Published in final edited form as:
J Neurosci. 2010 February 10; 30(6): 2340–2355. doi:10.1523/JNEUROSCI.1730-09.2010.

A Probabilistic Strategy for Understanding Action Selection

Byounghoon Kim[1] and **Michele A. Basso**[1,2]

[1]Department of Physiology, University of Wisconsin, Madison Medical School, Madison, WI 53706 USA

[2]Department of Ophthalmology and Visual Sciences, University of Wisconsin, Madison Medical School, Madison, WI 53706 USA

Abstract

Brain regions involved in transforming sensory signals into movement commands are the likely sites where decisions are formed. Once formed, a decision must be read-out from the activity of populations of neurons to produce a choice of action. How this occurs remains unresolved. We recorded from four superior colliculus (SC) neurons simultaneously while monkeys performed a target selection task. We implemented three models to gain insight into the computational principles underlying population coding of action selection. We compared the population vector average (PVA), winner-takes-all (WTA) and a Bayesian model, maximum *a posteriori* estimate (MAP) to determine which predicted choices most often. The probabilistic model predicted more trials correctly than both the WTA and the PVA. The MAP model predicted 81.88% whereas WTA predicted 71.11% and PVA/OLE predicted the least number of trials at 55.71 and 69.47%. Recovering MAP estimates using simulated, non-uniform priors that correlated with monkeys' choice performance, improved the accuracy of the model by 2.88%. A dynamic analysis revealed that the MAP estimate evolved over time and the posterior probability of the saccade choice reached a maximum at the time of the saccade. MAP estimates also scaled with choice performance accuracy. Although there was overlap in the prediction abilities of all the models, we conclude that movement choice from populations of neurons may be best understood by considering frameworks based on probability.

Keywords

saccade; eye movement; movement control; choice; decision; population coding; multiple neuron recording

INTRODUCTION

How perceptions, thoughts, decisions and actions arise from the activity of populations of neurons is arguably the most vexing question in cognitive neuroscience. A number of lines of evidence from experimental work in monkeys indicate that perceptual decisions leading to eye movements (saccades) evolve within sensorimotor centers of the brain such as the lateral intraparietal area (LIP), parietal reach region (PRR), frontal eye field (FEF) and the superior colliculus (SC) in the midbrain (Gold and Shadlen, 2000; Roitman and Shadlen, 2002; Ratcliff et al., 2003; Horwitz et al., 2004; Ratcliff et al., 2007; Scherberger and Andersen, 2007; Kim and Basso, 2008). A critical, unresolved issue is how the activity of neurons signaling targets and distractors is combined to contribute to a choice and then, how the combined activity is

Address correspondence: Michele A. Basso, Ph.D., Department of Physiology, University of Wisconsin, Madison, Medical School, 1300 University Ave., Room 127 SM1, Madison, WI 53706 USA, michele@physiology.wisc.edu.

read-out to result in a saccade. In other words, a key question remains unknown, what is the read-out rule that underlies movement choice?

Simultaneous recordings from multiple neurons within the monkey superior colliculus (SC) made during performance of a task in which one, differently-colored target appears in an array with three, same-colored stimuli, reveal that when the discriminability between the level of target and distractor neuronal activity is high, saccade choices are likely to be accurate. In contrast, when the discriminability between the level of activity of target and distractor neurons is reduced, choice performance is likely to be poor. This result is consistent with the suggestion that SC neuronal activity signals an eye movement decision. It also reveals that the choice of which eye movement to make depends upon the combined activity of neurons representing targets and distractors.

Because SC neurons are tuned broadly for target locations and saccade endpoints (Schiller and Koerner, 1971; Wurtz and Goldberg, 1972; Sparks, 1975, 1978; McIlwain, 1986; McIlwain, 1991), it is believed that the activity of large numbers of SC neurons are pooled to compute a vector average which determines the saccade direction (Ottes et al., 1986; Van Gisbergen et al., 1987; McIlwain, 1991); Groh, 2001 in much the same way as arm movement directions are coded by motor cortex neurons (Georgopoulos et al., 1986; Schwartz et al., 1988). Indeed, simultaneous electrical activation of two regions of the SC results in saccades with vectors that are averages of the saccade vectors produced by stimulation of each site independently (Robinson, 1972). Furthermore, inactivation of regions of the SC produces inaccuracies in saccade directions and lengths that are largely consistent with predictions of a population vector averaging (Lee et al., 1988; Quaia et al., 1998; Hanes and Wurtz, 2001). Whereas these experiments relied on measures of saccades made to single spots of light, more recent experiments requiring the identification of one target from an array of distractor stimuli (Basso and Wurtz, 1998; McPeek and Keller, 2004) or choosing between two, simultaneously or sequentially appearing stimuli (Port and Wurtz, 2003; Li and Basso, 2005; Kim and Basso, 2008) suggest that winner-takes-all or probabilistic strategies may more accurately reflect the information in SC neuronal populations.

Based on our previous work and that of others showing that SC neuronal activity scales with the likelihood of a correct saccade choice (Basso and Wurtz, 1998; Dorris and Munoz, 1998; Kim and Basso, 2008), here we explored whether SC neurons could formally encode information about saccade choices probabilistically. We implemented three different models to reveal principles underlying how SC neuronal activity might be encoded by the population and then interpreted by downstream structures during the performance of a target selection task. Unique to our experiments is that we recorded from four SC neurons simultaneously and each neuron contained one element of the visual display within its response field (RF). We determined the best estimate of the saccade choice by implementing a probabilistic (Bayesian) model, the maximum *a posteriori* estimate (MAP), winner-takes-all (WTA) and a population vector average (PVA/OLE) model. To assess the models, we compared how well each estimate predicted the saccade choice on a trial by trial basis for correct and error trials. Furthermore, we examined for the first time, the temporal development of the maximum *a posteriori* estimate. We found that the MAP model provided the best estimate of saccade choices across all trials, took time to develop and scaled with the monkeys' choice performance. These results are consistent with a probabilistic coding strategy underlying movement choice.

EXPERIMENTAL PROCEDURES

Physiological and Eye Movement Monitoring Procedures

For electrophysiological recording of SC neurons and monitoring eye movements, cylinders and eye loops were implanted in two rhesus monkeys (*Macaca mulatta*) using documented

procedures (Judge et al., 1980; Kim and Basso, 2008; Li and Basso, 2008). We recorded from 120 neurons within the intermediate layers of the superior colliculus (SC). We used a subset of the same data set used for a previous report (Kim and Basso, 2008). Neurons were recorded simultaneously in sets of four. In monkey m we recorded 13 sets of four SC neurons (n=52). In monkey c we recorded 17 sets of four SC neurons (n=68). Of the total 120 neurons in both monkeys, all neurons were defined statistically as buildup/prelude (Munoz and Wurtz, 1995; Basso and Wurtz, 1998; McPeek and Keller, 2002; Li and Basso, 2005, 2008) except three which were defined statistically as visual-tonic (McPeek and Keller, 2002; Li and Basso, 2008).

Neurons were recorded with four independently moveable, tungsten microelectrodes (*Frederick Haer*, Bowdoin, ME) with impedances between 0.3 and 1.0MΩ measured at 1 kHz. Four electrodes were aimed at the SC, each through different stainless steel guide tubes held in place by a plastic grid secured to the cylinder (Crist et al., 1988). Two were aimed at one SC and two were aimed at the other SC. Electrodes were introduced independently and neurons (action potential waveforms) were isolated on each electrode sequentially. Response fields (RF) of SC neurons were mapped online. Mapping was done by moving a spot around the screen and having monkeys make saccades to the different spots. We listened for maximal discharge and also monitored raster plots of the discharge on line. We considered the center of the RF to be the location at which a saccade was associated with maximal discharge of the neuron (audibly and visually). When recorded in the single target condition, we ensured that each stimulus drove only one of the recorded neurons. In other words, the RFs of each of the four neurons were non-overlapping when recorded in the single target condition. Action potential waveforms were filtered and amplified by a differential amplifier (*Alpha Omega*, Nazareth, Israel; MCP-Plus) and then sampled and digitized (*Measurement Computing*, Norton, MA; PCI-DAS4020/16). The digitized waveforms were identified and sorted with an interactive computer program (*Mex*; National Eye Institute, Bethesda, MD) allowing the experimenter to sort waveforms in real time. Neuronal data were also saved to disk as waveforms and sorted offline to confirm the adequacy of the on-line discrimination. For offline analysis we used custom software (written and compiled in Delphi 5.0) that sorted spikes based on time-voltage criteria.

Using the magnetic induction technique (C.N.C Engineering, Seattle WA.Fuchs and Robinson, 1966), voltage signals proportional to horizontal and vertical components of eye position were filtered (8 pole Bessel −3dB, 180 Hz), digitized at 16-bit resolution and sampled at 1 kHz (*National Instruments*; Austin, TX; PCI-6036E). The data were saved for off-line analysis using an interactive computer program (*Dex*, National Eye Institute Bethesda, MD) designed to display and measure eye position and calculate eye velocity. We used an automated procedure to define saccadic eye movements by applying velocity and acceleration criteria of 20°/s and 8000°/s^2, respectively. The adequacy of the algorithm was verified and adjusted as necessary on a trial-by-trial basis by the experimenter.

All experimental protocols were approved by the University of Wisconsin - Madison Institutional Animal Care and Use Committee and complied with and generally exceeded the standards set by the Public Health Service policy on the humane care and use of laboratory animals.

Behavioral Procedures

We used a real-time experimental data acquisition and visual stimulus generation system Rex, Vex and Mex, developed and distributed by the Laboratory of Sensorimotor Research National Eye Institute Bethesda MD (Hays et al., 1982) to create the behavioral paradigm and acquire two channels of eye position and four channels of neuronal data. Trained monkeys sat in a custom primate chair with head stabilized during the experimental session (typically 3–5h).

Visual stimuli were rear-projected onto a screen at 51cm distance from the subject using a projector (LP130, Infocus, Wilsonville, OR) with a native resolution of 1,024 × 768 and operating at 60Hz. A photocell secured to the screen sent a transistor-transistor logic pulse to the experimental PC providing an accurate measure of stimulus onset. The fixation spot at the center of the screen had a (mean of three measurements) luminance of 1.52cd/m^2. Visual stimuli each had luminance values of 5.8cd/m^2 (mean of three measurements). The background luminance was 0.58cd/m^2 (mean of three measurements). The PC for the visual stimulus display was a slave device to the PC used for experimental control and data acquisition.

After fixating on a centrally-located spot (0.6° diameter) for a random time of 1,800 – 2,300ms, four spots (1.0° diameter) appeared and the central spot disappeared. Each spot was located in the center of each empirically defined RF of the four SC neurons (Figure 1a; see text above). The task required monkeys to choose the differently colored target within ~300ms by making a saccade to the differently colored spot immediately after the disappearance of the fixation spot (coincident with the array onset). The target could be either red among green distractors or green among red distractors. The color arrangement of the display was fixed each day of recording but varied across recording days. After making a choice, monkeys maintained fixation at the target spot for a random time of 500–600ms and then received fluid reward. The location of the target spot was randomized (with replacement) among the four possible locations. On interleaved trials a single spot appeared in each of the four possible locations. Two spots appeared in each hemifield although the exact location of the visual spots depended upon the location of the four electrodes within the SC (Figure 1b).

Data Analysis

We implemented three broad classes of model to recover the population estimate of saccade choices in the selection task. The simplest model was a winner-takes-all (WTA) (Feldman, 1982). Next was a population vector average (PVA) model similar to that implemented by Georgopoulos and colleagues in motor cortex (1986) and Port and Wurtz in SC (2003). We also implemented an optimized vector averaging method developed by Salinas and Abbott (1994) referred to as the optimal linear estimator (OLE). The final model we implemented was a likelihood estimator based on Bayesian inferential statistics, the maximum *a posteriori* estimate (MAP). This model arises from ideas formulated previously (Sanger, 1996; Oram et al., 1998; Sanger, 2002; Sanger, 2003) and recently extended to sensory processing in MT (Jazayeri and Movshon, 2006) and decision-making in LIP (Beck et al., 2008). Our implementation of this model is similar but has important extensions of this recent work that are discussed below. In the text that follows, we describe how each model was implemented. We assessed the quality of the model prediction by comparing the saccade choice recovered by the model to the actual saccade choice monkeys made, regardless of whether the trial was correct or in error.

Bayesian model maximum a posteriori estimate (MAP—To compute the posterior probability distribution over the four possible saccade choices we first consider Bayes' rule which states:

$$P(s|r)=\frac{p(r|s)P(s)}{P(r)} \quad (1).$$

$P(s|r)$ is the conditional probability of observing a particular saccade choice given a particular discharge rate, also called the posterior. $P(r|s)$ is the conditional probability of observing a particular discharge rate when a particular saccade occurs. This value is known as the likelihood. $P(s)$ is the probability of a saccade choice or the prior. $P(r)$ is the probability of a particular discharge rate in (sp/s). A Bayesian framework provides a way to quantify guesses

about events when faced with uncertainty. The probabilities in Bayes' rule indicate the strength of a belief from 0 to 1. Since the probability of the discharge rate *P(r)* is independent of the saccade choice, we can restate the posterior as proportional to the product of the likelihood and the prior (Földiak, 1993; Oram et al., 1998):

$$P(s|r) \propto P(r|s)P(s) \quad (2).$$

In words, the probability of a saccade choice given the observation of a particular discharge {*P(s|r)*} is proportional to the conditional probability of the discharge given the saccade choice {*P(r|s)*} multiplied by the prior {*P(s)*}. Thus, the posterior is proportional to the product of the likelihood and the prior. Note that for display, we include the normalization factor: *P(r)* so the scaling ranges from 0 – 1 and so that bona fide probabilities can be compared across conditions (Oram 1998).

We implemented two prior {*P(s)*} functions. In one, we used a discrete uniform prior with four Dirac delta functions at each of four possible target choices:

$$P(s)=0.25\sum_{i=1}^{4}\delta(s-(i-1)\pi/2) \quad (3).$$

P(s) is a discrete prior function describing the four possible saccade target choices (*s*) separated by 90°. The four delta functions (δ) for each choice are defined by shifts from the first by 90° ($\pi/2$). We summed the four delta functions and multiplied by 0.25 (four possible choices) so the prior function had uniform probabilities for the four possible target/saccade choices (Figure 2h). This reflects the experimental situation used. Each of the four possible saccade choices represents one of the four possible target locations and each of these occurred with an equal (25%) probability.

