301: 229-232, 2003.

19. **Matsuzaka Y, Aizawa H and Tanji J**. A motor area rostral to the supplementary motor area (presupplementary motor area) in the monkey: neuronal activity during a learned motor task. *J Neurophysiol.* 68: 653-662, 1992.

20. **Matsuzaka Y and Tanji J**. Changing directions of forthcoming arm movements: neuronal activity in the presupplementary and supplementary motor area of monkey cerebral cortex. *J Neurophysiol.* 76: 2327-2342, 1996.

21. **Maunsell JHR**. Neuronal representations of cognitive state: reward or attention? *Trends in Cognitive Sciences* 8: 261-265, 2004.

22. **Mazzoni P, Andersen RA and Jordan MI**. A more biologically plausible learning rule that backpropagation applied to a network model of cortical area 7a. *Cereb. Cortex* 1: 293-307, 1991.

23. **Mitz A and Wise S**. The somatotopic organization of the supplementary motor area: intracortical microstimulation mapping. *J Neurosci.* 7: 1010-1021, 1987.

24. **Musallam S, Corneil BD, Greger B, Scherberger H and Andersen RA**. Cognitive Control Signals for Neural Prosthetics. *Science* 305: 258-262, 2004.

25. **Nakamura K, Sakai K and Hikosaka O**. Neuronal Activity in Medial Frontal Cortex During Learning of Sequential Procedures. *J Neurophysiol* 80: 2671-2687, 1998.

26. **Parthasarathy H, Schall J and Graybiel A**. Distributed but convergent ordering of corticostriatal projections: analysis of the frontal eye field and the supplementary eye field in the macaque monkey. *J Neurosci.* 12: 4468-4488, 1992.

27. **Platt ML and Glimcher PW**. Neural correlates of decision variables in parietal cortex. *Nature* 400: 233-238, 1999.

28. **Roesch MR and Olson CR**. Impact of Expected Reward on Neuronal Activity in Prefrontal Cortex, Frontal and Supplementary Eye Fields and Premotor Cortex. *J Neurophysiol* 90: 1766-1789, 2003.

29. **Roesch MR and Olson CR**. Neuronal Activity Related to Reward Value and Motivation in Primate Frontal Cortex. *Science* 304: 307-310, 2004.

30. **Russo G and Bruce C**. Effect of eye position within the orbit on electrically elicited saccadic eye movements: a comparison of the macaque monkey's frontal and supplementary eye fields. *J Neurophysiol.* 69: 800-818, 1993.

31. **Satoh T, Nakai S, Sato T and Kimura M**. Correlated Coding of Motivation and Outcome of Decision by Dopamine Neurons. *J. Neurosci.* 23: 9913-9923, 2003.

32. **Schall JD, Morel A and Kaas JH**. Topography of supplementary eye field afferents to frontal eye field in macaque: Implications for mapping between saccade coordinate systems. *Visual Neurosci.* 10: 385-393, 1993.

33. **Schlag J and Schlag-Rey M**. Evidence for a supplementary eye field. *J Neurophysiol* 57: 179-200, 1987.

34. **Schlag J and Schlag-Rey M**. Unit activity related to spontaneous saccades in frontal dorsomedial cortex of monkey. *Exp Brain Res.* 58: 208-211, 1985.

35. **Schultz W, Dayan P and Montague PR**. A Neural Substrate of Prediction and Reward. *Science* 275: 1593-1599, 1997.

36. **Shidara M and Richmond BJ**. Anterior Cingulate: Single Neuronal Signals Related to Degree of Reward Expectancy. *Science* 296: 1709-1711, 2002.

37. **Shima K and Tanji J**. Neuronal Activity in the Supplementary and Presupplementary Motor Areas for Temporal Organization of Multiple Movements. *J Neurophysiol* 84: 2148-2160, 2000.

38. **Stuphorn V, Taylor T and Schall J**. Performance monitoring by the supplementary eye field. *Nature* 408: 857-860, 2000.

39. **Sugrue LP, Corrado GS and Newsome WT**. Matching Behavior and the Representation of Value in the Parietal Cortex. *Science* 304: 1782-1787, 2004.

40. **Suri RE and Schultz W**. A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience* 91: 871-890, 1999.

41. **Sutton RS and Barto AG**. *Reinforcement learning : an introduction*. Cambridge, Mass: MIT Press, 1988.

42. **Tehovnik E and Sommer M**. Compensatory saccades made to remembered targets following orbital displacement by electrically stimulating the dorsomedial frontal cortex or frontal eye fields of primates. *Brain Res.* 727: 221-224, 1996.

