

NORTHWESTERN UNIVERSITY

Machine Learning Approaches Towards Understanding Movement
Planning in Naturalistic Settings

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Neuroscience

By

Joshua I. Glaser

EVANSTON, ILLINOIS

March 2018

© Copyright by Joshua I. Glaser 2018

All Rights Reserved

ABSTRACT

Machine Learning Approaches Towards Understanding Movement Planning in Naturalistic Settings

Joshua I. Glaser

A central question in neuroscience is how the brain plans movements. Here, I apply neural data analysis and machine learning methods to better understand both eye and arm movement planning, in particular focusing on naturalistic settings. First, I built encoding models to investigate the factors that led to neural activity in macaque Frontal Eye Field (FEF) during a natural scene search task (Ch. 2,3,4). One central finding was that FEF neurons did not represent task-relevant visual features within natural scenes. Another central finding was that separate populations of neurons represented preliminary and definitive plans for movement. The neurons that represented preliminary plans represented the probabilities of potential upcoming saccades. I found similar characteristics in dorsal premotor cortex, where populations of neurons represented the probability distributions of possible upcoming reaches (Ch. 5). Finally, I compared many different methods for neural decoding to demonstrate that modern machine learning methods lead to performance improvements, even for limited amounts of data (Ch. 6). Overall, I have

provided insights into neural activity across a wide range of motor behaviors in more naturalistic settings, and have demonstrated the value of using machine learning methods within neuroscience.

Acknowledgments

There are many faculty who have helped make my PhD a very positive experience. First and foremost, I would like to thank my advisor, Konrad Kording, for all his guidance, mentoring, advice, and support. I would like to thank Mark Segraves and Lee Miller, who both acted like co-advisors for much of my PhD. They both provided valuable guidance, along with giving me an appreciation for the experimental side of things. Finally, I would like to thank all the other faculty members, including Sara Solla, Eric Perreault, and CJ Heckman, who provided additional mentorship throughout my PhD.

My colleagues have also been extremely important in creating an environment full of collaboration, learning, and interesting conversations. I'd like to thank all the Kording lab members throughout the years, plus the many members of the Miller and Segraves labs I have worked with. I'd like to especially thank several collaborators I worked on projects closely with: Pat Lawlor, Ted Cybulski, Pavan Ramkumar, Dan Wood, Hugo Fernandes, Matt Perich, Raeed Chowdhury, Stephanie Naufel.

Last but not least, I want to thank my friends and family for all their support, their love, and for occasionally giving me a hard time. I especially want to thank my wife Mallory,

my parents Bob and Sharon, and my sister Beth, for all their help and love over the years.

Preface

Chapter 1 (Introduction): This work is my own.

Chapter 2: A version of this work was published in the Journal of Neurophysiology [1], and was also part of Pat Lawlor's thesis. Pat Lawlor and Pavan Ramkumar were co-first authors of this publication. I was third author. Daniel Wood and Adam Phillips were also middle authors. Mark Segraves and Konrad Kording were the senior authors.

Chapter 3: A version of this work was published in the Journal of Neurophysiology [2]. I was a co-first author of this publication with Daniel Wood. Pat Lawlor and Pavan Ramkumar were middle authors. Konrad Kording and Mark Segraves were the senior authors.

Chapter 4: A version of this work has been submitted as of January 22, 2018. It is also posted on bioRxiv [3]. I am the first author of this work. Daniel Wood and Patrick Lawlor are middle authors of this work. Mark Segraves and Konrad Kording are the senior authors.

Chapter 5: A version of this work was under review as of January 22, 2018. It is also posted on bioRxiv [4]. I am the first author of this work. Matthew Perich and Pavan Ramkumar are middle authors of this work. Lee Miller and Konrad Kording are the senior authors.

Chapter 6: A version of this work was under review as of January 22, 2018. It is also posted on arXiv [5]. I am the first author of this work. Raees Chowdhury, Matthew

Perich, and Lee Miller are middle authors of this work. Konrad Kording is the senior author.

Chapter 7 (Discussion): This work is my own.

Table of Contents

ABSTRACT	3
Acknowledgments	5
Preface	7
List of Tables	12
List of Figures	14
Chapter 1. Introduction	44
Chapter 2. Feature-based attention and spatial selection in frontal eye fields during natural scene search	48
Abstract	48
Introduction	49
Methods	51
Results	78
Discussion	92
Chapter 3. The role of expected reward in frontal eye field during natural scene search	99
Abstract	99

	10
Introduction	100
Methods	101
Results	114
Discussion	127
Chapter 4. From preliminary to definitive plans: two classes of neurons in frontal eye field	135
Abstract	135
Introduction	135
Results	137
Discussion	150
Methods	157
Chapter 5. Population coding of conditional probability distributions in dorsal premotor cortex	173
Abstract	173
Introduction	174
Results	175
Discussion	194
Methods	198
Chapter 6. Machine learning for neural decoding	222
Abstract	222
Introduction	222
Methods	223

	11
Results	231
Discussion	233
Chapter 7. Discussion	242
References	248

List of Tables

- 2.1 **List of units characterized by cell type and list of modeled units.** The H or V in brackets indicates the orientation of the target for the Gabor search tasks. V, M, and VM stand for visual, movement, and visuomovement cell types. We could not determine cell types for a small fraction of the cells in the Fly task (undetermined cells). We were able to model the majority of characterized units (modeled units), although certain units needed to be discarded due to low firing rates. 59
- 2.2 **Statistics of search behavior for each monkey and task summarized across sessions** (mean \pm SEM). The trial duration is the entire duration from scene onset to scene offset. The success rate is the percentage of trials in which the monkey successfully located the target. The number of saccades to find the target is given only for successful trials, and includes the last, target-finding saccade. The fixation duration is averaged across all saccades in successful and failed trials. 80
- 2.3 **Correlation between success rate and ROC values across sessions.** 83

2.4 **GLM analysis summary statistics.** Number of neurons that were significantly modulated by relevance, energy or movement in different models. Neurons were deemed to be significantly tuned if the 4 confidence intervals of the (relative) pseudo- R^2 s exceeded zero. We used a strict 4 threshold to sufficiently correct for multiple comparisons across neurons and models. Column 3: number of significant neurons for the relevance-only model. Column 4: number of neurons significantly tuned for movement by comparing a leave-movement-out model against a joint model with movement and relevance. Column 5: number of neurons significantly tuned for relevance by comparing a leave-relevance-out model against a joint model with movement and relevance. Column 6: number of neurons significantly tuned for relevance by comparing a leave-relevance-out model against a comprehensive full model (see text). Column 7: number of significant neurons for the energy-only model. Column 8: number of neurons significantly tuned for energy by comparing a leave-energy-out model against a comprehensive full model (see text).

List of Figures

- 2.1 **Example stimuli used in natural scene search tasks.** Targets were blended into the natural scenes and monkeys were given a water reward for finding the target within a fixed number of saccades. For the sake of illustration, targets are encircled in red; the fly is shown unblended since it is difficult to see at this resolution. 54
- 2.2 **Operational definition of relevance (target similarity) and edge-energy.** (A) Relevance maps were obtained by convolving the target and its quadrature phase shift with the natural scene and then taking the sum of their squares. (B) Edge-energy maps were computed by taking the sum of squares of horizontal and vertical edge gradients. Before computing the relevance and energy maps, the image was degraded in accordance with decreasing visual acuity in the periphery, with respect to an example fixation location shown as a yellow crosshair (adapted from [6]). 61
- 2.3 **Schematic illustration of the comprehensive generative model of neural spikes using a GLM framework.** The model comprises visual features: saliency, relevance and energy from a neighborhood around fixation location after the saccade, untuned responses aligned

to saccade and fixation onsets, and the direction of the saccade. The features are passed through parameterized spatial filters (representing the receptive field) and temporal filters. The model also comprises spike history terms (or self terms). All these features are linearly combined followed by an exponential nonlinearity, which gives the conditional intensity function of spike rate, given model parameters. Spikes are generated from this model by sampling from a Poisson distribution with mean equal to the conditional intensity function. Brown: basis functions modeling temporal activity around the saccade onset; Green: basis functions modeling temporal responses around the fixation onset; Blue: basis functions modeling temporal responses after spike onset. 65

2.4 **Parameterization of the generative model.** Untuned saccade- and fixation-related temporal responses were modeled using linear combinations of raised-cosine temporal basis functions. Spatiotemporal tuning to saliency, relevance, energy, and saccade direction were modeled using bilinear models with left multipliers representing temporal basis function loadings, and right multipliers representing spatial basis function (cosine and sine) loadings. Additionally, a spike-history (self) term was modeled using a linear combination of temporal basis functions causally aligned to spike events. Parameters of the model were fit using maximum-likelihood with elastic net regularization. 67

2.5 **Prediction of gaze using visual features at fixation.** We compared bottom-up IK-saliency, top-down relevance and energy at fixated (above:

left panel) and non-fixated, i.e. shuffled control (above: right panel) targets by computing the area under the ROC curves (below). The star indicates statistically significant difference from a chance level of 0.5 at a significance level of $p < 0.05$.

81

2.6 **Example neuron fit using a movement-only model.** The model was fit to an independent held-out half of the data that is not visualized here **(A)** Raster plots, peri-saccadic time histograms (PSTHs) and model predictions, separated into 8 categories according to the direction of the upcoming saccade (color-coded in the central glyph), and aligned to saccade onset (vertical line). **(B)** Data and model PSTHs aligned to the saccade onset show tuning to upcoming movement direction.

85

2.7 **Example neuron fit using a relevance-only model.** The model was fit to an independent held-out half of the data that is not visualized here **(A)** Raster plots, fixation-aligned PSTHs, and model predictions, separated into 8 categories according to the direction of the maximally relevant octant (color-coded in the central glyph). **(B)** Data and model PSTHs aligned to the fixation onset do not show a clear tuning to direction of maximum relevance during the fixation onset, but do show a modest tuning at around 200 milliseconds after fixation, as indicated by the gray panel.

86

2.8 **Apparent tuning to relevance is explained away.** **(A)** Data PSTHs (black) and corresponding model predictions (red) for a single example neuron, of (from L to R): relevance-only model, relevance

covariates of the joint model comprising relevance and movement, and movement covariates of the joint model, overlaid on top of each other. Dots show spike rasters. Both rasters and model predictions are aligned to fixation onset **(B)** Scatter plots of pseudo- R^2 (goodness of fit) values of univariate against multivariate models. Top panel: relevance-only model vs. a joint model comprising relevance and movement. Bottom panel: movement-only model vs. a joint model comprising movement and relevance.

88

2.9 **Relevance and edge-energy do not modulate FEF firing rates around saccade onset (above) or fixation onset (below).** We divided saccades into the top and bottom 50 percentiles of the respective visual features either at the saccade landing (above) or within the pre-saccadic receptive field (below), and calculated the saccade- (above) or fixation- (below) aligned PSTHs for each neuron separated in eight directional bins. To combine these across neurons, we then selected the directional bin with highest peak firing rate to represent the receptive field, and averaged the within-RF PSTHs across neurons to compute the saccade-aligned PSTHs (see Methods). For fixation-aligned PSTHs, we first averaged responses around fixation around the directional bins representing out-of RF saccades and then averaged these single-neuron PSTHs (see Methods).

91

2.10 **Power analysis.** Both plots show marginal predictive power of relevance as a function of its simulated modulation depth. Red points

indicate that the marginal predictive power is significantly greater than zero; blue points indicate that it is not significantly greater than zero. Error bars are bootstrapped 95% confidence intervals. **(A)** Injected noise analysis. After simulating the neural data, we corrupted the relevance model with varying degrees of noise, and then fit the model. Each line represents the model fits for a different level of added noise. The r values indicate the correlation between the original model covariates and the corrupted model covariates. **(B)** Incorrect model analysis. We simulated neural data using edge energy, and fit the model using the relevance model (yellow line). For comparison, we include the model fits when the correct model, relevance, was used to simulate the spikes (green line).

93

3.1 Experimental Setup. **(A, B)** In our experiments, monkeys freely searched for targets in natural scenes. In the fly search task (left; done by M14 and M15), the target was an embedded fly. In the Gabor search task (right; done by M15 and M16), the target was an embedded Gabor wavelet. Monkeys were rewarded only after fixating the target for a specified duration. **(C)** Functional characterization of FEF in M16 based on stimulation results for the semi-chronic array. Colors indicate current intensities at (or below) which saccades were reliably elicited. Only neurons at locations where current intensity was equal to or less than $50 \mu\text{A}$ were used in our analyses. **(D)** Characterization of FEF in M14. A subset of locations for acute recordings are marked by blue

circles. All recording locations were characterized as FEF by elicitation of saccades with $\leq 50\mu A$ during stimulation. 103

3.2 Behavioral differences of saccades due to expected reward.

(A) Saccades that land near the target followed by fixation are defined as target, expected reward (T+/ER+), and are shown in blue in subsequent panels. Saccades that do not land near the target are defined as non-target, not expected reward (T-/ER-), and are shown in red in subsequent panels. (B) Latencies are compared between T+/ER+ (blue) and T-/ER- (red) saccades. Mean +/- SEM are shown. For this, and subsequent, panels, each column is behavior from a different monkey. (C) Latency distributions. (D) The mean latencies (+/- SEM) of T+/ER+ and T-/ER- saccades are shown as a function of saccade amplitude. (E) Velocities are compared between T+/ER+ (blue) and T-/ER- (red) saccades. Mean velocities (+/- SEM) as a function of amplitude. Note that differences in the magnitude of velocities are due to differences in eye tracking technology across monkeys (IR camera vs. eye coil; see *Methods*). (F) Because velocity is dependent on amplitude, we define vigor as the velocity divided by the expected velocity for that amplitude (for all saccades). Shading is based on whether scaled velocities are greater or less than 1. (G) Vigor distributions. (H) Vigor as a function of saccade latency for all monkeys. All saccades, regardless of expected reward, were combined, as they had the same trends. 117

3.3 **Neural activity reflects expected reward.** (A) A schematic of T+/ER+ and T-/ER- saccades into the RF. While the RF is shown as an angular wedge, the RFs generally had some amplitude dependence. (B) Normalized PSTHs averaged across neurons, aligned to saccade onset. We plot T+/ER+ saccades into the RF (blue) and T-/ER- saccades into the RF (red). Error bars represent the SEM. (C) PSTHs of example neurons, aligned to saccade onset. Error bars represent the SEM of that neuron's firing rate. (D) Normalized PSTHs averaged across neurons (for each monkey), aligned to saccade onset. (E) Normalized PSTHs of T-/ER- saccades that have subsequent fixation times of greater than (purple) and less than (brown) 200 ms. Only saccades into the RF are used. (F) We fit separate tuning curves to T+/ER+ and T-/ER- saccades using a Von-Mises function. An example tuning curve fit is shown. Each dot represents a single saccade. (G) Tuning curves were fit to all neurons. The gains (**left**), widths (**middle**), and baselines (**right**) are compared between T+/ER+ and T-/ER- saccades. The medians of these ratios (across neurons) +/- the standard error of the median (computed by bootstrapping) are shown. 119

3.4 **Increased firing rates due to expected reward, short latency, and high velocity saccades: PSTHs.** (A) As latency and vigor are correlated, we look at the effect of expected reward across the behavioral space of latency and velocity. On the **top**, the relative likelihood of a saccade having a certain latency and vigor for T+/ER+

compared to T-/ER- saccades, averaged across monkeys. More yellow areas mean that a particular latency and vigor are more likely for T+/ER+ saccades. **Below**, a threshold is put on the upper panel to determine behavioral regions that are more likely to contain saccades accompanied by expected reward (yellow) or no expected reward (blue). **(B)** Normalized PSTHs averaged across neurons (from all monkeys) are constructed from 3 conditions: 1) T+/ER+ saccades in the yellow behavioral region of the above panel (cyan); 2) T-/ER- saccades in the yellow behavioral region (orange); 3) T-/ER- saccades in the blue behavioral region (pink). **(C)** Same as panel B, for individual monkeys. 121

3.5

Increased firing rates due to expected reward, short latency, and high velocity saccades: GLMs. **(A)** Using a GLM, we used the latency, vigor, and expected reward of the saccade (whether it was to the target) to explain the peak firing rate of each neuron. Histograms of the regression coefficients for each covariate are shown, colored based on their significance (dark blue is not significant). Insets (below) show the percentage of neurons that are significant. For target (left), having positive regression coefficients signifies that going to the target (T+/ER+ saccade) increased the firing rate. For latency (middle), having positive coefficients signifies that longer latency saccades increased the firing rate. For vigor (right), having positive coefficients signifies that higher vigor increased the firing rate. **(B)** Same as panel A, with individual monkeys shown in each row. **(C)** Bars show conditional

probabilities (the probability of a neuron representing one feature, e.g. expected reward, conditional on the neuron representing another feature, e.g. latency) from the GLM results across all monkeys. Dashed lines show the general probabilities of a feature being significant in the GLM (not conditioned on the neuron representing any other features). Error bars are SEMs from bootstrapping. For example, on the left we show the probability that neurons' firing rates significantly increase due to expected reward (dashed line) conditioned on the neurons' having significantly higher firing rates due to shorter latencies (left bar) and higher vigor (right bar).

123

3.6 Modulation in FEF due to expected reward rather than target presence. **(A)** We compared T+/ER+ (blue), T+/ER- (green), and T-/ER- (red) saccades into the RF. **(B)** PSTHs show the average normalized firing rate \pm SEM of saccades across all neurons (from all monkeys). **(C)** Same as panel B, but for individual monkeys. **(D)** For an example monkey, M14, we show the difference in the distributions of landing distances from the target for T+/ER+ and T+/ER- saccades. **(E)** We control for the possibility that T+/ER+ saccades only have a higher peak firing rate than T+/ER- saccades because T+/ER+ saccades land closer to the target on average (panel D). We plot PSTHs for T+/ER+ and T+/ER- saccades that land close and far from the target. Close and far are defined as less and more (respectively) than

the median distance from the target for all T+ saccades. Error bars are not shown to ensure that the means are visible.

126

4.1 **Experiment and behavior.** **(A)** Monkeys freely searched for a Gabor target embedded in natural scenes. **(B)** The probability of the direction of the upcoming saccade is dependent on the eye position on the screen. This is an example where the eye position is to the right of the screen. **(C)** We quantify the relationship between the upcoming saccade direction and position using ϕ , the angle between the position vector (relative to center), and the upcoming saccade vector. **(D)** The distribution of ϕ 's for all saccades (blue), and split according to whether the starting eye position was close (purple) or far (orange) from the border. The close/far distinction was based on being less/more than the median distance from a border. **(E)** The mean latency of saccades as a function of ϕ . **(F)** The latency difference between saccades towards the center ($|\phi - 180^\circ| < 60^\circ$) and saccades away from the center ($|\phi - 180^\circ| > 120^\circ$), as a function of the starting eye position's distance from a border. The distance from the border was divided into quartiles. 140

4.2 **Different times of saccade selectivity for E-Sel and L-Sel neurons, and neural differences related to saccade latencies.** Peri-event time histograms (PETHs), aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). **First Row of PETHs:** An example late selection neuron. **Second Row:** Normalized averages of late selection neurons. **Third**

Row: An example early selection neuron. **Bottom Row:** Normalized averages of early selection neurons. **(A)** PETHs of saccades near the preferred direction (PD; black) versus opposite the PD (brown). **(B)** PETHs of saccades near the PD, for saccade latencies less than 150 ms (orange) versus latencies greater than 150 ms (green). 142

4.3 **E-Sel, but not L-Sel, neurons have higher activity in positions that are more likely to result in saccades near the PD.** Peri-event time histograms (PETHs), aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). **First Row of PETHs:** An example late selection neuron. **Second Row:** Normalized averages of late selection neurons. **Third Row:** An example early selection neuron. **Bottom Row:** Normalized averages of early selection neurons. **(A)** PETHs of saccades near the PD, with a starting position near the PD (unlikely that upcoming saccade will be near PD; blue) versus a position opposite the PD (likely that upcoming saccade will be near PD; red). **(B)** PETHs of saccades opposite the PD, with a starting position near the PD (blue) versus a position opposite the PD (red). 145

4.4 **GLM results - importance of upcoming saccade and eye position parameters.** Importance of parameters in the generalized linear model, across time, aligned to fixation. **(A)** The mean relative pseudo- R^2 of the upcoming saccade (green) and eye position (purple)

covariates for late selection neurons. **(B)** Same plot for early selection neurons. Shaded areas represent SEMs. 146

4.5 **E-Sel population activity reflects the probabilities of upcoming saccades.** On the left, we plot the population activity of E-Sel neurons as a function of the relative position angle. On the right, with a gray background, we show how this relates to the behavioral distribution of ϕ 's. **(A)** The relative position angle is the difference between a neuron's preferred direction and the eye position vector (from the center). **(B)** Copied from Fig. 1c, ϕ is the difference between the upcoming saccade and the eye position vector. **(C)** On the left, a heat map of normalized activity over time, as a function of relative position angle, averaged across neurons. On the right, the normalized average activity in the 100 ms surrounding fixation, plotted as a function of the relative position angle. Only saccades away from the PD are included. **(D)** The distribution of ϕ 's across all saccades combined across monkeys. **(E)** Same as panel C, but now separated for initial eye positions close to the borders (top left, right in purple) and far from the borders (bottom left, right in orange). **(F)** Same as panel D, but now separated for initial eye positions close to the borders (purple) and far from the borders (orange). 148

4.6 **L-Sel population activity does not reflect the probabilities of upcoming saccades.** We plot the population activity of L-Sel neurons (left) and E-Sel neurons (right) as a function of the relative position

angle, as in Figs. 5 and S4. For the polar plots on the right, we use activity in the 100 ms around fixation for E-Sel neurons, as in Fig. 5. For L-Sel neurons, we use activity from 150-250 ms following fixation. Note that the color bars for E-Sel neurons differ from those in Fig. 5. **(A)** All saccades are included. **(B)** Only saccades away from the PD are included. **(C)** We control for the correlation between the eye position and upcoming saccade direction. We resampled saccades to create a distribution of saccade directions relative to position angles that were centered on 45° . This plot ensures that our main results are not simply caused by a correlation with the true upcoming saccade (in which case the activity after resampling would become peaked closer to 45°), but rather reflect the distribution of upcoming saccades (in which case the activity after resampling would remain peaked at $\sim 135^\circ$).

150

4.1 **PETHs based on saccade direction and latency, for individual monkeys.** Peri-event time histograms (PETHs), aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). **First Row of PETHs:** Normalized averages of late selection neurons from Monkey J. **Second Row:** Normalized averages of late selection neurons from Monkey K. **Third Row:** Normalized averages of early selection neurons from Monkey J. **Bottom Row:** Normalized averages of early selection neurons from Monkey K. **(A)** PETHs of saccades near the preferred direction (PD; black) versus opposite the PD (brown). **(B)** PETHs of saccades near

the PD, for saccade latencies less than 150 ms (orange) versus latencies greater than 150 ms (green). 168

4.2 **PETHs based on eye position, for individual monkeys.** Peri-event time histograms (PETHs), aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). **First Row of PETHs:** Normalized averages of late selection neurons from Monkey J. **Second Row:** Normalized averages of late selection neurons from Monkey K. **Third Row:** Normalized averages of early selection neurons from Monkey J. **Bottom Row:** Normalized averages of early selection neurons from Monkey K. **(A)** PETHs of saccades near the PD, with a starting position near the PD (blue) versus a position opposite the PD (red). **(B)** PETHs of saccades opposite the PD, with a starting position near the PD (blue) versus a position opposite the PD (red). 169

4.3 **PETHs based on eye position, for different saccade latencies.** Peri-event time histograms (PETHs) of E-Sel neuron averages, aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). As in Fig. 3A, we compare PETHs of saccades near the PD, with a starting position near the PD (blue) versus a position opposite the PD (red). Now, we show these plots for **(A)** saccade latencies less than 150 ms, **(B)** saccade latencies from 150-250 ms, and **(C)** saccade latencies greater than 250 ms. 170

- 4.4 **GLM results, for individual monkeys.** Importance of parameters in the generalized linear model, across time, aligned to fixation. Results for Monkeys J and K are in the top and bottom rows, respectively. **(A)** The mean relative pseudo- R^2 of the upcoming saccade (green) and eye position (purple) covariates for late selection neurons. **(B)** Same plot for early selection neurons. Shaded areas represent SEMs. 171
- 4.5 **E-Sel population activity reflects the probabilities of upcoming saccades, for individual monkeys.** Population activity of E-Sel neurons as a function of relative position angle, for Monkeys J (left) and K (right). **(A)** On the left, a heat map of normalized activity over time, as a function of relative position angle, averaged across neurons. On the right, the normalized average activity in the 100 ms surrounding fixation, as a function of the relative position angle. Only saccades away from the PD are included. **(B)** Same as panel A, but now separated for initial eye positions close to the borders (top left, right in purple) and far from the borders (bottom left, right in orange). **(C)** We control for the correlation between the previous and upcoming saccade directions. We only included saccades in which the previous and upcoming saccade directions were less than 90° apart (while in the actual data, previous and upcoming saccades are more likely to be in opposite directions). Unlike in panels A and B, we do not exclude saccades towards the PD. 172
- 5.1 **Experimental design and statistics.** **(a)** If the arm is outstretched, the only possible arm movements are back toward the body (**left**). In

other limb postures, it may be possible to move the arm in any direction (**right**). In blue, circular probability distributions are shown for the possible upcoming movements based on the current arm posture. (**b**) Experimental design. The monkey makes sequences of four reaches, briefly holding within each target box before the next target appears. (**c**) The current hand position limits the range of possible locations of the next target, due to the borders of the workspace and target presentation algorithm. (**d**) The probability distributions of upcoming reach directions (blue) from different areas of space (x and y divided into quartiles). Green arrows point toward the circular means of the distributions. (**e**) ϕ is the angular difference between the upcoming movement vector (the vector that brings the hand to the target) and the current angular hand position (relative to the center of the workspace). (**f**) The probability distribution of ϕ 's from all hand positions. (**g**) The probability distribution of ϕ 's from initial hand positions within 2 cm of the center of the workspace.