For the second implementation, we used a simulated *P(s)* and determined the distribution that maximized the MAP model's performance. To do this, we generated four random values that summed to 1.0 to simulate probability distributions:

$$P(s)=\sum_{i=1}^{4}\left[\mathrm{rand}_i/\sum_{j=1}^{4}\mathrm{rand}_j\right]\delta(s-(i-1)\pi/2) \quad (4).$$

Equation 4 defines a discrete prior function as a summation of four delta functions as in equation 3. Each value in equation 4 (rand_{1-4}) varied independently from 0.01 to 0.97 with an interval of 0.01. To ensure that the sum of this distribution was 1.0, we divided each of four random values (rand_i) by the sum of all values $(\sum_{j=1}^{4}\mathrm{rand}_j)$.This was then multiplied by the 90° shifted delta functions yielding a prior probability distribution that is a discrete, non-uniform distribution with four values, one for each of four possible saccade choices (Figure 2i). Because they were generated randomly, equation 4 produced the same combinations in some cases. Therefore we selected only the unique combinations. This left a total of 156,941 unique combinations of four values. With the simulated prior distributions in hand, we then recomputed the MAP estimate using each one of these 156,941 simulated prior distributions. We indentified the simulated prior function that when used to recover the MAP estimate, resulted in the same or better prediction accuracy as the MAP estimate with the non-uniform prior function. The actual prior function that monkeys might use is unknowable. The prior

function used, however, is likely to be related to the final choice behavior (distribution of saccade choices). To test this we calculated the distribution of saccade choices using the actual behavior of the monkeys on a trial by trial basis. In Figure 4a we show through simulations of multiple possible prior functions, that there is a strong relationship between the simulated prior distribution and the distribution of saccade choices. These correlations validate our use of the simulated prior to recover the MAP estimate of the saccade choice.

A critical aspect of computing the MAP estimate is how to determine $P(r|s)$. In recent work it was shown that a good characterization of $P(r|s)$ can be obtained by assuming a Poisson probability distribution or any distribution of the exponential family with linear sufficient statistics (Ma et al., 2006; Beck et al., 2007). So our first approach was to use a Poisson probability distribution constrained by the tuning properties of our SC neurons to estimate $P(r_{1-4}|s)$. Indeed as a first approximation, our neurons behaved in a linear sufficient fashion as determined by assessing the relationship of the variance of action potential counts across trials to the mean of the action potential counts (see Supplemental Figure 1a). We used the Poisson probability density function in place of the likelihood, $P(r_{1-4}|s)$ where lambda is the expected number of action potential occurrences in a Poisson probability distribution and in our implementation was the tuning curve { $fi(s)$ } of the i_{th} neuron. The exponent and the denominator (r_i) are the numbers of action potentials measured in a 20ms time epoch (28ms before to 8 ms before the onset of the saccade):

$$P(s|r_{1-4}) \propto \left[\prod_{i=1}^{4} \frac{f_i(s)^{r_i}}{r_i!} e^{-f_i(s)}\right] P(s) \tag{5}$$

The posterior probability of a saccade choice (s) given the discharge of all four neurons $\{P(s|r_{1-4})\}$ was estimated by computing the conditional probability for each of the four neurons and multiplying by the prior probability. In order to combine the neuronal activity linearly we made the reasonable assumption that the neurons in our sample were statistically independent. We describe how we deal with this assumption in Supplemental Figure 1b, Supplemental Figure 2 and in the results. To allow summation rather than multiplication, we took the logarithm of equation (5)

$$\log(P(s|r_{1-4})) = \sum_{i=1}^{4} r_i \log(f_i(s)) - \sum_{i=1}^{4} f_i(s) - \sum_{i=1}^{4} \log(r_i!) + \log(P(s)) \tag{6}$$

$\sum_{i=1}^{4} \log(r_i!)$ can be ignored because it is independent of the saccade choice. However, we maintained the $\sum_{i=1}^{4} f_i(s)$ term because across our neurons, the tuning curves were different (Figure 2d). As a result this term does not sum to a constant and must remain in the model. The calculation simplifies to:

$$\log(P(s|r_{1-4})) = \sum_{i=1}^{4} r_i \log(f_i(s)) - \sum_{i=1}^{4} f_i(s) + \log(P(s)) \tag{7}$$

Note that for *fi(s)*we also implemented a version of the MAP model in which we used identical Gaussian functions (Edelman and Keller, 1998) peak shifted by 90° to simulate SC tuning curves (Figure 2b). In this case, the $\sum_{i=1}^{4} f_i(s)$ term was omitted from the model because summing over these functions is a constant. To avoid overestimation of the model, we implemented a leave-one-out cross validation procedure. For this, we extracted one trial from each data set and used the remaining trials to estimate *P(r|s)* from each set. We then recovered the posterior from the extracted trial. This procedure was repeated for all trials for each data set. Equation 7 returns one value for each of the possible saccade choices. Computing this value for each of the four possible saccade choices (s_j) where $j = 1–4$, defines the posterior distribution across the four possible saccade choices. In the case of the uniform prior function, the result is a Bayesian estimator that yields the same result as a maximum likelihood estimator as formalized by others (Sanger, 1996; Sanger, 2002; Jazayeri and Movshon, 2006). In the case of the non-uniform prior, the result is a Bayesian estimator distinct from a maximum likelihood. To determine how well the posterior distribution predicted monkeys' actual choices we compared the maximum *a posteriori* estimate (MAP) with the saccade choice on a trial by trial basis:

$$\begin{aligned} MAP(s_j) &= \arg\max(\log(P(s_j|r_{1-4})) \\ &= \arg\max\left[\sum_{i=1}^{4} r_i \log(f_i(s_j)) + \log(P(s_j))\right] \end{aligned} \quad (8).$$

When the saccade choice and the maximum *a posteriori* estimate corresponded, we considered the model to have a correct prediction. Figure 2 provides a graphic depiction of the MAP model along with the different *P(r|s)* and *P(s)* implementations.

Determining the likelihood using a Poisson probability density function relies on two assumptions (Földiak, 1993; Sanger, 1996; Oram et al., 1998; Sanger, 2002; Jazayeri and Movshon, 2006; Ma et al., 2006; Beck et al., 2008). The first is that the occurrence of each action potential in a spike train is independent of the occurrence of other action potentials in the train. If time between successive action potentials is random we can consider the train of action potentials as a Poisson process. A common way to assess whether discharge statistics can be described as a Poisson process is to determine the index of proportionality, also referred to as the fano factor which is the ratio of the variance of the number of action potentials in an epoch to the number of action potentials in an epoch across trials. On a linear plot, a slope of 1.0 indicates linearity. To determine the fano factor of SC neurons, we counted the number of action potentials within the 28 to 8ms epoch before the onset of a saccade for each trial in a data set. We then determined the trial to trial variance of the action potential counts by subtracting individual trial counts from the mean count and squaring that quantity. This was done for the set of trials across all neurons. We then computed the mean of the difference measure and the mean of the action potential count and plotted these values for each neuron. Supplemental Figure 1a shows the action potential count variance against the mean count across all 120 SC neurons when the stimulus in the neurons' RF was either a target or a distractor. The fano factor for neurons when targets were in their RFs was 1.44 (n=120). The fano factor for neurons when distractors were in their RFs was 1.03 (n=360). These observations are consistent with the assumption of linear sufficient statistics, at least for distractor activity.

Supplemental Figure 1 shows that when targets appeared in the RF, the variance to mean relationship diverged from linearity. This is because SC neurons are exhibiting rapid increases in discharge associated with the saccade to the target in the RF. To deal with this deviation from linearity, we extended our probabilistic model to eliminate the Poisson probability

distribution to estimate $P(r_{1-4}|s)$. Instead, we determined $P(r_{1-4}|s)$ directly by using a non-parametric density estimation procedure (Optican and Richmond, 1987; Scott, 1992). Panels e and f of Figure 2 show graphically how this was performed. Non-parametric density estimation is simply smoothing a frequency histogram. This procedure is similar to that used to calculate spike density functions from raster plots (MacPherson and Aldridge, 1979). We first plotted the distribution of discharge rates measured in the four possible target condition during the 20ms epoch measured 28 to 8ms before saccade onset. We applied a smoothing kernel ($k[\ ]$):

$$k[\ ]=\frac{e^{(x-r)^2/-2h^2}}{\sqrt{2\pi}}\cdot h \tag{9}$$

where h is the number of bins, r is the discharge rate measured in the 20ms epoch and the domain of x is the set of all numbers defined by the discharge rate. From here, the Gaussians are summed over the discharge rates and the sum is weighted by the number of bins in the frequency distribution (n). Assuming a normal probability density, h can be estimated by minimizing the (averaged) mean integrated squared error AMISE,(Scott, 1992):

$$h=\left(\frac{4}{3}\right)^{1/5}\sigma n^{-1/5}\approx 1.06\sigma n^{-1/5} \tag{10}$$

Convolving the histograms with the smoothing kernel in equation 9 yields the empirical probability density distribution:

$$P(r|s)=\frac{1}{n}\sum_{i=1}^{n}k\left[\frac{x-r_i}{h}\right] \tag{11}$$

This procedure was done to obtain a probability density function for each neuron. From here we could extract the $P(r|s)$ directly to compute the posterior distribution over the four possible saccade choices again on a trial by trial basis:

$$P(s|r_{1-4})=\left[\prod_{i=1}^{4}P(r_i|s)\right]P(s)/P(r) \tag{12}$$

As in the model shown in equation 8, when the saccade choice and the maximum *a posteriori* estimate determined from equation 12 agreed, we considered the model to have a correct prediction.

The second assumption that is required to compute the posterior probability is that the noise correlations between the four neurons should be independent. Because we were careful to record from neurons with non-overlapping RFs, we assumed independence of the neuronal discharge. We calculated the noise correlation coefficients between neuronal responses to confirm our assumption (Averbeck et al., 2006). Since we have four neurons, combining each into unique pairs resulted in six pairs allowing us to test all possible noise correlations between the four neurons *(total conditions = 6(pairs)×4(target conditions)×30(data sets) = 720).* Neuronal activity was measured 28 to 8 ms before saccade onset for all six pairs. Across our sample of 720 pairs, only 9.86% (71/720) of the pairs had statistically significant noise

correlations (Supplemental Figure 1b). To confirm that the noise correlations were accounted for in our model or did not contribute much to the result of the model we performed a shuffled analysis of our data (as shown in Supplemental Figure 2).

It is important to note that although our implementation is Bayesian in the sense that we calculated a posterior probability by combining likelihoods and prior functions, there are some differences between our model and true Bayesian estimator. First, the likelihood functions in our model are discrete and are estimated from four individual neurons. Any additional variability that may be conveyed leading to a saccade choice is ignored. Second, the prior functions we implemented are also discrete and deterministic. Third, and as noted above, since we cannot ever know the true prior function, we simulated it. As shown in Figure 4a the simulated prior function correlates with the distribution of saccade choices made by the monkeys. This validates our approach and indicates that the simulated prior function we used to recover the MAP estimates of saccade choice was a good approximation to the actual prior used by the monkeys while performing this task.

Winner-takes-all (WTA)—We implemented a WTA model by computing the mean discharge rate during an interval 28ms to 8ms before the onset of a saccade (Miyashita and Hikosaka, 1996) on a trial by trial basis from each of the four neurons across all 30 sets. For each set of four neurons, the neuron with the highest discharge rate was defined as the 'winner'. We then compared the RF location of the winner neuron on each trial to the location of the saccade choice on that trial. Because there were four neurons, each representing one possible location, a correct prediction occurred when the neuron corresponding to the saccade choice had the highest discharge rate.

Population Vector Average (PVA)—To compute the population vector average ($\vec{V}_{population}$) we considered each of the four neurons simultaneously recorded as one of a larger population of neurons representing one of the possible saccade choices. We computed the ($\vec{V}_{population}$) for each trial using the four neurons, one representing the target (target neuron) and the other three representing the distractors (distractor neurons). We implemented a similar procedure for computing the ($\vec{V}_{population}$) as used previously in motor cortex (Georgopoulos et al., 1986) and SC (Port and Wurtz, 2003). However, we adopted a normalization procedure suggested by Salinas and Abbott (1994) to avoid obtaining negative vectors:

$$\vec{V}_{population} = \frac{\sum_{i=1}^{4}\left(r_i\left(\sqrt{\sum_{j=1}^{4} r_j^2}\right)^{-1}\vec{S}_i\right)}{4} \quad (13).$$

We computed the neuronal population vector average using equation (13) where r_i was the mean discharge rate for each i_{th} neuron measured during the 20ms interval immediately before the onset of the saccade during the selection task (28ms to 8ms before saccade onset; note that this is the same interval as used for MAP and WTA). Since each neuron had different discharge rates and different baseline rates, it was necessary to normalize the neuronal responses to avoid arbitrary biases in ($\vec{V}_{population}$). The normalized neuronal response was determined by calculating $\{r_i\left(\sqrt{\sum_{j=1}^{4} r_j^2}\right)^{-1}\}$. This term was then multiplied by the unit vector $\vec{S}_i$ which was defined as the average saccade direction determined by the actual eye movements made by the monkeys. This result was then summed and divided by four to obtain ($\vec{V}_{population}$). The saccade choice with the smallest angular difference between it and the ($\vec{V}_{population}$) was considered a

correct model prediction. One assumption of this version of the PVA is that the SC contains a homogenous representation of all possible saccades. So we also implemented a model developed by Salinas and Abbott (1994) that optimizes the neuronal population vector to take into account inhomogeneous distributions. For this model we determined ($\vec{V}_{population}$) by summing the discharge rates weighted by an optimized saccade vector:

$$\vec{V}_{population} = \left(\sum_{i,j=1}^{4} r_i \vec{D}_j \right) /4$$
$$\vec{D}_{j(optimized)} = \left[\left(\sum_{i=1}^{4} C_{r_i} \right) /4 \right]^{-1} \cdot \left(\sum_{i=1}^{4} r_{ij} \cdot \vec{S}_j \right) /4 \quad (14)$$

where i = neuron number and j = saccade choice. The optimized saccade vector ($\vec{D}_{j(optimized)}$ was determined by multiplying an inverse correlation matrix ($\left[\left(\sum_{i=1}^{4} C_{r_i} \right) /4 \right]^{-1}$ Supplemental Figure 3a) of the discharge rates for all possible saccade choices by another correlation matrix (Supplemental Figure 3b) between four saccade unit vectors ($\vec{S}_j$) and neuronal responses (r_{ij}). Once we obtained the optimized saccade vectors, we multiplied the neuronal responses (r_i) by the optimized saccade vectors and averaged them. This calculation resulted in an optimized $\vec{V}_{population}$ that took into account the inhomogeneity of the saccade choice locations represented by the sample of neuronal discharges. As for the PVA, we considered the optimal linear estimator (OLE) as correctly predicting the saccade choice when the angle between the $\vec{V}_{population}$ and the actual saccade choice was smallest (Supplemental Figure 4).