43. **Thompson K, Hanes D, Bichot N and Schall J**. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J Neurophysiol.* 76: 4040-4055, 1996.

44. **Tremblay L, Hollerman JR and Schultz W**. Modifications of Reward Expectation-Related Neuronal Activity During Learning in Primate Striatum. *J Neurophysiol* 80: 964-977, 1998.

45. **Tremblay L and Schultz W**. Reward-Related Neuronal Activity During Go-Nogo Task Performance in Primate Orbitofrontal Cortex. *J Neurophysiol* 83: 1864-1876, 2000.

46. **Watanabe K, Lauwereyns J and Hikosaka O**. Neural Correlates of Rewarded and Unrewarded Eye Movements in the Primate Caudate Nucleus. *J. Neurosci.* 23: 10052-10057, 2003.

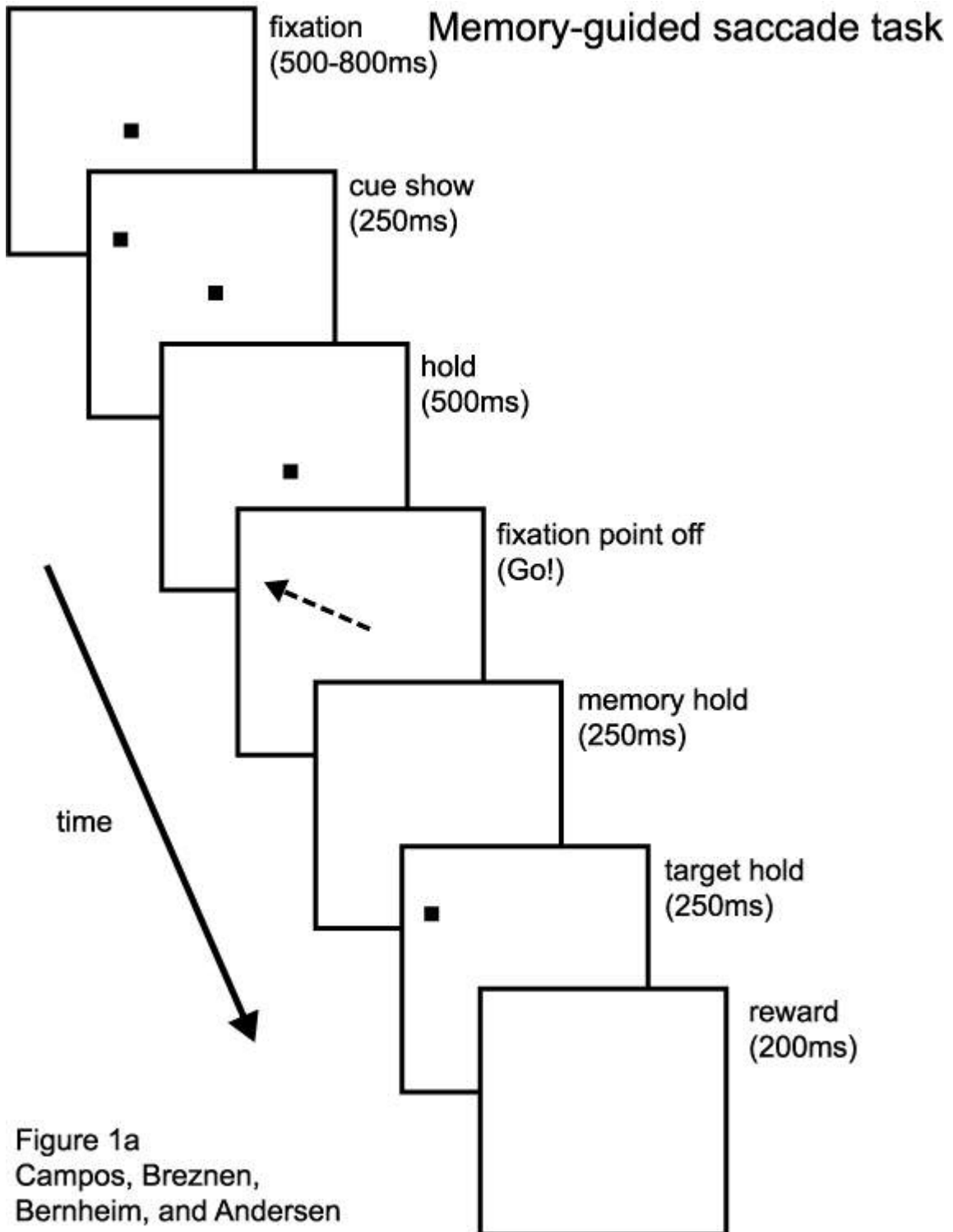47. **Watanabe M**. Reward expectancy in primate prefrontal neurons. *Nature* 382: 629-632, 1996.