178

5.2 **Behavior.** (**a**) An example trajectory. The initial direction of the reach (green) starts toward the expected direction of the target, given the current hand position. It later moves in the actual direction of the target. The inset shows an enlarged view of the beginning of the reach. (**b**) The median bias of the trajectory over time. A bias of 1 signifies that the direction of the trajectory is toward the expected target direction, while a bias of 0 signifies that the direction of the trajectory

is toward the actual target direction. Negative biases signify movement away from the expected direction. Different traces are shown for hand positions at varying distances from the center of the workspace. Error bars are standard errors of the median. **(c)** The mean latency of reaches as a function of the angular difference between the actual and expected target directions. **(d)** The difference in mean latency between expected and unexpected reaches (expected minus unexpected), depending on the hand's distance from the center. "Expected" reaches are those that had an angular difference between the actual and expected target directions of less than 60° . "Unexpected" reaches had an angular difference of more than 120° . In panels c and d, error bars represent SEMs. In panels b and d, distances from the center are divided as follows: "closest" is 0-20% of distances from the center, "mid-close" is 20-40%, "mid-far" is 40-60%, and "farthest" is 60-100%. We used these divisions for plotting, rather than standard quartiles, to ensure that there were "unexpected" reaches in each bin.

180

5.3 **PMd PSTHs and GLM Results. First Row:** A selected-response (SR) neuron. **Second Row:** Normalized averages of SR neurons. **Third Row:** A potential-response (PR) neuron. **Bottom Row:** Normalized averages of PR neurons. **(a-c)** Peristimulus time histograms (PSTHs) for PMd neurons, aligned to target onset. Shaded areas represent SEMs. **(a)** PSTHs of reaches near the preferred direction (PD; black) versus opposite the PD (brown). **(b)** PSTHs of reaches near the

PD, with a starting hand position near the PD (lower probability of moving near the PD; blue) versus a position opposite the PD (higher probability of moving near the PD; red). **(c)** PSTHs of reaches opposite the PD, with a starting hand position near the PD (blue) versus a position opposite the PD (red). **(d)** We utilized a generalized linear model (GLM) to control for confounds in the PSTHs, including different distributions of starting positions, upcoming movements, and previous movements. Here, we show the importance of parameters in the GLM, across time, for PMd neurons. We show mean relative pseudo- R^2 over time, of the upcoming movement (green) and hand position (purple) covariates. For the 2nd and bottom row, shaded areas represent SEMs across neurons. For individual neurons, shaded areas represent the standard deviation across bootstraps.

182

5.4 **PMd population jointly represents the distribution of upcoming movements.** **a) Left:** The average normalized firing rate of all PR neurons, over time, as a function of relative angular hand position. For each neuron, the relative angular position is the preferred direction of the neuron minus the angular hand position. **Right:** The average normalized firing rate in the 100 ms prior to target onset, as a function of the relative angular hand position. **b)** The distribution of upcoming movement directions relative to position angles (duplicated from Fig. 1f). **c,d)** Same as panels a and b, but for only for reaches starting near (within 2 cm of) the center. **e) Left:** Position activity maps for example

PR neurons with preferred movement directions oriented upwards. Position activity maps show the neurons' activity as a function of hand position (blue is low; yellow is high) from -100 to 50 ms surrounding target onset. **Right:** The sum of position maps for all PR neurons, when their preferred directions are oriented upwards. **f)** A map showing the probability that the next movement will be upwards, as a function of initial hand position. **g)** Preferred reach directions for all PR neurons, when space is rotated so that their preferred hand position angle is oriented to be at the bottom (270°). **h)** A histogram of preferred reach directions of all PR neurons relative to their preferred angular hand position (the reach PD minus the preferred angular hand position). 185

5.5

PMd population activity represents the distribution of upcoming movements- single reach decoding. **(a)** The distribution of decoded reach directions (blue) from the population of PMd neurons for two example reaches (left and right), before and after target onset (top and bottom). The purple circle is the current hand position, and the green square is the target location. **(b)** The distribution across all reaches of pre-target decoded reach directions relative to the hand's angular position (this is ϕ from Fig. 1, except with decoded reach directions instead of actual reach directions). **(c)** The distribution across reaches starting within 2 cm of the center, of pre-target decoded reach directions relative to the hand's angular position. **(d)** The predictions of two hypotheses (left and right), shown

for two different hand positions (example 1 vs. example 2). **(e)** Average pre-target decoded reach direction distributions as a function of hand position. These distributions are constructed to have the average width and peak angle of all decoded distributions from the hand positions within the grid square. **(f)** The full width at half maximum of the pre-target decoded distribution as a function of hand distance from the center of the workspace. Distances from the center are binned as in Fig. 2. Error bars represent SEMs. **(g)** The width of the decoded distributions over time, for starting positions that are the closest (blue) and farthest (purple) from the center. **(h)** The latency of the reach as a function of the angular difference between the pre-target decoded direction and the actual target direction. Error bars represent SEMs. **(i)** The bias of the initial trajectory of the reach (100-150ms from target onset) toward the pre-target decoded direction. 95% confidence intervals, computed via bootstrapping, are shown.

189

5.6

Visuomotor Rotation Control Task. **(a)** The visuomotor rotation (VR) task. Movements on the screen (in the workspace) are rotated 30° counterclockwise relative to the hand movement. **(b)** The distribution of hand movements relative to the position angle in the workspace, i.e. ϕ 's, for the baseline (blue) and VR (orange) tasks. Arrows point toward the distributions' circular means. **(c)** The difference between the initial reach direction (100-150ms from target onset) and the expected reach direction during the baseline task (blue) and different periods during

the VR task (orange). The expected reach direction was the most likely upcoming reach direction given the current workspace position and movement statistics (panel b). Positive values mean the initial reach direction was counterclockwise of the expected target direction, meaning the monkey had not adapted. **(d)** Position activity maps (activity as a function of position in the workspace) of example PR neurons in the baseline (top) and VR task in the second 2/3 of trials (bottom), as in Fig. 4c. In the middle, we show the direction (clockwise or counterclockwise) and magnitude of change of the preferred angular position. **(e)** The change in preferred angular position of all PR neurons (VR minus baseline). Positive means a counterclockwise shift. **(f)** The distribution of pre-target decoded reach directions relative to the hand's angular position (decoded ϕ 's) for positions not near the center (greater than the median distance). Decoding from the VR task used the second 2/3 of trials. **(g)** The difference between the circular mean of the distributions of decoded ϕ 's in panel f, between the baseline and VR tasks (VR minus baseline). Error bars represent 95% confidence intervals from bootstrapping.

192

5.7

M1 does not reflect the probability of upcoming movements.

(a-d) PSTHs and GLM results for M1 neurons. Columns have the same schematics as Fig. 3. **First Row of PSTHs:** Normalized averages of “reach” neurons, defined as those neurons significant for movement during the late period, but not position in the early period

of the GLM. This was the same criteria as for SR neurons in PMd.

Second Row: Normalized averages of “reach and position” neurons, defined as those neurons significant for movement during the late period, and position in the early period of the GLM. This was the same criteria as for PR neurons in PMd. Note that we did not use the same “SR/PR” nomenclature as PMd, because there was no evidence in the PSTHs of M1 neurons that position was used to represent potential upcoming movements. **(e)** Same schematic as Fig. 4a, but for M1 reach and position neurons. **Left:** The normalized average firing rate, as a function of time and relative angular position. Activity is averaged across all reach and position neurons. **Right:** The normalized average firing rate in the 100 ms prior to target onset, plotted as a function of the relative angular hand position.

193

5.1 **Trajectory Biases.** We compare **(a)** the biases in the reach direction towards the expected direction of the reach given the statistics of target presentation (identical to Fig. 2b) versus **(b)** the biases in the reach direction towards the center of the workspace.

213

5.2 **PMd PSTHs and GLM results, for individual monkeys.**

Peristimulus time histograms (PSTHs) and GLM results for PMd neurons, for individual monkeys. All columns are the same as in Fig. 3. **First and Second PSTH Rows:** Normalized averages of selected-response (SR) neurons from Monkey M and Monkey T, respectively. **Third and Bottom Rows:** Normalized averages of

potential-response (PR) neurons from Monkey M and Monkey T, respectively. 214

5.3 **Explaining pre-target activity for SR neurons.** In Fig. 3, for SR neurons, prior to target onset there began to be a slight separation between activity traces depending whether the reach will be near vs. opposite the PD. Given that SR neurons are supposed to only respond after target onset, this is initially surprising. However, there are two likely reasons for SR neurons' apparent pre-target activity. The first reason is our classification criteria of SR and PR neurons. PR neurons, unlike SR neurons, were significantly modulated by hand position prior to target onset (see Methods for details). That is, if a neuron was modulated by hand position with $> 95\%$ (e.g. 96%) confidence, then it was a PR neuron, but if it was modulated by hand position with $\leq 95\%$ (e.g. 94%) confidence, it would be an SR neuron. Thus, using this "conservative classification" (as we do in Fig. 3), we are likely including some PR neurons in the SR category. A PSTH using this conservative classification, copied from Fig. 3, is shown in the top row. Instead, if we use a "relaxed classification" that includes neurons as PR neurons if they are modulated by hand position with $> 50\%$ confidence, then SR neurons should not include any true PR neurons. When we plot SR neurons using this relaxed classification (bottom row), the differential activity prior to target onset diminishes, demonstrating that some PR neurons being included as SR neurons was a cause of the differential

activity. Note that Supplementary Fig. 4 gives more details about different “conservative” and “relaxed” classification types. A second reason for the pre-target-onset differentiation of SR neurons in Fig. 3 is jitter in the time of target onset. While we subtracted the average delay for the target to be displayed on screen, there was some jitter in this timing (see Methods). Thus, some activity aligned to target onset could appear slightly earlier than it occurred.

215

5.4 **Neuron Classification.** The proportion of potential-response (PR) and selected-response (SR) neurons using different classification criteria. **a,b)** We include all sessions (as in the main text) **a)** We defined SR neurons as those significantly modulated by upcoming movement in the late period in the GLM, but were not significantly modulated by hand position in the early period. PR neurons were significantly modulated by upcoming movement in the late period and by hand position in the early period. Significantly modified means that the 95% confidence intervals of pseudo- R^2 and relative pseudo- R^2 values were greater than 0 (see Methods). This was the criteria used for all parts of the main text with the exception of decoding. **b)** A more relaxed criteria is to look at all neurons that had a median pseudo- R^2 and relative pseudo- $R^2 > 0$. In the main text, we used this relaxed criteria to include PR neurons for decoding. **c,d)** These are the same as panels a and b, respectively, except we now only include sessions that were recorded in separate

weeks, to decrease the number of “repeat” neurons that were recorded in multiple sessions.

216

5.5

PMd population activity represents the distribution of upcoming movements- accumulated across reaches: individual monkeys and controls. As in Fig. 4a, all heat maps show normalized average activity over time as a function of relative angular position (the angular hand position relative to neurons’ PDs). To the right of each heat map, normalized activity in the 100 ms prior to target onset, as a function of the relative angular position. Each row calculates activity from a different set of reaches. **Top row:** From all reaches (as in Fig. 4a). **Second row:** Control for the correlation between the previous and upcoming movement directions, which tended to be in opposite directions. We examined neural activity in the infrequent cases when pairs of sequential reaches were in similar directions (less than 90° away from each other). **Third row:** We resampled reaches to create a distribution of reach directions relative to angular hand position that were centered on 180° (rather than 150°). This plot ensures that our main results were not simply caused by a correlation between the angular hand position and the true upcoming movement (in which case the activity after resampling would become centered on 180°), but rather reflected the distribution of upcoming movements (in which case the activity after resampling would remain centered at 150°). **Bottom row:** From reaches starting near the center (as in Fig. 4c).

217

5.6 **Visuomotor Rotation Control Task.** (a) The distribution of pre-target decoded reach directions relative to the hand's angular position (decoded ϕ 's) for all positions. Arrows point toward the circular means of the distributions. (b) The difference between the circular means of the distributions of decoded ϕ 's in panel a (rotation minus baseline). Error bars are 95% confidence intervals from bootstrapping. (c) In both the baseline task and visuomotor rotation task, we display on the **left**: The average normalized firing rate of all PR neurons, over time, as a function of relative angular hand position (as in Fig. 4a). For each monkey, we display on the **right**: The average normalized firing rate in the 100 ms prior to target onset, as a function of the relative angular position (baseline in blue; rotation task in orange). Arrows point toward the peak angles of activity. (d) For each session, the difference between the angles corresponding to peak activity in panel c, between the baseline and rotation tasks (rotation minus baseline). In all panels, results from the VR task used the second 2/3 of trials. 218

5.7 **Visuomotor Rotation Task - changes in PDs after target onset.** On the **left**, we plot a histogram of changes in the visual PD for PR neurons, in the time period 50-200 ms after target onset. If the neurons were representing the movement rather than the target, then the neurons would be most active when the target is +30 degrees (counterclockwise) in the VR task. If they were representing the visual location of the target, then we would see a change of 0. The median

change across PR neurons is -2.1 degrees. On the **right**, we have the same plot for SR neurons, with a median change of 0.4 degrees. Thus, the visuomotor rotation does not change PMd's representation of the target after target onset.

219

5.8 **M1 does not reflect the probability of upcoming movements, for individual monkeys. (a-d)** PSTHs and GLM results for M1 neurons. Columns have the same schematics as Fig. 3. **First and Second Rows of PSTHs:** Normalized averages of “reach” neurons. **Third and Fourth Rows:** Normalized averages of “reach and position” neurons. **(e)** Same schematic as Fig. 4a, but for M1 neurons. **Left:** The normalized average firing rate, as a function of time and relative angular hand position. Activity is averaged across all reach and position neurons. **Right:** The normalized average firing rate in the 100 ms prior to target onset, as a function of the relative angular position.

220

5.9 **Relationship between previous movement and hand position.**

(a) We define the angle between the previous movement vector and the hand position vector (relative to the center) as ψ . More specifically, ψ is the previous movement vector direction minus the position vector direction. **(b)** The distribution of ψ 's, across all reaches. Note the slight counter-clockwise bias from 0° (it is peaked at about 10° and has a circular mean at 23°), that may explain a portion of the counter-clockwise bias in M1 activity in Fig. 7e.

221

6.1 **Tasks and Decoding Schematic.** **a)** In the task for decoding from motor and somatosensory cortices, monkeys continuously reached to new targets that were presented, with a brief hold period between reaches [4]. **b)** In the task for decoding from hippocampus, rats chased rewards on a square platform [7, 8]. **c)** To decode (predict) the output in a given time bin, we used the firing rates of all N neurons in B time bins. In this schematic, $N = 4$ and $B = 3$ (one bin preceding the output, one concurrent bin, and one following bin). In our data, we predicted two outputs from each brain region (x and y components of velocity from motor and somatosensory cortex, and x and y components of position from hippocampus). For each region, the number of neurons and time bins used for decoding are described in *Methods*. Also, note that this schematic does not apply for the Kalman Filter decoder. **d)** For the non-recurrent decoders (Wiener Filter, Wiener Cascade, SVR, XGBoost, and Feedforward Neural Network), this is a standard regression problem where $N \times B$ features (the firing rates of each neuron in each relevant time bin) are used to predict the output. **e)** For the recurrent decoders (simple RNN, GRUs, LSTMs), to predict an output, we used N features, with temporal connections across B bins. A schematic of an RNN predicting a single output is on the right. 232

6.2 **Example Decoder Results.** Example decoding results from motor cortex (left), somatosensory cortex (middle), and hippocampus (right),

for all ten methods (top to bottom). Ground truth traces are in black, while decoder results are in various colors. 234

6.3 **Decoder Result Summary.** R^2 values are reported for all decoders (different colors) for each brain area (top to bottom). Error bars represent the mean \pm SEM across cross-validation folds. X's represent the R^2 values of each cross-validation fold. Note the different y-axis limits for the hippocampus dataset. 235

6.4 **Example results with limited training data.** Using only 2 minutes of training data for motor cortex and somatosensory cortex, and 15 minutes of training data for hippocampus, we trained two traditional methods (Wiener filter and Kalman filter), and two modern methods (feedforward neural network and LSTM). Example decoding results are shown from motor cortex (left), somatosensory cortex (middle), and hippocampus (right), for these methods (top to bottom). Ground truth traces are in black, while decoder results are in the same colors as previous figures. 236

6.5 **Decoder results with varying amounts of training data.** Using varying amounts of training data, we trained two traditional methods (Wiener filter and Kalman filter), and two modern methods (feedforward neural network and LSTM). R^2 values are reported for these decoders (different colors) for each brain area (top to bottom). Error bars are 68% confidence intervals (meant to approximate the SEM) produced via bootstrapping, as we used a single test set. Values with negative

R^2 s were not shown. Also note the different y-axis limits for the hippocampus dataset. 237

6.1 **Kalman Filter Versions.** R^2 values are reported for different versions of the Kalman Filter for each brain area (top to bottom). On the left (in more transparent cyan), the Kalman Filter is implemented as in [9]. On the right (in more opaque cyan), the Kalman Filter is implemented with an extra parameter that scales the noise matrix associated with the transition in kinematic states (see *Methods*). This version with the extra parameter is the one used in the main text. Error bars represent the mean +/- SEM across cross-validation folds. X's represent the R^2 values of each cross-validation fold. Note the different y-axis limits for the hippocampus dataset. 241

CHAPTER 1

Introduction

How do we plan our upcoming movements? There has been much research on how the brain plans movements, both for saccades (eye movements) and reaching movements. However, most research has been in simplified scenarios that greatly differ from the natural world. This is largely due to researchers' desires to isolate how neural activity relates to specific environmental or behavioral variables, while there are often many simultaneous stimuli and behaviors in the real world. Machine learning techniques can be particularly useful in disentangling the effect of many real-world features on neural activity. Here, I aimed to understand movement planning in more naturalistic settings, and demonstrate the use of machine learning for neuroscience.

Several brain regions, including the frontal eye fields (FEF), supplementary eye fields (SEF), superior colliculus (SC), and lateral intraparietal cortex (LIP), have been shown to be involved in saccade planning. Here, I focus on the FEF. Activity in FEF neurons is predictive of the upcoming saccade direction and amplitude [10]. FEF activity is not purely related to saccade execution; when there is a delay period between the onset of the saccade target and a go-cue, many FEF neurons are active during this delay period [10]. When there are multiple objects to choose from, FEF neurons have been shown to be involved in target selection. FEF neurons have higher activity when a target is placed in their receptive fields compared with a distractor [11, 12]. Plus, they have higher activity for distractors that are more similar to the target [12, 13]. Additionally,

some FEF neurons have greater activity when there is an increased probability that the target will be placed in the neurons' receptive fields [14]. In decision-making tasks, FEF neurons' activities are modulated by the strength of evidence, the reaction time, and the final decision [15]. There is much past research showing FEF's role in saccade planning.

Several brain regions, including the dorsal premotor cortex (PMd), parietal reach region (PRR), and supplementary motor area (SMA), have been shown to be involved in reach planning. Here, I focus on PMd. Activity in PMd neurons is predictive of the upcoming reach direction [16]. Like in FEF, when there is a delay period between target onset and movement, PMd neurons are active during this delay period [17]. When there are several movement possibilities, PMd activity relates to all the possible movements [18, 19]. Plus, neurons have higher activities when there is a higher probability the upcoming movement will be in the neurons' preferred directions [20]. In decision-making tasks, PMd neurons' activities are modulated by the current strength of sensory evidence and the final decision [21]. In fact, the neurons' activities dynamically track the sensory evidence about the movement options [21]. There is much past research showing PMd's role in reach planning.

Most past studies have experimental setups that differ in many ways from the real world, which could have a large effect on movement planning. For instance, conventional eye movement experiments typically saliently display a small number of targets, while natural scenes can have hundreds or thousands of objects that are not displayed as saliently. This difference was a focus of chapter 2, in which I explored how FEF activity was modulated by target-related visual features in natural scenes. Additionally, conventional eye movement experiments often require single saccades instructed by a target, while in the

natural world we make many self-guided saccades to explore a scene. This difference was a focus of chapter 3, in which I investigated how FEF activity differed between saccades to targets and those when exploring a scene. Additionally, most conventional studies do not vary the initial eye position, while the eye's position can affect saccade planning in natural settings (e.g. being more likely to look left when currently looking at the right side of the screen). This difference was a focus of chapter 4, in which I investigated how FEF neurons made preliminary plans for upcoming saccades based on the eye position on the screen. A final difference is that conventional studies often have a small number of discrete targets to choose from, while visual scenes often have unlimited and continuous possibilities of saccade locations. This difference was another focus of chapter 4, in which I investigated how FEF neurons represented a continuous probability distribution of possible upcoming saccades. Across these chapters (2-4), I studied many factors that modulate FEF activity in more naturalistic conditions.

For reaching experiments, many of the same differences exist between traditional experiments and real-world conditions. In chapter 5, I studied how PMd could represent a continuous probability distribution of upcoming possible reaches, while most past studies have looked at how PMd represents a small number of discrete movement options. In this study, the possible movements were based on the hand's location within the workspace, as possible movements in the real world are often based on the state of the body within the environment. Moreover, this study had a sequence of movements every trial, rather than a single movement. Thus, this one study allowed me to investigate many aspects of reach planning that occur in the real world.

Machine learning tools can be very useful for neural data analysis, and can be especially valuable in understanding data from more naturalistic experiments. This is because, in naturalistic experiments, it is often necessary to disentangle how several variables relate to neural activity. All the machine learning methods discussed in this thesis are forms of “supervised learning”, meaning that we are essentially learning a mapping from one set of variables to another. In chapters 2-5, I used Poisson Generalized Linear Models (GLMs) to predict neurons’ firing rates from external variables that may be correlated. GLMs are an extension of multiple linear regression for non-Gaussian outputs, and allowed me to determine whether individual variables independently related to the neural activity. For example, I often wanted to determine whether neural activity was related to variables related to movement planning or the movements themselves. In chapter 6, I compared many different machine learning methods for neural decoding, i.e. predicting behavioral outputs such as velocity, from the activity of many neurons. This allowed me to demonstrate the additional predictive power of modern machine learning techniques for neuroscience.

CHAPTER 2

Feature-based attention and spatial selection in frontal eye fields during natural scene search**Abstract**

When we search for visual objects, the features of those objects bias our attention across the visual landscape (feature-based attention). The brain uses these top-down cues to select eye movement targets (spatial selection). The frontal eye field (FEF) is a prefrontal brain region implicated in selecting eye movements and is thought to reflect feature-based attention and spatial selection. Here, we study how FEF facilitates attention and selection in complex natural scenes. We ask whether FEF neurons facilitate feature-based attention by representing search-relevant visual features, or whether they are primarily involved in selecting eye movement targets in space. We show that search-relevant visual features are weakly predictive of gaze in natural scenes and additionally have no significant influence on FEF activity. Instead, FEF activity appears to primarily correlate with the direction of the upcoming eye movement. Our result demonstrates a concrete need for better models of natural scene search and suggests that FEF activity during natural scene search is explained primarily by spatial selection.

Introduction

A central question in neuroscience is how the brain selects eye movement targets. Eye movements align objects with the high-acuity fovea of the retina, making it possible to gather detailed information about the world. In this way, eye movements inform critical everyday behaviors, including gathering food and avoiding danger. However, from amongst a torrent of visual stimuli, only some are important for the task at hand. How does the brain prioritize information for eye movement selection?

Prioritizing visual information becomes easier when we search for a target object known in advance. This allows us to use target features such as shape and color to guide our search and filter irrelevant information [22]. How exactly the brain performs this computation remains unclear. A leading hypothesis posits that feature-based attention guides the deployment of spatial selection [23]. During fixation, parts of the visual periphery similar to the search target are assigned high priority. We refer to this as feature-based attention. This spatial priority map may then be used to select a part of the visual field to make an eye movement to. We refer to this as spatial selection. In this way, feature-based attention and spatial selection cooperate to influence eye movements.

One brain region thought to be important for selecting eye movements is the frontal eye field (FEF), a prefrontal area on the anterior bank and fundus of the arcuate sulcus [24, 10, 25]. An extensive literature has highlighted the FEF's role in attention and eye movement planning [13, 26, 27, 28, 29, 30]. One central finding is that the FEF appears to be selective for both bottom-up saliency [31, 32] and task-relevant visual features [11, 12, 13]. Crucially, distractors that share a greater number of features with the target (e.g., shape, color) lead to higher firing rates. These results implicate the FEF

in feature-based attention because they demonstrate selectivity to target-relevant features in FEF independent of eye movement selection.

However, many of these studies have been conducted in artificial settings that lack the richness of a more naturalistic environment. First, typical artificial stimuli contain only a handful of distracting objects, whereas natural scenes contain hundreds or thousands. Second, conventional tasks used to study eye movements require a small number of task-instructed saccades, whereas in the natural world we often make tens of self-guided saccades to explore a scene. Therefore, it is important to ask how FEF’s selectivity for bottom-up saliency and task relevance generalizes to tasks in more complicated natural scenes. Recent work by our lab has shown that the apparent encoding of bottom-up saliency by the FEF can be explained away by the direction of spatial selection as defined by the upcoming eye movement [33]. This discrepancy in findings between studies using artificial and natural stimuli suggests that the way the brain selects targets for eye movements during natural vision remains an open question.

Here we investigated the role of the FEF in feature-based attention and spatial selection during natural scene search. In particular, we asked if FEF activity is driven by visual features of the search target, the upcoming eye movement (as a result of spatial selection), or by both. To this end, we recorded from FEF neurons while macaque monkeys searched for known targets embedded in naturalistic stimuli. We then modeled neural activity as a function of target features and the direction of upcoming eye movements. We found that the direction of the upcoming eye movement explained a considerable amount of neural variability, whereas task-relevant visual features did not. Therefore, the reflection of feature-based attention in FEF activity appears to be explained by spatial selection.