RESULTS

Trained monkeys performed a pop-out selection task in which one stimulus from among four was uniquely identified as the target because of its color. In some arrays the target appeared red and the three distractors appeared green. In other arrays the target appeared green and the three distractors appeared red. The color of the target varied from experimental day to experimental day. The position of the target within the array of four stimuli appeared at a random location within experimental days. Therefore, for each trial, the monkeys knew the color of the target but they did not know the position of the target. Figure 1a shows an example of the task in which the target appeared red. We recorded from four neurons simultaneously while monkeys performed this task. Each stimulus position appeared in the empirical center of the response field (RF) of the recorded neuron. As a result, the positions of the stimuli in the array were constrained by the positions of the electrodes within the superior colliculus (SC) map, although two were always located in each SC and we excluded neurons with overlapping RFs (Figure 1b, see **EXPERIMENTAL PROCEDURES**). Figure 1a shows the idealized case in which the stimuli appeared 90° from one another. For data presentation we normalized the positions so that each of four spots appeared at 45°, 135, 225 and 315°.

Because of the variability in the positions of the targets there was variability in choice performance (Kim and Basso, 2008). In the example shown in Figure 1a the monkey performed with 100% accuracy on trials when the target appeared at location 45°. On the same experimental day, performance was between 70 and 99% accurate on trials when the target appeared at the 315° position. Performance was poorer (<70% accurate) when the target appeared at position 225°. Figure 1c shows an example of the recordings from four neurons in this task. We demonstrated recently that the relative level of activity of SC target neurons and SC distractor neurons predicts saccade accuracy in a manner consistent with the interpretation that SC neuronal activity encodes the saccade choice (Kim and Basso, 2008). Furthermore, the

range of choice probabilities we and others obtained from SC neurons was similar to the range reported in decision making tasks in SC and other brain regions such as frontal eye field (FEF) and lateral intraparietal area (Horwitz and Newsome, 2001; Shadlen and Newsome, 2001; Gold and Shadlen, 2002; Horwitz et al., 2004). This suggests that similar numbers of neurons in these areas are pooled to determine the choice.

SC Population Activity and Saccade Choice

Previous work in SC indicates that buildup/prelude neuronal activity scales with the likelihood of saccade occurrence (Basso and Wurtz, 1998; Dorris and Munoz, 1998; Kim and Basso, 2008), suggesting that saccade choice may be encoded across the population of SC buildup neurons. We formulated three models to reveal insights into the computational principles underlying saccade choice in SC. For each model, we compared the results of the model with the saccade choice made by the monkeys to determine the models' accuracy. We implemented a simple winner-takes-all model (WTA), two variants of the population vector average model (PVA and OLE), and a probabilistic model based on Bayesian inference called the maximum *a posteriori* estimate (MAP). These models are explained in **EXPERIMENTAL PROCEDURES**. In what follows, we describe the results of each model. In parallel, we point out important extensions of the probabilistic model we implemented compared to that implemented by others.

Probabilistic and Winner-Takes-All (WTA) models—One important point about computing the MAP is the manner in which $P(r|s)$ (the likelihood) is estimated. We did this in two different ways. First, we used a Poisson probability density function to characterize variability for $P(r|s)$ as depicted in Figure 2a–d and as done recently by others (Jazayeri and Movshon, 2006; Beck et al., 2008). Second, we determined $P(r|s)$ directly using non-parametric density estimation (see **EXPERIMENTAL PROCEDURES**, Figure 2e–f). When estimating $P(r|s)$ using a Poisson probability density function to characterize variability, we constrained the distribution in three different ways. In one way, we used the log of Gaussian functions that were designed to simulate the tuning curves of SC neurons with parameters, $\sigma = 20.6°$ and baseline discharge = 7sp/s (Figure 2b; Edelman and Keller, 1998). Note that recent work used von Mises curves for which the logs are cosine functions (Jazayeri and Movshon, 2006; Beck et al., 2008). In a second way, we estimated the tuning curves by measuring the discharge rates of neurons during performance of saccades to the different target positions when only a single target appeared. In the third way we used the discharge of SC neurons when saccades were made in the four possible target conditions to estimate the tuning curves (Figure 2d). Importantly, for the tuning curves measured from the four possible target trials, we used the same data for estimating tuning curves and predicting movement choice. Therefore, we implemented a leave-one-out cross validation procedure to avoid model overestimation (see **EXPERIMENTAL PROCEDURES**).

Figure 2 illustrates these different approaches to estimating (log) likelihoods. In these examples individual neurons are indicated by different colors and the condition shown is when the spot located in the 45° position was the target. For this condition, the red neuron contributes little discharge (Figure 2a and c, red dot). Likewise, the green and black neurons contribute little since these neurons are largely inactive for this target position. The blue neuron contributes maximally since the target is located within the center of its RF (Figure 2a and c, blue dot). Note that in Figure 2a and c we interpolated discharge rate points in between the four target positions to generate smooth functions for the discharge rates. This was done for display only. In reality there are only four points, one for each of the four possible target conditions. These discharge rates are then weighted by the logs of the individual tuning curves, either modeled as Gaussian functions (Figure 2b), estimated from the real data (Figure 2d) or determined directly from the probability density estimation procedure using the real data (Figure 2e, f).

Combining the (log) likelihoods (Figure 2g) with the prior probability (Figure 2h, i; see **EXPERIMENTAL PROCEDURES**) yields the posterior probability across all four possible saccade choices (Figure 2j and k). The same computation excluding the prior yields a log likelihood distribution (Figure 2g).

All models used simultaneously recorded neuronal data measured in 20ms epochs 28ms to 8ms before the onset of the saccade on a trial by trial basis for all 120 neurons (30 data sets of 4 neurons) from two monkeys. Figure 3a shows the result of the MAP model with simulated Gaussian tuning curves. The black line shows the mean of the MAP estimates of all the trials for each of the four possible target positions in which the saccade choice was predicted correctly by the maximum of the posterior probability distribution. The grey line shows the same for the trials in which the saccade choice was not predicted by the MAP estimates. Overall the MAP estimates predicted saccade choices on 2870/4035 (71.11%) trials. Having the neuronal response of each neuron weighted by the log of a uniform Gaussian (or uniform cosine functions) means that the contribution of each neuron was directly proportional to its discharge rate. Since the Gaussian tuning curves were shifted by 90° they were minimally overlapping and therefore the distractor neurons contributed an identical and minimal amount to all conditions. Thus, the neuron with the highest discharge rate dominates the prediction of saccade choice in the posterior distribution. This result is in principle, identical to the prediction of a WTA model.

Figure 3b plots frequency histograms of the mean discharge rate in each of the 4035 trials from all 120 neurons. In the WTA model the neuron with the highest discharge rate predicts saccade choice. As is evident from Figure 3b, the distribution of discharge rates measured from target neurons overlaps with the distribution of discharge rates measured from distractor neurons (Figure 3b, cf., black lines and gray bars). This indicates that the distractor neurons were as likely to have the highest discharge rate as the target neurons for some trials. This occurred even though the saccade choice corresponded to the RF location of the target neuron on all of the trials from which these data were taken. Overall 71.11% (2870/4035) of saccade choices were predicted correctly by the WTA model.

The homogeneous and non-overlapping tuning curves used in this version of the MAP model do not provide any benefit over the WTA model in predicting saccade choice. This result also reveals the possibility that simulated, non-overlapping and homogeneous tuning curves may not be optimal for population coding. To explore whether using the actual neuronal tuning curves would improve the model predictions over simulated ones, we implemented a version of the MAP model using tuning curves estimated from actual neuronal responses. In Figure 3c we show the result of the MAP model when the tuning curves were estimated from the SC neuronal discharge in the four possible target conditions. In this case, the tuning curves 'overlap' in the sense that there is discharge in the distractor neurons because there is always a stimulus in their RF, whether or not that stimulus will ultimately be chosen for a saccade. This implementation of the MAP model improved accuracy by ~10% over the previous version in that it predicted saccade choice on 3294/4035 (81.64%) of trials compared to 71.11% of the trials. This provides some indication that the properties and amount of overlap in the distributions used to estimate the likelihood (logs of tuning curves in this case) is an important variable when considering population coding of movement choice.

Although we used the tuning curves estimated from the neuronal data the implementation of the MAP model described above still made the assumption that the variability used to estimate $P(r|s)$ could be characterized by a Poisson probability distribution. However, the Poisson assumption is unrealistic biologically for single neurons. This is unrealistic particularly for SC neurons where discharges are characterized by robust bursts shortly before and during the generation of saccades (see Supplemental Figure 1a; Sparks, 1986; Moschovakis et al.,

1996). In light of this, we implemented an additional version of the MAP in which we determined *P(r|s)* directly from our recorded neuronal data by generating a probability distribution of the neuronal discharges (Figure 2e and f; see **EXPERIMENTAL PROCEDURES**). Estimating *P(r|s)* directly from the raw discharges improved the result of the MAP model minimally. Overall it predicted saccade choices correctly in 81.88% (3304/4035) of trials (Figure 3d). That there was little difference between the result using this method and the result using a Poisson probability distribution is consistent with recent theoretical work (Ma et al., 2006; Beck et al., 2007) and suggests that even though individual neuronal variability is not well-explained by Poisson statistics, a MAP estimate based on Poisson variability performs well.

MAP estimates with the non-uniform priors—Up to now we recovered the MAP estimates using a uniform prior and the empirical PDFs to estimate *P(r|s)*. We did this because the target for the saccade choice was equally likely to appear in each of the four possible locations on every trial and monkeys did not have to use a prior to perform this task. However, in spite of the correct target occurring with a 25% probability, monkeys did not always perform with this accuracy indicating that they likely incorporate biases into their choices. These biases must be based on something other than the sensory information (Kim and Basso, 2008). Therefore, we implemented a non-uniform prior distribution into our MAP model to determine whether we could improve the MAP estimates of saccade choice. As described in the **EXPERIMENTAL PROCEDURES**) we simulated many possible prior distributions and identified one for each data set that resulted in the maximum number of correct MAP estimates. To validate our use of these simulated distributions, we compared them to the distributions of saccade choices the monkeys actually made. Note that the prior distributions and the distributions of saccade choices are not identical but they should be related. Although the actual prior distribution used by monkeys is unknown, we reasoned that correlations between the simulated prior distributions and the distributions of saccade choices would indicate a reasonable approximation to the actual prior. In Figure 4a, we show that across the sample, the simulated prior distributions were correlated to the distributions of saccade choices. An example of one such correlation is shown in the inset of Figure 4a. Each point in Figure 4a is a pair of points, one from the distribution of saccade choices and one from one of the simulated prior distributions. The arrow drawn from the inset to the point in Figure 4a shows the two points from the inset remapped to the plot of points shown in Figure 4a. The total number of points in Figure 4a is 120 because we had 30 data sets and each data set has four possible saccade choices. One simulated prior function was used for each data set.

Figure 4b shows the result of adding the non-uniform prior function to recover the MAP estimate using the empirical PDF to characterize *P(r|s)*. The black line shows the mean of the MAP estimates for trials correctly predicted by the MAP model. The grey line shows the mean of the MAP estimates for trials that were incorrectly predicted by the model. Overall the MAP estimates with the non-uniform prior predicted saccade choices on 3420/4035 (84.76%) trials. By using the non-uniform prior, the performance of the MAP estimates improved 2.88%, from 81.88% to 84.76%.

The influence of priors on model performance is highly dependent on task demands. Therefore, we next asked whether the change in prediction accuracy occurred for correct and error trials similarly. We reasoned that monkeys might make more errors because they relied more on their priors than on the sensory information to inform their choice. Therefore, we guessed that MAP estimates of saccade choice might improve preferentially for error trials over correct trials when we used the non-uniform prior. Figure 4c plots the percent of trials correctly predicted by the MAP estimates using the uniform versus the non-uniform priors sorted by whether the trial was correct or in error. For correct trials, the MAP estimates with the uniform prior predicted choice in 84.91% of the total trials (Figure 4c; second gray bar). Using the non-

uniform prior increased the prediction accuracy to 87.22%. This is an improvement in the MAP estimate of 2.31%. For error trials, the MAP estimates determined with the uniform prior accurately predicted saccade choices on 71.49% of trials (Figure 4c; third gray bar). The MAP estimates determined with the non-uniform prior accurately predicted 76.32% of the error trials. This represents an improvement in the model accuracy by 4.83% (Figure 4c; third black bar). We performed a re-sampling procedure in which we randomly selected correct trials and error trials and assessed whether they were predicted by the MAP model with the uniform or the non-uniform model. Repeating this sampling procedure 1000 times and comparing the resulting distributions indicated that the improvement in prediction accuracy that occurred by implementing a non-uniform prior to recover the MAP estimate occurred more for error trials than for correct trials. The differences in improvements were statistically significant (X^2 = 1000, $p<0.001$). Based on these results we conclude that the prediction accuracy of the MAP estimates using a non-uniform prior exceeds that of the uniform prior. Furthermore, using a non-uniform prior improves the MAP estimates preferentially for error trials over correct trials. This suggests that errors in choice may occur because monkeys base their choices on the prior information rather than the sensory information.

Population vector average (PVA) and optimal linear estimator (OLE)—In the same way that motor cortex is considered to encode arm movement direction (Georgopoulos et al., 1986; Schwartz et al., 1988), it is considered that the SC encodes the direction of a saccade by averaging across the population of active neurons, each of which contributes a minivector to determine saccade direction. Although supported by lesion experiments (Lee et al., 1988; Quaia et al., 1998; Hanes and Wurtz, 2001) and argued on theoretical grounds (Van Gisbergen et al., 1987; McIlwain, 1991; Groh, 2001), recent lesion experiments (McPeek and Keller, 2002) and dual neuron recording experiments (Port and Wurtz, 2003) suggest that a vector average may be too simplistic, at least when considering saccades made in the presence of more than one visual stimulus. Therefore, we implemented a PVA model as well as an improved version, the optimal linear estimator (OLE) to assess whether these models could predict saccade choice as well as the MAP model.