fixation
(500-800ms)

Memory-guided saccade task

cue show
(250ms)

hold
(500ms)

fixation point off
(Go!)

memory hold
(250ms)

time

target hold
(250ms)

reward
(200ms)

Figure 1a
Campos, Breznen,
Bernheim, and Andersen

Object-based saccade task

fixation
(500-800ms)

object show
(500ms)

8deg

cue L/R
(500ms)

hold
(500-800ms)

fixation off, target on
(Go!)

time

hold
(250ms)

reward
(250ms)
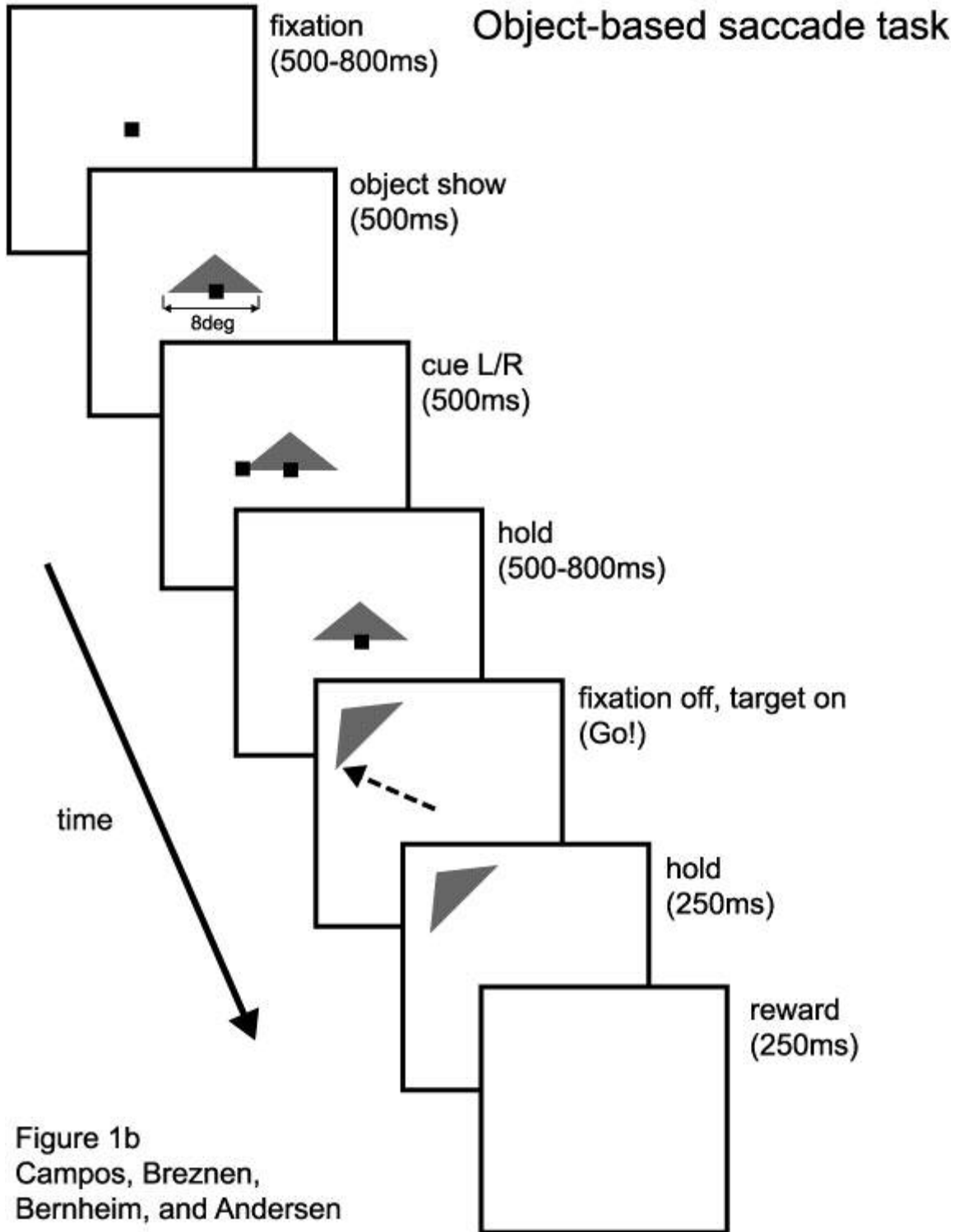
Figure 1b
Campos, Breznen,
Bernheim, and Andersen