Methods

Animals and Surgery

We used three adult female rhesus monkeys (*Macaca mulatta*), ages 14 - 17 years, and weight 5 - 6 kg. We refer to them as M14, M15, and M16. Northwestern University's Animal Care and Use Committee approved all procedures for training, surgery and experiments. Each monkey received preoperative training followed by an aseptic surgery to implant a recording cylinder above the FEF, as well as a titanium receptacle to allow head fixation. Surgical anesthesia was induced with thiopental (5 - 7 mg/kg IV) or propofol (2 - 6 mg/kg IV), and maintained using isoflurane (1.0% - 2.5%) inhaled through an endotracheal tube. For single electrode recordings performed in M14 and M15, an FEF cylinder was centered over the left hemisphere at stereotaxic coordinates anterior 25 mm and lateral 20 mm. Chronic recording of multiple single units was performed with an individually adjustable 32-channel micro drive in monkeys M15 and M16 (Gray Matter Research, Bozeman, MT). The recording chambers for these arrays were centered and mounted at stereotaxic coordinates anterior 24 mm and lateral 20 mm (M15 left hemisphere; M16 right hemisphere).

Behavioral Tasks

We analyzed data from two different experiments involving visual search in natural scenes: the fly search task and the Gabor search task (Fig. 1). Importantly, both tasks were designed to generate large numbers of purposeful, self-guided saccades. Across all sessions and tasks, monkeys performed approximately 300 - 1500 trials per session. Thus, for a typical task comprising ~20 sessions, approximately 6000 - 30000 images were shown.

To control experimental stimuli and data collection, we used the PC-based REX system (Hays Jr et al., 1982), running under the QNX operating system (QNX Software Systems, Ottawa, Ontario, Canada). Visual stimuli were generated by a second, independent graphics process (QNX - Photon) and rear-projected onto a tangent screen in front of the monkey by a CRT video projector (Sony VPH-D50, 75 Hz non-interlaced vertical scan rate, 1024x768 resolution). The distance between the front of the monkey's eye and the screen was 109 cm (43 in). Each natural scene spanned $48^\circ \times 36^\circ$ of the monkey's visual field.

Fly search task

In this task, monkeys (M14 and M15) were trained to locate a picture of a small fly embedded in photographs of natural scenes. Monkeys initiated each trial by fixating a central red dot for 500 – 1000 ms, after which the scene and fly target appeared on the screen, and the fixation dot disappeared. The fly target was placed pseudo-randomly so that its location was balanced across eight 45° sectors. Within these sectors, the fly target was pseudo-randomly placed between 3 and 30 degrees of visual angle from the center of the screen. The trial ended when either the monkey fixated a 2° window around the target for 300 ms, or failed to find the target after 25 saccades. When the target was found and successfully fixated, the monkey was rewarded with water (for details see: [33, 34]).

The natural scene images used for this task were drawn from a library of over 500 images, originally used for [34]. Photographs were taken with a digital camera, and included scenes with animals, people, plants, or food. Image order was chosen pseudo-randomly so that images were repeated only after all others had been shown. Although each unique scene was repeated approximately 10 times over the course of the search

task, since the locations of the targets were randomized, memorization was not likely to be useful and the monkey had to search visually on each trial to successfully find the target.

Since the monkeys quickly learned to perform the task using only a small number of saccades, we made the task more difficult by blending the fly image with the background photographs. We did this using a standard alpha-blending technique (see [33] for details). Even for targets with a transparency of approximately 65%, the average success rate across animals and sessions approached $\sim 85\%$.

Gabor search task

In this task, monkeys (M15 and M16) were trained to locate a Gabor wavelet embedded in photographs of natural scenes. Gabor wavelets are oriented gratings convolved with a local Gaussian, and have been used extensively in studies of visual search [35]. Here, we used Gabor wavelets because their properties can be easily manipulated (e.g., orientation), and because natural scenes often contain oriented textures. Taken together, this made it possible to quantify relevant features in the environment in the sense that image patches sharing features (e.g., orientation) with the Gabor wavelet could be expected to draw the monkey’s gaze. In these experiments, we used Gabor wavelets with the same orientation (either vertical or horizontal) within each session. Task parameters and background images were similar to those used in the Fly search task. This task was significantly more difficult than the Fly search task and both monkeys did not exceed an average success rate of $\sim 50\%$.



Figure 2.1: **Example stimuli used in natural scene search tasks.** Targets were blended into the natural scenes and monkeys were given a water reward for finding the target within a fixed number of saccades. For the sake of illustration, targets are encircled in red; the fly is shown unblended since it is difficult to see at this resolution.

Data Acquisition and Preliminary Characterizations

Tracking

To track eye gaze behavior, we recorded monkeys' eye position with a precision of up to 0.1° resolution. M14 and M15 received an aseptic surgery to implant a subconjunctival wire-search coil to record eye movements for the fly search task. The coil was sampled at 1 kHz. Eye movements of M15 and M16 were measured using an infra-red eye tracker for the Gabor search tasks (ISCAN Inc., Woburn, MA, <http://www.iscaninc.com/>), which samples eye position at 60 Hz.

Saccade detection

We detected saccade onsets and offsets from the kinematics of recorded eye position. We used a threshold of 80 deg/s for start velocity, and marked a saccade starting time when the velocity increased above this threshold. Likewise, saccade-ending times were marked when the velocity fell below 100 deg/s at the end of this period of decrease. Saccades

longer than 80 deg or with duration longer than 150 ms were discarded as eye-blinks or other artifacts. Fixation locations were computed as the median (x, y) gaze coordinate in the inter-saccadic interval.

Neural Recording

We analyzed experiments that used two different electrophysiological recording setups. One set of experiments (the Fly search task) used single-electrode recordings, whereas the other sets of experiments (Gabor search tasks) used adjustable electrode arrays to simultaneously record multiple single units.

Single-unit recording

For monkeys M14 and M15 in the Fly search task, single-unit activity was recorded using tungsten microelectrodes (A-M Systems, Inc., Carlsborg, WA, USA). Electrode penetrations were made through stainless steel guide tubes that just pierced the dura. Guide tubes were positioned using a Crist grid system (Crist et al., 1988, Crist Instrument, Co., Hagerstown, MD, USA). Recordings were made using a single electrode advanced by a hydraulic microdrive (Narashige Scientific Instrument Lab, Tokyo, Japan). The inter-electrode distance was 1.0 mm. On-line spike discrimination and the generation of pulses marking action potentials were accomplished using a multichannel spike acquisition system (Plexon, Inc., Dallas, TX, USA). This system isolated a maximum of two neuron waveforms from a single electrode. Pulses marking the time of isolated spikes were transferred to and stored by the REX system. During the experiment, a real-time display generated by the REX system showed the timing of spike pulses in relation to selected behavioral events.

Recordings were confirmed to be in the FEF by the ability to evoke saccades with current intensities of $\leq 50 \mu\text{A}$ [10]. To stimulate electrically, we generated 70 ms trains of biphasic pulses, negative first, 0.2 ms width per pulse phase, delivered at a frequency of 330 Hz.

Chronic recording

For monkeys M15 and M16 in the Gabor search tasks, recordings were performed using a 32 channel chronically implanted electrode micro drive (Gray Matter Research, Bozeman, MT, USA). The depth of each individual tungsten electrode (Alpha-Omega, Alpharetta, GA) was independently adjustable over a range of 20 mm. The inter-electrode distance was 1.5 mm.

All electrodes were initially lowered to pierce the dura. Individual electrodes were then gradually lowered until a well-isolated unit was located. In general, only a subset of electrodes was moved on any given day, and electrodes were left in place for at least three days before further lowering.

Both spikes and local field potentials (LFPs) were recorded with a multichannel acquisition system (Plexon, Inc., Dallas, TX, USA) based on a separate PC. Spike waveforms, sampled at 40 kHz were stored for offline sorting. In addition, a continuous analog record of electrode signals sampled at 20 kHz was saved for offline LFP analysis. Automatic spike sorting was performed offline using the Plexon Offline Sorter (Plexon, Inc., Dallas, TX, USA).

Since any given electrode was often left in place for multiple days, we likely recorded from the same neuron across sessions. Therefore, we combined data from units that persisted across recording sessions on different days by manually comparing spike waveforms

from units recorded at the same site on different days. Generally, we merged units sharing waveform shape (rise/fall characteristics, concavity/convexity, etc.), and time course. Ambiguous cases were not combined, and waveforms that did not have a single characteristic shape were considered to represent multi-unit activity, which were also included for analysis. See Table 1 (below) for the entire set of animals, tasks, and units analyzed.

To verify that our recording sites were in the FEF, we used microstimulation to evoke saccades in two of three monkeys (M14 and M16) for both tasks and M15 for the Fly search task. For the single unit recordings in the Fly search task (M14, M15) we stimulated at the end of each recording session and only used data from sessions that reliably evoked saccades with thresholds of ≤ 50 A (Bruce et al., 1985). For the implanted electrode array (M16), we only used units isolated from electrodes for which saccades could be evoked with thresholds of ≤ 50 μ A. Since M15 was required for future experiments and stimulation quickly degrades the recording fidelity of the tungsten electrodes, we were unable to stimulate to verify the location of our recording sites for the Gabor search tasks in this monkey. As a result, we were able to verify FEF location with stimulation in only one (M16) of the two chronically implanted monkeys. However, in M15 the chronic array was centered at the stereotaxic location matching maximum FEF evoked saccade sites in M16 and 3 other monkeys used in previous studies. Furthermore, we limited our analyses to units which had properties (receptive field structure and response characteristics) expected in the FEF (see FEF cell characterization below). This decreased the chance of analyzing units from nearby brain regions that were not part of the FEF.

FEF cell characterization

In the Fly search task experiments, cell type characterization was performed for all cells using a standard battery of tests (memory-guided delayed saccade task and visually-guided delayed saccade task; for details see [33, 34]). We excluded cells not meeting criteria for having either visual-related activity or movement-related activity. From this dataset, we analyzed 46 neurons (21 visual cells, 4 movement cells, 13 visuomovement cells, and 8 other cells that did not pass any of these criteria; see Table 1).

For the Gabor search task experiments, we did not use the standard battery due to the large number of simultaneously recorded cells resulting from chronic recordings. Instead, we used activity from the natural scene search to estimate the degree of visual and motor activity. In particular, we labeled cells as having visual activity (visual cells) if the firing rate changed significantly from baseline when the natural scene flashed on (baseline interval: 100-0 ms before scene onset; test interval: 0-100 ms after scene onset; Wilcoxon rank-sum test, $p < 0.005$). Similarly, we labeled cells as having movement activity (movement cells) if the firing rate at saccade initiation exceeded that at baseline in any of 8 (45 degree) binned directions (baseline interval: 300-200 ms before saccade initiation; test interval: 50 ms before to 50 after saccade; Wilcoxon rank-sum test, $p < 0.005$). If they passed both tests, we considered them as visuomovement cells. From these datasets, we analyzed a total of 218 cells (49 cells that passed the visual test only, 52 cells that passed the movement test only, and 117 cells that passed both the visual and movement tests; see Table 1).

Animal	Task	Session	# units	V	M	VM	U	# used
M14	Fly	19	19	10	1	9	7	15
M15	Fly	27	27	11	3	4	1	25
M15	Gabor (H)	27	93	13	24	56	-	81
M16	Gabor (H)	36	76	18	21	37	-	57
M16	Gabor (V)	11	49	18	7	24	-	49

Table 2.1: **List of units characterized by cell type and list of modeled units.** The H or V in brackets indicates the orientation of the target for the Gabor search tasks. V, M, and VM stand for visual, movement, and visuomovement cell types. We could not determine cell types for a small fraction of the cells in the Fly task (undetermined cells). We were able to model the majority of characterized units (modeled units), although certain units needed to be discarded due to low firing rates.

Behavioral Data Analysis

Relevance map

This study asks whether FEF activity reflects feature-based attention as a means to select eye movements. In our visual search tasks, we operationalized feature-based attention as a bias for visual features similar to the search target. We therefore define these target-similar visual features as relevant for the search task. To examine whether relevance influences search behavior and FEF activity, we needed to precisely define relevance. We generated relevance maps by performing a two-dimensional convolution of the visual scene with the search target. If an image patch is similar to the target, their convolution will yield a large value. In practice, to avoid sensitivity to the precise phase of the Gabor, we convolved the target as well as its 90-deg phase-shifted version with the scene. We then took the sum of squares of the convolutions (Ramkumar et al., 2015; see Fig. 2A). This operation effectively measures the low-level visual feature overlap between image regions

and the search target. Because our search tasks used different search targets (horizontal Gabor wavelet, vertical Gabor wavelet, fly image), we generated relevance maps separately for each task. For the fly search task, which used color images, we summed the relevance map over the three color channels (RGB).

Edge-energy map

Edge-energy of natural scenes is known to influence the fixation choice of both humans and monkeys in visual search (Rajashekar et al., 2003); Ganguli et al., 2010); Ramkumar et al., 2015). Therefore, we also computed edge energy as a potential feature that may influence fixation choice, defined as the sum of squares of the vertical and horizontal edge gradients (Fig. 2B, for details, see [6]). For the Fly search task, we calculated the energy maps for each of the RGB color channels and summed them.

Itti-Koch saliency map

Although this study emphasizes the effect of feature-based attention (relevance) on behavior and FEF activity, we also analyzed bottom-up saliency, as it has been influential in eye movement behavior and FEF electrophysiology [33]. We operationally defined bottom-up saliency using the Itti-Koch model [36]. This model defines saliency in terms of contrast of luminance, color, and orientation at multiple spatial scales.

Analysis of the effect of visual features on fixation selection

Before examining the electrophysiological data, we asked whether visual features (relevance, energy and saliency) did indeed guide eye movements in behavior. More specifically, we asked whether the visual features predicted eye movements to those locations. To do this, we performed a Receiver Operator Characteristic (ROC) analysis. To construct the

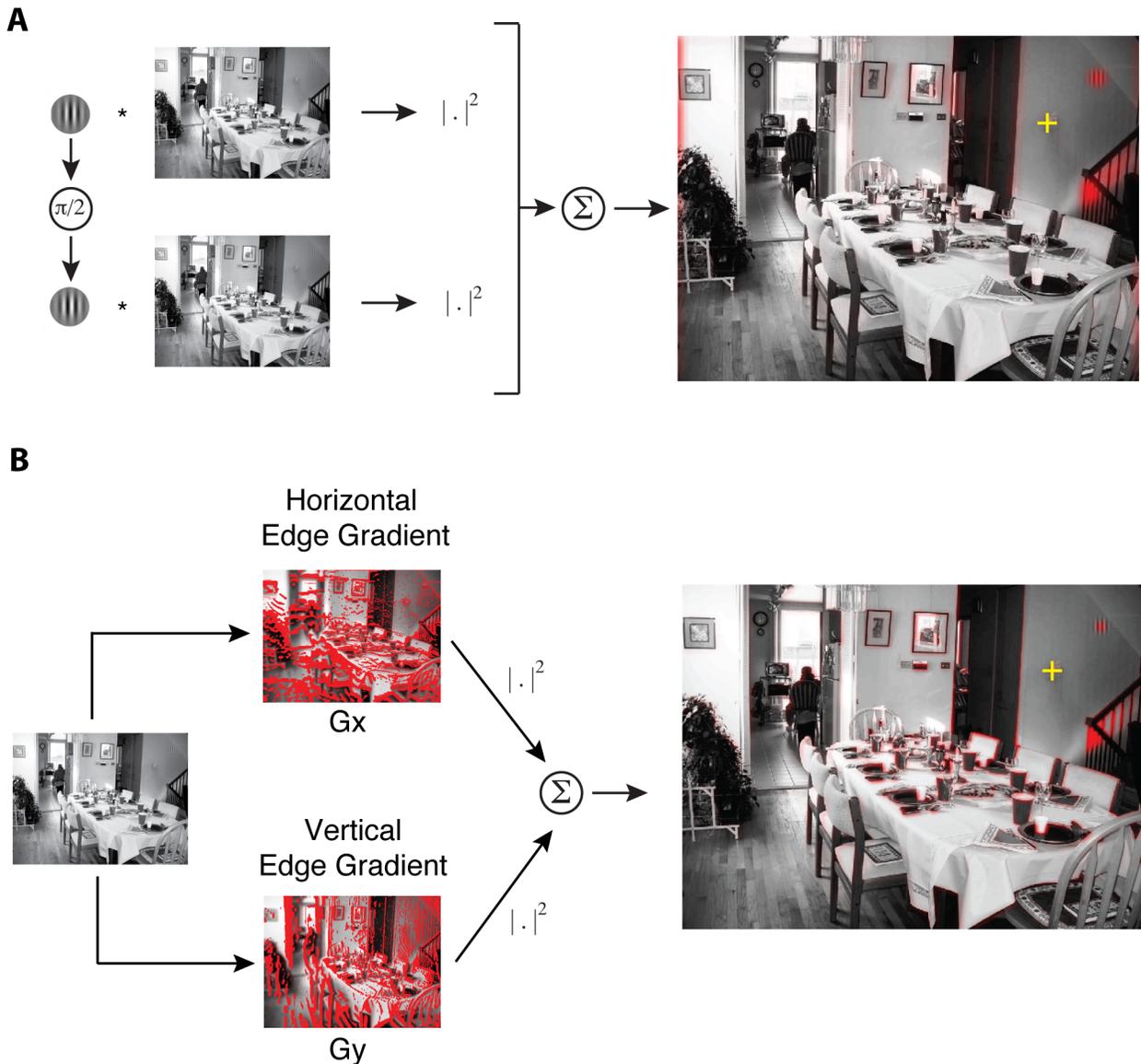


Figure 2.2: **Operational definition of relevance (target similarity) and edge-energy.** (A) Relevance maps were obtained by convolving the target and its quadrature phase shift with the natural scene and then taking the sum of their squares. (B) Edge-energy maps were computed by taking the sum of squares of horizontal and vertical edge gradients. Before computing the relevance and energy maps, the image was degraded in accordance with decreasing visual acuity in the periphery, with respect to an example fixation location shown as a yellow crosshair (adapted from [6]).

ROC curves, we used the experimentally measured locations of saccade endpoints (fixations), in conjunction with the model-based feature maps (described above) for each image viewed by the monkeys. If feature maps indeed predict fixation locations, saccade endpoints should fall on image patches with higher values of the respective maps than those not drawing saccades.

To control for center bias in our scenes, we used a shuffle control approach. We did this because human-photographed scenes often include interesting objects (those likely to be relevant) in the center of the image [37, 38]. This makes it possible, for example, to misattribute fixation selection to a certain visual feature when the better explanation is simply that the observer tends to look towards the center of the image/screen. To control for this possibility, we asked whether the predictive power of visual features at fixated patches (for a given image) was greater than that of the features of the same patches (same fixation locations) superimposed on a randomly chosen image. If the predictive power of feature maps was only due to center bias, the visual features at fixated patches in true images should not be more predictive of saccades than the features of image patches in randomly chosen images on average.

Although it is well known that visual acuity is strongest at the fovea and falls off with eccentricity, predictive models of gaze behavior have not taken it into account. Indeed, we have recently shown that modeling visual acuity has enabled the discovery of gaze strategies at different time scales during visual search [6]. Therefore, to model decreasing visual acuity with peripheral distance, we processed the stimuli prior to computing the feature maps. More specifically, we applied a peripheral degradation filter with respect to the previous fixation location (see [6] for details).

To compare the distribution of feature (relevance, edge energy, or saliency) values for fixated patches in the true images and shuffled-control images, we generated the feature probability distribution functions (PDFs) by aggregating behavioral data across days for each monkey. We then generated the cumulative distribution function (CDF) of these PDFs for true and shuffled fixation patterns. To compare these two distributions, we plotted the true CDF versus the shuffled CDF, effectively yielding an ROC curve. We then computed the area under the ROC curve for each behavioral session from the Gabor search tasks.

Neural Data Analysis

The goal of this study is to ask what factors influence FEF activity during natural scene search. Because there may be multiple such factors, we used a multiple-regression approach: the Poisson Generalized Linear Model (GLM). We modeled neural activity using factors that potentially influence FEF spiking: upcoming saccade direction (as a proxy for saccadic motor command) and visual features (relevance and saliency). To quantify the extent to which each of these factors explains FEF activity, we fit the model to the experimental data.

Generalized linear modeling

Here we model extracellularly recorded spiking activity in the FEF as a Poisson process with a time-varying firing rate. The spiking activity of most cortical neurons follow Poisson statistics and we specifically verified that this was a reasonable assumption by checking that the variance was equal to the mean spike count over a wide range of spike counts across tasks, animals, and neurons. In general, it is reasonable to assume that the

variability around the mean spike count is Poisson distributed for sufficiently narrow time bins within which the neuron’s firing rate can be assumed to be constant (homogeneous). We used 10 ms time bins in our models. To model non-negative, time-varying firing rates, explanatory features are linearly combined and then passed through an exponential nonlinearity (the Poisson inverse link function) (Fig. 3). The number of spikes in each 10-ms time bin is then drawn randomly from a Poisson distribution with the mean given by the estimated firing rate in that bin.

To accurately model FEF neurons, we included the known characteristics of FEF RFs. FEF cells are typically influenced by movement and/or visual features in a particular part of retinocentric space, and have a stereotypical temporal response. In particular, firing rates of classical visually-tuned neurons are typically modulated by visual features within the receptive field (RF). Likewise, classical movement-tuned neurons are modulated by upcoming/current eye movements in the direction of the preferred movement direction (movement field). Thus, FEF RFs have both spatial and temporal components. We jointly estimated the spatial visual receptive field, the movement receptive field, as well as the temporal response to visual and movement features.

For mathematical tractability, we also assume that the spatial and temporal parts of the RF are multiplicatively separable. Details of the parameterization and fitting process are provided below.

Spatial Receptive Field Parameterization

Spatially, FEF RFs are retinocentric in nature, with centers ranging from foveal to eccentric [10]. This is true for classically defined movement, visual, and visuomovement

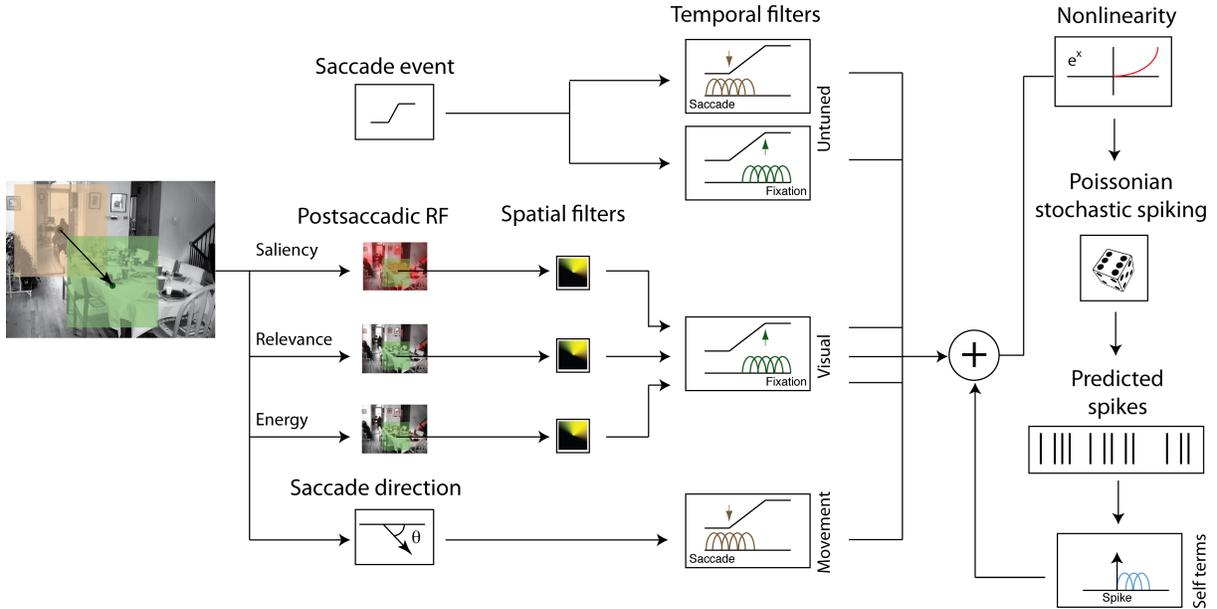


Figure 2.3: **Schematic illustration of the comprehensive generative model of neural spikes using a GLM framework.** The model comprises visual features: saliency, relevance and energy from a neighborhood around fixation location after the saccade, untuned responses aligned to saccade and fixation onsets, and the direction of the saccade. The features are passed through parameterized spatial filters (representing the receptive field) and temporal filters. The model also comprises spike history terms (or self terms). All these features are linearly combined followed by an exponential nonlinearity, which gives the conditional intensity function of spike rate, given model parameters. Spikes are generated from this model by sampling from a Poisson distribution with mean equal to the conditional intensity function. Brown: basis functions modeling temporal activity around the saccade onset; Green: basis functions modeling temporal responses around the fixation onset; Blue: basis functions modeling temporal responses after spike onset.

cells. For both movement and visual features, we parameterize space using polar coordinates (angle and eccentricity). More specifically, we use cosine tuning for the angular coordinate, and flat tuning for the eccentric coordinate. We chose flat eccentric tuning for mathematical tractability, but this assumption is realistic for many FEF neurons, which tend to have large RFs [10].

Temporal Receptive Field Parameterization

Because of the unconstrained nature of visual search, modeling the temporal responses of FEF neurons is complex. To simplify the problem, we chose to model only the neural activity in a fixed temporal window surrounding each eye movement (200 ms both before and after saccade initiation, as well as fixation onset) rather than all neural activity in each trial. This interval is large enough to contain both pre-saccadic activity and post-saccadic fixation-related activity.