For the PVA and OLE the same data set was used as that used for the MAP model. Figure 5a shows the result of the PVA. Figure 5b shows the result of the OLE. Each line is the neuronal population vector for each trial (n = 4035). The black lines show the result when the direction of the population vector and the direction of the saccade had the smallest angular difference (see Supplemental Figure 4). We considered these trials to have a correct prediction. The grey lines show the result when the difference in the angle of the direction of the population vector and the angle of the direction of one of the distractor stimuli was the smallest. We considered these trials to have an incorrect prediction. Overall the PVA accurately predicted 2248/4035 (55.71%) of saccade choices (Figure 5a).

The PVA predicted the saccade choice correctly for many trials but also failed quite often. Next, we optimized the neuronal vectors with the correlation matrix. This maximizes the PVA performance by taking into account the fact that our sample of neurons did not contain a homogenous representation of saccade space (see **EXPERIMENTAL PROCEDURES**). Figure 5b shows the prediction results of the OLE. Overall the OLE predicted 2803/4035 (69.47%) of saccade choices accurately which represents a 13.76% improvement over the PVA. However, when compared to the WTA and MAP models, the OLE showed the lowest prediction accuracy of saccade choice. Note that this is in spite of the fact that like the MAP, the OLE uses information from all four neurons to determine the saccade choice.

Figure 6 provides a direct comparison of the results from all the models and their different implementations. The MAP using a non-parametric density estimation procedure (from data recorded in the four possible target condition) for determining $P(r|s)$ and a non-uniform prior

predicts saccade choices very well. This version of MAP predicted saccade choices correctly in 84.76% (3420/4035) of all the trials (first bar; Figure 6a). The next best prediction occurred for the same implementation with the uniform prior. This version predicted 81.88% (3304/4035) of all trials (second bar; Figure 6a). When we estimated $P(r|s)$ using a Poisson probability density function constrained by the tuning curves measured from the four stimulus condition data, the MAP estimate predicted saccade choices equally well at 81.64% (3294/4035) of all trials (third bar; Figure 6a). The MAP with Gaussian tuning curves and WTA had identical results, both correctly predicting saccade choices in 71.11% (2870/4035) of all trials (fourth and fifth bar; Figure 6a). When we used the data from the single target condition for the tuning curves and the Poisson probability density, MAP predicted saccade choices in 69.94% (2822/4035) of the trials. This was a drop in model performance by 14.82% compared to the best MAP prediction (69.94% versus 84.76%%; first bar and sixth bar, Figure 6a). The model performance degraded further when we determined $P(r|s)$ from the non-parametric density function built from the data collected during performance of the single target condition. In this case, 58.64% (2366/4035) of all trials were predicted from the trials (seventh bar, Figure 6a).

When the OLE was optimized using the data from the four stimuli condition it predicted saccade choice for 69.47% (2803/4035) of all trials (eighth bar, Figure 6a). This result is about as good as the WTA (69.47% versus 71.11%). However, when the OLE was optimized using the data recorded during the single target condition, its performance dropped to 60.25% (2431/4035; ninth bar Figure 6a). As expected, the PVA performed least well, predicting 55.71% (2248/4035) of all trials (tenth bar, Figure 6a). Taken together these results point toward two important conclusions. First, the model of the population code from the SC build up neurons that can be used to predict choice improves when it combines all the information about the neurons in the population such as their full tuning curves and not just the peak of the tuning curve as is the case for the traditional population vector average. Second, a critical aspect that determines the performance of the model of the population code is the distribution from which $P(r|s)$ is drawn. The reason that the accuracy of the models using the single target neuron data performed so poorly is because these data are not accurate characterizations of $P(r|s)$ in the four possible target condition.

The Venn diagram in Figure 6b shows the percentages of trials predicted by each model and how the predictions overlapped. Determining whether there are substantial numbers of trials predicted exclusively by one model rather than another provides important information about the computational principles underlying population coding and choice in the SC. We selected three models that showed the best prediction accuracy: the MAP with non-parametric PDF estimation from the four stimuli condition and the non-uniform prior, the OLE optimized with the four stimuli condition and the WTA. Saccade choices were predicted successfully in 92.0% of all trials by any of the models. 8.0% of all trials were unpredictable by any of the models. 53.5% of all trials fell into the intersection of the three models - that is each model did a good job at predicting saccade choice. The MAP predicted 6.7% of trials exclusively whereas OLE predicted 2.1% of trials exclusively and WTA predicted 3.2% of trials exclusively. Thus, although all models perform reasonably well, the probabilistic model overall performs slightly better than both of the others. Future experiments with target locations represented by overlapping neuronal RFs will provide further and better tests of the ability of MAP estimates to predict choices relative to these other models.

The posterior probability distribution scales with choice performance accuracy

Previously we showed that for buildup neurons encoding targets and distractors the levels of activity scaled with performance accuracy (Kim and Basso, 2008). When performance accuracy was high the differences in discharge rates between target and distractor neurons were

highly discriminable. When performance accuracy was poor, the differences in discharge rates were less discriminable. In light of this, we were also interested in determining whether the posterior probability from the MAP model and the angular difference from the OLE varied with the variability in behavioral performance. Since these two models are based on combining activity from multiple neurons, we expected that the output of these models would scale with performance accuracy. In the case of the OLE, we expected to see a small angular difference between the population vector and the saccade choice when performance accuracy was high. We expected to see a larger angular difference when performance accuracy was poor. For the MAP we expected to see the peak of the posterior distribution centered on the saccade choice with a higher probability when performance accuracy was high. When performance accuracy was low, we still expected to see the peak of the posterior distribution centered on the saccade choice but with a lower probability. Note that the area under the distribution would remain 1, but the relative probabilities associated with each saccade choice would differ with choice accuracy.

To explore the relationship between the variability of behavioral performance and the model predictions we sorted all of the correct trials (n = 3317) from the 30 data sets into three bins of performance accuracy, <70% correct (n=737), 70–99% correct (n=1255) and 100% correct (n=1325). We then fed the neuronal discharge data measured from 28 to 8ms before the onset of the saccade from these sorted trials into both the MAP and the OLE models. Figure 7a shows the angular difference in the directions of the population vector and the four possible saccade choices. Because we sorted these trials from only correct trials for this analysis, the target direction shows a much smaller angular difference than any of the other locations. However, when we compared the angular difference for the target location across the different performance conditions, there were small changes in the difference but they failed to reach statistical significance (ANOVA; $F(2,117) = 1.02$, $p=0.363$; Figure 7a cf., black and gray lines). In contrast to the OLE, the posterior distribution showed scaling with performance accuracy and these differences were statistically significant (Figure 7b cf., black and grey lines; ANOVA $F(2,117) = 3.98$, $p<0.05$). Previously, we found that the activity of SC neurons encodes saccade choice as well as the certainty of the choice (Basso and Wurtz, 1997, 1998; Kim and Basso, 2008). The findings described here corroborate and extend that result showing that the posterior distribution recovered from combining likelihoods obtained from SC neuronal activity and prior information predicts saccade choices and scales with choice accuracy. Inherent in the posterior distribution is the certainty of the choice indicated by the performance accuracy.

In one implementation of the MAP model we used a uniform prior probability distribution. Because of this, the result of this model is mathematically equivalent to that of likelihood models (Sanger, 2002; Sanger, 2003) most recently implemented in MT (Jazayeri and Movshon, 2006). Figure 7c plots the log likelihood distributions for each of the three performance conditions so that we could compare directly the MAP to the log likelihood. Although not intuitive, the distributions shown in Figure 7c reveal a pattern. When performance accuracy was low the distribution did not have a clear peak (Figure 7c, lightest gray line). Whereas when performance accuracy was high, the log likelihood distribution was centered on the saccade choice (Figure 7c, black line). However, because these are likelihoods and not bona fide probability distributions, they cannot be compared directly across conditions unless they are normalized. Note that the light grey line is for the poorer performance trials even though it has the highest overall log likelihood. A difference between the maximum and minimum log likelihood can be taken for this purpose. Using a posterior distribution however, which is a bona fide probability distribution, this difference step is not required.

The posterior probability distribution develops over time

Given that decisions and saccade choices likely develop over time (Carpenter and Williams, 1995; Gold and Shadlen, 2000; Gold and Shadlen, 2007), we extended our analysis to determine whether the information encoded in the posterior probability and the OLE developed over time to reflect a single saccade choice. For this, we aligned the trials on the time beginning 50ms (average SC visual latency) after the onset of the stimulus array and computed the model result for each 1ms until the saccade onset. Figure 8a plots the angular difference between the neuronal and saccade direction measured using the OLE. Initially, the angular difference fluctuated and began settling on a small difference value ~100ms after the onset of the stimulus array. Consistent with the results of the stationary analysis shown in Figure 7a, the dynamic OLE prediction did not scale very well with performance accuracy. In contrast, the developing choice as encoded by the posterior was more obvious in the MAP estimate. Figure 8b shows the developing MAP estimate of the saccade choice in the three performance accuracy conditions (<70%, 70~99% and 100%). The MAP estimate developed rapidly in the 100% performance accuracy condition and reached a maximum probability of 0.95 at the mean time of the saccade onset (Figure 8b, black line). In the 70–90% and the <70% performance accuracy conditions, the MAP estimate rose less rapidly and reached a probability of 0.90 and 0.83 respectively, at the mean time of saccade onset (Figure 8b, dark grey and grey lines). The differences in the peak probability at the time of the mean saccade onset were statistically significant (ANOVA, $F(2,117) = 5.15$, $p < 0.01$). This result is interesting because it reveals that across the different performance accuracy conditions, the height of the posterior distribution is lower for the poorer performance trials (Figure 8b light grey line) than for the better performance trials (Figure 8b black line). In each case monkeys made the correct saccade choice. This is evidence that the decision or choice threshold varies for these different performance accuracy trials (Hanes and Schall, 1996; Parè and Hanes, 2003).

Figure 9 shows the evolution of the posterior for the target location (TG) and for the three distractor locations (D1 D2 D3) using both the uniform prior (Figure 9a,c and e) and the non-uniform prior (Figure 9b,d and f). In each panel the MAP estimate for the saccade choice diverged from the MAP estimates for the distractors locations as saccade onset approached (Figure 9 c,f., thick, thin, dashed and dotted lines). As previously reported in LIP (Beck et al. 2008), these results show the development of the posterior probability of the saccade choice and the decrease of the posterior probabilities of the distractors over time. The results show for the first time that when the non-uniform prior function is applied to recover the MAP estimate, the posterior distribution favors the saccade choice slightly, even before the stimulus array appears (cf., Figure 9a, c, e and Figure 9b, d, f). Finally, as would be predicted from a structure that signals the saccade choice to be made, the posterior distribution almost collapses around the saccade choice at the time of the saccade. This is evident from the MAP estimates because the saccade choice probability is close to 1.0 and the probability for the distractors is close to 0. This behavior is unclear from the dynamic analysis of the OLE shown in Figure 8a.

DISCUSSION

In this report we show for the first time that the relationship between SC buildup neuronal population activity and saccade choice is well-described by a probabilistic scheme. Here, we considered SC buildup neurons as encoders of likelihood functions. When the likelihood were combined with prior information we could construct a posterior distribution over four possible saccade choices whose maximum predicted saccade choices well. Somewhat astonishing to us, was that combining the activity of only four simultaneously recorded neurons and recovering the maximum value of the posterior distribution predicted saccade choices accurately on as many as 84.76% of the trials. The MAP outperformed two well-known algorithms, population vector average (PVA/OLE) and the winner-takes-all (WTA). For the

first time, we also showed that the posterior distribution across saccade choices develops over time and reached a maximum around the saccade choice at the time of the saccade onset. The posterior distribution almost collapses around the saccade choice at the time of the movement, as would be predicted for a structure so close to the motoneurons (Miyashita and Hikosaka, 1996). We also showed that computing the posterior distribution across saccade choices by estimating the likelihood directly from the data using a probability density estimation procedure maximizes performance of the model. However, consistent with theoretical predictions, the MAP estimates recovered using the empirical PDF were little improved over those using a Poisson distribution to characterize variability. Finally, by incorporating a non-uniform prior, we found that MAP estimates improved by ~3% across all trials. This suggests that incorporating prior information with the likelihood information provided by SC buildup neurons is a way population neuronal activity could encode saccade choices. In what follows we first describe how these new results extend previous work on the SC and saccade choice. Then we describe how our MAP model complements and extends previous work on probabilistic approaches to sensory encoding and decision-making.

Relationship to previous work in SC

It is well-accepted that the SC employs a population code to determine saccadic eye movements (McIlwain, 1986; McIlwain, 1991). Experimental evidence shows that individual neurons encode saccade vectors, with each neuron having broad tuning for particular saccade directions and amplitudes (Robinson, 1972; Schiller and Stryker, 1972; Wurtz and Goldberg, 1972; Sparks, 1975, 1978). A weighted sum of the activity of SC neurons across the map is considered to determine the saccade direction(Ottes et al., 1986; Van Gisbergen et al., 1987; Lee et al., 1988; Quaia et al., 1998; Groh, 2001; Hanes and Wurtz, 2001) in much the same way that motor cortical neurons encode the direction of arm movements (Georgopoulos et al., 1986). However, the experiments leading to this conclusion for the SC are based largely on simulations or were performed using only a single saccade target (Lee et al., 1988; Quaia et al., 1998; Hanes and Wurtz, 2001).

Recent experiments in the saccadic system using more complex displays such as when multiple visual stimuli appear, suggest that a WTA strategy is used (Port and Wurtz, 2003; McPeek and Keller, 2004). Thus, we are left with the conclusion that for single targets the SC operates using PVA whereas for multiple stimuli the SC operates as WTA. This conundrum is evident from behavioral studies too. For example, it is well-known that when two visual stimuli appear in close proximity, saccades land in a location between the two stimuli - a phenomenon called the global effect or averaging saccades (Findlay, 1982; Glimcher and Sparks, 1993; Kowler and Blaser, 1995; Edelman and Keller, 1998; McGowan et al., 1998; Melcher and Kowler, 1999). However, if the targets appear further apart or more time is provided, a saccade can be made to one or the other stimulus (Ottes et al., 1984). This phenomenon is not unique to the SC. In the middle temporal area (MT) electrical stimulation and recording experiments support PVA or WTA or both, leading to the idea that perceptual decisions rely on a WTA scheme whereas movement decisions rely on a PVA scheme (Salzman et al., 1990; Salzman et al., 1992; Ferrera and Lisberger, 1997; Groh et al., 1997; Recanzone et al., 1997; Britten and Heuer, 1999; Churchland and Lisberger, 2001). Population coding schemes that are task or time-dependent require mechanisms to switch between them. How this switch would be implemented biologically is unclear.