Monkey S

a
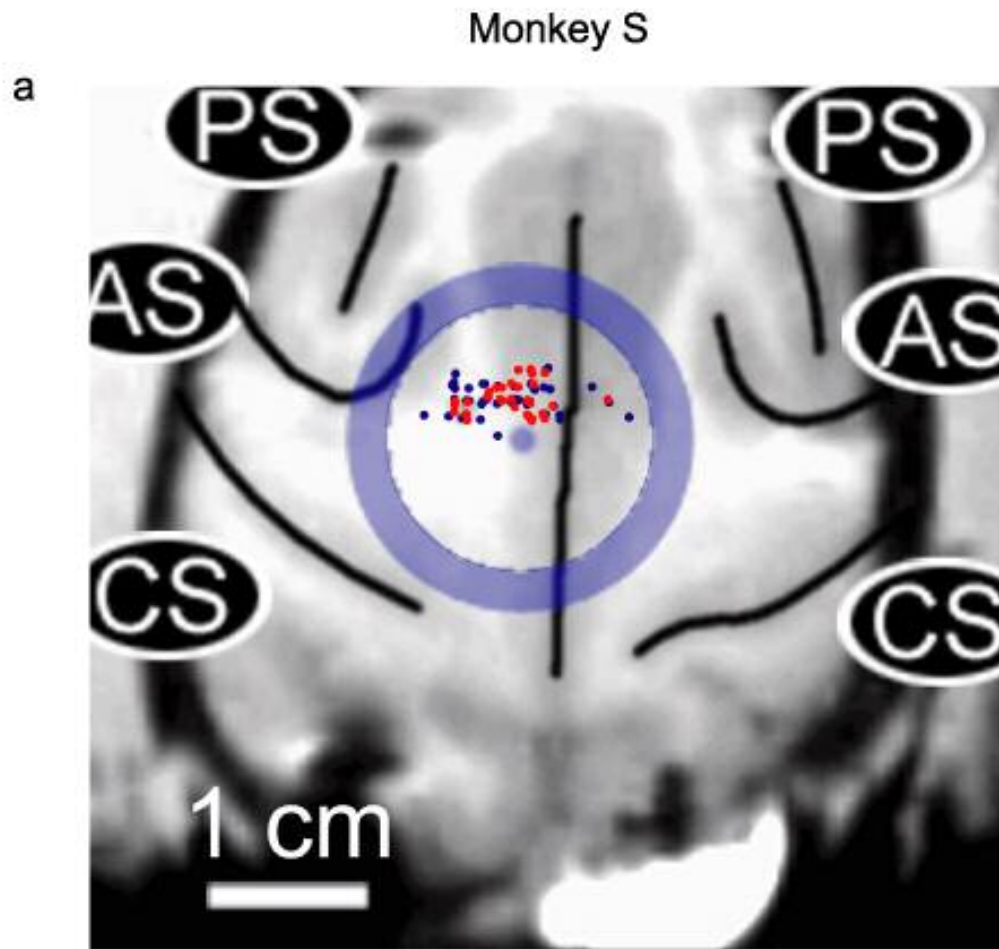
Figure 2 a
Campos, Breznen, Bernheim, and Andersen