The temporal responses of FEF neurons are heterogeneous, so we allowed for sufficient variability in their shape. To do this, we allowed both saccadic motor activity and visual responses to be explained by a range of temporal scales spanning the 400 ms temporal window. More specifically, we convolved the spatial receptive field features with a temporal basis set, g_i , consisting of 5 fixed-width ($w_i = 40$ ms) raised cosine functions (Eqn. (1)) centered at times $t_i = \{-140, -70, 0, 70, 140\}$ ms with respect to saccade onset (for upcoming eye movement related activity) or fixation onset (for visual activity). Ultimately, these temporal basis functions allow us to explain a wide range of neural response templates related to both saccades and fixation.

$$(2.1) \quad g_i = \left(1 + \cos \left(\frac{t_i - c_i}{2w_i} \right) \right)$$

Generative model

Using the above parametric structure, the entire generative model of spike activity is built using the following components (Fig. 4).

$$\begin{aligned}
\log \lambda(t) = & \alpha_0 \text{ Baseline firing rate} \\
& + \sum_{j=1}^G \alpha_j^{\text{sac}} g_j(t) * E_{\text{sac}}(t) \\
& + \sum_{j=1}^G \alpha_j^{\text{fix}} g_j(t) * E_{\text{fix}}(t) \\
& + \beta_{\text{sal}}^T \begin{bmatrix} \langle \cos(\psi)S(t) \rangle g_1(t) * E_{\text{fix}}(t) & \langle \sin(\psi)S(t) \rangle g_1(t) * E_{\text{fix}}(t) \\ \vdots & \vdots \\ \langle \cos(\psi)S(t) \rangle g_j(t) * E_{\text{fix}}(t) & \langle \sin(\psi)S(t) \rangle g_j(t) * E_{\text{fix}}(t) \\ \vdots & \vdots \\ \langle \cos(\psi)S(t) \rangle g_G(t) * E_{\text{fix}}(t) & \langle \sin(\psi)S(t) \rangle g_G(t) * E_{\text{fix}}(t) \end{bmatrix} \begin{bmatrix} \beta_c \\ \beta_s \end{bmatrix} \\
& + \beta_{\text{rel}}^T \begin{bmatrix} \langle \cos(\psi)R(t) \rangle g_1(t) * E_{\text{fix}}(t) & \langle \sin(\psi)R(t) \rangle g_1(t) * E_{\text{fix}}(t) \\ \vdots & \vdots \\ \langle \cos(\psi)R(t) \rangle g_j(t) * E_{\text{fix}}(t) & \langle \sin(\psi)R(t) \rangle g_j(t) * E_{\text{fix}}(t) \\ \vdots & \vdots \\ \langle \cos(\psi)R(t) \rangle g_G(t) * E_{\text{fix}}(t) & \langle \sin(\psi)R(t) \rangle g_G(t) * E_{\text{fix}}(t) \end{bmatrix} \begin{bmatrix} \beta_c \\ \beta_s \end{bmatrix} \\
& + \beta_{\text{rel}}^T \begin{bmatrix} \langle \cos(\psi)E(t) \rangle g_1(t) * E_{\text{fix}}(t) & \langle \sin(\psi)E(t) \rangle g_1(t) * E_{\text{fix}}(t) \\ \vdots & \vdots \\ \langle \cos(\psi)E(t) \rangle g_j(t) * E_{\text{fix}}(t) & \langle \sin(\psi)E(t) \rangle g_j(t) * E_{\text{fix}}(t) \\ \vdots & \vdots \\ \langle \cos(\psi)E(t) \rangle g_G(t) * E_{\text{fix}}(t) & \langle \sin(\psi)E(t) \rangle g_G(t) * E_{\text{fix}}(t) \end{bmatrix} \begin{bmatrix} \beta_c \\ \beta_s \end{bmatrix} \\
& + \gamma_{\text{mov}}^T \begin{bmatrix} \cos(\theta(t))g_1(t) * E_{\text{sac}}(t) & \sin(\theta(t))g_1(t) * E_{\text{sac}}(t) \\ \vdots & \vdots \\ \cos(\theta(t))g_j(t) * E_{\text{sac}}(t) & \sin(\theta(t))g_j(t) * E_{\text{sac}}(t) \\ \vdots & \vdots \\ \cos(\theta(t))g_G(t) * E_{\text{sac}}(t) & \sin(\theta(t))g_G(t) * E_{\text{sac}}(t) \end{bmatrix} \begin{bmatrix} \gamma_c \\ \gamma_s \end{bmatrix} \\
& + \sum_{j=1}^H \delta_j^{\text{self}} h_j(t) * y(t)
\end{aligned}$$

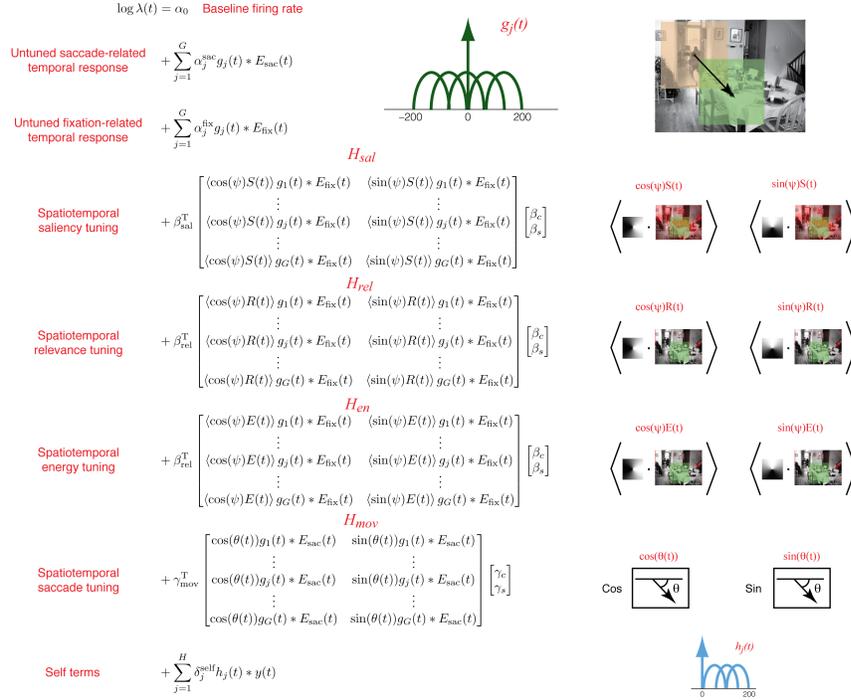


Figure 2.4: **Parameterization of the generative model.** Untuned saccade- and fixation-related temporal responses were modeled using linear combinations of raised-cosine temporal basis functions. Spatiotemporal tuning to saliency, relevance, energy, and saccade direction were modeled using bilinear models with left multipliers representing temporal basis function loadings, and right multipliers representing spatial basis function (cosine and sine) loadings. Additionally, a spike-history (self) term was modeled using a linear combination of temporal basis functions causally aligned to spike events. Parameters of the model were fit using maximum-likelihood with elastic net regularization.

- i) Baseline firing rate: We use a scalar term α_0 to model a constant baseline firing rate.
- ii) Untuned temporal responses: Neurons may have temporal responses to saccade and fixation events that do not depend on direction. We account for this possibility using separate untuned temporal responses aligned to both saccade onset and fixation onset. The untuned saccade-related response is given by $\sum_{j=1}^G \alpha_j^{sac} g_j(t) * E_{sac}(t)$, where α_j^{sac} are the free parameters, $g_j(t)$ are the basis functions, and $E_{sac}(t)$ is an event function (a delta function) specifying saccade onsets. Similarly, the untuned fixation-related response is given by $\sum_{j=1}^G \alpha_j^{fix} g_j(t) * E_{fix}(t)$. Each untuned response is specified by 5 free parameters.
- iii) Spatiotemporal visual (saliency, relevance and energy) tuning: We modeled the neural activity around a fixation event as a function of the visual features (relevance and saliency) in the RF with respect to that fixation location. For each fixation, we began by extracting the visual features from an image patch (400 x 400 pixels; or $\sim 13^\circ \times 13^\circ$) centered on the fixation location. The values for each visual feature were taken from their respective feature maps (see above). To simulate the effect of resolution decreasing with eccentricity, we then applied a blurring transform using image pyramids (Geisler and Perry, 1998; see [6] for details).

We do not know the visual RF a priori; it has to be estimated from the data. To do this, we first constructed the x and y components for each visual feature by applying a spatial cosine and sine mask, respectively, to the visual feature maps of the extracted image patch (Fig. 4). The RF center (preferred angle, Ψ^*) can then be inferred from the data by fitting the corresponding weights for the sine and cosine masked images, and using the trigonometric identity:

$$(2.2) \quad \cos(\Psi - \Psi^*) = \cos \Psi^* \cos \Psi + \sin \Psi^* \sin \Psi, \text{ and } \Psi^* = \tan^{-1} \left(\frac{\sin \Psi^*}{\cos \Psi^*} \right)$$

where $\cos(\Psi^*)$ and $\sin(\Psi^*)$ are linear model parameters of the GLM.

Along with estimating a spatial RF (via the above parameterization), we also wanted to model the temporal response around the fixation event. To this end, we convolved each spatial covariate (sine and cosine) with each of the 5 allowed raised-cosine temporal basis functions. Combining all of these parts, the complete spatiotemporal visual response can be succinctly specified by two bilinear models as:

$$(2.3) \quad \beta_{sal}^T H_{sal} \begin{bmatrix} \beta_c \\ \beta_s \end{bmatrix} + \beta_{rel}^T H_{rel} \begin{bmatrix} \beta_c \\ \beta_s \end{bmatrix} + \beta_{en}^T H_{en} \begin{bmatrix} \beta_c \\ \beta_s \end{bmatrix}$$

where β_{sal} , β_{rel} and β_{en} are free parameters specifying the temporal response to saliency, relevance, and energy, β_c and β_s specify the angular position of the visual RF, and H_{sal} , H_{rel} and H_{en} are matrices representing spatial sine and cosine covariates convolved with the temporal basis functions (Fig. 4). Note that we assume the same spatial RF for both saliency and relevance. (i.e., β_c and β_s are the same for saliency, relevance and energy). Thus, we have 5 free parameters each for the saliency, relevance and energy temporal response, and two free parameters for the spatial RF, resulting in a total of 17 free parameters.

- iv) Spatiotemporal saccade tuning: We modeled the neural activity around a saccade event as a function of the upcoming saccade direction (upcoming saccadic motor

command). To do this, we constructed the movement covariates as the sine and cosine projections of the upcoming eye movement direction. Note that we do not incorporate previous knowledge of the neuron’s RF from RF mapping tasks. As with the visual RF estimation, the angular position of the neuron’s movement field can be inferred from the data by fitting the model parameters corresponding to sine and cosine covariates.

As with the visual response, to simultaneously estimate a temporal movement response along with the movement field, we can specify a bilinear model as:

$$(2.4) \quad \gamma^T H_{mov} \begin{bmatrix} \gamma_c \\ \gamma_s \end{bmatrix}$$

Altogether, this response is specified by 5 temporal and two spatial parameters, resulting in 7 free parameters.

- v) Spike history terms: To further explain variability in neural activity, we included a spike history term. This feature is not central to the logic of our argument, but serves to improve our model of neural activity. To model the effect of spike history, we simply convolve the spike train with 3 raised-cosine temporal basis functions, $h_j(t)$, spanning a range of $[0, 200]$ ms with respect to each spike event. The basis functions were centered at 60, 100, and 140 ms. The temporal response for the spike history term is not coupled to any external events (i.e., fixation or saccade onset), and is given by $\sum_{j=1}^H \delta_j^{self} h_j(t) * y(t)$, where $y(t)$ is the spike train that we are modeling with the GLM. This response is specified by 3 free parameters, δ_j^{self} , $j = 1, 2, 3$.

Spatiotemporal RF fitting algorithm

To summarize our model parameterization: we model space with polar coordinates, but dispense with eccentricity for mathematical tractability. We model time using raised-cosine temporal basis functions.

For the spatiotemporal model terms, each model feature (saliency, relevance, energy, and upcoming movement direction) is initially parameterized by two covariates: the sine and cosine projections of that feature. We then model temporal responses of each spatial feature with five raised cosine functions. Thus, the sine and cosine projection of each spatial feature is convolved with five temporal basis functions. This leads to a total of 40 covariates for the model of saccadic motor and visual activity ($4 \times 2 \times 5 = 40$, four features, two spatial coordinates for each feature, five temporal basis functions for each spatial covariate).

For the temporal model terms, the untuned temporal responses aligned to saccade and fixation onsets have 5 covariates each, and the spike history term provide 3 covariates. Including the baseline term with all of these gives us a total of 53 covariates.

Using all 53 convolved covariates as an independent variable would make the maximum-likelihood estimation problem linear and convex, but could lead to different temporal responses for the sine and cosine terms, making the estimate hard to interpret. For example, it would be difficult to rationalize differing time courses for horizontal and vertical saccades.

To keep the model interpretable, we adopted the bilinear formulations of the spatiotemporal terms given in Eqns. (34), resulting in a total of 37 (17 for saliency, relevance,

and energy, 7 for movement, 10 for untuned, and 3 for spike history terms) free parameters. However, since the log-likelihood of this bilinear formulation is no longer convex, estimating it could result in local minima and would in general suffer from the difficulties of optimizing non-convex functions.

Therefore, to estimate the parameters of this model, we adopted an iterative algorithm in which we alternatively held the spatial parameters $(\beta_c, \beta_s, \gamma_c, \gamma_s)$, or the temporal parameters $(\beta_{sal}, \beta_{rel}, \beta_{en}, \gamma_{mov})$ fixed while fitting all the others. In this approach, each iteration step is a convex optimization problem. This method guarantees that the temporal response of each spatial covariate will be the same. We alternated between fitting stages until the model parameters converged.

Model Fitting

We trained and tested our model using non-overlapping 2-fold cross validation. To avoid overfitting, we estimated model parameters using elastic net regularization [39, 40, 41] (Glmnet implemented in Matlab). Regularization helps to select for simpler models by penalizing models with large or many parameter values. Elastic net regularization includes two free parameters: α , which determines the strength of L1 relative to L2 penalization, and λ , which determines the strength of regularization. We selected the values of these parameters ($\alpha = 0.01$, $\lambda = 0.05$) on a different data set using cross validation.

Model Comparison

The main goal of modeling spike trains using GLMs was to determine whether FEF activity significantly encodes visual or movement features. To address the main scientific question of whether the FEF encodes task relevant visual features, we fit partial models to

the data using relevance covariates only or movement covariates only, and compared them against joint models comprising both relevance and movement covariates. To maximally explain variance and address possible confounding factors, we also fit a more comprehensive and more complex model including saliency, energy and spike history terms. The partial models fell into two categories: leave-out models, and leave-in models. Leave-out models leave out the main feature of interest; the idea is that by comparing a leave-out model against a full model, we can quantify the marginal predictive power of the left out feature. Leave-in models only include the features of interest; the idea is to characterize the apparent encoding of these features by neural activity when other features are not considered. We used leave-out models to assess statistical significance, and leave-in models as an interpretive tool for apparent encoding.

To measure the quality of our model fits, we used two metrics: pseudo- R^2 , and conventional- R^2 . Pseudo- R^2 is related to the likelihood ratio and extends the idea of linear R^2 to non-Gaussian target variables. The idea of the pseudo- R^2 metric is to map the likelihood ratio into a $[0, 1]$ range, thus offering an intuition similar to the conventional R^2 used with normally distributed data. Many definitions exist for the pseudo- R^2 , but we used McFadden’s formula [42].

$$(2.5) \quad R_D^2(\text{model}) = 1 - \frac{\log L(n) - \log L(\hat{\lambda})}{\log L(n) - \log L(\bar{n})}$$

Where $\log L(n)$ is the log likelihood of a perfect model, $\log L(\hat{\lambda})$ is the log likelihood of the model in question, and $\log L(\bar{n})$ is the log likelihood of a model using only the average firing rate. More specifically, it can be interpreted as the relative improvement

that a given model offers above and beyond the simplest possible model (constant firing rate).

We also use a variant of pseudo- R^2 , relative pseudo- R^2 , to compare nested models (e.g., a partial model with the full model). The relative pseudo- R^2 quantifies the increase in model likelihood as a result of adding back the left-out features from the partial model (also maps to the $[0, 1]$ range). A relative pseudo- R^2 significantly greater than zero indicates that the left-out feature is a statistically significant explanatory feature.

$$(2.6) \quad R_D^2(\text{model 1, model 2}) = 1 - \frac{\log L(n) - \log L(\hat{\lambda}_2)}{\log L(n) - \log L(\hat{\lambda}_1)}$$

Where $\log L(\hat{\lambda}_2)$ is the log likelihood of the full model, and $\log L(\hat{\lambda}_1)$ is the log likelihood of the nested, partial model. It can be interpreted as the relative improvement due to the model components left out by model 1.

We computed pseudo- R^2 for all models and relative pseudo- R^2 for all partial models on test sets of both cross-validation folds. We obtained 95% confidence intervals on these metrics using bootstrapping. A left-out feature was deemed significant at four-sigma ($p < 0.006$; uncorrected for multiple comparisons) if the minimum of the lower bounds of the relative pseudo- R^2 s was greater than zero (a conservative measure).

Pseudo- R^2 values are not directly comparable to (are much smaller than) conventional R^2 , and are thus more difficult to interpret. We also computed conventional R^2 values by calculating the correlation coefficient between (1) the saccade- and fixation-averaged PSTHs and (2) the saccade- and fixation-averaged model predictions. Averaging across

hundreds of fixations renders Poisson spiking into a smooth curve, which can be compared with the smooth firing rate predictions of the model.

Explaining Away

By using a multivariate modeling approach, we are able to compare the relative contributions of different model components (e.g., visual features and upcoming eye movements). In this section, we elaborate on the nuances of interpreting the results of a multivariate analysis.

Even if two model components are both individually correlated with neural activity, multiple situations related to marginal explanatory power can arise in theory:

- i) If two components explain the same neural variability, they will have overlapping explanatory power. In this case, a multivariate model with both components is unlikely to be significantly better than the single best univariate model.
- ii) If two components explain different neural variability, they will have non-overlapping explanatory power. In this case, a multivariate model with both components is likely to be significantly better than both univariate models.
- iii) If two components explain neural variability that is similar but is neither completely identical nor distinct, an intermediate situation arises. In this case, a multivariate model is likely to be significantly better than both univariate models as the overlap between the two decreases.

PSTH Analysis

In addition to modeling individual neurons using GLMs, we also analyzed them in a conventional way using peri-saccade time histograms (PSTHs). We chose to perform this

analysis for one set of sessions for which relevance and energy were maximally predictive of fixation choice (i.e. for animal M16, vertical Gabor search task). We computed these PSTHs for two different sets of saccades as follows.

First, we selectively analyzed saccades into the movement RF as follows. We categorized each saccade into one of eight directional bins with bin centers at $0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \frac{-3\pi}{4}, \frac{-\pi}{2}, \frac{-\pi}{4}$, and averaged the firing rates in 10 ms time bins across saccades, separately within each directional bin. We then considered the movement-field (movement RF) to be along the directional bin having highest peak firing rate around the saccade. We took this within-RF PSTH for each neuron and max-normalized it, i.e. set the peak firing rate within this bin to 1. We then averaged these normalized PSTHs across neurons to produce a population-averaged PSTH. We separately calculated such a population-averaged PSTH for saccades categorized according to high and low relevance (top 50% of saccades to high- and low-relevance locations) at saccade landings, as well as high and low edge-energy.

Second, we selectively analyzed saccades out of the movement RF as follows. As before, we categorized saccades into one of eight directional bins and computed the PSTHs for each directional bin within 10 ms time bins. We then defined saccades out of the RF to be those that were not into the directional bin defined by the maximum peak firing rate (i.e. the movement RF) as well as the two neighboring directional bins. We averaged the PSTHs across these five out-of-RF directional bins. As before, we max-normalized it before averaging them across neurons. Crucially, we calculated these out-of-RF PSTHs

separately for high and low relevance (top and bottom 50% of saccades) within the pre-saccadic RFs (not those at saccade landings). We also calculated these PSTHs separately for high and low edge-energy, again within pre-saccadic RFs, not saccade landings.

Power Analysis

Since we found that relevant visual features did not explain FEF activity, we wanted to ask whether our approach had sufficient power to detect such an effect. To do this, we simulated neural data with known parameters (e.g., a weak relevance representation), and attempted to detect the effects of those parameters. We used behavioral data from a single experimental session (eye movements and image features), and simulated spiking activity according to our Poisson model.

To ask whether we could detect a weak effect of relevance, we first simulated neural activity with a range of representation strengths for relevance. As a measure of representation strength, we defined modulation depth as the relative change in firing rate due to a 1 standard deviation change in the given feature. For example, if a fixated image patch had a relevance value that was 1 standard deviation above average, and led to a 10% increase in firing rate from baseline, we would say the modulation depth was 10% (alternatively, 1.1). We explored ten modulation depths that were evenly spaced between 1.05 and 1.30. We used a baseline firing rate of 20 spikes/s, and ensured that average firing rates were the same across modulation depths. To do this, we randomly removed spikes to achieve an average of 20 spikes/s in each condition. A given representation was said to be detectable if the marginal predictive power of relevance was statistically greater than zero (if its 95% confidence interval did not overlap zero).

To explore the consequences of using noisy or inaccurate relevance models, we performed two additional analyses. In the first, we simulated neural data according to our relevance model as above. We then corrupted the relevance model covariates with Gaussian noise before fitting the model.

$$(2.7) \quad x_{fitted} = x_{simulated} + \alpha N$$

Where $x_{simulated}$ is the covariate used for data simulation with $\mu = 1, \sigma = 1$, x_{fitted} is the covariate used for fitting, N is Gaussian noise with $\mu = 1, \sigma = 1$, and α tunes the degree of noise added. In addition to the uncorrupted covariates, we used three levels of noise. These corrupted model covariates used for fitting were correlated with the model covariates used for simulation with $r = 0.8, 0.5, 0.2$. This procedure allowed us to explore when the effect of relevance was detectable, even if the definition of relevance used for fitting was a noisy version of the correct definition.

In the second analysis, we characterized the consequences of simulating and fitting the model with qualitatively different definitions of relevance. Specifically, we simulated the data using edge-energy, a visual feature that is correlated with relevance ($r=0.61$) and fit the model using relevance [6]. This allowed us to characterize the effects of fitting the model with an inaccurate version of the correct definition of relevance.

Results

When we search for a known target, we can use properties of that target to guide our search. In this study, we ask whether the Frontal Eye Field (FEF), a region heavily

implicated in eye movement selection, facilitates feature-based attention by biasing gaze towards target features, or whether it reflects spatial selection of subsequent eye movement targets. We recorded a heterogeneous population of FEF neurons from three macaques while they performed several variations of a natural scene search task. In each task, a target known to the monkeys ahead of time was embedded in natural scenes, and the monkeys were rewarded for successfully fixating this target (see Methods for details). Since the spatial distribution of targets in the scene was extremely broad and there was no contextual information about their location in the scene, spatial attention would not help them find the target. Furthermore, since the targets were blended into the scene, they did not stand out with respect to their local background. Thus, looking for salient objects was not a viable strategy either. However, because the targets were known ahead of time, the monkeys could use target features – task-relevant features – to guide their search. To disentangle possible influences of feature-based attention and spatial selection on FEF activity, we then analyzed the data with a multiple-regression approach.

First, we found that monkeys were able to perform the tasks to varying levels of success (see Table 2 for the search performance of individual monkeys averaged across sessions). The Fly task was somewhat easier because it was a large, mostly black and white target blended into a colored natural scene, and because its position was restricted to a 3 - 30° range around the center. By comparison, the Gabor task with a grayscale background and a grayscale target uniformly distributed around the scene was significantly harder.