Advantages of probabilistic schemes for understanding action choice

Probabilistic strategies offer a solution to this conundrum. Because the posterior distribution uses all of the information contained within the tuning curve and it combines activity across all tuning curves, it naturally represents multiple stimuli simultaneously. Furthermore, because the peak and the variance of tuning curves (when considered as likelihood functions - or the

estimated probability density functions) signal the certainty of the encoded parameter, the posterior distribution provides a normalized likelihood (conditional probability) for each of the alternatives. This eliminates the need for a switch between population coding schemes. Both PVA and WTA use only the peak activity. WTA further disregards information from distractor neurons, leaving much of the information in the population activity unused. Even when more information is provided for the PVA as in the case of the OLE, it still predicts fewer saccade choices as accurately as the MAP. This is because the OLE does not incorporate variability as does the MAP model. The OLE however, was almost as good as the MAP with the Gaussian tuning curves which in turn, was identical to the WTA. This results because the correlation matrix that we used to optimize the vector estimation was determined from neuronal activities that were largely non-overlapping. Only the empirical probability density functions appear to represent the true population variability. Thus we see improvement from the OLE (69.47%) and WTA (71.11%) to the MAP with the empirical PDF (81.88%). As a result of ignoring much of the information in the population, variations in behavior, uncertainty, or even attentional modulation (Spitzer et al., 1988; McAdams and Maunsell, 1999; Pouget et al., 1999) cannot be resolved using WTA or PVA/OLE approaches. Future experiments pushing the amount of overlap in the visual display, the RF of the recorded neurons and thus the empirical PDFs, will further distinguish these models.

Probabilistic approaches have the additional advantage in that they are easily extended to the domain of decision-making (Smith and Ratcliff, 2004; Gold and Shadlen, 2007; Beck et al., 2008). Current models of decision-making or eye movement selection rely on taking the difference of activity from two populations of neurons representing independent alternatives (Shadlen and Newsome, 2001; Roitman and Shadlen, 2002; Ratcliff et al., 2003; Huk and Shadlen, 2005; Boucher et al., 2007; Ratcliff et al., 2007). Extending these models to more than two choices or to continuous decisions as opposed to discrete choices, is difficult (Ratcliff et al., 2007; Beck et al., 2008; Churchland et al., 2008; Niwa and Ditterich, 2008). Furthermore, these models have difficulty incorporating changing evidence ‘on the fly’. A neuronal population representing the posterior distribution across choices has the unique advantage of naturally representing multiple possibilities, discrete or continuous variables, the ability for recursive computation (our Figure 8, Montagnini et al., 2007) and incorporation of prior evidence. These features make changing a mind ‘on the fly’ seamless.

We think of the posterior distribution as describing the probability of a saccade choice. We interpret our result as providing evidence that the saccade choice may be encoded as the posterior probability distribution across all possible saccade choices or the uncertainty associated with each of the possible saccade options. Whereas conceptual models of the decision-process suggest that uncertainty underlying decisions occurs primarily within the sensory system (Sugrue et al., 2005), our results suggest that there is uncertainty associated with the choice of action that may be distinguished from sensory uncertainty.

Relationship to previous and current models

Here we tested some of the theoretical assumptions of similar probabilistic models. We made a number of important findings. First, assuming a Poisson probability distribution to characterize variability to estimate likelihood is as good as determining the likelihood directly from empirical probability density estimation. Therefore, even though individual neuronal variability is not well-explained by Poisson statistics, a MAP estimate based on Poisson statistics performs well (Ma et al., 2006; Beck et al., 2007). Second, vonMises or Gaussian tuning may not be the best representation of neuronal activity across the population. We show here that vonMises tuning curves (logs of von Mises curves are cosine functions) are not required for the MAP to perform well. Similarly, using Gaussian tuning curves resulted in model performance that was only as good as WTA. Estimating the likelihoods empirically

improved model performance. This result is consistent with results seen in the arm movement literature (Amirikian and Georgopoulos, 2000; Serruya et al., 2002; Taylor et al., 2002; Carmena et al., 2003). That von Mises tuning curves and the minimally-overlapping Gaussian tuning curves result in performance identical to WTA is expected since these approaches minimize or negatively weight (in the case of cosine functions) the activity of distractor neurons. Indeed, in recent work extending a Bayesian approach to decision-making (Beck et al. 2008), they assumed von Mises curves (the logs of which are cosine functions) for the LIP neurons. We suspect they would obtain the same results as they did if they also implemented a WTA scheme.

Third, we implemented both a uniform and a non-uniform prior probability distribution. Therefore, the MAP estimate in the former case is a Bayesian estimator that is identical to a maximum likelihood estimator and the MAP estimate in the latter case is a Bayesian estimator distinct from a maximum likelihood. Although the addition of the prior did not improve the model enormously, we suspect this is because in our task it was not to the monkeys' advantage to use a prior. Consistent with this, we found more improvement for error trials than for correct trials when the non-uniform prior was added.

In net, we hypothesize that SC buildup neurons encode likelihood functions as demonstrated here. These in turn may be integrated by burst neurons in SC together with priors (or biases) from other sources to determine the posterior probability distribution across saccade choices. We propose that using *bona fide* probability density functions (such as the empirical PDF) for characterizing $P(r|s)$ and encoding saccade choice as a posterior distribution across saccade choices, is a simple way for the brain to represent and compute choices within the population (Deneve et al., 1999). Since the peak of the posterior would be associated with the highest chance of activating the downstream neurons responsible for driving a particular saccade, decoding the choice is implicit in the population activity. Furthermore, the width of the posterior distribution is an implicit way to represent the uncertainty of the choice.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH EY13692 (MAB). We also acknowledge the support of NCRR P51 RR000167 to the Wisconsin National Primate Research Center. We are especially grateful for invaluable discussions with Drs. Alexander Pouget, Anthony Movshon and Lance Optican. We thank Drs. Merhdad Jazayeri, Alexander Grunewald and Jochen Ditterich for comments on previous aspects of this work, Dr. Emilio Salinas for sharing matlab code for computing the OLE and the two anonymous reviewers for their helpful critiques.

REFERENCES

Amirikian B, Georgopoulos AP. Directional tuning profiles of motor cortical cells. Neuroscience Research 2000;36:73–79. [PubMed: 10678534]

Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. Nature Reviews Neuroscience 2006;7:358.

Basso MA, Wurtz RH. Modulation of neuronal activity by target uncertainty. Nature 1997;389:66–69. [PubMed: 9288967]

Basso MA, Wurtz RH. Modulation of neuronal activity in superior colliculus by changes in target probability. Journal of Neuroscience 1998;18:7519–7534. [PubMed: 9736670]

Beck, J.; Ma, WJ.; Latham, PE.; Pouget, A. Probabilistic population codes and the exponential family of distributions. In: Cisek, P.; Drew, T.; Kalaska, JF., editors. Progress in Brain Research. Elsevier; 2007. p. 509-519.

Beck JM, Ma WJ, Kiani R, Hanks T, Churchland AK, Roitman J, Shadlen MN, Latham PE, Pouget A. Probabilistic population codes for Bayesian decision making. Neuron 2008;60:1142–1152. [PubMed: 19109917]

Boucher L, Palmeri TJ, Logan GD, Schall JD. Inhibitory control in mind and brain: an interactive race model of countermanding saccades. Psychological Review 2007;114:376–397. [PubMed: 17500631]

Britten KH, Heuer HW. Spatial summation in receptive fields of MT neurons. Journal of Neuroscience 1999;19:5074–5084. [PubMed: 10366640]

Carmena JM, Lebedev MA, Crist RE, O'Doherty JE, Santucci DM, Dimitrov DF, Patil PG, Henriquez CS, Nicolelis MAL. Learning to control a brain-machine interface for reaching and grasping by primates. PLoS Biology 2003;1(2):e42. [PubMed: 14624244]

Carpenter R, Williams M. Neural computation of log likelihood in control of saccadic eye movements. Nature 1995;377:59–62. [PubMed: 7659161]

Churchland AK, Kiani R, Shadlen MN. Decision-making with multiple alternatives. Naturet Neuroscience 2008;11:693–702.

Churchland MM, Lisberger SG. Shifts in the population response in the middle temporal visual area parallel perceptual and motor illusions produced by apparent motion. Journal of Neuroscience 2001;21:9387–9402. [PubMed: 11717372]

Crist CF, Yamasaki DSG, Komatsu H, Wurtz RH. A grid system and a microsyringe for single cell recording. Journal of Neuroscience Methods 1988;26:117–122. [PubMed: 3146006]

Deneve S, Latham P, Pouget A. Reading population codes: a neural implementation of ideal observers. Nature Neuroscience 1999;2(8):740–745.

Dorris MC, Munoz DP. Saccadic probability influences motor preparation signals and time to saccadic initiation. Journal of Neuroscience 1998;18:7015–7026. [PubMed: 9712670]

Edelman JA, Keller EL. Dependence on target configuration of express saccade-related activity in the primate superior colliculus. Journal of Neurophysiology 1998;80:1407–1426. [PubMed: 9744949]

Feldman JA. Dynamic connections in neural networks. Biological Cybernetics 1982;46:27–39. [PubMed: 6307398]

Ferrera VP, Lisberger SG. Neuronal responses in visual areas MT and MST during smooth pursuit target selection. Journal of Neurophysiology 1997;78:1433–1446. [PubMed: 9310434]

Findlay JM. Global visual processing for saccadic eye movements. Vision Research 1982;22:1033–1045. [PubMed: 7135840]

Földiak, P. Computation and Neural Systems. Eeckman, F.; Bower, J., editors. Norwell, MA: Kluwer Academic Publishers; 1993. p. 55-60.

Fuchs AF, Robinson DA. A method for measuring horizontal and vertical eye movement chronically in the monkey. Journal of Applied Physiology 1966;21:1068–1070. [PubMed: 4958032]

Georgopoulos AP, Schwartz AB, Kettner RE. Neuronal population coding of movement direction. Science 1986;233:1416–1419. [PubMed: 3749885]

Glimcher PW, Sparks DL. Representation of averaging saccades in the superior colliculus of the monkey. Experimental Brain Research 1993;95:429–435.

Gold JI, Shadlen MN. Representation of a perceptual decision in developing oculomotor commands. Nature 2000;404:390–394. [PubMed: 10746726]

Gold JI, Shadlen MN. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. Neuron 2002;36:299. [PubMed: 12383783]

Gold JI, Shadlen MN. The neural basis of decision making. Annual Review of Neuroscience 2007;30:535–574.

Groh JJ. Converting neural signals from place codes to rate codes. Biological Cybernetics 2001;85:159–165. [PubMed: 11561817]

Groh JM, Born RT, Newsome WT. How is a sensory map read out? effects of microstimulation in visual area MT on saccades and smooth pursuit eye movements. Journal of Neuroscience 1997;17:4312–4330. [PubMed: 9151748]

Hanes DP, Schall JD. Neural control of voluntary movement initiation. Science 1996;274:427–430. [PubMed: 8832893]

Hanes DP, Wurtz RH. Interaction of the frontal eye field and superior colliculus for saccade generation. Journal of Neurophysiology 2001;85:804–815. [PubMed: 11160514]

Hays AV, Richmond BJ, Optican LM. A UNIX-based multiple process system for real-time data acquisition and control. WESCON Conf Proc 1982;2:1–10.

Horwitz GD, Newsome WT. Target selection for saccadic eye movements: prelude activity in the superior colliculus during a direction-discrimination task. Journal of Neurophysiology 2001;86:2543–2558. [PubMed: 11698541]

Horwitz GD, Batista AP, Newsome WT. Representation of an abstract perceptual decision in macaque superior colliculus. Journal of Neurophysiology 2004;91:2281. [PubMed: 14711971]

Huk AC, Shadlen MN. Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. Journal of Neuroscience 2005;25:10420–10436. [PubMed: 16280581]

Jazayeri M, Movshon JA. Optimal representation of sensory information by neural populations. Nature Neuroscience 2006;9:690.

Judge SJ, Richmond BJ, Chu FC. Implantation of magnetic search coils for measurement of eye position: An improved method. Vision Research 1980;20:535–538. [PubMed: 6776685]

Kim B, Basso MA. Saccade target selection in the superior colliculus: a signal detection theory approach. Journal of Neuroscience 2008;28:2991–3007. [PubMed: 18354003]

Kowler E, Blaser E. The accuracy and precision of saccades to small and large targets. Vision Research 1995;35:1741–1754. [PubMed: 7660582]

Lee C, Rohrer WH, Sparks DL. Population coding of saccadic eye movements by neurons in the superior colliculus. Nature 1988;332:357–360. [PubMed: 3352733]

Li X, Basso MA. Competitive stimulus interactions within single response fields of superior colliculus neurons. Journal of Neuroscience 2005;25:11357–11373. [PubMed: 16339031]

Li X, Basso MA. Preparing to move increases the sensitivity of superior colliculus neurons. Journal of Neuroscience 2008;28:4561–4577. [PubMed: 18434535]

Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. Nature Neuroscience 2006;9:1432.

MacPherson JM, Aldridge JW. A quantitative method of computer analysis of spike train data collected from behaving animals. Brain Research 1979;175:183–187. [PubMed: 487149]

McAdams CJ, Maunsell JHR. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. Journal of Neuroscience 1999;19:431–441. [PubMed: 9870971]

McGowan JW, Kowler E, Sharma A, Chubb C. Saccadic localization of random dot targets. Vision Research 1998;38:895. [PubMed: 9624439]

McIlwain JT. Point images in the visual system: New interest in an old idea. Trends in Neuroscience 1986;9:354–358.

McIlwain JT. Distributed spatial coding in the superior colliculus: A review. Visual Neuroscience 1991;6:3–13. [PubMed: 2025610]

McPeek RM, Keller EL. Saccade target selection in the superior colliculus during a visual search task. Journal of Neurophysiology 2002;88:2019–2034. [PubMed: 12364525]

McPeek RM, Keller EL. Deficits in saccade target selection after inactivation of superior colliculus. Nature Neuroscience 2004;7:757–763.