Monkey R
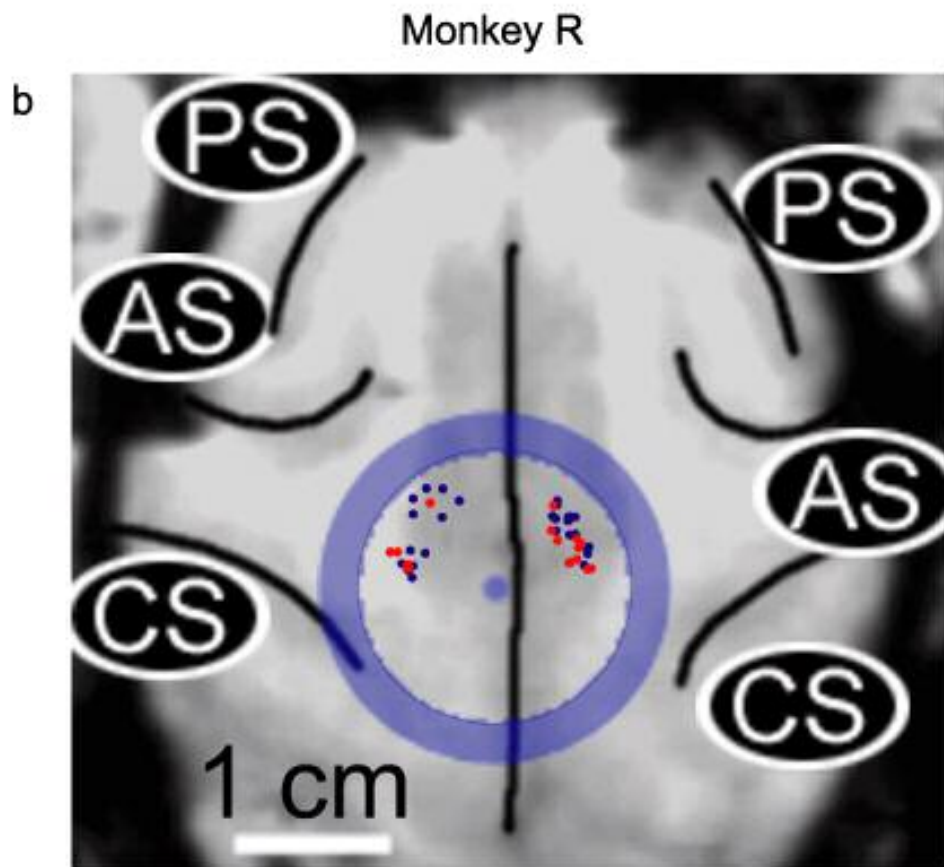
Figure 2 b
Campos, Breznen, Bernheim, and Andersen
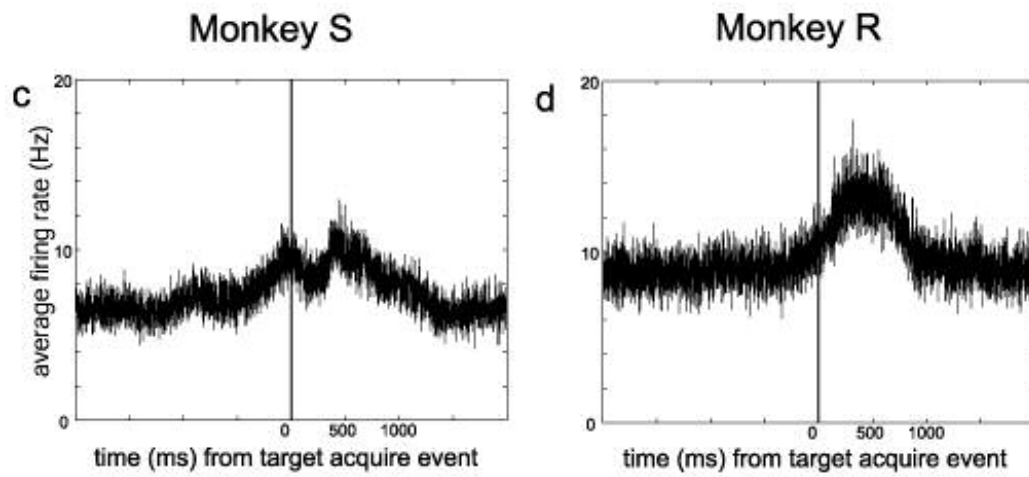
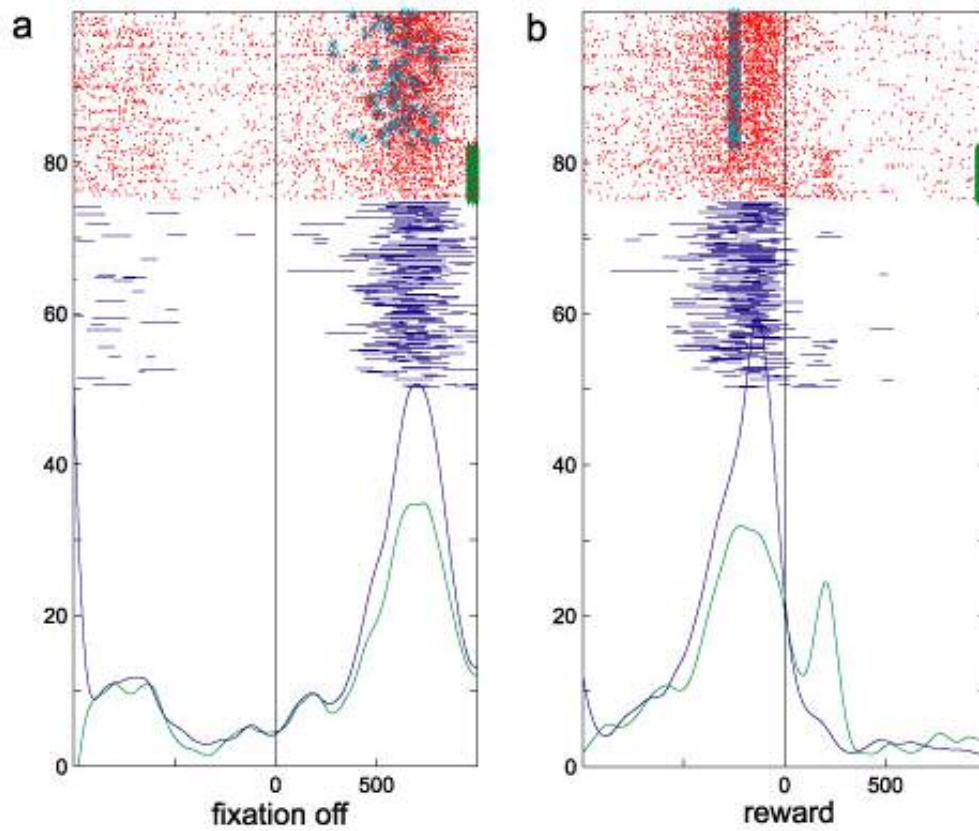Figure 2 c,d
Campos, Breznen, Bernheim, and Andersen