Next, we asked if monkeys indeed use visual features to guide their search. To do this, we performed a Receiver Operator Characteristic (ROC) analysis to examine whether image patches with higher task relevance predicted fixations. Since human-photographed

Animal	Task	Trial duration (s)	Success rate (%)	# Saccades to locate target	Fixation duration in successful trials (ms)	Fixation duration in failed trials (ms)
M14	Fly	1.4 ± 0.3	72.8 ± 7.5	4.7 ± 0.3	185 ± 17	221 ± 47
M15	Fly	1.4 ± 0.4	98.8 ± 6.0	3.3 ± 0.6	302 ± 38	465 ± 317
M15	Gabor (H)	6.1 ± 0.7	49.0 ± 13.1	5.1 ± 0.7	351 ± 45	321 ± 81
M16	Gabor (H)	7.8 ± 1.1	38.8 ± 15.8	7.2 ± 0.9	297 ± 19	303 ± 34
M16	Gabor (V)	6.1 ± 1.2	75.2 ± 19.0	5.4 ± 0.8	316 ± 11	275 ± 49

Table 2.2: **Statistics of search behavior for each monkey and task summarized across sessions** (mean \pm SEM). The trial duration is the entire duration from scene onset to scene offset. The success rate is the percentage of trials in which the monkey successfully located the target. The number of saccades to find the target is given only for successful trials, and includes the last, target-finding saccade. The fixation duration is averaged across all saccades in successful and failed trials.

natural scenes are known to have a center bias, we compared the feature distribution of fixated image patches to the feature distribution of the same fixation pattern superimposed on a randomly chosen image (see Methods for details). None of the visual features were able to predict fixation choice significantly above chance (AUC of 0.5) for the Fly search task (Fig. 5). For the Gabor task, both relevance and energy predicted fixations only for monkey M16 (Fig. 5). The overall predictive power of relevance was weak (AUC $<$ 0.55) but energy was more strongly predictive of fixations than relevance for the same subset of tasks (Fig. 5; AUC $>$ 0.6). Although the effect size of relevance appears modest, it is reasonable in the context of most predictive models of gaze behavior (for a recent survey see [43]; the best performing models have an AUC of under 0.6). Thus, at the very least,

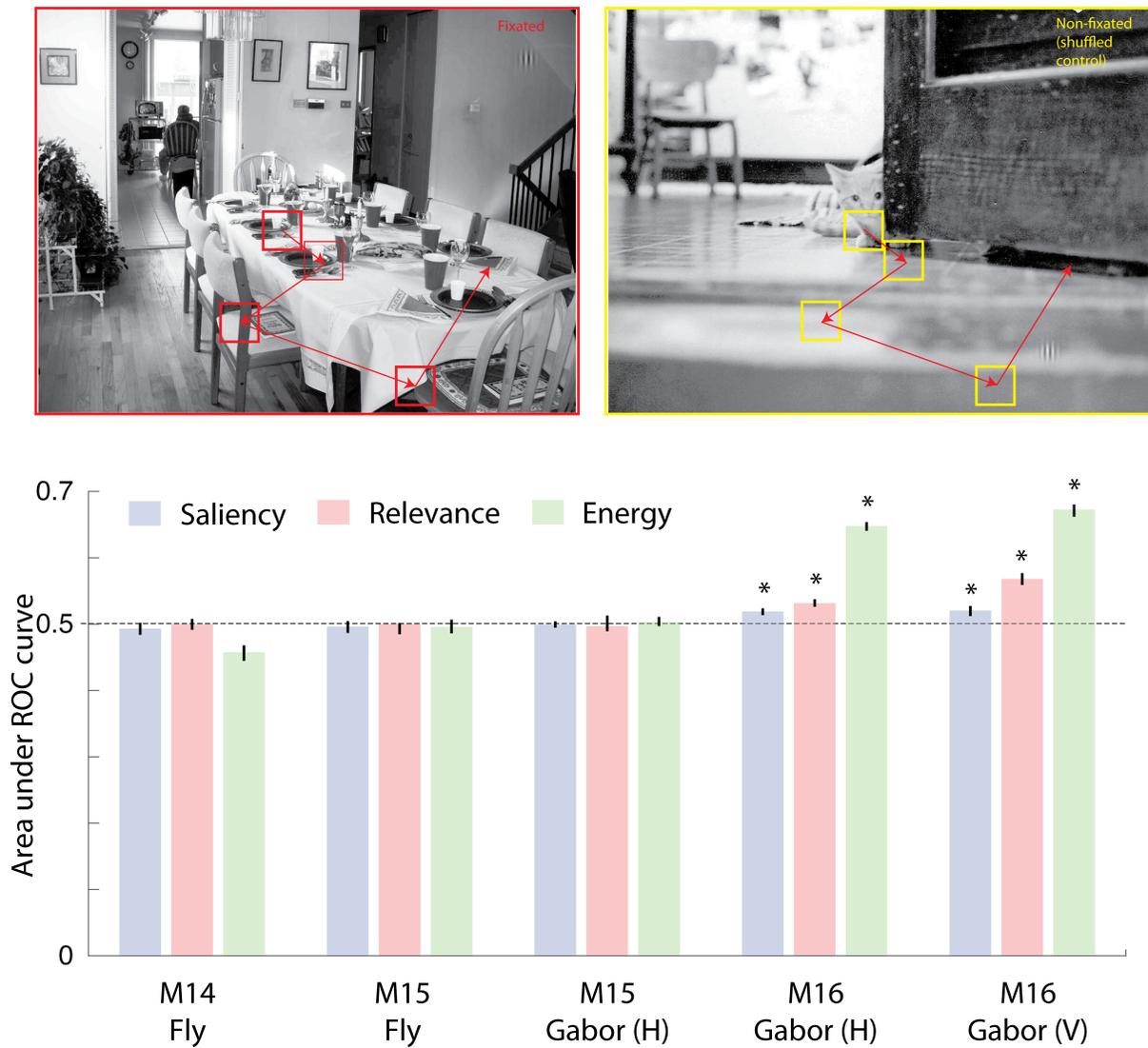


Figure 2.5: **Prediction of gaze using visual features at fixation.** We compared bottom-up IK-saliency, top-down relevance and energy at fixated (above: left panel) and non-fixated, i.e. shuffled control (above: right panel) targets by computing the area under the ROC curves (below). The star indicates statistically significant difference from a chance level of 0.5 at a significance level of $p < 0.05$.

correlates of task relevance and energy may be encoded by brain regions responsible for saccade selection.

The ROC analysis only provides us with a session-by-session summary statistic of the influence of search-target related features on fixation selection. To ask if the influence of these features on fixation choice was modulated by search performance, we performed three different analyses. First, we separately analyzed the fixations from successful and unsuccessful trials. However, we found no significant differences between the AUCs across these two types of trials (not shown). We did not analyze the Fly search task in this way because we did not have a sufficient number of failed trials for reliable estimation of ROC curves.

Second, we asked whether the predictive power of visual features (AUC) in a given session was correlated with search performance (percentage of trials in which the monkey found the target). We found a strong correlation between relevance and search performance, but only for the vertical Gabor search task performed by M16, for which relevance was significantly predictive of fixation choice (Table 3). Surprisingly, we did not find such a correlation between energy and search performance even though energy was predictive of fixation choice (Table 3). This strongly suggests that whenever relevance has an effect on fixation choice, it also has an effect on search performance. Unlike relevance, energy predicts fixations but not search performance. Thus, although energy is predictive of saccade targets, it may be a bottom-up feature and therefore not an important factor in feature-based attention.

Third, the natural distribution of saccade velocities is likely to reflect the distribution of urgency with which the animal selects fixations. If the peak velocities of saccades were correlated with relevance or energy, it would suggest that these features influence the conscious choice of fixations. However, we did not find any correlation between peak

Animal	Task	Relevance vs. Success Rate	Energy vs. Success Rate
M14	Fly	$r = -0.18, p = 0.52$	$r = -0.49, p = 0.06$
M15	Fly	$r = 0.21, p = 0.31$	$r = 0.32, p = 0.11$
M15	Gabor (H)	$r = 0.14, p = 0.49$	$r = -0.02, p = 0.90$
M16	Gabor (H)	$r = 0.20, p = 0.25$	$r = 0.21, p = 0.22$
M16	Gabor (V)	$r = 0.96, p < 10^{-4}$	$r = 0.18, p = 0.60$

Table 2.3: Correlation between success rate and ROC values across sessions.

saccade velocity and relevance or energy (on average across sessions and animals, these correlations did not exceed 0.05 ($-0.05 \leq \text{Pearson's } r \leq 0.05$)).

Taken together, these behavioral analyses suggest that saliency is not predictive, relevance is weakly but significantly predictive, and energy is strongly predictive of fixation locations for a subset of animals and tasks. There appears to be no correlation between behavioral parameters such as peak saccadic velocity and visual features of saccade landings. Importantly, even though relevance was weakly predictive, when it was predictive of saccade choice, its predictive power was correlated with success rate of search behavior.

Based on the weak behavioral effect alone, it is still very possible that the FEF could represent feature-based attention as operationally defined by relevance. Indeed, just as cortical area V1 represents visual information regardless of whether such information directly informs behavioral choices, the FEF could represent relevance in order to consider and reject potential saccades to locations that are similar to the target, but not sufficiently similar as to warrant a saccade. In such a circumstance, feature-based attention would inform saccadic decisions, but would not manifest in measurable fixation behavior, since

saccades would only be made to locations that have a very high target similarity. Therefore it is important to examine neural activity no matter the result of the behavioral analysis.

Next, we asked which features best explain FEF activity during natural scene search. We used a multiple-regression approach to model neural activity while monkeys performed a target search task in natural scenes. More specifically, we used the Poisson Generalized Linear Model framework to explain spiking events in terms of behavior and task variables. We modeled spiking in terms of visual feature maps of relevance and energy, in addition to other features thought to be encoded by the FEF: upcoming eye movements (as a proxy for spatial selection or planning), and visual feature maps. In this way, we tested the potential influence of feature-based attention on FEF activity during natural scene search.

We found that some neurons appeared to be well explained by a simple model of saccadic motor activity aligned to saccade onsets, based on the direction of upcoming movement alone (Fig. 6). Others initially appeared to be explained by a simple model of relevant visual features alone, aligned to the onset of fixation (Fig. 7).

Much previous work has characterized the activity of FEF cells as movement-related or visually related using simple, artificial stimuli and tasks. Therefore, having first established that univariate models can explain both visually-related and movement-related activity, we used our multiple-regression approach to characterize neurons in the same way: to what extent is the activity of FEF neurons predicted by upcoming movement, relevance, or some combination thereof? We addressed this question by comparing univariate models of relevance against multivariate models of relevance and movement tuning.

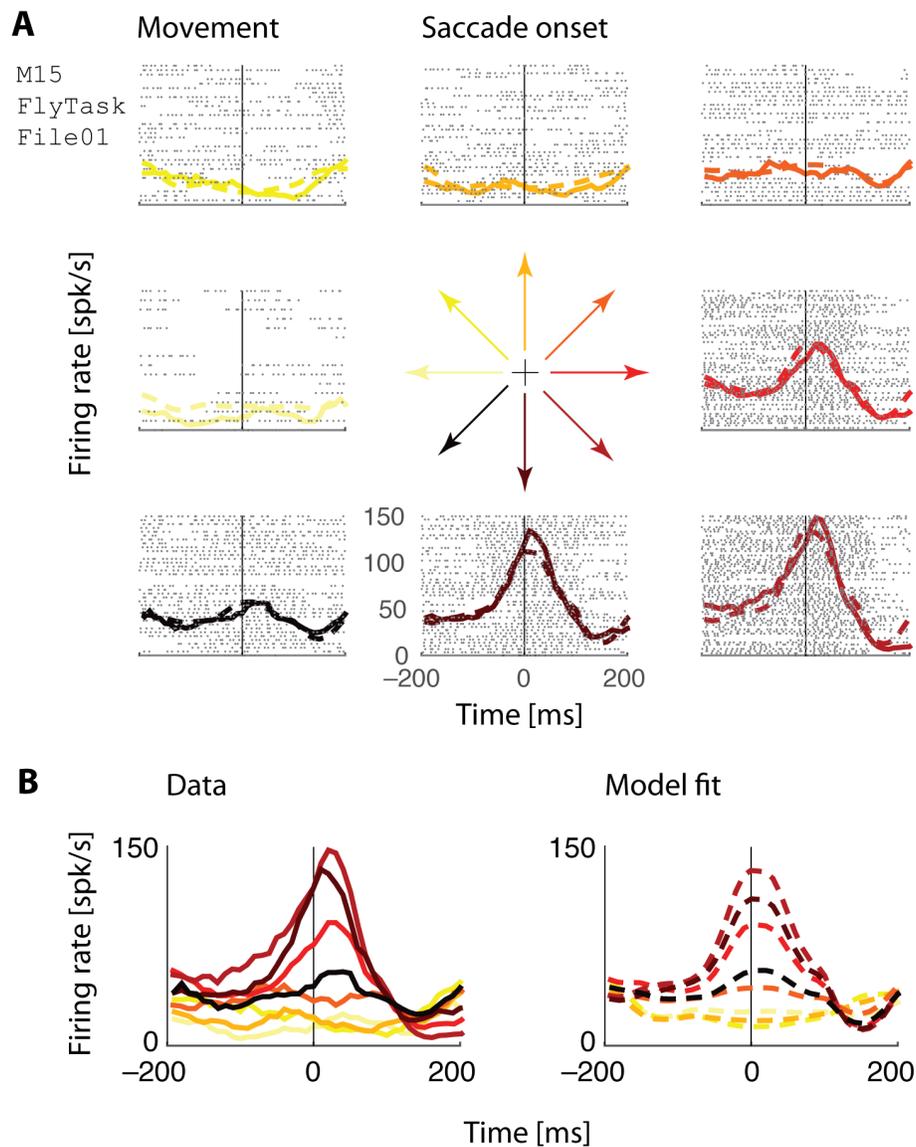


Figure 2.6: **Example neuron fit using a movement-only model.** The model was fit to an independent held-out half of the data that is not visualized here (**A**) Raster plots, peri-saccadic time histograms (PSTHs) and model predictions, separated into 8 categories according to the direction of the upcoming saccade (color-coded in the central glyph), and aligned to saccade onset (vertical line). (**B**) Data and model PSTHs aligned to the saccade onset show tuning to upcoming movement direction.

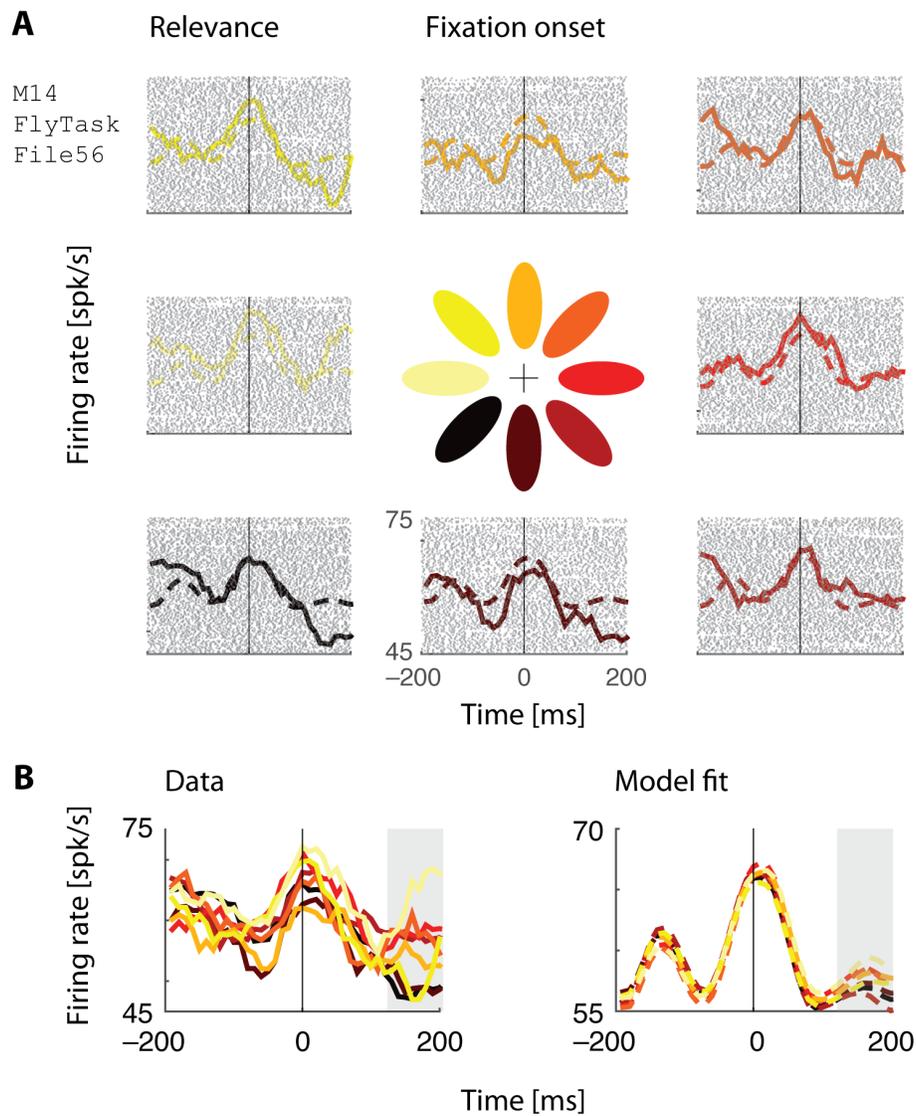


Figure 2.7: **Example neuron fit using a relevance-only model.** The model was fit to an independent held-out half of the data that is not visualized here (A) Raster plots, fixation-aligned PSTHs, and model predictions, separated into 8 categories according to the direction of the maximally relevant octant (color-coded in the central glyph). (B) Data and model PSTHs aligned to the fixation onset do not show a clear tuning to direction of maximum relevance during the fixation onset, but do show a modest tuning at around 200 milliseconds after fixation, as indicated by the gray panel.

We found that although some neurons initially appeared to encode relevant visual stimuli, an upcoming movement was a better predictor of neural activity aligned to fixation (example neuron in Fig. 8A). For most neurons, we found that a relevance-only model was significantly improved by adding a movement covariate (Fig. 8B, top panel), but a movement-only model was not significantly improved by adding a relevance covariate (Fig. 8B, bottom panel).

We formally quantified the predictive power of relevance and movement using a relative pseudo- R^2 analysis, which compares a model leaving out the covariate of interest (either relevance or movement) against a joint model comprising both relevance and movement, as well as a full model comprising additional covariates including bottom-up saliency, edge-energy, and self terms (see Methods). We found that apparent tuning to relevance (Table 4, column 3) was progressively explained away when compared against the joint model and the more comprehensive full model (Table 4, columns 5 and 6). A nearly identical effect was observed for edge-energy (Table 4, columns 7 and 8). However, movement does not get explained away when comparing a model that left movement out against the joint model comprising relevance and movement (Table 4, column 4). Therefore, neural activity is correlated with both saccades and relevant image patches, even though only one of the features (upcoming saccades) is truly encoded by neural activity.

Although the GLM analysis convincingly suggests that movement explains away any apparent effect of relevance or energy, it may be limited by the specific assumptions of the linear-nonlinear Poisson model. Therefore, we analyzed the data using a more conventional technique by visualizing PSTHs, only for the set of sessions for which relevance and edge-energy were maximally predictive of fixation choice (M16, vertical Gabor search).

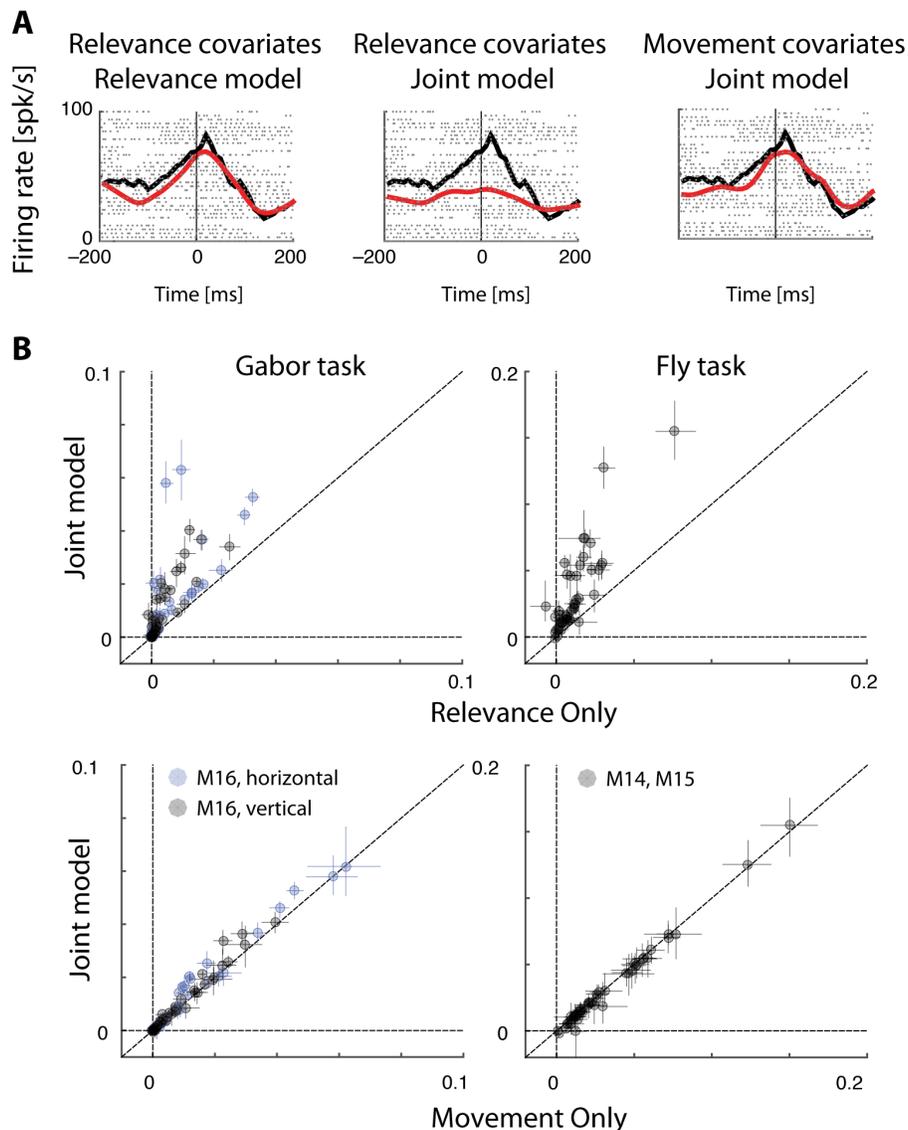


Figure 2.8: **Apparent tuning to relevance is explained away.** (A) Data PSTHs (black) and corresponding model predictions (red) for a single example neuron, of (from L to R): relevance-only model, relevance covariates of the joint model comprising relevance and movement, and movement covariates of the joint model, overlaid on top of each other. Dots show spike rasters. Both rasters and model predictions are aligned to fixation onset (B) Scatter plots of pseudo- R^2 (goodness of fit) values of univariate against multivariate models. Top panel: relevance-only model vs. a joint model comprising relevance and movement. Bottom panel: movement-only model vs. a joint model comprising movement and relevance.

Model		Rel. only	Mvt. + Rel.	Mvt. + Rel.	Full	En. only	Full
Animal	Task (# nrn)	# Rel. nrn	# Mvt. nrn	# Rel. nrn	# Rel. nrn	# En. nrn	# En. nrn
M14	Fly (15)	5	9	0	0	4	0
M15	Fly (25)	8	8	1	0	8	0
M15	Gabor H (81)	57	55	1	1	54	1
M16	Gabor H (57)	24	25	1	0	28	3
M16	Gabor V (49)	23	24	1	0	22	1

Table 2.4: **GLM analysis summary statistics.** Number of neurons that were significantly modulated by relevance, energy or movement in different models. Neurons were deemed to be significantly tuned if the 4 confidence intervals of the (relative) pseudo- R^2 s exceeded zero. We used a strict 4 threshold to sufficiently correct for multiple comparisons across neurons and models. Column 3: number of significant neurons for the relevance-only model. Column 4: number of neurons significantly tuned for movement by comparing a leave-movement-out model against a joint model with movement and relevance. Column 5: number of neurons significantly tuned for relevance by comparing a leave-relevance-out model against a joint model with movement and relevance. Column 6: number of neurons significantly tuned for relevance by comparing a leave-relevance-out model against a comprehensive full model (see text). Column 7: number of significant neurons for the energy-only model. Column 8: number of neurons significantly tuned for energy by comparing a leave-energy-out model against a comprehensive full model (see text).

Specifically, we computed population-averaged normalized PSTHs for saccades into the RF (see Methods), separated by the top and bottom 50% of relevance (or energy) of saccade landings (Fig. 9, above). We also computed these PSTHs for saccades out of the RF (see Methods), separated by the top and bottom 50% of relevance (or energy) in the pre-saccadic RF (Fig. 9, below). We found no significant firing rate differences between

these saccades, suggesting yet again that relevance and energy have no aggregate effect on population FEF firing rates during natural scene search.

In summary, we found that upcoming movements, rather than relevant visual features, best explained neural activity in many neurons examined across all tasks. Could this result simply be due to a lack of statistical power? Neural representations of visual features are likely more complicated and weaker in natural scenes than in simple, artificial scenes. To address this possibility, we performed a power analysis based on simulated neural data. In short, we used behavioral data from a real experimental session to simulate many versions of neural data with different modulatory influences. We then fit this data with our model to ask if these influences were detectable (see Methods). We found that, even when the simulated modulation depth (relative increase in firing rate; see Methods) of relevance was low, we were able to detect its influence on neural activity (Fig. 10A, green line). Therefore, assuming our model of relevance representation is accurate, we would be able to detect relevance representations even when they were weak.

But what if our model of relevance is not accurate? The monkeys may use visual features to locate the target that are similar but not identical to our convolution-based relevance measure. To explore this issue, we performed two additional power analyses to ask whether we could detect the influence of relevance on neural activity. Both analyses make use of the following approach: We use one set of covariates to simulate the neural data, and a different set of covariates to fit the model to the data. This procedure mimics the situation in which we have only approximate information about the relevance model that the monkey is using. In the first analysis, we used our standard relevance model covariates to simulate the neural data, but fit the model using relevance model covariates

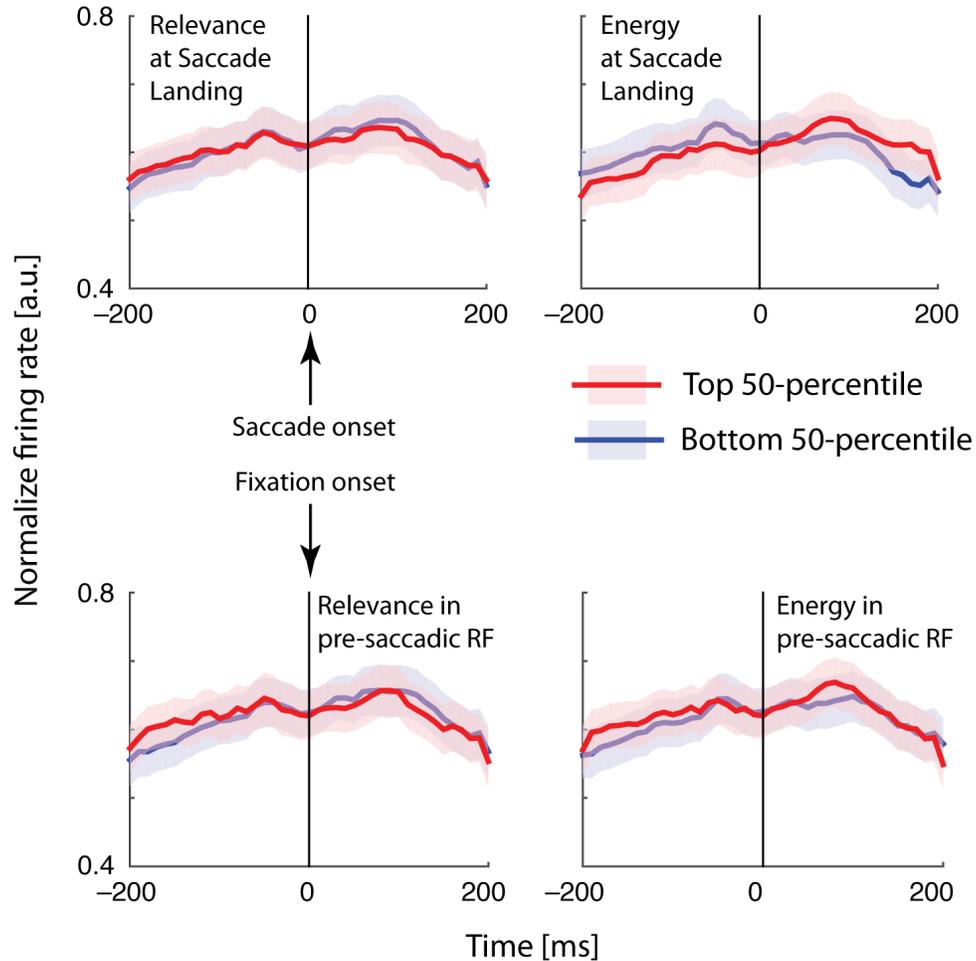


Figure 2.9: **Relevance and edge-energy do not modulate FEF firing rates around saccade onset (above) or fixation onset (below).** We divided saccades into the top and bottom 50 percentiles of the respective visual features either at the saccade landing (above) or within the pre-saccadic receptive field (below), and calculated the saccade- (above) or fixation- (below) aligned PSTHs for each neuron separated in eight directional bins. To combine these across neurons, we then selected the directional bin with highest peak firing rate to represent the receptive field, and averaged the within-RF PSTHs across neurons to compute the saccade-aligned PSTHs (see Methods). For fixation-aligned PSTHs, we first averaged responses around fixation around the directional bins representing out-of RF saccades and then averaged these single-neuron PSTHs (see Methods).

that were corrupted with noise (see Methods). Although our ability to detect the influence of relevance was decreased, detection required a stronger modulation depth of relevance; this modulation depth was still physiologically plausible (Fig. 10A). For example, even when the correlation between the relevance models for simulating and fitting was only .5, we could detect relevance when its modulation depth was ~ 1.13 (compared with ~ 1.08 when no noise was added) (Fig. 10A, red line).