Melcher D, Kowler E. Shapes, surfaces and saccades. Vision Research 1999;39:2929. [PubMed: 10492819]

Miyashita N, Hikosaka O. Minimal synaptic delay in the saccadic output pathway of the superior colliculus studied in awake monkey. Experimental Brain Research 1996;112:187–196.

Montagnini A, Mamassian P, Perrinet L, Castet E, Masson GS. Bayesian modeling of dynamic motion integration. Journal of Physiology-Paris 2007;101:64.

Moschovakis AK, Scudder CA, Highstein SM. The microscopic anatomy and physiology of the mammalian saccadic system. Progress in Neurobiology 1996;50:133–254. [PubMed: 8971981]

Munoz DP, Wurtz RH. Saccade-related activity in monkey superior colliculus. I. Characteristics of burst and buildup cells. Journal of Neurophysiology 1995;73:2313–2333. [PubMed: 7666141]

Niwa M, Ditterich J. Perceptual decisions between multiple directions of visual motion. Journal of Neuroscience 2008;28:4435–4445. [PubMed: 18434522]

Optican LM, Richmond BJ. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. Journal of Neurophysiology 1987;57:162–178. [PubMed: 3559670]

Oram MW, Földiák P, Perrett DI, Oram MW, Sengpiel F. The 'Ideal Homunculus': decoding neural population signals. Trends in Neurosciences 1998;21:259. [PubMed: 9641539]

Ottes FP, Van Gisbergen JAM, Eggermont JJ. Metrics of saccade responses to visual double stimuli: two different modes. Vision Res 1984;24:1169–1179. [PubMed: 6523740]

Ottes FP, Van Gisbergen JAM, Eggermont JJ. Visuomotor fields of the superior colliculus: a quantitative model. Vision Res 1986;26:857–873. [PubMed: 3750869]

Parè M, Hanes DP. Controlled movement processing: superior colliculus activity associated with countermanded saccades. Journal of Neuroscience 2003;23:6480. [PubMed: 12878689]

Port NL, Wurtz RH. Sequential activity of simultaneously recorded neurons in the superior colliculus during curved saccades. Journal of Neurophysiology 2003;90:1887–1903. [PubMed: 12966180]

Pouget A, Deneve S, Ducom J-C, Latham P. Narrow vs wide tuning curves: what's best for a population code? Neural Computation 1999;11:85–90. [PubMed: 9950723]

Quaia C, Aizawa H, Optican LM, Wurtz RH. Reversible inactivation of monkey superior colliculus: II. Maps of saccadic deficits. Journal of Neurophysiology 1998;79:2097–2110. [PubMed: 9535971]

Ratcliff R, Cherian A, Segraves M. A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. Journal of Neurophysiology 2003;90:1392. [PubMed: 12761282]

Ratcliff R, Hasegawa YT, Hasegawa RP, Smith PL, Segraves MA. Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. Journal of Neurophysiology 2007;97:1756–1774. [PubMed: 17122324]

Recanzone GH, Wurtz RH, Schwarz U. Responses of MT and MST neurons to one and two moving objects in the receptive field. Journal of Neurophysiology 1997;78:2904–2915. [PubMed: 9405511]

Robinson DA. Eye movements evoked by collicular stimulation in the alert monkey. Vision Research 1972;12:1795–1808. [PubMed: 4627952]

Roitman JD, Shadlen MN. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. Journal of Neuroscience 2002;22:9475–9489. [PubMed: 12417672]

Salinas E, Abbott LF. Vector reconstruction from firing rates. Journal of Computational Neuroscience 1994;1:89–107. [PubMed: 8792227]

Salzman CD, Britten KH, Newsome WT. Cortical microstimulation influences perceptual judgements of motion detection. Nature 1990;346:174–177. [PubMed: 2366872]

Salzman CD, Murasugi CM, Britten KH, Newsome WT. Microstimulation in visual area MT: effects on direction discrimination performance. Journal of Neuroscience 1992;12:2331–2355. [PubMed: 1607944]

Sanger TD. Probability density estimation for the interpretation of neural population codes. Journal of Neurophysiology 1996;76:2790–2793. [PubMed: 8899646]

Sanger TD. Decoding neural spike trains: calculating the probability that a spike train and an external signal are related. Journal of Neurophysiology 2002;87:1659–1663. [PubMed: 11877538]

Sanger TD. Neural population codes. Current Opinion in Neurobiology 2003;13:238–249. [PubMed: 12744980]

Scherberger H, Andersen RA. Target selection signals for arm reaching in the posterior parietal cortex. Journal of Neuroscience 2007;27:2001–2012. [PubMed: 17314296]

Schiller PH, Koerner F. Discharge characteristics of single units in superior colliculus of the alert rhesus monkey. Journal of Neurophysiology 1971;34:920–936. [PubMed: 4999593]

Schiller PH, Stryker M. Single-unit recording and stimulation in superior colliculus of the alert rhesus monkey. Journal of Neurophysiology 1972;35:915–924. [PubMed: 4631839]

Schwartz AB, Kettner RE, Georgopoulos AP. Primate motor cortex and free arm movements to visual targets in three-dimensional space. I. Relations between single cell discharge and direction of movement. Journal of Neuroscience 1988;8:2913–2927. [PubMed: 3411361]

Scott, DW. Theory practice and visualization. New York: John Wiley and Sons, Inc; 1992. Multivariate density estimation.

Serruya MD, Hatsopoulos NG, Paninski L, Fellows MR, Donoghue JP. Instant neural control of a movement signal. Nature 2002;416:141–142. [PubMed: 11894084]

Shadlen MN, Newsome WT. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. Journal of Neurophysiology 2001;86:1916. [PubMed: 11600651]

Smith PL, Ratcliff R. Psychology and neurobiology of simple decisions. Trends in Neurosciences 2004;27:161–168. [PubMed: 15036882]

Sparks DL. Response properties of eye movement-related neurons in the monkey superior colliculus. Brain Research 1975;90:147–152. [PubMed: 1131686]

Sparks DL. Functional properties of neurons in the monkey superior colliculus: coupling of neuronal activity and saccade onset. Brain Research 1978;156:1–16. [PubMed: 100173]

Sparks DL. Translation of sensory signals into commands for control of saccadic eye movements: Role of primate superior colliculus. Physiological Reviews 1986;66:118–171. [PubMed: 3511480]

Spitzer H, Desimone R, Moran J. Increased attention enhances both behavioral and neuronal performance. Science 1988;240:338–340. [PubMed: 3353728]

Sugrue LP, Corrado GS, Newsome WT. Choosing the greater of two goods: neural currencies for valuation and decision-making. Nature Reviews Neuroscience 2005;6:363–375.

Taylor DM, Tillery SIH, Schwartz AB. Direct Cortical Control of 3D Neuroprosthetic Devices. Science 2002;296:1829–1832. [PubMed: 12052948]

Van Gisbergen JAM, Van Opstal AJ, Tax AAM. Collicular ensemble coding of saccades based on vector summation. Neuroscience 1987;21:541–555. [PubMed: 3614643]

Wurtz RH, Goldberg ME. Activity of superior colliculus in behaving monkey: III. Cells discharging before eye movements. Journal of Neurophysiology 1972;35:575–586. [PubMed: 4624741]

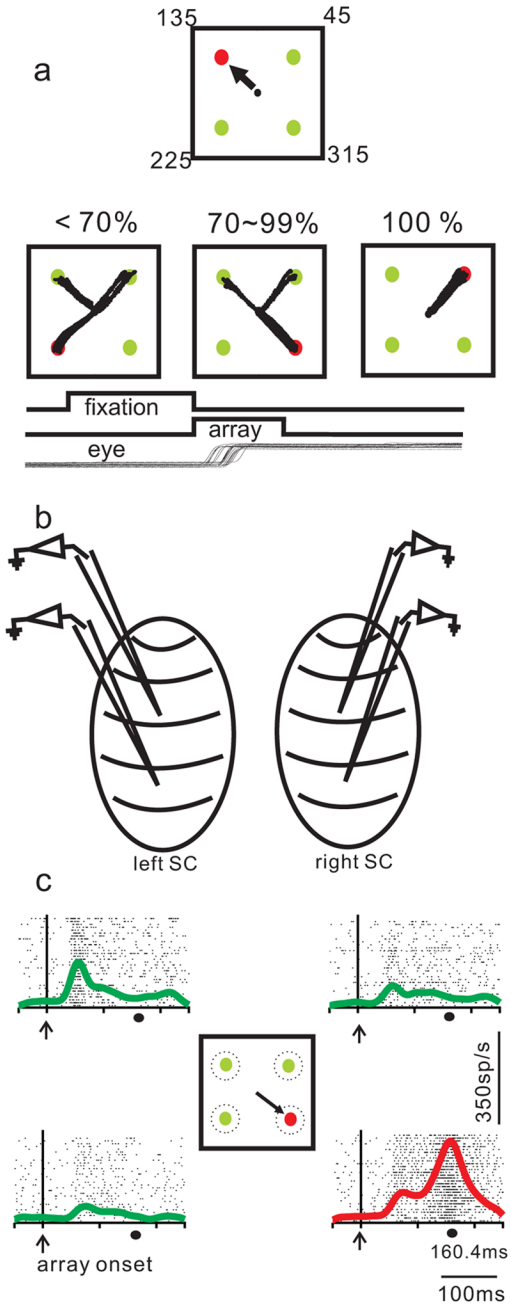

Figure 1.

Saccade choices in a target selection task. a. Each square shows the spatial arrangement of the task. The spot that is colored differently indicates the target whereas the same colored spots indicate distractors (either green among red or red among green). In this example, the target is red. The black circle in the center of each square indicates the location of the fixation spot. In the task, the array appeared and the fixation spot disappeared simultaneously, cueing monkeys to make a saccade to the target (black arrow). The black lines in each of the three bottom squares show actual saccade trajectories. Each of the three squares show examples of choice performance of <70%, 70–99% and 100% correct. Note that all spots were normalized to 45°, 135°, 225° and 315° positions for display. The temporal sequence of the task appears as labeled,

black lines below the spatial depiction. b. Schematics of the two SCs and the electrodes for multiple neuron recordings. c. Examples of four neurons recorded simultaneously from the SC. Each tick in a row indicates the time of occurrence of an action potential and each row indicates a trial. The four raster plots correspond to the neurons with RF centers located at the positions shown in the center square. The spike density function (SDF; σ = 10ms) is superimposed on the raster plots. Green lines show SDFs for distractor neurons and red lines show SDFs for target neurons. The mean reaction time of the saccade is shown by the black circle at the bottom of the traces labeled 160.4ms. Each trace is aligned to the onset of the array marked by the upward arrow, vertical line and labeled “array onset”.

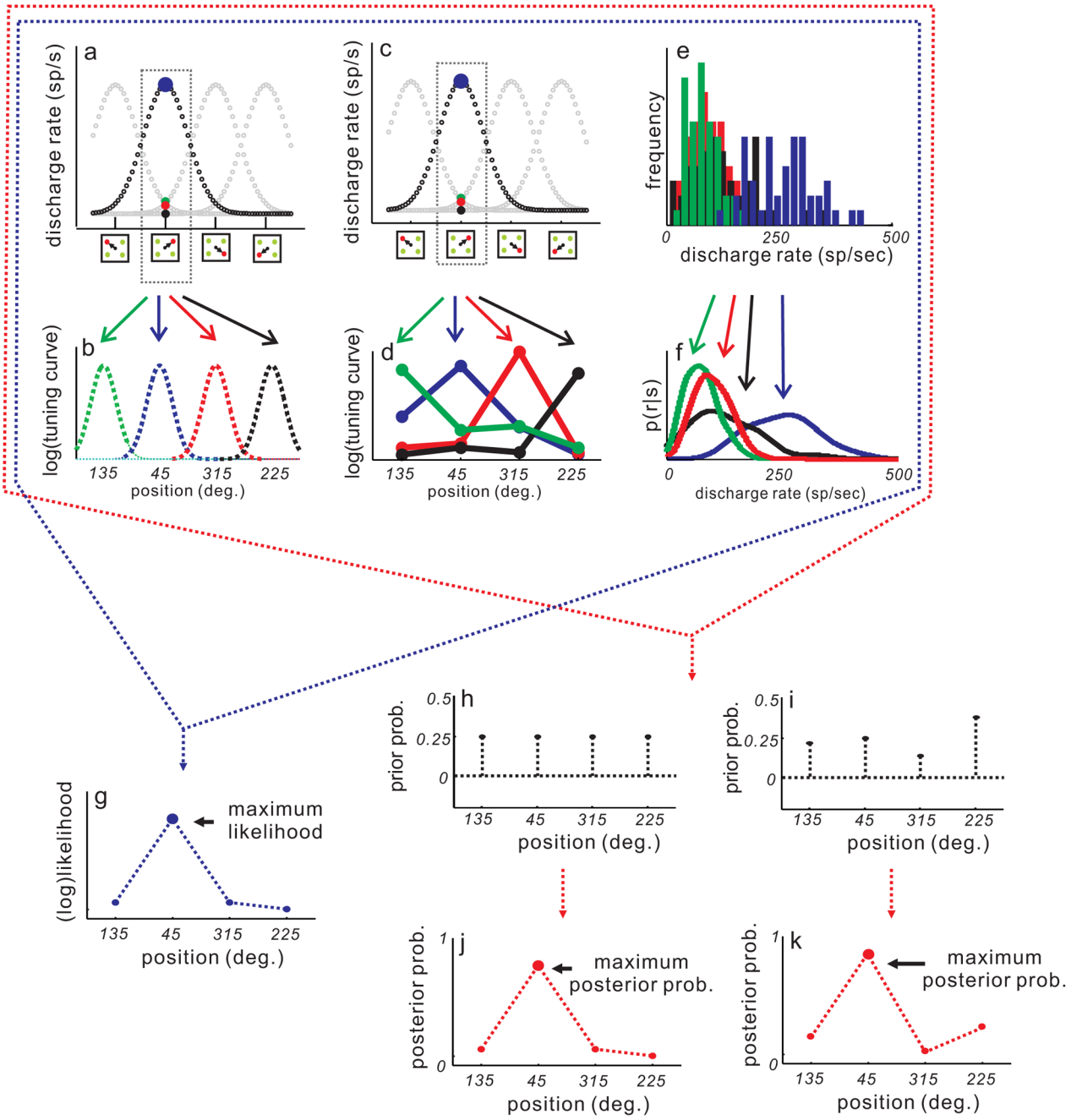

Figure 2.