memory task

Figure 3 a,b
Campos, Breznen, Bernheim, and Andersen
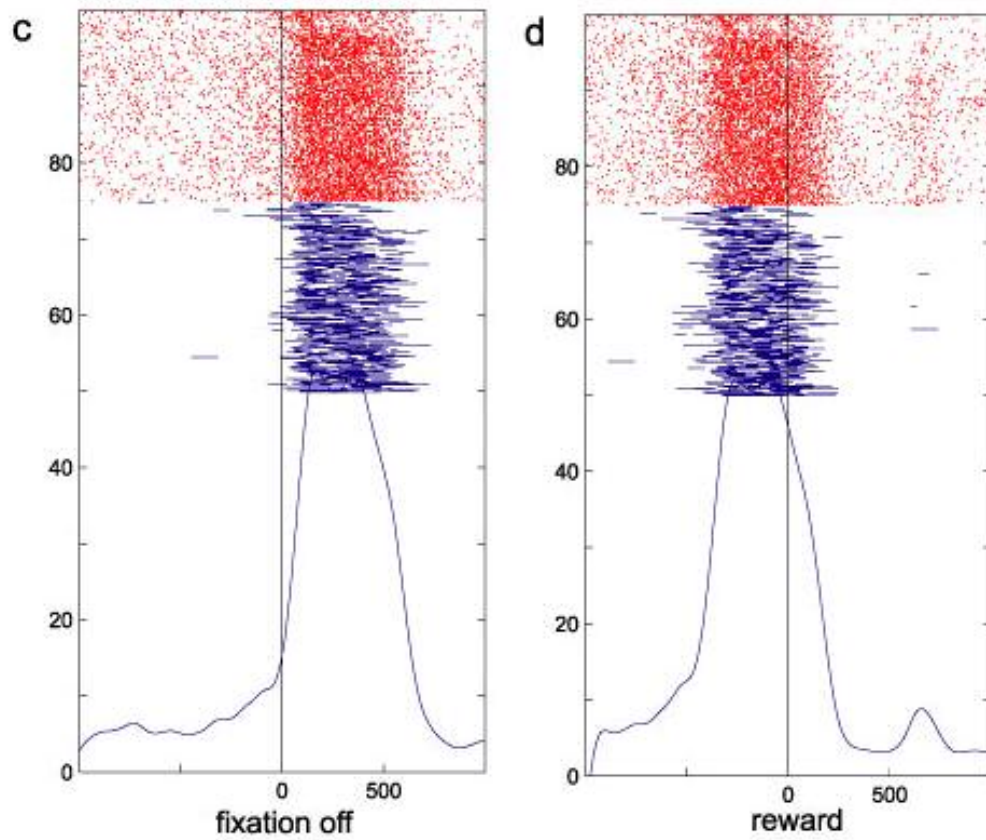
object task