In the second analysis, we used two qualitatively different models for simulating and fitting the model. We simulated neural data using edge energy (see Methods for definition), and fit the model using our convolution-based measure of relevance. Although not identical, edge energy and relevance are correlated (e.g., $r = 0.61$ for the Gabor search task), meaning that we should nonetheless be able to detect its influence. Our ability to detect relevance was indeed decreased, but not beyond physiological plausibility. More specifically, we could detect relevance when its modulation depth was ~ 1.13 , even though a qualitatively different model (edge energy) produced the data. Therefore, we find that our method of estimating the influence of relevant visual features on neural activity is robust to both weak effects and inaccurate models of relevance. Our inability to find relevance representations in the real data suggests that it is represented very weakly, or that it is represented in a more complex manner than our bilinear spatiotemporal model can describe.

Discussion

In this study, we used a modeling-based approach to analyze the frontal eye field (FEF) of monkeys while they searched complex, natural scenes. They searched for a target

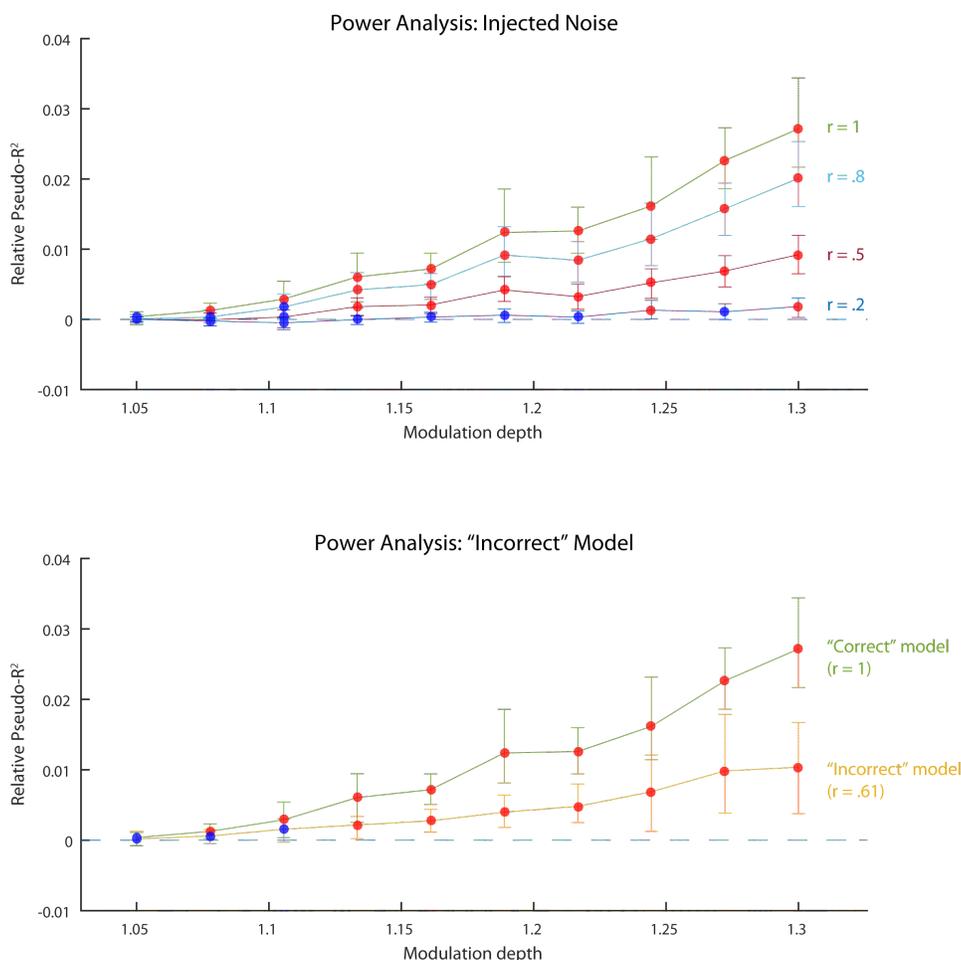


Figure 2.10: **Power analysis.** Both plots show marginal predictive power of relevance as a function of its simulated modulation depth. Red points indicate that the marginal predictive power is significantly greater than zero; blue points indicate that it is not significantly greater than zero. Error bars are bootstrapped 95% confidence intervals. **(A)** Injected noise analysis. After simulating the neural data, we corrupted the relevance model with varying degrees of noise, and then fit the model. Each line represents the model fits for a different level of added noise. The r values indicate the correlation between the original model covariates and the corrupted model covariates. **(B)** Incorrect model analysis. We simulated neural data using edge energy, and fit the model using the relevance model (yellow line). For comparison, we include the model fits when the correct model, relevance, was used to simulate the spikes (green line).

known ahead of time, making it possible for them to guide their saccades using target-similar features (task relevance). We then asked if FEF reflected feature-based attention, i.e., whether neural activity was explainable using task relevance and edge energy. We found that FEF activity was explained primarily by upcoming eye movements (a proxy for spatial selection or planning), and not by task relevance (a proxy for feature-based attention), or bottom-up influences such as saliency and energy.

Studies investigating the neural basis of feature-based attention have implicated the FEF in feature-based attention [13]. However, a recent study by the same group [44] implicates a relatively unexplored region (but see Kennerley and Wallis, 2009) in the prefrontal cortex the ventral pre-arcuate area (VPA) as the primary source of feature-based attention. They showed that pharmacologically inactivating the VPA interfered with feature-based search behavior, and also eliminated the signature of feature-based attention in FEF activity, while leaving the signature of spatial selection unaffected. This revised understanding of the role of FEF in attention may explain why we did not find a clear reflection of feature-based attention in FEF activity.

Our study thus raises the question of how results from simplified stimuli and tasks generalize to complex, natural vision. Beginning with the pioneering studies of Bruce and Goldberg [45, 10, 25], many studies have implicated the FEF in planning saccades [46, 47, 48], employing covert attention [13, 27], and selecting salient [31, 32, 49] and task-relevant objects from distractors [50]. While these studies have provided the foundation of our understanding about FEF function, they have typically used simplified tasks (involving a single cued saccade) in the context of artificial stimuli (that present limited choices against a homogeneous background). By contrast, natural scene search requires

navigating hundreds or thousands of distracting stimuli, and often requires making tens of self-initiated saccades. In support of these differences being important, a recent study from our group found that FEF activity is better explained by upcoming eye movements than visual saliency [33]. The current study extends this work by examining the influence of top-down, rather than bottom-up, visual features on FEF activity. A central component of feature-based attention is a top-down bias for target-like objects, which is exactly what we failed to detect. Our study thus provides another line of evidence that FEF function may differ in the context of natural behavior and stimuli.

One important limitation of this study is the weak effect of relevant visual features on search behavior. The main implication of this weak effect is that if our relevance metric does not accurately model the search strategy used by monkeys to plan their saccades, then the brain may not reflect relevance. Although we found that the predictive power of relevance on search behavior varied across animals and tasks, relevance failed to predict firing rate changes. Despite this, we found, using simulations, that both noisy and incorrect definitions of relevance in our model were able to estimate the simulated effects on firing rates. Furthermore, despite this weak effect of relevance, when monkeys used relevance to inform saccade selection, they were more likely to be successful in finding the target (Table 3). Therefore, it is important to test the possibility that relevance modulates neural activity during saccade planning.

Our difficulty in predicting fixation choice with high accuracy suggests that better models of behavior are likely to be more successful in discovering the role of FEF in feature-based attention during naturalistic vision. However, even cutting-edge behavioral models of gaze achieve modest predictive power (for instance, see [43] for a review of

contemporary saliency models; best AUC ; 0.6). These numbers highlight the difficulty of modeling complex, natural behaviors. Indeed, fixation choice is likely driven by many factors beyond relevant and salient visual features. Furthermore, not all saccades in a natural scene are made to locations that maximize immediate expected reward. Some saccades are corrective saccades, bridging the discrepancy between intended and current gaze locations in a sequence [51, 52, 53]. Other saccades are exploratory or information-gathering in intent, which are useful to maximize long-term expected reward [35, 54]. These possibilities suggest that improved behavioral models of gaze as well as improved models of neural coding of behavioral variables might yield more success in understanding the computational role of the FEF during search. For example, more sophisticated models could incorporate the need to balance exploring the scene with exploiting particular image patches, or take into account the shifting spatial spotlight of covert attention.

Recent studies with natural scenes, including our own [55, 33], have suggested that FEF may not encode visual information that is not targeted by an upcoming saccade. Such studies have called into question the conventional understanding that FEF represents a feature-based priority map. These findings need to be reconciled with findings from artificial tasks. What explains the discrepancy between results from artificial scenes and our findings using natural scenes? One possible explanation arises from the number of potential saccade targets in complex natural stimuli. In artificial search tasks with few saccade targets (typically fewer than eight), it may be possible to deploy covert attention to all of them. Therefore, selecting the saccade target based upon its similarity to the search target is a feasible strategy, and FEF activity might reflect this similarity. In crowded natural scenes, by contrast, the space of possible saccade targets is continuous

(infinite). In these contexts, it might only be feasible to attend to a local region around the point of fixation using feature-based attention. If this were true, the animal is more likely to be successful by making several saccades to new areas in order to maximize the likelihood of finding the target within these parafoveal regions. Therefore, during natural scene search, FEF activity might primarily reflect spatially selected saccade landings.

Another important discrepancy between artificial and natural search tasks arises from the natural behavior elicited by our tasks. Several studies with artificial stimuli have described an evolution of FEF activity from stimulus selection to saccade selection, suggesting distinct roles and classes of neurons in the continuum between representing visual and movement variables [56, 32, 49]. In our study, rather than holding fixation, the monkeys were allowed to freely move their eyes to locate targets. Hence, they are likely to immediately saccade to a target (or relevant stimulus) once it is detected. This means that target detection and making saccades to the target are likely to be tightly coupled during natural behavior. This close natural overlap between saccade decision and execution makes it challenging to disambiguate neural activity specifically related to each.

In summary, our results suggest that, during natural vision, the FEF does not reflect feature-based attention. We emphasize that upcoming eye movements explain away the effects of visual features on FEF activity. It is not the case that we failed to explain FEF activity, but rather that upcoming eye movements provide a better explanation than task-relevant visual features. These results suggest that FEF activity more closely reflects movement than sensory variables during natural scene search. They also expose an important need to develop better models of gaze behavior during natural scene search.

More broadly, our study illustrates the complexity of understanding the role of higher-order brain areas during unconstrained, natural behaviors.

CHAPTER 3

The role of expected reward in frontal eye field during natural scene search**Abstract**

When a saccade is expected to result in a reward, both neural activity in oculomotor areas and the saccade itself (e.g. its vigor and latency) are altered (compared to when no reward is expected). As such, it is unclear whether the correlations of neural activity with reward indicate a representation of reward beyond a movement representation; the modulated neural activity may simply represent the differences in motor output due to expected reward. Here, to distinguish between these possibilities, we trained monkeys to perform a natural scene search task while we recorded from the frontal eye field (FEF). Indeed, when reward was expected (i.e., saccades to the target), FEF neurons showed enhanced responses. Moreover, when monkeys accidentally made eye movements to the target, firing rates were lower than when they purposively moved to the target. Thus, neurons were modulated by expected reward rather than simply the presence of the target. We then fit a model that simultaneously included components related to expected reward and saccade parameters. While expected reward led to shorter latency and higher velocity saccades, these behavioral changes could not fully explain the increased FEF firing rates. Thus, FEF neurons appear to encode motivational factors such as reward expectation, above and beyond the kinematic and behavioral consequences of imminent reward.

Introduction

When an animal expects a movement to result in a reward, activity in motor circuits of the brain is modulated [57, 58, 59, 60]. The result is movements with greater vigor (higher velocity), lower latency, and greater accuracy [61, 62, 63]. In other words, the movement itself reflects the expectation of reward. Hence, the muscles and motor units driving the movement reflect reward expectancy, but only to the extent that the movement itself is different when reward is expected. As the motor pathway is traced farther back from motor units to the more complex cortical and subcortical circuits that plan the movements, it becomes more likely that reward expectancy influences variables other than movement alone [64]. In any given brain area, to test whether reward is represented beyond its obvious influence on movement, it is essential to consider movement and reward parameters simultaneously.

The oculomotor system has been an important model for understanding the effects of reward on neural activity and movement [65, 66, 67]. However, much is still unknown about reward processing in the frontal eye field (FEF), a prefrontal cortical area involved in the control of saccades [10, 68] and spatial attention [44, 27, 28]. In the past, when researchers have discovered expected reward modulation in FEF [69, 70]), the underlying cause has been unclear. Was activity only modulated because it led to a modulated saccadic output? Or did expected reward affect neural activity above and beyond its effects on motor output, due to factors such as motivation and attention?

Here, we investigated reward expectancy in FEF while monkeys searched for a target embedded in natural scene images. This natural search behavior comprised a broad range of saccadic velocities and latencies, which allowed us to disentangle motor output and

expected reward as drivers of neural modulation. We found that saccadic vigor and latency were correlated with FEF activity. After accounting for this, reward expectancy still contributed unique variance to the neural response, suggesting that reward modulation in FEF plays a role that is not limited to the programming of eye movements.

Methods

Animals and Surgery

Northwestern University's Animal Care and Use Committee approved all procedures for training, surgery and experiments. We used three adult female rhesus monkeys (*Macaca mulatta*), which we referred to as M14, M15, and M16. Each monkey received preoperative training followed by an aseptic surgery to implant a recording cylinder above the FEF, as well as a titanium receptacle to allow head fixation. Surgical anesthesia was induced with thiopental (5-7 mg/kg IV) or propofol (2-6 mg/kg IV), and maintained using isoflurane (1.0%-2.5%) inhaled through an endotracheal tube. For single electrode recordings performed in M14 and M15, an FEF cylinder was centered over the left hemisphere at stereotaxic coordinates anterior 25 mm and lateral 20 mm. A chronic array was used to record from multiple units in monkeys M15 (right hemisphere) and M16 (left hemisphere). The recording chambers for these arrays were centered and mounted over the arcuate and principal sulci at stereotaxic coordinates anterior 24 mm and lateral 20 mm.

Behavioral Paradigms

Setup

To control experimental stimuli and data collection, we used the PC-based REX system [71], running the QNX operating system (QNX Software Systems, Ottawa, Ontario, Canada). Visual stimuli were generated by a second, independent graphics process (QNX - Photon) and rear-projected onto a tangent screen in front of the monkey by a CRT video projector (Sony VPH-D50, 75 Hz non-interlaced vertical scan rate, 1024 x 768 resolution). The distance between the front of the monkey's eye and the screen was 109.22 cm (43 in). Each natural scene spanned $48^\circ \times 36^\circ$ of the monkey's visual field.

Search Tasks

Monkeys M14 and M15 searched for a fly embedded in a natural scene (Fig. 1a). The monkey made saccades around the scene until a saccade landed near the fly target (within 5 degrees), and she held fixation there for 200 ms, at which point she received a water reward. If she made 25 saccades without finding the target, the trial ended. Each trial was initiated by the monkey fixating on a red dot in the center of the blank screen.

Monkeys M15 and M16 performed a Gabor search task, which was identical to the fly search task with the exception that a Gabor target (rather than a fly) was embedded in a natural scene (Fig. 1b).

Typically, between 1000 and 2500 saccades were made each day. The first saccade from every trial was excluded from our main analyses due to the confounding visual onset effects.

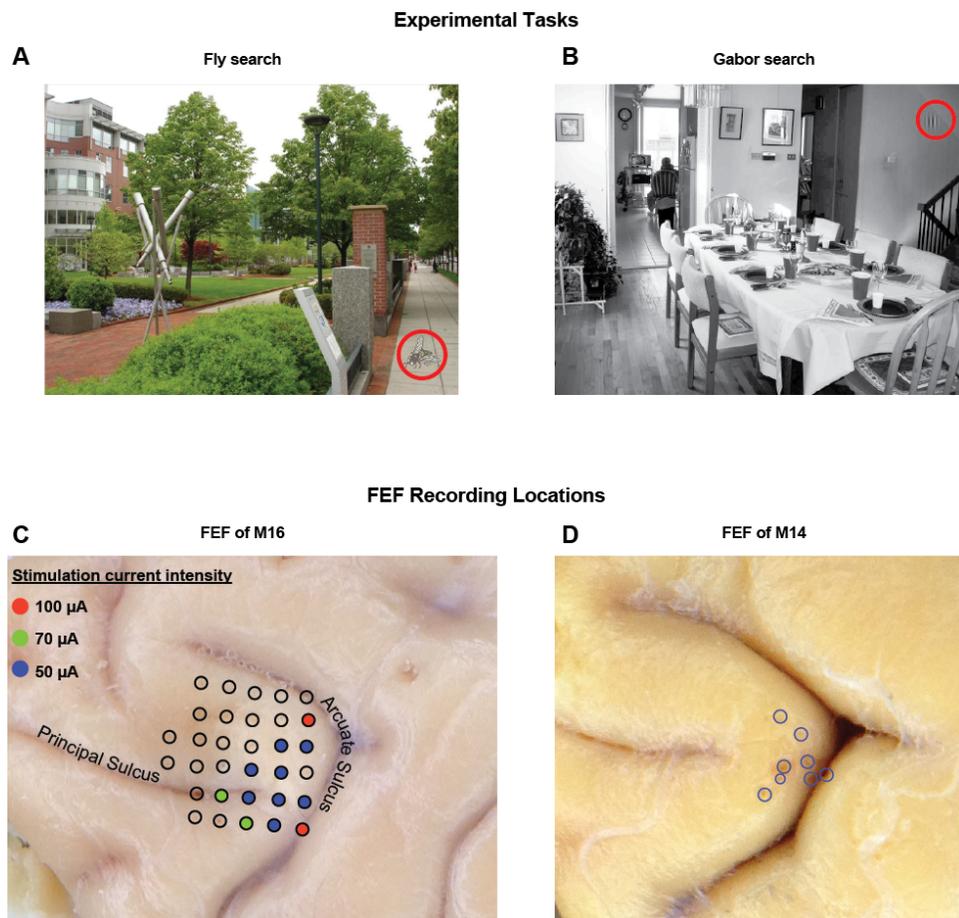


Figure 3.1: **Experimental Setup.** (A, B) In our experiments, monkeys freely searched for targets in natural scenes. In the fly search task (left; done by M14 and M15), the target was an embedded fly. In the Gabor search task (right; done by M15 and M16), the target was an embedded Gabor wavelet. Monkeys were rewarded only after fixating the target for a specified duration. (C) Functional characterization of FEF in M16 based on stimulation results for the semi-chronic array. Colors indicate current intensities at (or below) which saccades were reliably elicited. Only neurons at locations where current intensity was equal to or less than $50 \mu\text{A}$ were used in our analyses. (D) Characterization of FEF in M14. A subset of locations for acute recordings are marked by blue circles. All recording locations were characterized as FEF by elicitation of saccades with $\leq 50 \mu\text{A}$ during stimulation.

Data Acquisition

Eye Tracking

In the fly search task (M14 and M15), eye movements were tracked with a subconjunctival wire search coil, sampled at 1 kHz [72, 73]. In the Gabor search task (M15 and M16), eye movements were tracked with an infrared eye tracker (ISCAN Inc., Woburn, MA, <http://www.iscaninc.com/>) at 60 Hz.

Single-unit recording

During the fly search task, single-unit activity was recorded using tungsten microelectrodes (A-M Systems, Inc., Carlsborg, WA, USA). Electrode penetrations were made through stainless steel guide tubes that just pierced the dura. Guide tubes were positioned using a Crist grid system (Crist 88) (Crist Instrument, Co., Hagerstown, MD, USA). Recordings were made using a single electrode advanced by a hydraulic microdrive (Narashige Scientific Instrument Lab, Tokyo, Japan). On-line spike discrimination and the generation of pulses marking action potentials were accomplished using a multichannel spike acquisition system (Plexon, Inc., Dallas, TX, USA).

Recordings were confirmed to be in the FEF by ability to evoke low-threshold saccades with current intensities of $\leq 50\mu A$. To stimulate electrically, we generated 70 ms trains of biphasic pulses, negative first, 0.2 ms width per pulse phase, delivered at a frequency of 330 Hz.

Chronic recording

During the Gabor search task, recordings were performed using a 32 channel chronically implanted electrode array (Gray Matter Research, Bozeman, MT, USA). The depth

of each individual tungsten electrode (Alpha-Omega, Alpharetta, GA) was independently adjustable over a range of 20 mm.

All electrodes were initially lowered to pierce the dura. Individual electrodes were then gradually lowered until a well-isolated unit was located. In general, only a subset of electrodes was moved on any given day, and electrodes were left in place for at least three days before further lowering.

Spikes were recorded at 40 kHz with a multichannel acquisition system (Plexon, Inc., Dallas, TX, USA) on a separate PC. Automatic spike sorting was performed offline using the Plexon Offline Sorter (Plexon, Inc., Dallas, TX, USA).

Because any given electrode was often left in place for multiple days, we often recorded from the same neuron across sessions. To make use of this, we combined data from units that persisted across recording sessions on different days. To do this, we manually compared spike waveforms from units recorded at the same site on different days. Generally, we merged units sharing waveform shape (rise/fall characteristics, concavity/convexity, etc.), and time course. Ambiguous cases were not combined. Additionally, we included multi-unit activity in our present analysis.

As stimulation quickly degrades the recording fidelity of the tungsten electrodes in the array, we were unable to stimulate until the array was ready to be removed. Because of this, we were able to verify FEF location with stimulation (using the same parameters described for single-unit recording) in only one (M16) of the two chronically implanted monkeys (Fig. 1c). In M16, only units recorded on electrodes where saccades were reliably elicited at or below 50 A at any depth were included in our analyses. The cluster of electrodes where stimulation evoked saccades in M16 roughly matched the

location of maximum FEF evoked saccade sites in M14 (Fig. 1d). However, many of the recordings were more superficial than the depths at which we observed consistent elicitation of saccades. This does not rule out these cells being part of FEF, but it does increase the probability that some of the cells that we recorded were not from FEF.

In M15, the chronic array was centered at the same stereotaxic location as M16 and a sample of 3 monkeys used in earlier studies. Further, the spatial pattern of electrode grid locations where FEF-like responses were observed closely resembled the pattern of FEF locations in the grid of M16. We note here that, similar to M16, chronic recordings for M15 were at relatively shallow depths (e.g., 3 to 5 mm below dura).

In M15 and M16, we limited our analyses to units that had presaccadic properties expected in the FEF (see following section). This decreased the chance of analyzing units from nearby brain regions that were not part of the FEF. However, we cannot rule out the possibility that some of the units included in our analysis were from prefrontal regions other than FEF, especially in the case of M15.

Neuron Characterization and Selection

Cell Selection

We analyzed units that had presaccadic activity. To determine whether there was presaccadic activity, we looked at peri-saccadic time histograms (PSTHs) aligned to the start of the upcoming saccade, binned into 8 angular directions (according to saccade direction), with each bin subtending 45 degrees. In any bin, if there was a significant difference between a firing rate and baseline, and the rise started before the saccade onset, then there was presaccadic activity. Additionally, we only analyzed cells with

average firing rates of at least 5 spikes per second during the task. These criteria gave us 142 units to analyze: 27 from M14, 79 from M15, and 36 from M16.

Most neurons recorded during the fly task were classified using a memory-guided saccade task, as having visual, delay and/or movement activity. The neurons recorded during the Gabor task were not classified.

Importantly, a neuron having presaccadic activity in our natural scene search does not mean that it is classically a movement neuron. It has previously been shown that purely visual cells (according to the classic nomenclature) have presaccadic activity during a natural scene search task [55]. Moreover, from our 38 classified neurons, 4 were purely visual, 18 were visual/delay, 11 were visual/delay/movement, 2 were visual/movement, and 3 were purely movement. Within this small number of (mostly visual) neurons, we did not observe a noticeable difference in results based on classification. However, as the majority of our neurons were not classified in a traditional manner, we were unfortunately unable to do a thorough analysis of the effects of cell type on our results.

Receptive Field Characterization

To map receptive fields, we first looked at the binned (by saccade angle) PSTHs. We initially assigned the receptive field (RF) as the bin that had the largest difference between baseline and the peak. We then found the time point corresponding to the peak firing rate in this bin.