Computing the posterior distribution across saccade choices. a. Simulated, idealized responses (sp/s) of four SC neurons in the four different target conditions of the task. Small squares on the abscissa show the four possible stimulus configurations of the task. The four curves simulate the activity of each neuron in each of the four possible conditions. For display, these data were simulated as Gaussian functions with σ = 40°. Each Gaussian peak was shifted by 90° to represent the four possible locations. Each simulated response shows a peak discharge when the target corresponds to the preferred location for the neuron and shows a reduced level of discharge as the target moves away from the preferred location. The dotted rectangle around the black simulated response shows an example response whose peak discharge (blue circle)

appears for targets located at 45° and a reduced response for target locations elsewhere. The green dot, red dot and black dot show the simulated discharge rate for the other three neurons when the target appears outside of their RF. b. Simulated weights for SC neurons. Each tuning curve represents the probability of neuronal discharge given a stimulus position from 0 – 359° (in steps of 0.1°). The tuning curves were simulated as Gaussian functions of equal amplitude (arbitrary units) and widths of 20.6° (Edelman and Keller, 1998). The logs of these Gaussian functions were used as weights. The green curve shows the weight for the neurons with the highest discharge rate in the 135° target condition. The blue curve shows the weight for the neurons with peak rates in the 45° target condition. The red and cyan curves show weights for the neurons with peak rates in the 315° and 225° target conditions respectively. c. The same as in a. d. Each point is the measured discharge from an SC neuron when monkeys made saccades to one of each of the four possible locations. e. Frequency histogram of discharge rates measured from the sets of four neurons in the four possible target conditions across all data sets. The activity of target neurons is shown in the blue bars whereas the activity of distractor neurons is shown in the green, red and black bars. These histograms formed the basis for the density estimation (see **EXPERIMENTAL PROCEDURES** for details). f. The result of the density estimation (smooth histograms in e - see **EXPERIMENTAL PROCEDURES**) is the likelihood of a response given a particular saccade choice. Distractor neurons are green, red and black. The target neurons are blue. g. Without incorporating a prior probability distribution the result is a (log) likelihood function. h. The uniform prior probability distribution. Each dot shows the probability of one saccade choice. Since the target could appear in each of the four locations with equal probability the function is uniform. i. Same as in h but using a simulated, non-uniform prior function (see **EXPERIMENTAL PROCEDURES**) . j. Estimated posterior probability distribution for one example target condition (45°). Multiplying all four conditional probability across the four neurons and the prior, results in the posterior probability for the example target condition across the four neurons. k. Same as in j for the distribution using a non-uniform prior.

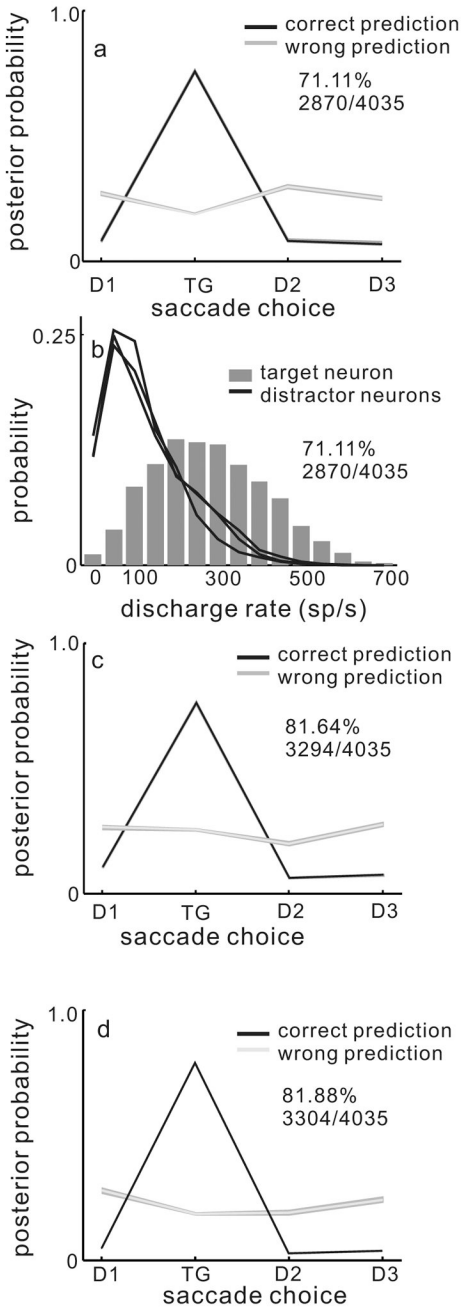

Figure 3.

WTA and MAP models of saccade choice. a. The mean of the MAP estimates across all trials, computed using the logs of fixed Gaussians as weights is plotted for the four possible saccade choices. The black line shows the mean of the MAP estimates for the trials in which the model estimate correctly predicted the saccade choice. The gray line shows the mean of MAP estimates across all the trials in which the model estimate did not predict saccade choice correctly. The shading around the lines indicates 1 standard error. 71.11% (2870 out of 4035 trials) of all saccade choices could be predicted by this version of the MAP model. b. The probability of measuring a particular discharge rate in the interval 28 - 8ms before saccade onset in the selection task is plotted for all neurons across all 30 data sets. Black lines show

probability of non-selected neuron discharge rates. The grey bars show probability of selected neuron discharge rates. 71.11% (2870 out of 4035 trials) of all saccade choices could be predicted by the WTA model. c. The same as in a except the posterior probability was computed using the log of the tuning curves of SC neurons to characterize $P(r|s)$. 81.64% (3294/4035) of saccade choices were predicted correctly. d. Same as in a and c except the posterior was computed using the empirical probability distribution to characterize $P(r|s)$. 81.88% (3304/4035) of saccade choices were correctly predicted.

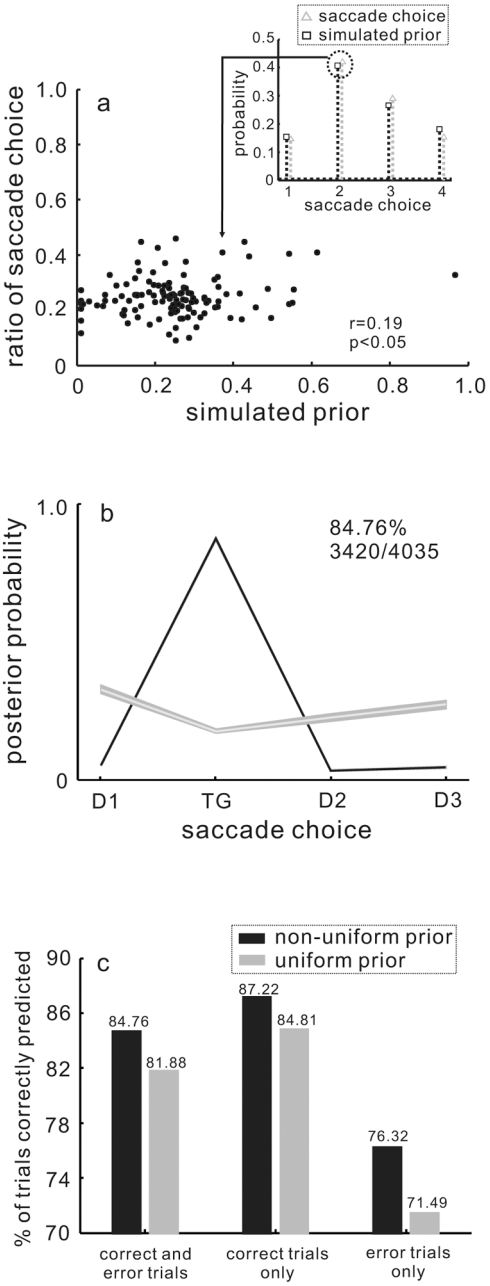

Figure 4.

The influence of the prior in predicting a saccade choice. a. The ratio of the saccade choice is plotted against the simulated prior probability. One example of choice distribution and the simulated prior distribution appears in the inset (black dotted line and squares = simulated prior; gray dotted line and triangles = the distribution of saccade choices). To show the relationship between the distribution of saccade choices and the simulated prior distribution, two points from the inset were remapped to the plot of points shown in a. This was done for all 120 data sets so n = 120 points. The Pearson correlation coefficient was determined, r = 0.19. This relationship was small but statically significant ($p<0.05$) b. The mean of the MAP estimates across all trials, computed using the empirical PDF and the non-uniform prior

function is plotted for the four possible saccade choices. The black line shows the mean of the MAP estimates for the trials in which the model estimate correctly predicted the saccade choice. The gray line shows the mean of MAP estimates across all the trials in which the model estimate did not predict saccade choice correctly. Shading around the lines indicates 1 standard error. This version of MAP model predicted the saccade choice correctly on 3420/4035 (84.76%) trials. c. The percent of trials predicted correctly from the MAP model is plotted for the models using the uniform (grey bars) and non-uniform (black bars) prior distribution. The trials are sorted by whether they were correct or in error. The percentages in each condition appear on the top of each bar.

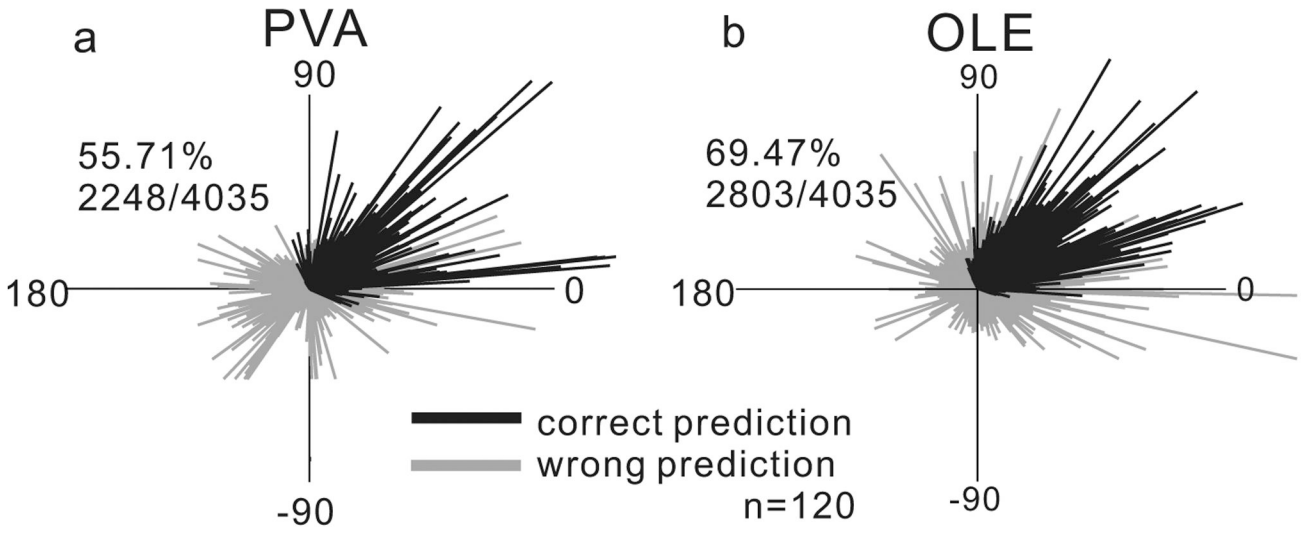

Figure 5.

Population vector average and saccade choice. a. The neuronal vector determined by equation (13). The direction of each line indicates the direction of the population vector and the length indicates the strength of activity. The black lines show the trials in which the population vector correctly predicted saccade choice. The grey lines show the incorrectly predicted choice. 55.71% (2248/4035) of trials were correctly predicted by PVA. b. Same as in a for the OLE computed as described in equation (14). 69.47% (2803/4035) trials were predicted correctly by the OLE.

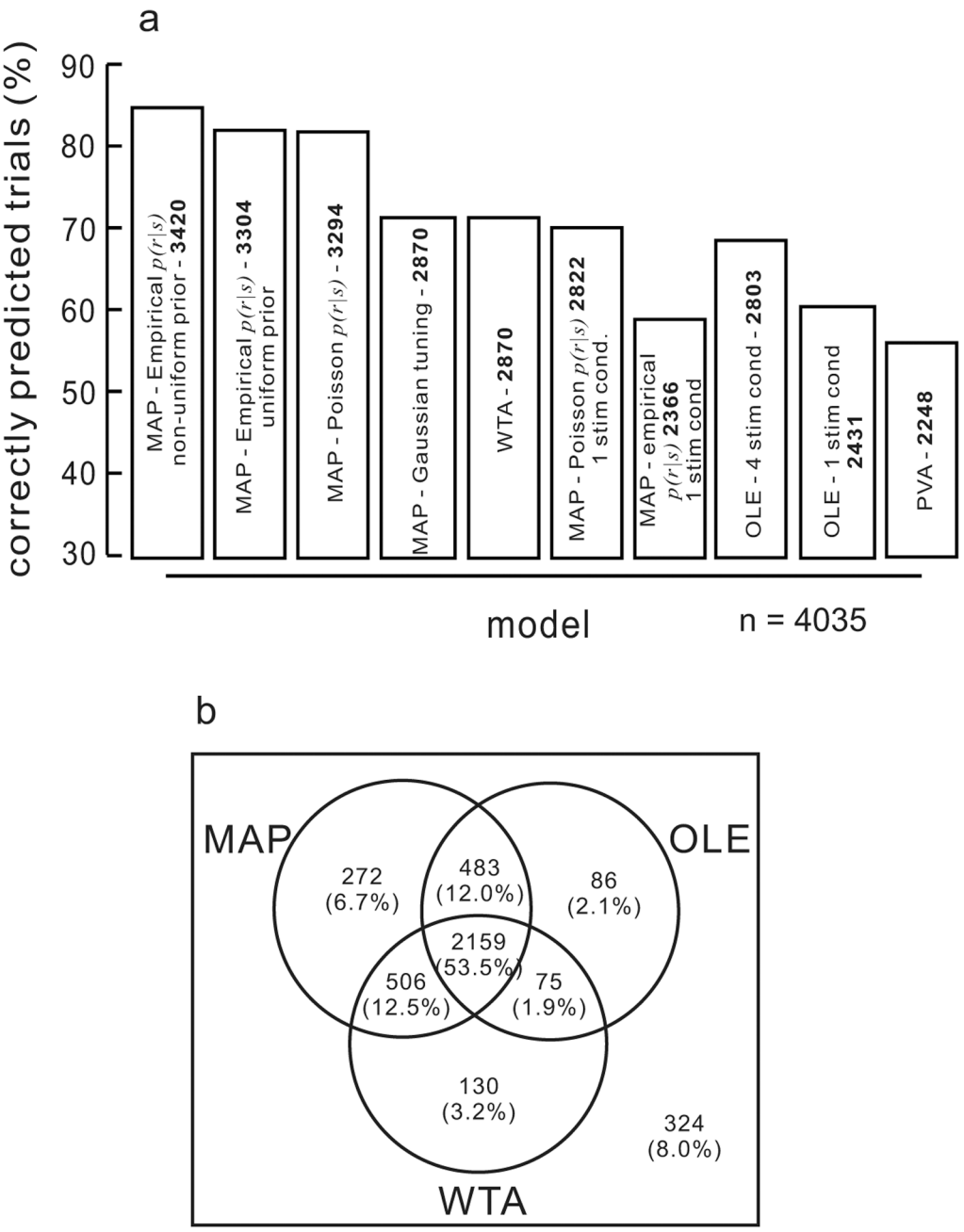

Figure 6.