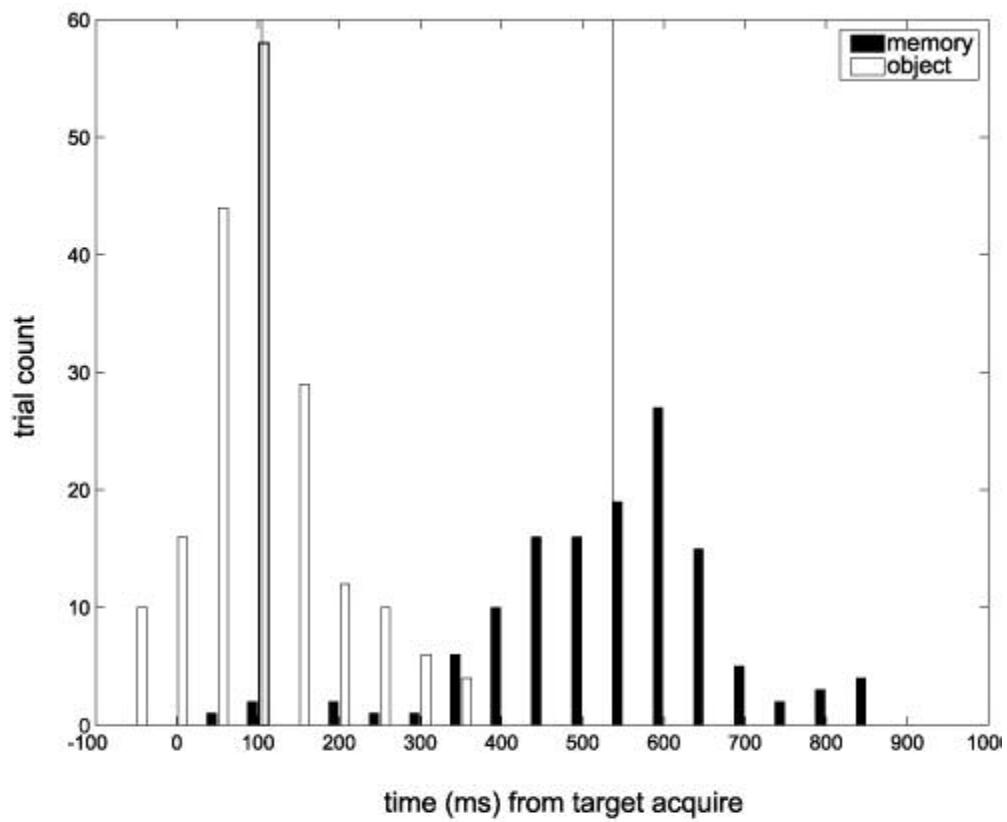Figure 3 c,d
Campos, Breznen, Bernheim, and Andersen
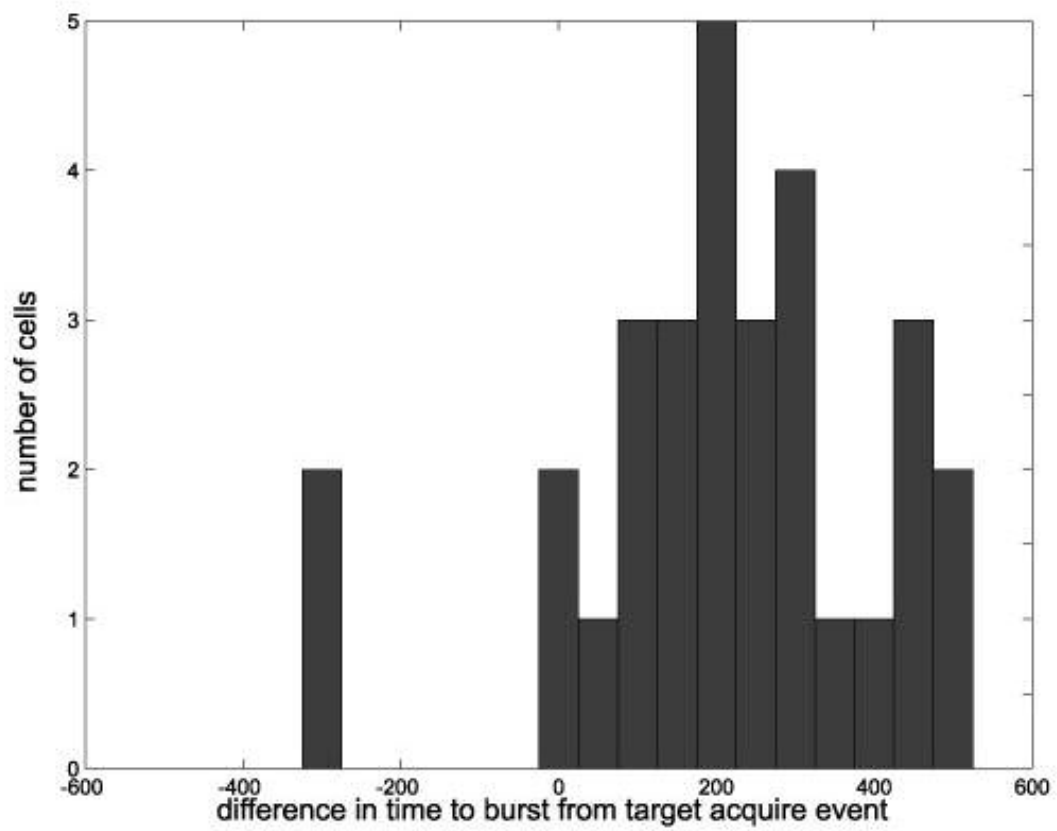
Figure 4a
Campos, Breznen, Bernheim, and Andersen