We then calculated a better estimate of the RF at this time point. To do this, we calculated a smoothed version of the average firing rate across space. More specifically, we tiled space with square pixels that were $3/4$ of a degree. Each pixel was given the average firing rate (at the given time) of the 200 non-target saccades that landed nearest

to that pixel. After creating this smoothed firing rate map, we said that the RF was all of the pixels with a smoothed firing rate more than 50% of the way from the minimum to the maximum smoothed firing rate.

Note that when we previously ran all our analyses using a simpler (but less accurate) RF characterization that only consisted of angle bins, we obtained the same general results.

Behavioral Data Analysis

Saccades were classified into 3 types. First, target, expected reward (T+/ER+) saccades were those that landed near the target followed by a fixation and reward, indicating knowledge of the target. Second, target, no expected reward (T+/ER-) saccades were those that landed near the target but were *not* followed by a fixation long enough (200 ms) to receive a reward. We assumed that this indicated lack of knowledge regarding the target, and thus a lack of expected reward. Third, we defined non-target, no expected reward (T-/ER-) saccades as those not landing near the target.

To statistically compare the latencies between T+/ER+ and T-/ER- saccades, we used a two-tailed Wilcoxon Rank-sum test. Note that latencies below 90 ms were excluded in this analysis under the assumption that anything below this cutoff would be an express saccade. Additionally, we exclude latencies above 1000 ms, as it is likely the next saccade was not detected. For M15, this statistical test was done separately for the fly and Gabor task. A single result was reported, as it was the same for both tasks. When plotting the distribution of latencies (Fig. 2c), we put latencies into 40 ms bins between 90 and 490 ms. The small number of saccades with latencies greater than 490 ms were not plotted.

Since saccade velocities are strongly dependent on saccade magnitude, we plotted the peak saccade velocity as a function of saccade magnitude for both T+/ER+ and T-/ER- saccades. We put saccades with magnitudes between 5 and 25 degrees into bins with a size of 1 degree. Note that for the fly task, M14 and M15 had eye movements tracked with an eye coil, while for the Gabor task, M15 and M16 had eye movements tracked with an IR camera (see *Methods*). The recorded peak velocities were consistently smaller when using the IR camera, due to its limited 60Hz resolution (Fig. 2e). As we are comparing velocities of saccade types, rather than being concerned with absolute velocities, we do not view these differences as problematic. Note that for M15, we averaged the plots created using the fly and Gabor task. In order to statistically compare the peak velocities between T+/ER+ and T-/ER- saccades, for each saccade magnitude bin, we computed the relative change in peak velocity $\frac{v_T - v_{NT}}{v_T}$, where v_T and v_{NT} are the peak velocities for T+/ER+ and T-/ER- saccades, respectively. This created a vector of 21 proportions of change. We did a two-tailed Wilcoxon signed-rank test to determine whether this proportion was significantly greater than 0. We defined vigor as the velocity of a saccade divided by the expected velocity of that saccade, given its amplitude (Fig. 2f). The expected velocity of a saccade was calculated by averaging the velocities of the 25 saccades with the closest amplitudes to that of the given saccade for that monkey (regardless of whether they were T+/ER+ or T-/ER-). When plotting the distribution of vigor (Fig. 2g), we put vigor into bins of 0.1. Vigor less than 0.5 or greater than 1.5 was not plotted. When plotting vigor versus saccade latency (Fig. 2h), latencies were put into 40 ms bins and scaled velocities into bins of 0.02 (M14) or 0.01 (M15, M16).

To examine whether latency and vigor were correlated, we tested whether the correlation coefficient was significantly different from 0 using a two-tailed t-test. For M15, this statistical test was done separately for the fly and Gabor task. A single result was reported, as it was the same for both tasks.

In order to look at the effect of expected reward as a function of latency and vigor simultaneously (Fig. 4a), we put latencies into 40 ms bins, and scaled velocities into bins of size 0.1. We plotted the logarithm of the ratio of the probability distribution of T+/ER+ saccades versus the probability distribution of T-/ER- saccades (over all shown behavioral conditions). Each probability distribution was smoothed (in a 3 row x 3 column window with a Gaussian filter of $\sigma = 0.5$), prior to taking the ratio. We didn't show bins with less than 0.1% of saccades for both conditions. When averaging across monkeys, the plot for M15 was first averaged across the fly and Gabor tasks.

Neural Data Analysis

For all subsequent analyses, we put spikes into 10 ms bins. We excluded the first saccade in each trial from the analysis due to the confounding effects of the image onset. We only included saccades that started and ended on the screen. In general, we focus on analyzing activity around the time neurons are most active. This is because 1) our heterogeneous population of neurons are active at different times in our natural scene search task (e.g. see Fig. 3c); and 2) averaging activity over a large, static time window can nullify effects (e.g., if a condition increases the firing rate at one time, and decreases it at another; as in Fig. 3c, M14 cell 22).

Comparing Peak Firing Rates Between Conditions

To compare firing rates between conditions, we computed perisaccadic time histograms (PSTHs) of spiking activity. When showing PSTHs of individual neurons (Fig. 3c), we show the mean firing rate across saccades. The error bars on shown PSTHs are the standard error of the mean (SEM) across saccades. When showing the PSTHs averaged across neurons, we first normalize the mean firing rate for each neuron by dividing by the peak firing rate when saccades are made into the RF, but not to the target (peak of the red trace in Fig. 3c). We then show the average of these normalized firing rates across neurons. Error bars are the SEM across neurons. All traces are smoothed using a 50 ms sliding window.

To determine whether there was a significant difference in the peak firing rate of the PSTHs between two conditions for a monkey (averaged across neurons), we 1) for a given condition found the time bin that had the highest firing rate in the smoothed PSTH; 2) for that condition, for every neuron, computed the normalized average firing rate in a 50 ms interval around that bin (i.e., the previous two 10 ms bins, that 10 ms bin, and the next two 10 ms bins). This creates a vector of firing rates for each neuron in a given condition (e.g. firing rates around each T+/ER+ saccade); 3) used a two-tailed Wilcoxon signed-rank test to compare peak firing rates between the two conditions.

Tuning Curves

For our tuning curve analysis (Fig. 3f-g), we calculated firing rates in the 50 ms window that contained the highest firing rate in the PSTH of all saccades into the RF. The time window was independently determined for each neuron.

To fit tuning curves, we created a scatterplot showing the firing rate (y-axis) versus direction (x-axis) of each saccade. We then fit a Von-Mises function to these data points, to model how firing rate varies as a function of direction (Fig. 3f). We fit a standard four-parameter model [74]

$$(3.1) \quad \lambda = \alpha + \beta \exp(\kappa \cos(\theta - \theta^*))$$

where λ is the firing rate for each saccade, θ is the saccade direction for the trial, and $[\alpha, \beta, \kappa, \theta^*]$ are the parameters. We estimated these parameters using nonlinear least-squares fitting, constraining the minimum of the tuning function to be non-negative, $\kappa \leq 10$, $\alpha > 0$, and $[-\pi \leq \theta^* \leq \pi]$. We then computed the following quantities from the estimated parameters:

Baseline: The minimum of the tuning curve, i.e. $\alpha + \beta \exp(-\kappa)$,

Gain: The difference between the maximum and minimum, i.e. $\beta \exp(\kappa) - \beta \exp(-\kappa)$,

Width: The empirical full-width half-maximum of the tuning curve,

Preferred direction: θ^*

In order to determine whether a parameter was significantly different across neurons between T+/ER+ and T-/ER- saccades, we used a two-tailed Wilcoxon signed-rank test. For each monkey, we plotted (Fig. 3g) the median of the ratio between the parameters for T+/ER+ and T-/ER- saccades. For error bars, we plotted the standard error of the median, computed by bootstrapping. The median and standard error of the median (as opposed to the mean) were used due to skewed distributions with outliers.

Generalized Linear Models

We used a Poisson generalized linear model (GLM) in order to determine whether neurons uniquely encoded saccade latency, saccade vigor, and expected reward (whether the saccade was to the target). In this analysis, we only considered saccades into the receptive field, and those with latencies less than 1000 ms.

For each neuron, we aimed to predict the number of spikes during each saccade (into the RF) during the 50 ms of peak activity for that neuron. In other words, we found the 50 ms window with the highest firing rate (averaged across saccades into the RF; as described in Comparing Peak Firing Rates Between Conditions), and then found the number of spikes in this window for each saccade. This yielded a vector of spike counts, \mathbf{Y} . Note that we also ran a GLM where we always used a 100 ms time window prior to saccade onset, and found the same general results.

In order to explain the spike counts, we used 4 covariates: 1) saccade latency, $\boldsymbol{\theta}_L$ (a vector of the latencies for each saccade); 2) saccade vigor, $\boldsymbol{\theta}_V$; 3) expected reward (a binary variable for whether the saccade was to the target), $\boldsymbol{\theta}_T$; and 4) the expected firing rate given the saccade vector, $\boldsymbol{\theta}_{RF}$ (the smoothed average firing rate at the location of the saccade vector see *Receptive field characterization*). This last covariate was included because the proximity of the saccade vector to the preferred vector of the cell could help explain the firing rate.

Overall, the model that generates the firing rate ($\boldsymbol{\lambda}$) can be written as:

$$(3.2) \quad \boldsymbol{\lambda} = \exp(\beta_0 + \beta_L \boldsymbol{\theta}_L + \beta_V \boldsymbol{\theta}_V + \beta_T \boldsymbol{\theta}_T + \beta_{RF} \boldsymbol{\theta}_{RF})$$

where the β 's are the weights for each covariate (β_0 is a baseline term) that we fit. Note that the covariates are passed through an exponential nonlinearity, ensuring that firing rates are positive. The model assumes that spikes are generated from the firing rate, λ , according to a Poisson distribution. We fit the weights to the data using maximum likelihood estimation. That is, we found the β 's that were most likely to produce the true spike output (assuming spikes were generated from the firing rate in a Poisson nature).

For each neuron, we tested whether β_L , β_V , and β_T were significantly different from 0 using a two-tailed t-test.

Multiple Comparisons Testing

When listing the number of significant neurons in our GLM analysis, we listed uncorrected (for multiple comparisons) statistics. This was because we were not concerned with whether specific individual neurons were significant; rather, we wanted to show an uncorrected comparison of the numbers of neurons that had significant positive and negative effects. Nonetheless, there were many individual neurons that do survive Bonferroni corrections.

Results

In two separate search tasks, three head-fixed monkeys (M14, M15, and M16) freely searched for a target embedded in a natural scene (Fig. 1). Monkeys were rewarded for locating and holding fixation on the target. Targets were blended into the background, making the task difficult enough that the monkeys typically had to make several saccades (on average, 5 to 7) in order to locate the target.

We defined target, expected reward (T+/ER+) saccades as those that landed near the target followed by a fixation and reward, indicating knowledge of the target (Fig. 2a). The trial ended following these saccades. Similarly, we defined non-target, no expected reward (T-/ER-) saccades as those not landing near the target (although this could also include a small number of saccades where reward was expected; Fig. 2a). The trial continued after these saccades, unless the maximum number of saccades allotted per trial was reached. Our main comparison was between these two saccade types.

Expected reward alters saccade latencies and velocities

Prior to testing for neural differences resulting from expected reward, we investigated whether there were behavioral differences between T+/ER+ and T-/ER- saccades. Specifically, we tested for differences in latency and velocity of saccades, as previous studies have shown that expected reward can decrease the latency and increase the velocity of saccades [75, 61, 76, 58].

Indeed, while there was a large range of latencies (Fig. 2c), T+/ER+ saccades had significantly shorter latencies than T-/ER- saccades (Fig. 2b; $p < 0.001$ for all monkeys). Interestingly, for T-/ER- saccades, the latencies systematically depended on the saccade amplitude, and this pattern was different for each monkey, likely reflecting individual search strategies (Fig. 2d; red). However, for T+/ER+ saccades, latencies were consistent across saccade amplitudes (Fig. 2d; blue). Monkeys appear to have idiosyncratic exploration strategies but share a common, fast, strategy for making saccades to the target (for which they expect reward).

One might also predict that reward expectation leads to higher saccade velocities. We observed that T+/ER+ saccades had a significantly higher peak velocity than T-/ER- saccades (Fig. 2e; $p < 0.001$ for all monkeys), especially for larger amplitude saccades. For both types of saccades, however, velocity increased with amplitude, consistent with the main sequence for saccades [77]. Therefore, in order to ask whether saccade velocity varies with reward expectancy, we needed a velocity measure that accounts for the fact that velocity varies as a function of amplitude. We use vigor, [61] which describes how much the velocity of a saccade is above or below the expected velocity (for all saccades) given the saccade amplitude (Fig. 2f,g; see *Methods* for calculation of vigor). Furthermore, vigor and latency are unlikely to be independent; for example, a high degree of motivation may lead to both increased velocity and decreased latency. Indeed, vigor and latency are significantly negatively correlated (Fig. 2h; $p < 0.001$ for all monkeys), suggesting a related neural mechanism.

Increased firing rates for saccades into the receptive field, when reward is expected

We aimed to understand whether FEF activity is modulated by the expected reward of saccades. To do this, we first computed PSTHs (aligned to saccade onset) for T+/ER+ and T-/ER- saccades made into the RF for each neuron (see *Methods* for RF characterization). To examine the average effect, we then normalized and averaged these PSTHs across neurons. Average peak firing rates were significantly higher for T+/ER+ saccades versus T-/ER- saccades (Fig. 3a-d) for all monkeys (M14, $p < 0.001$; M15, $p < 0.001$; M16, $p = 0.0218$). This effect was not due to the increased fixation times that followed

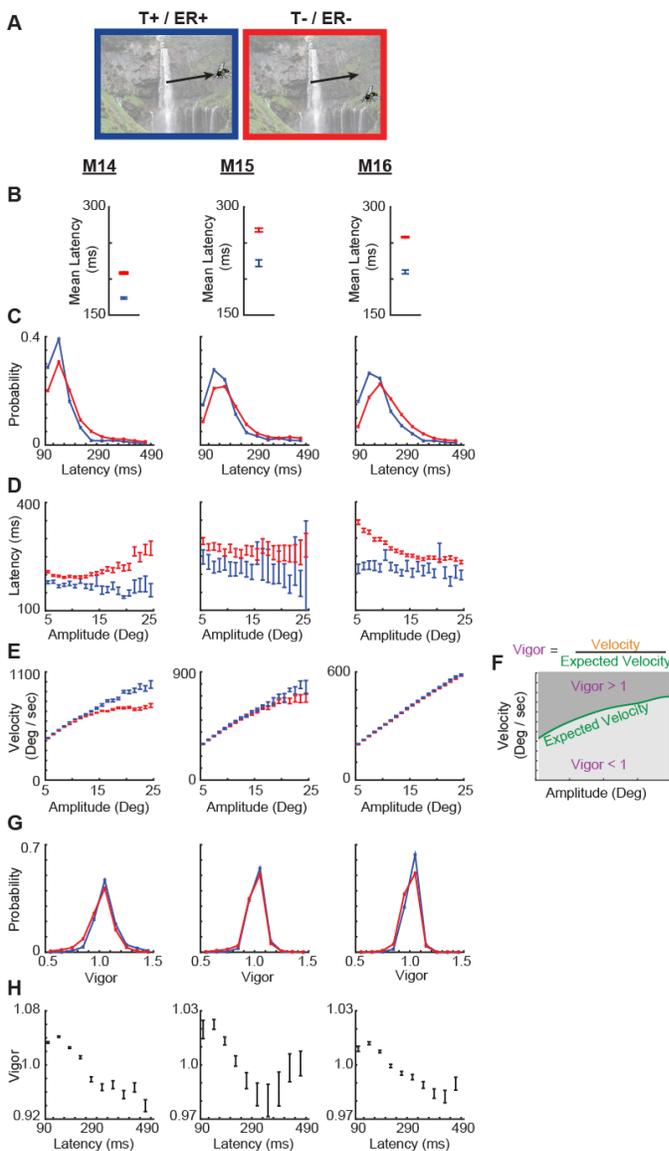


Figure 3.2: **Behavioral differences of saccades due to expected reward.** (A) Saccades that land near the target followed by fixation are defined as target, expected reward (T+/ER+), and are shown in blue in subsequent panels. Saccades that do not land near the target are defined as non-target, not expected reward (T-/ER-), and are shown in red in subsequent panels. (B) Latencies are compared between T+/ER+ (blue) and T-/ER- (red) saccades. Mean \pm SEM are shown. For this, and subsequent, panels, each column is behavior from a different monkey. (C) Latency distributions. (D) The mean latencies (\pm SEM) of T+/ER+ and T-/ER- saccades are shown as a function of saccade amplitude. (E) Velocities are compared between T+/ER+ (blue) and T-/ER- (red) saccades. Mean velocities (\pm SEM) as a function of amplitude. Note that differences in the magnitude of velocities are due to differences in eye tracking technology across monkeys (IR camera vs. eye coil; see *Methods*). (F) Because velocity is dependent on amplitude, we define vigor as the velocity divided by the expected velocity for that amplitude (for all saccades). Shading is based on whether scaled velocities are greater or less than 1. (G) Vigor distributions. (H) Vigor as a function of saccade latency for all monkeys. All saccades, regardless of expected reward, were combined, as they had the same trends.

saccades to the target (fixation was required for reward), as long fixation times did not increase firing rates for T-/ER- saccades (M14, $p = 0.029$ for short fixation times increasing firing rates; M15, $p > 0.05$; M16, $p > 0.05$; Fig. 3e). Therefore it appears that expected reward does modulate FEF neurons' activities during saccades into the RF.

In order to more generally understand how expected reward modulates FEF activity during saccades across visual space (rather than just into the RF), we fit directional tuning curves to both T+/ER+ and T-/ER- saccades (Fig. 3f,g). We found that T+/ER+ saccades led to significantly larger tuning curve gains (maximum minus minimum of the tuning curve) in all monkeys (M14, $p < 0.001$; M15, $p < 0.001$; M16, $p = 0.031$), but neither the width nor the baseline of the tuning curves changed systematically across all monkeys ($p > 0.05$ for all except $p = .047$ for M14 width). These results support an increase in neural activity that is specific for the upcoming saccade direction, rather than a general enhancement (e.g. due to arousal). This gain increase could be beneficial because it increases information [78] about the rewarded target location for downstream brain areas [79, 80, 81, 82, 83].

Velocity and latency differences explain some, but not all, of the neural differences due to expected reward

Because expected reward affects motor output (e.g., the latency and velocity of saccades), any brain structure that leads to motor output can be modulated by expected reward. For example, neurons in the brain stem that affect the velocity of saccades will be modulated by expected reward. Thus, it is important to ask whether FEF's modulation due to

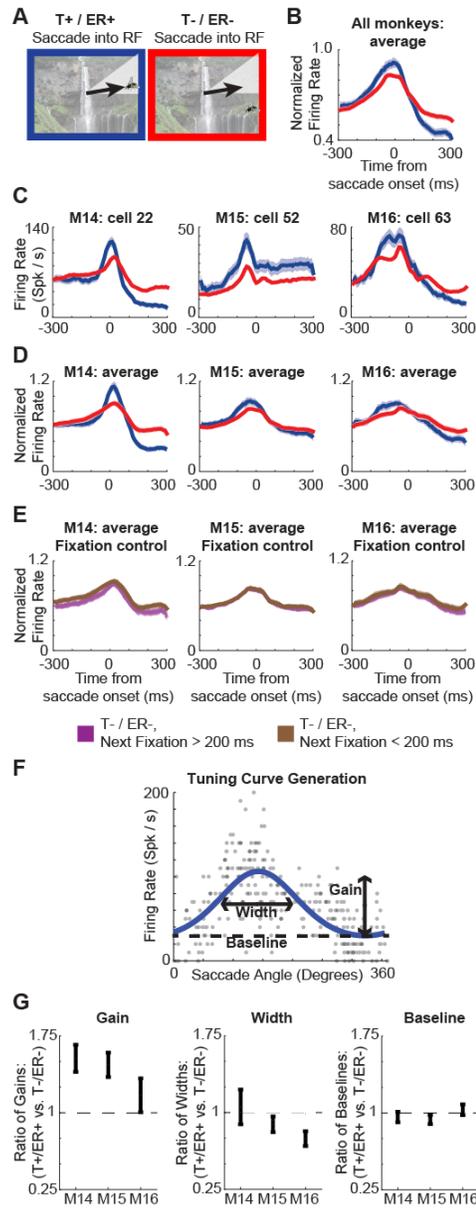


Figure 3.3: Neural activity reflects expected reward.

(A) A schematic of T+/ER+ and T-/ER- saccades into the RF. While the RF is shown as an angular wedge, the RFs generally had some amplitude dependence. (B) Normalized PSTHs averaged across neurons, aligned to saccade onset. We plot T+/ER+ saccades into the RF (blue) and T-/ER- saccades into the RF (red). Error bars represent the SEM. (C) PSTHs of example neurons, aligned to saccade onset. Error bars represent the SEM of that neuron's firing rate. (D) Normalized PSTHs averaged across neurons (for each monkey), aligned to saccade onset. (E) Normalized PSTHs of T-/ER- saccades that have subsequent fixation times of greater than (purple) and less than (brown) 200 ms. Only saccades into the RF are used. (F) We fit separate tuning curves to T+/ER+ and T-/ER- saccades using a Von-Mises function. An example tuning curve fit is shown. Each dot represents a single saccade. (G) Tuning curves were fit to all neurons. The gains (left), widths (middle), and baselines (right) are compared between T+/ER+ and T-/ER- saccades. The medians of these ratios (across neurons) +/- the standard error of the median (computed by bootstrapping) are shown.

expected reward is simply because FEF is programming appropriate saccade outputs, or, whether there are higher level, cognitive effects of expected reward in FEF, or both.

It is important to first determine whether the latency and velocity of saccades explain FEF variability. To determine whether latency and velocity could influence firing rates independent of reward expectation, we quantified the latency and vigor dependence of responses during T-/ER- saccades (Fig. 4b,c). We found that higher vigor and shorter latency did increase the firing rate (M14, $p < 0.001$; M15, $p = 0.0064$; M16, $p = 0.0017$). Behavioral factors thus explain some variability *within* T-/ER- saccades. Given this finding, we asked if these movement variables could explain all the neural differences caused by differences in expected reward.

To test this possibility, we compared T+/ER+ and T-/ER- saccades while controlling for latency and vigor. Specifically, we sub-selected saccades (both T+/ER+ and T-/ER-) with behavioral markers of expected reward (namely, high vigor and short latency saccades) so that the behavior matched between T+/ER+ and T-/ER- saccades (Fig. 4a; see *Methods* for details). T+/ER+ saccades had a higher peak firing rate than T-/ER- saccades even when matched for latency and vigor (Fig. 4b-c; M14, $p < 0.001$; M15, $p = 0.0057$; M16, $p = 0.039$). Thus, the enhanced response during T+/ER+ saccades is only partially explained by behavioral differences in motor output.

To more rigorously control for velocity and latency, we fit a multiple regression model (the Poisson generalized linear model (GLM)) to explain the peak firing rate of each neuron. The model can tell us which factors uniquely contribute to the firing rate, even when these factors are themselves correlated. We found that many neurons independently encoded expected reward (i.e., that the saccade would land on the target), latency, and

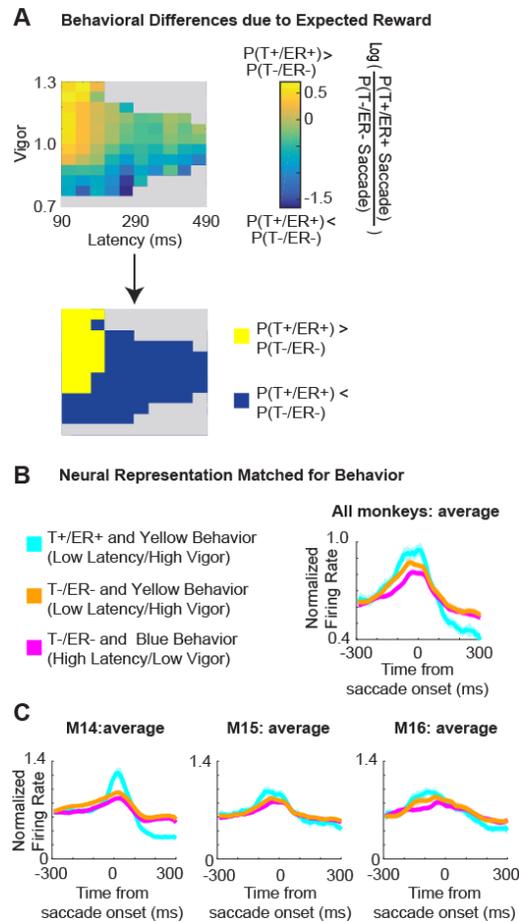


Figure 3.4: **Increased firing rates due to expected reward, short latency, and high velocity saccades: PSTHs.** (A) As latency and vigor are correlated, we look at the effect of expected reward across the behavioral space of latency and velocity. On the **top**, the relatively likelihood of a saccade having a certain latency and vigor for T+/ER+ compared to T-/ER- saccades, averaged across monkeys. More yellow areas mean that a particular latency and vigor are more likely for T+/ER+ saccades. **Below**, a threshold is put on the upper panel to determine behavioral regions that are more likely to contain saccades accompanied by expected reward (yellow) or no expected reward (blue). (B) Normalized PSTHs averaged across neurons (from all monkeys) are constructed from 3 conditions: 1) T+/ER+ saccades in the yellow behavioral region of the above panel (cyan); 2) T-/ER- saccades in the yellow behavioral region (orange); 3) T-/ER- saccades in the blue behavioral region (pink). (C) Same as panel B, for individual monkeys.

vigor (Fig. 5). We found higher peak firing rates for saccades when reward was expected (22% of neurons significantly higher vs. 8% lower firing rates), short latency saccades (39% vs. 11%), and high vigor saccades (30% vs. 8%). These GLM findings support our previous PSTH-based analysis by showing that while FEF neurons are modulated by latency and vigor, expected reward has an effect beyond the motor output.