Comparing models of saccade choice. a. The percent of trials predicted correctly from each of the models is plotted. The number inside each bar indicates the number of trials (out of the total 4035) that were correctly predicted by each model. Each model result is labeled inside each bar. From left to right the bars are: MAP with empirical PDF characterization of *P(r|s)* determined from four stimulus condition data with the non-uniform prior; MAP with Poisson PDF characterization of *P(r|s)* determined from four stimulus condition data with the uniform prior; MAP with Poisson probability density function constrained by the tuning curves measured from the four stimulus condition data; MAP with Gaussian tuning curves; WTA; MAP with Poisson PDF characterization of *P(r|s)* determined from single stimulus condition

data; MAP using empirical PDF characterization of $P(r|s)$ using single stimulus condition data; OLE optimized with four stimulus condition data; OLE optimized with single stimulus condition data; PVA. b. Venn diagram of the percentage of trials predicted by each of the models. The MAP model in this plot is the one using the empirical PDF characterization of $P(r|s)$ determined from the four stimulus condition data and the non-uniform prior (the first bar in the top panel). 8.0% (324/4035) of the trials were not predicted by any of the models.

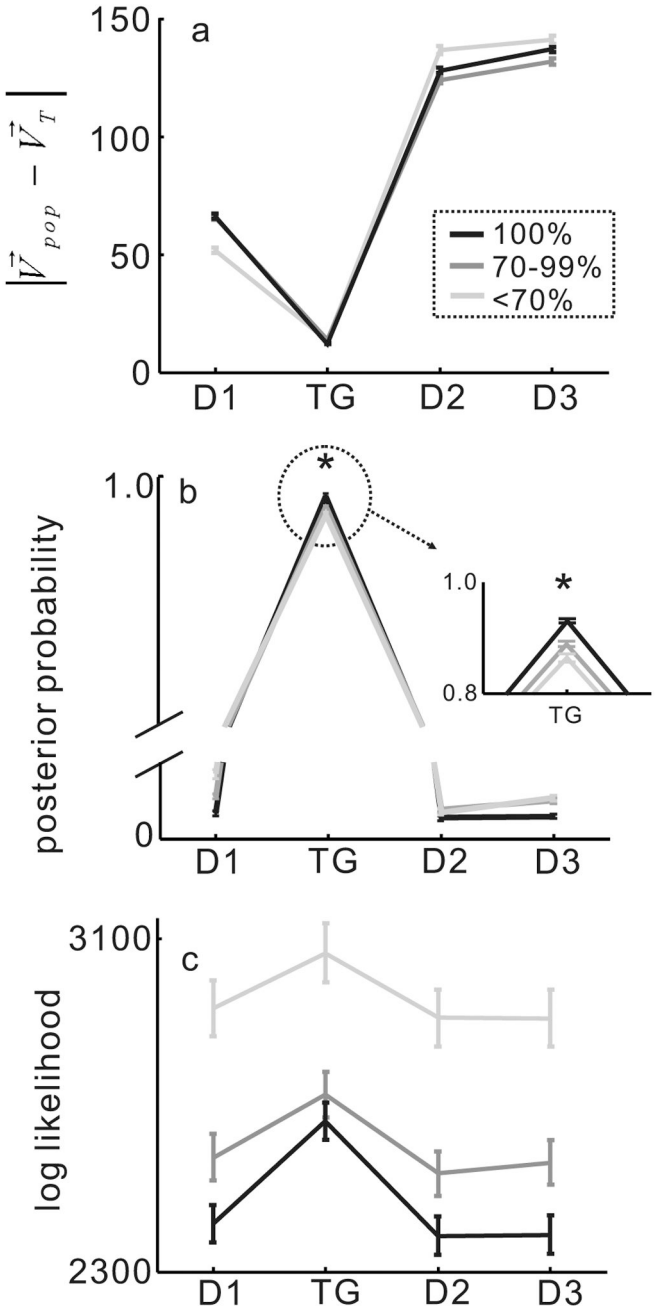

Figure 7.

MAP estimates scale with performance accuracy. The black lines show mean angular difference for the trials in which performance was 100%. The dark grey lines show the mean angular difference for trials in which performance was 70–99% accurate. The light grey lines show the mean angular difference for trials in which performance accuracy was <70%. D1 = distractor 1; D2 = distractor 2; D3 = distractor 3; TG = saccade target choice. a. The angular difference between the direction of the prediction resulting from the OLE calculation and the actual saccade direction for trials in which performance accuracy varied. b. The mean of the MAP estimate with non-uniform prior for the four possible saccade choices, D1, TG, D2 and D3 as in a is plotted for each of the three accuracy conditions. The inset shows the region of

the distribution in the dashed circle. * = $p < 0.05$. c. Log likelihood for each of the four possible saccade choices. The arrangement is the same as in a and b.

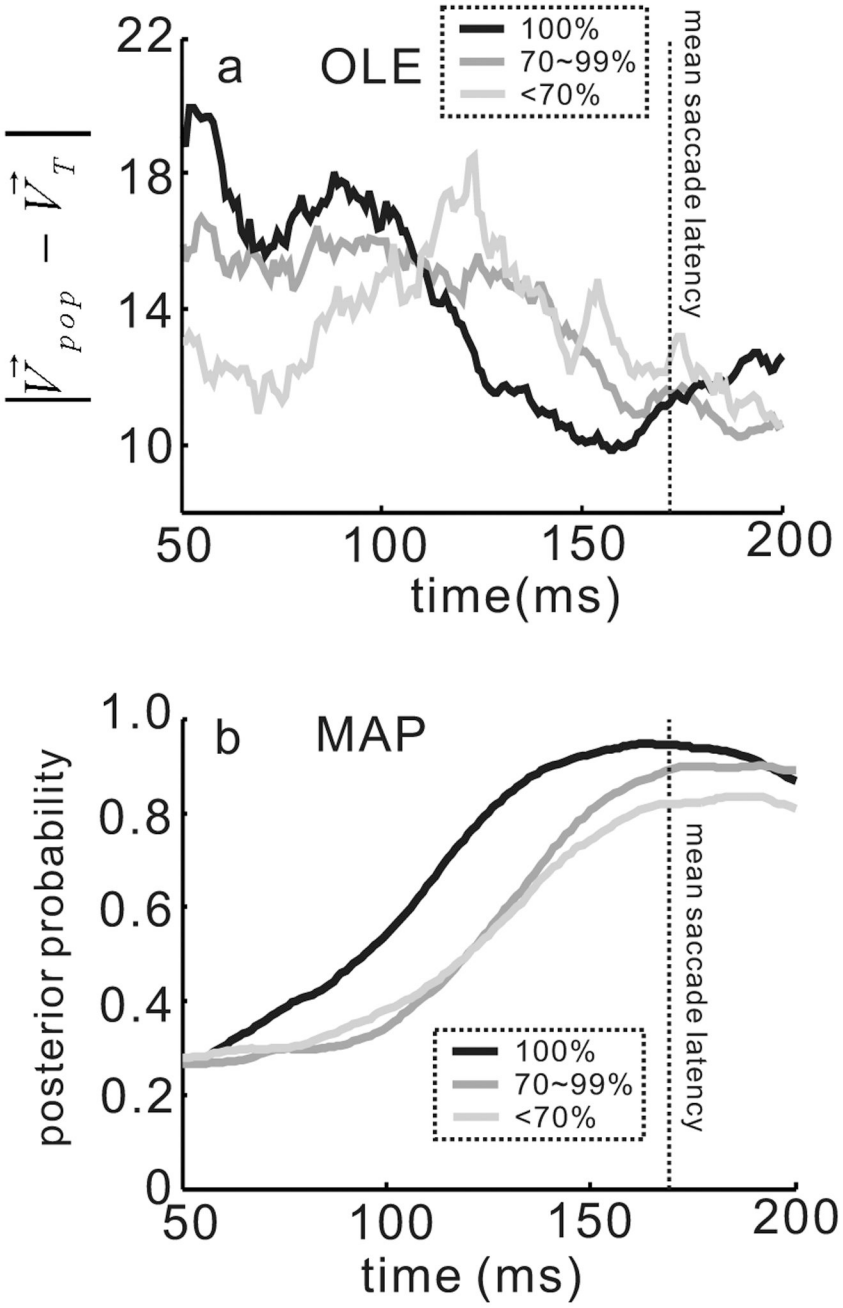

Figure 8.

OLE and MAP predictions develop over time. a. The angular difference in the direction of the OLE vector and the saccade vector is plotted against time for the three choice accuracy levels. The traces are aligned to the time point 50ms after the onset of the stimulus array. Black lines are from the 100% performance accuracy trials, dark grey lines are the 70–90% performance accuracy trials, and the light grey lines are the <70% performance accuracy trials. b. The mean of the MAP estimates for the saccade choice in the three levels of performance accuracy are plotted over time as in a. The vertical dotted line in both panels is the mean saccade reaction time (RT).

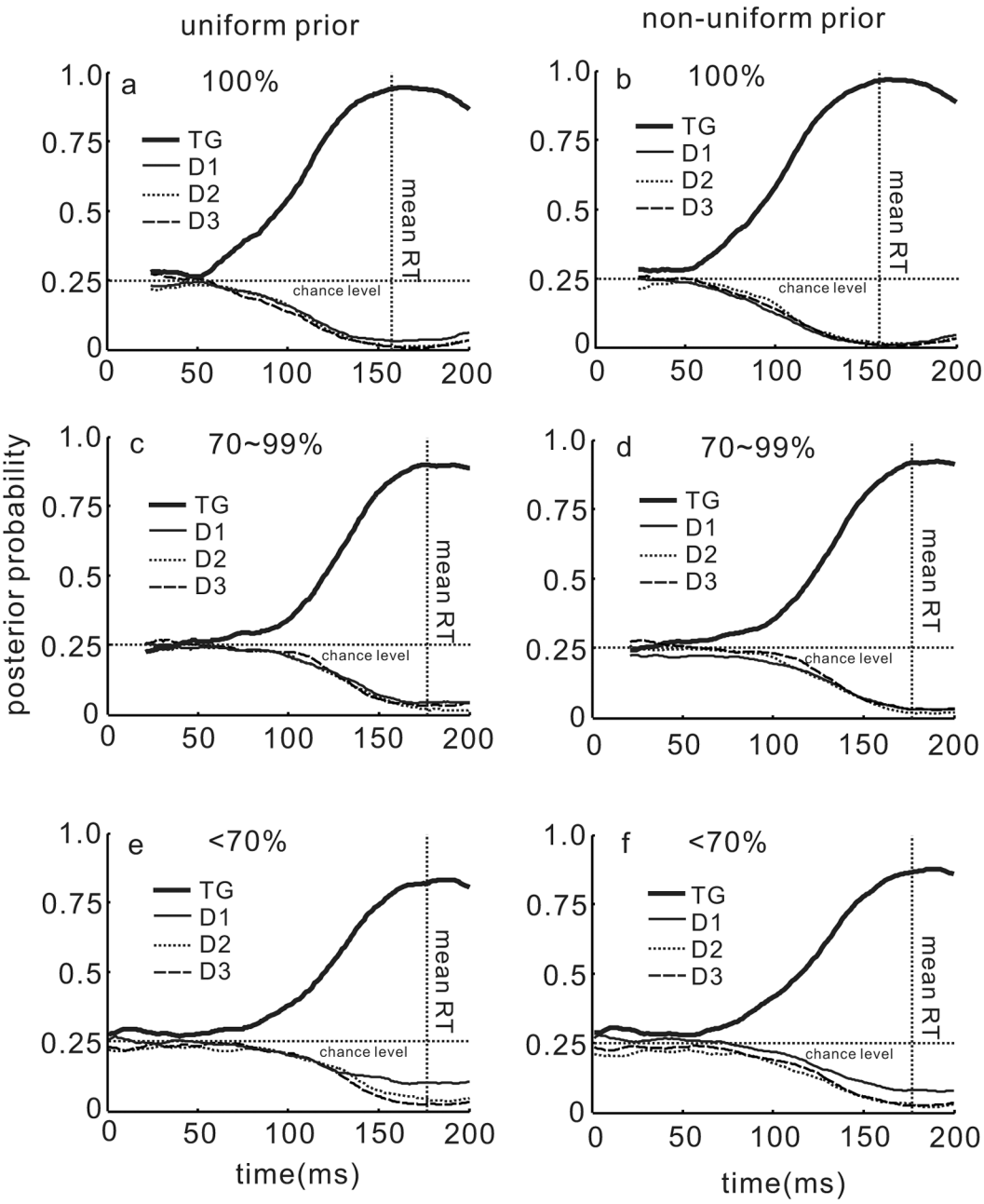

Figure 9.

Evolution of the posterior distribution. a. The posterior probabilities for the saccade choice (TG) and the three distractor locations (D1, D2, and D3) are plotted over time in the 100% performance accuracy condition and the uniform prior. b. The same as in a for the non-uniform prior. c. The same as in a for the 70–99% performance accuracy condition. d. The same as in b for the 70–99% performance accuracy condition. e. The same as in a and c for the <70% accuracy condition. f. The same as in b and d for the <70% accuracy condition. In all panels, the thick black line is the TG; D1 is the thin black line; D2 is the dotted line and D3 is the dashed line. The horizontal dotted line indicates chance probability (25%). The vertical dotted line indicates the mean saccade reaction time (RT).