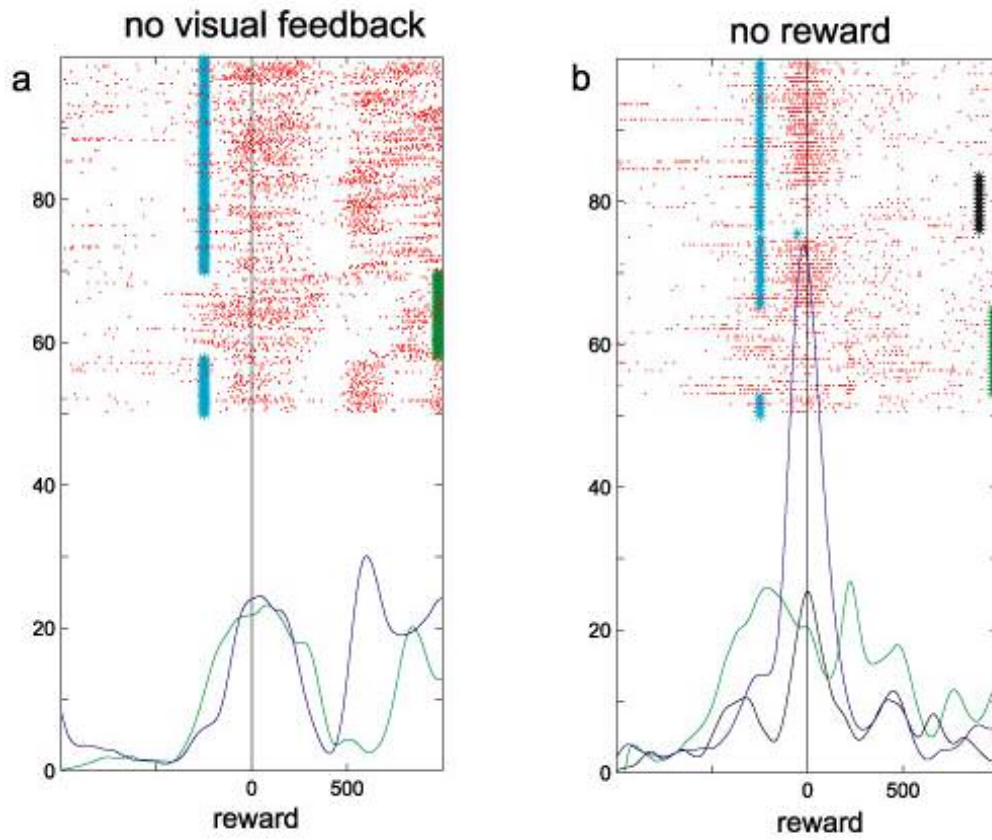Figure 4b
Campos, Breznen, Bernheim, and Andersen

Figure 5
Campos, Breznen, Bernheim, and Andersen

	Monkey S			Monkey R			Total		
	memory	object	both	memory	object	both	memory	object	both
Total recorded neurons	103	107	100	76	77	73	179	184	173
Task related modulation	71	77	55	42	44	29	113	121	84
Reward period activity	56	59	34	31	31	16	87	90	50
Post-saccadic bursts in > 30% of trials	32	27	18	18	21	12	50	48	30
Exclusive reward period activity	29	22	9	20	21	8	49	43	17
Response to no visual feedback control	16 (26)			6 (8)			22 (34)		
Response to no reward control	10 (10)			1 (1)			11 (11)		
Response to unexpected reward	2 (22)			0 (3)			2 (25)		

Table 1
Campos, Breznen, Bernheim, and Andersen

	cue	memory	saccade
Monkey S			
active above baseline	16	37	48
significant spatial and interval dependency	1	6	6
Monkey R			
active above baseline	7	14	23
significant spatial and interval dependency	0	1	0

Table 2
Campos, Breznen, Bernheim, and Andersen