Lastly, we asked whether expected reward, latency, and vigor were represented by separate populations of neurons (i.e., each neuron represents a single feature) or overlapping populations of neurons (i.e., neurons represent multiple features). To answer this, we determined the percentage of neurons that were significant for representing features (in the above GLM analysis), conditional on the neuron representing other features (Fig. 5C). If neurons represented only individual features, then the probability of a neuron representing one feature (e.g., expected reward), conditioned on that neuron already representing another feature (e.g., latency; bars in Fig. 5C), should be lower than the general probability of representing the original feature (expected reward; dashed lines in Fig. 5C). However, that was not the case. Neurons were more likely to represent a given feature if they already represented an additional feature (bars are above the lines in Fig. 5C). Thus, overlapping populations of neurons represent expected reward, latency, and vigor.

Enhanced FEF responses are due to expected reward, not simply due to moving to the target

In our previous analyses, all target saccades (T+) were associated with expected reward. It is possible that the neural differences based on expected reward were actually due to the visual presence of the target in the receptive field. The natural search task offered us

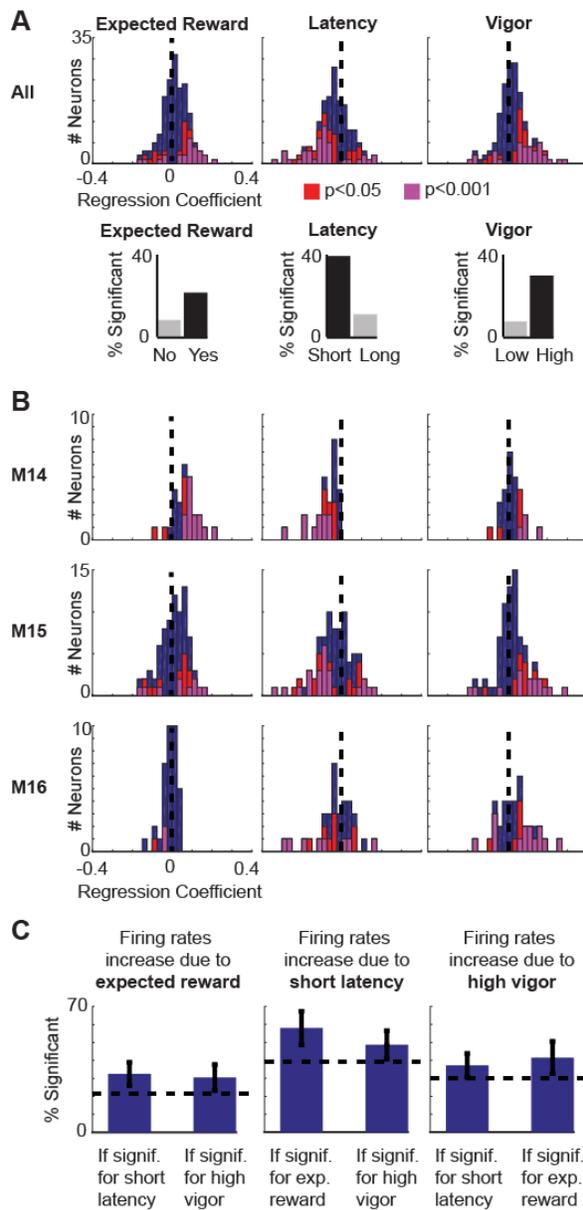


Figure 3.5: Increased firing rates due to expected reward, short latency, and high velocity saccades: GLMs. (A) Using a GLM, we used the latency, vigor, and expected reward of the saccade (whether it was to the target) to explain the peak firing rate of each neuron. Histograms of the regression coefficients for each covariate are shown, colored based on their significance (dark blue is not significant). Insets (below) show the percentage of neurons that are significant. For target (left), having positive regression coefficients signifies that going to the target (T+/ER+ saccade) increased the firing rate. For latency (middle), having positive coefficients signifies that longer latency saccades increased the firing rate. For vigor (right), having positive coefficients signifies that higher vigor increased the firing rate. (B) Same as panel A, with individual monkeys shown in each row. (C) Bars show conditional probabilities (the probability of a neuron representing one feature, e.g. expected reward, conditional on the neuron representing another feature, e.g. latency) from the GLM results across all monkeys. Dashed lines show the general probabilities of a feature being significant in the GLM (not conditioned on the neuron representing any other features). Error bars are SEMs from bootstrapping. For example, on the left we show the probability that neurons' firing rates significantly increase due to expected reward (dashed line) conditioned on the neurons' having significantly higher firing rates due to shorter latencies (left bar) and higher vigor (right bar).

a simple way of controlling for this possibility. Because the target is difficult to find in the complex background, it is likely that the monkey will occasionally saccade toward a target without awareness of the target. We thus looked at a third type of saccade: that which lands near the target, but is not followed by a fixation long enough (200 ms) to receive a reward. We assume that this indicates lack of awareness of the target, and thus a lack of expected reward. We defined these types of saccades as target, no expected reward (T+/ER-) saccades (Fig. 6a).

These T+/ER- saccades give us an opportunity to ask if target presence was represented, independent of expected reward. We observed lower firing rates for T+/ER- saccades compared to T+/ER+ saccades (Fig. 6b). Further, these differences were significant for each monkey (Fig. 6c; M14, $p < 0.001$; M15, $p = 0.0044$; M16, $p < 0.001$). In fact, the firing rate of T+/ER- saccades was generally the same as the firing rate of T-/ER- saccades (Fig. 6b,c; M14, $p < 0.021$; M15, $p > 0.05$; M16, $p > 0.05$). Note that for M14, the firing rate for T+/ER- saccades was slightly, but significantly, higher than the firing rate of T-/ER- saccades. We believe this is due to a small number of classification errors; that is, some of the T+/ER- saccades may have been reward expecting, but the monkey did not fixate long enough, or some of the T+/ER+ saccades may not have been reward expecting, but the monkey fixated long enough by chance. These results demonstrate that FEF indeed is modulated by the expected reward of saccades rather than the presence of the target in the RF or merely saccades to locations near the target.

A potential confound is that T+/ER+ saccades landed closer to the target than T+/ER- saccades (Fig. 6d). Thus, it could be possible that the landing distance from the target, rather than reward expectation, was responsible for the neural differences between

T+/ER+ and T+/ER- saccades. This does not appear to be the case, as T+/ER+ saccades landing far from the target (defined as greater than the median distance from the target) have higher firing rates than T+/ER- saccades landing far from the target (Fig. 6e; M14, $p = 0.0022$; M15, $p = 0.042$; M16, $p = 0.031$). This is also true for saccades landing near the target (Fig. 6e; M14, $p = 0.0025$; M15, $p = 0.032$; M16, $p = 0.0055$). In fact, T+/ER+ saccades landing far from the target even have higher firing rates than T+/ER- saccades landing near the target (Fig. 6e; M14, $p < 0.001$; M15, $p = 0.098$; M16, $p = 0.048$). Nonetheless, it is important to note that for T+/ER- saccades, landing closer to the target does increase firing rates. This is likely due to classification errors; that is, some of the T+/ER- saccades landing near the target were probably associated with expected reward. These results confirm that, when controlling for the landing distance from the target, FEF firing rates are enhanced by expected reward rather than target presence.

Differences between monkeys

It is worth noting that there were differences in results between monkeys. One main difference is the proportion of neurons that significantly represent features (expected reward, latency, or vigor) in the GLM analysis (or similarly, the effect size of the results in the PSTH analyses). M14 had the greatest proportion of neurons that represented features, while M16 had the fewest. Indeed, the GLM analysis yielded no cells for M16 where expected reward was significantly represented after accounting for latency and vigor. Another notable difference between monkeys is that the peak of the PSTHs happened near

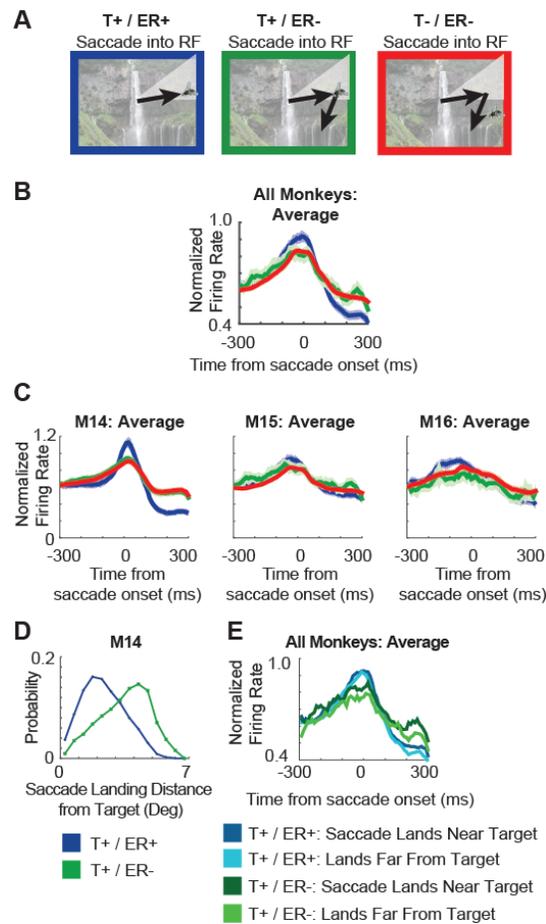


Figure 3.6: **Modulation in FEF due to expected reward rather than target presence.** (A) We compared T+/ER+ (blue), T+/ER- (green), and T-/ER- (red) saccades into the RF. (B) PSTHs show the average normalized firing rate +/- SEM of saccades across all neurons (from all monkeys). (C) Same as panel B, but for individual monkeys. (D) For an example monkey, M14, we show the difference in the distributions of landing distances from the target for T+/ER+ and T+/ER- saccades. (E) We control for the possibility that T+/ER+ saccades only have a higher peak firing rate than T+/ER- saccades because T+/ER+ saccades land closer to the target on average (panel D). We plot PSTHs for T+/ER+ and T+/ER- saccades that land close and far from the target. Close and far are defined as less and more (respectively) than the median distance from the target for all T+ saccades. Error bars are not shown to ensure that the means are visible.

saccade onset in monkey M14, but approximately 50 ms prior to onset in monkeys M15 and M16.

One reason for these differences is the way cells were selected. All units from M14, and some units from M15 were recorded using acute recordings (see *Methods*). In these cases, cells with strong visual or visuo-movement responses were almost always selected. However, all cells from M16 and most cells from M15 were recorded with chronic arrays. In those cases, cells were not carefully selected. Thus, these differences in selection could have led to differing results between monkeys.

Another reason for differences in results across monkeys is due to potential differences in anatomical location. Cells that were acutely recorded in M14 and M15 were confirmed to be in FEF. However, some of the cells that were chronically recorded in M15 and M16 may have been in nearby regions. Moreover, differences in recording location *within* FEF could have led to differences between monkeys, or between neurons in an individual monkey (see *Methods*).

Lastly, these differences in results could also be caused by different cell types being recorded across monkeys, as the majority of our neurons were not classically categorized into cell types (see *Methods*). Overall, we find it especially interesting that even though there may be differences in the relative distributions of cell types recorded from each monkey, most of the general trends are consistent.

Discussion

We recorded single-cell FEF activity while monkeys freely searched for targets in natural scenes. We found that expected reward modulated the activity of FEF neurons

during saccades into their receptive fields. This modulation was due to the expected reward, and not simply the presence of the target, as there was no modulation when monkeys accidentally made a saccade near the target. Importantly, expected reward altered saccadic parameters; velocity was increased and latency was decreased. Although FEF activity was modulated by these saccadic parameters, it additionally reflected the expected reward above and beyond those parameters.

Expected Reward

The modulation of FEF activity due to expected reward was a central finding in our study. We classified all non-target saccades, along with target saccades when the monkeys did not fixate long enough to receive a reward, as not expecting reward. This classification was not arbitrary, but was based on our assumption that the monkeys learned the required fixation time in order to gain the reward. Undoubtedly, some of the saccades were misclassified. For instance, if the monkey intended to fixate the target but ended up at a non-target location, the monkey expected a reward at the end of the saccade, but we misclassified the saccade as not expecting reward. Alternatively, if the monkey fixated near the target by chance and got rewarded, the saccade was misclassified as expecting reward. Importantly, errors of either kind in our data analysis would only lead us to *underestimate* the true effect size of the neural differences.

The effects of expected reward in the oculomotor circuits of the brain have been studied extensively. Aside from FEF, for example, many cortical areas have been investigated: lateral intraparietal sulcus [84, 85], supplementary eye fields [70, 86], dorsolateral prefrontal cortex [87], ventrolateral prefrontal cortex [88], orbitofrontal cortex [89, 90], and

premotor cortex [90], among others. The effects of expected reward also appear in sub-cortical areas: for example, superior colliculus [57, 91], the basal ganglia [92, 58], and ventral tegmental area [93]. Past studies that tested for reward size effects in FEF have yielded mixed results: no reward modulation [87], a general reward effect but no effect of reward size [69], and an effect of reward size [70]. In our study, while we did not manipulate reward size, we did observe a clear effect of expected reward on FEF activity while a monkey searched natural images to find a target.

Saccadic velocity and latency

Unlike in previous work [94], we observed that vigor (velocity relative to the main sequence, [77]) is reflected in FEF responses. Previous work may have not found this effect because classical tasks may elicit a more stereotyped relationship between saccade amplitude and velocity. Additionally, we observed that the greatest variability in velocity occurred in the 15-25 degree range of amplitudes (see Fig. 2e), whereas past work often looked at the lower end of this range. This demonstrates the importance of examining the full range of saccades in order to understand the neural encoding of saccade variables.

We have found a dependence of firing rate on latency. This is consistent with decision-making models that propose that there is a growing urgency signal [95, 21] or a decreasing threshold [96, 97] over time. These models suggest that a higher firing rate is necessary early on to make a decision, but lower firing rates are sufficient with increasing latencies. Thus, urgency models may help explain the neural mechanisms of sensorimotor decision-making not only during traditional multi-alternative choice tasks, but also during naturalistic behavior.

Simultaneously analyzing expected reward and motor variables

Given the effects of expected reward on saccadic parameters, it is important to understand the degree to which reward modulation in a particular area is a reflection of motor programming. While a small number of studies have used multiple regression to separate the variance accounted for by reward and movement [91, 98, 87], the vast majority of studies that examine reward in motor regions of the brain overlook this. Critically, in the cases where reward modulation has been observed in FEF [69, 70], no effort was made to test whether the effect of reward could be explained by movement variables like vigor or latency. The present study demonstrates that, while much of the influence of reward on FEF activity can be accounted for by changes in saccadic vigor and latency, there are still unique effects of reward on FEF above and beyond this.

We used two separate methods to demonstrate that the FEF reflects expected reward above and beyond saccadic parameters. Specifically, we (1) matched saccadic parameters for different saccade types, and (2) jointly modeled the effects of movement and reward on FEF activity using a GLM. Here we discuss potential weaknesses of these methods. In both methods, other saccade parameters, such as further derivatives of velocity (e.g., acceleration, jerk) could be uniquely driving neural activity. We used latency and velocity, as these have been previously shown to be affected by expected reward. In our GLM analysis, it could be the case that our model was unable to capture the true relationship between FEF firing and movement covariates. For example, some nonlinear combination of vigor and latency may better explain the neural activity. Additionally, there could be nonlinear effects (e.g., saturation) linking the movement covariates to the firing rate. It is possible that the neural activity that we are attributing to expected reward could be

explained by movement in a more accurate model. Nonetheless, these potential sources of error in the GLM analysis are addressed by the matching analysis, which does not make assumptions about the relationship between the movement variables and neural activity. Both methods yield the same results.

Interpretation limitations

While we distinguished between the neural representation of expected reward and motor variables, there are still many unanswered questions involving the neural mechanisms related to expected reward. First, we cannot be sure about causality. While it is possible that FEF activity is causally altering movement parameters, it is also possible that FEF activity is simply correlated with these motor outputs. In this scenario, expected reward would affect both the FEF and other oculomotor structures, but only a subset of the other oculomotor structures, exclusive of FEF, would be causally linked to the motor output. Observational studies like ours cannot ultimately answer such causal questions without further interventions.

Additionally, it is particularly difficult to determine the processes that are being modulated by expected reward in cortical oculomotor circuitry where there is a unique confluence of cognitive, sensory, and motor processing [99]. For example, expected reward modulates visuospatial working memory [88] and spatial attention [64, 100, 85]. In our study, the effect of expected reward that was observed above and beyond saccade parameters may thus have been due to an increase in attentional gain. While this is a likely explanation of our findings, in a naturalistic task, attention and expected reward are tied together, preventing mechanistic identification.

Lastly, neural activity may be related to the value of the reward itself, or the reward’s behavioral relevance. Indeed, human and non-human primate studies suggest that more anterior regions of the cortex (e.g., orbitofrontal, dorsolateral prefrontal) tend to be modulated by the value of the reward, while more posterior regions of frontal cortex (e.g., FEF, premotor) tend to be equally modulated by rewarding and aversive stimuli. This suggests that the latter are sensitive to the behavioral relevance, rather than the value, of the stimulus [101, 90, 102]. Thus, the effects of reward expectation in our study might more appropriately be attributed to the behavioral relevance of the target stimulus, rather than its reward value.

Lastly, we want to reiterate that, while we aimed to record only from FEF, it is possible that some recorded neurons may have been in nearby areas (see *Methods* and *Results: Differences between monkeys for details*). Given the past observation of differences in reward modulation between FEF and other prefrontal areas [87], it is important to consider the possibility that the individual differences in the reward-related effects we observed could be explained by differential contamination of the FEF recordings from other regions with different functional properties. Nonetheless, our main claim that there are neurons in FEF that represent expected reward after controlling for saccadic variables is strongest in monkey M14, whose recordings were confirmed to be in FEF.

Natural Scene Search

The interplay between the concepts of exploration and exploitation may provide meaningful insights into our natural scenes findings about expected reward. When searching for an object, we initially make exploratory saccades that aim to gather information about

the scene [35]. When we then find the object in the visual periphery, we make an exploitative saccade that aims to foveate the object in order to gain reward. The enhancement of neural activity associated with exploitative (expected reward) saccades could have several purposes. Along with allowing the subject to reach the object faster (due to increased vigor and decreased latency), the higher firing rates could lead to more precise saccades (due to the increased gains of tuning curves; Fig. 3g, [78]), which are likely more important for exploitative saccades. Tuning curves with higher gains could also more precisely allocate attention, which is useful to avoid distraction and ensure that this is the correct object (e.g. to make sure this is the berry you are looking for, not a poisonous one). Thus, the effects of expected reward during visual search can be understood in the context of exploration and exploitation in the real world.

Our results also highlight the value of using natural scenes in vision experiments with non-human primates [55, 103, 104, 34, 105, 106, 107]. For instance, we were able to analyze saccades that landed near the target accidentally (without reward expectation), which would not occur in an experiment with simple stimuli. The finding that these accidental target saccades did not increase firing rates can help explain previous discrepancies between experiments using artificial stimuli and natural scenes. While there have been studies using artificial stimuli that have found evidence that FEF encodes task-relevance or feature-based attention (i.e., visual similarity to a target) [11, 13], studies using natural scenes have not yielded this result [1]. Similarly, in our study, T+/ER- were not accompanied by an enhanced response, even when the target was contained within the RF. It may be the case that the ease of localizing stimuli in the artificial stimulus paradigms

more consistently leads to an awareness, and subsequent neural representation, of the target and objects similar to the target [107], as it did with the T+/ER+ saccades. These differences highlight the importance of using natural stimuli in order to test ecologically relevant behaviors.

CHAPTER 4

From preliminary to definitive plans: two classes of neurons in frontal eye field**Abstract**

Prior to selecting an action, we often consider other possibilities. How does the brain represent these preliminary plans prior to action selection? Here, we investigated this question in the oculomotor system during self-guided search of natural scenes. We found two classes of neurons in the frontal eye field (FEF): 1) “late selection neurons” that represented the selected action plan not long before the upcoming saccade, and 2) “early selection neurons” that became predictive of the upcoming saccade much earlier, often before the previous saccade had even ended. Crucially, these early selection neurons did not only predict the upcoming saccade direction; they also reflected the probabilities of possible upcoming saccades, even when they did not end up being selected. Our results demonstrate that during naturalistic eye movements, separate populations of neurons code for preliminary and definitive plans.

Introduction

Deciding where to look next in the real world is a complex process, as we must rapidly decide between countless options. Prior knowledge about the environment and past behavior can facilitate decisions by focusing limited computational resources on seemingly

good options. For example, if you are currently looking on the left side of the desk for a pencil, it will be most useful to look rightwards next. Using prior information to make preliminary plans about upcoming saccades could be an efficient use of neural resources in the oculomotor system.

Several previous studies have shown that oculomotor structures in the brain utilize prior information for planning saccades [108, 109, 14, 110, 12, 111]. In macaque superior colliculus (SC), burst neurons show increased pre-target activity [108, 109] when there is an increased probability that a target will be placed in the neurons' receptive fields. The same effect was found in corticotectal neurons in the frontal eye field (FEF) [14]. In both SC and FEF, the neurons with pre-target activity also responded after target onset, with greater activity when the target actually appeared in their receptive fields. That is, the same population of neurons appeared to be involved in the preliminary (pre-target) planning and final (post-target) action selection. However, it remains unclear how the oculomotor system accomplishes various stages of planning and execution, in particular in more naturalistic settings where targets are typically not flashed on and off.

Unlike unconstrained, natural eye movement behavior, the tasks used in previous studies imposed substantial limitations on the prior information available for making preliminary plans. Rather than eliciting self-guided search behavior, these tasks elicited single saccades instructed by a target. This approach eliminated the ongoing planning of sequences of saccades, which is a function of FEF neurons in natural search conditions [112, 13]. These tasks also removed the possibility of ruling out saccade targets based on previous saccades [110]. Moreover, the oculomotor system is modulated by eye position in a manner that favors movement towards the center of the oculomotor range [113].

By removing starting eye positions as a variable, these previous studies also removed a significant source of prior information for constraining the range of potential eye movements. In naturalistic settings, much remains unknown about how the oculomotor system represents preliminary plans, and how they evolve into definitive plans.

Here, to explore how the oculomotor system represents preliminary and definitive plans during naturalistic saccades, we recorded from macaque FEF during a natural scene search task. Rather than finding that the same neurons represented preliminary and definitive plans, we found two separate classes of neurons. “Early selection neurons” became predictive of the upcoming saccade prior to the current fixation, before they could have received new visual information. “Late selection neurons”, on the other hand, coded for the selected action plan shortly before the upcoming saccade. Importantly, the activity of early selection neurons related not only to the upcoming saccade; early selection neurons also reflected the probabilities of upcoming eye movements based on the current eye position, regardless of the actual selected saccade direction. We find that in naturalistic settings, two separate classes of neurons code for preliminary and definitive saccade plans.

Results

Behavior

To investigate how preliminary and definitive plans are represented during self-guided eye movements, we recorded single units from the frontal eye field (FEF) while head-fixed monkeys freely searched for a target embedded in natural scenes (Fig. 1A) [2, 1]. Trials either ended when the monkeys made 20 saccades without finding the target, or when

they made a saccade to the target and held gaze there to receive a reward. During such a self-guided search, monkeys could make preliminary plans for saccades before they have new detailed visual information at each upcoming fixation location.

One easily quantifiable factor that could contribute to preliminary saccade plans is the eye position on the screen. For instance, when the monkey is fixating on the right side of the screen, there are more possible saccadic opportunities to the left, and thus the monkey might make preliminary plans to go left (Fig. 1B).

To explore this idea, we defined a quantity ϕ , which was the angle between the eye position vector (relative to the center) and the upcoming saccade vector (Fig. 1C). When going back towards the center, $\phi = 180^\circ$, and when going away from the center, $\phi = 0^\circ$. We found that monkeys are more likely to look approximately opposite of their current position (away from the borders of the screen), and the effect is stronger when closer to the border of the screen; Fig. 1D). This is in line with the known finding of center bias in eye movement behavior [38, 114, 115]. Interestingly, the peak of ϕ is not at exactly 180° (when going back towards the center). In both monkeys, there is a higher probability of $\phi = 135^\circ$ or $\phi = 225^\circ$ than $\phi = 180^\circ$. This is because these statistics not only reflect the possible on-screen saccades (which would be centered on 180°); they also reflect any other strategies and biases of the monkeys, which could be involved in preliminary planning.

In addition to saccade directions, preliminary saccade planning may also be reflected in saccade latencies. Indeed, latencies were shorter for saccades made approximately opposite of the eye position, (at ϕ close to 180° ; both monkeys, $p < 1e - 10$; Fig. 1E), and the effect was stronger for eye positions closer to the border (Monkey J, $p = 4.9e - 4$; Monkey K, $p = 7.9e - 10$; Fig. 1F). These findings are consistent with several previous

studies showing that saccades back towards the center have shorter latencies [113, 116, 117]. Interestingly, latencies were actually slightly less for saccades at ϕ 's slightly offset from 180° (Fig. 1E). This is similar to the above finding that the distribution of ϕ was peaked at angles slightly offset from 180° . Overall, the monkeys' behaviors suggest that the oculomotor system became more prepared to look in a given direction as the probability of a saccade in that direction increased.

Saccade-related neural activity

To investigate the neural basis of preliminary and definitive saccade planning, we first looked at the time at which neurons' activities became predictive of the upcoming saccades. We focused on a subset of neurons (86/226, 38% of recorded neurons) whose activity we found to be modulated by the upcoming saccade direction (see *Methods*). Among this subset, we categorized these neurons as Early Selection (E-Sel) or Late Selection (L-Sel) depending on the time that their activity predicted the upcoming saccade direction. More specifically, we found the time that activity distinguished between upcoming saccades into neurons' preferred directions (PDs) versus opposite their PDs (Fig. 2; see *Methods*). L-Sel neurons (48/226, 21% of recorded neurons) became responsive to the upcoming movement direction following fixation (Figs. 2A and S1A, rows 1,2). E-Sel neurons (38/226, 17% of recorded neurons), however, became responsive to the upcoming saccade prior to fixation, i.e., during the previous saccade (Figs. 2A and S1A, rows 3,4). This agrees with the findings of Phillips and Segraves [112], who showed that during self-guided search, a subpopulation of FEF neurons became selective for the upcoming

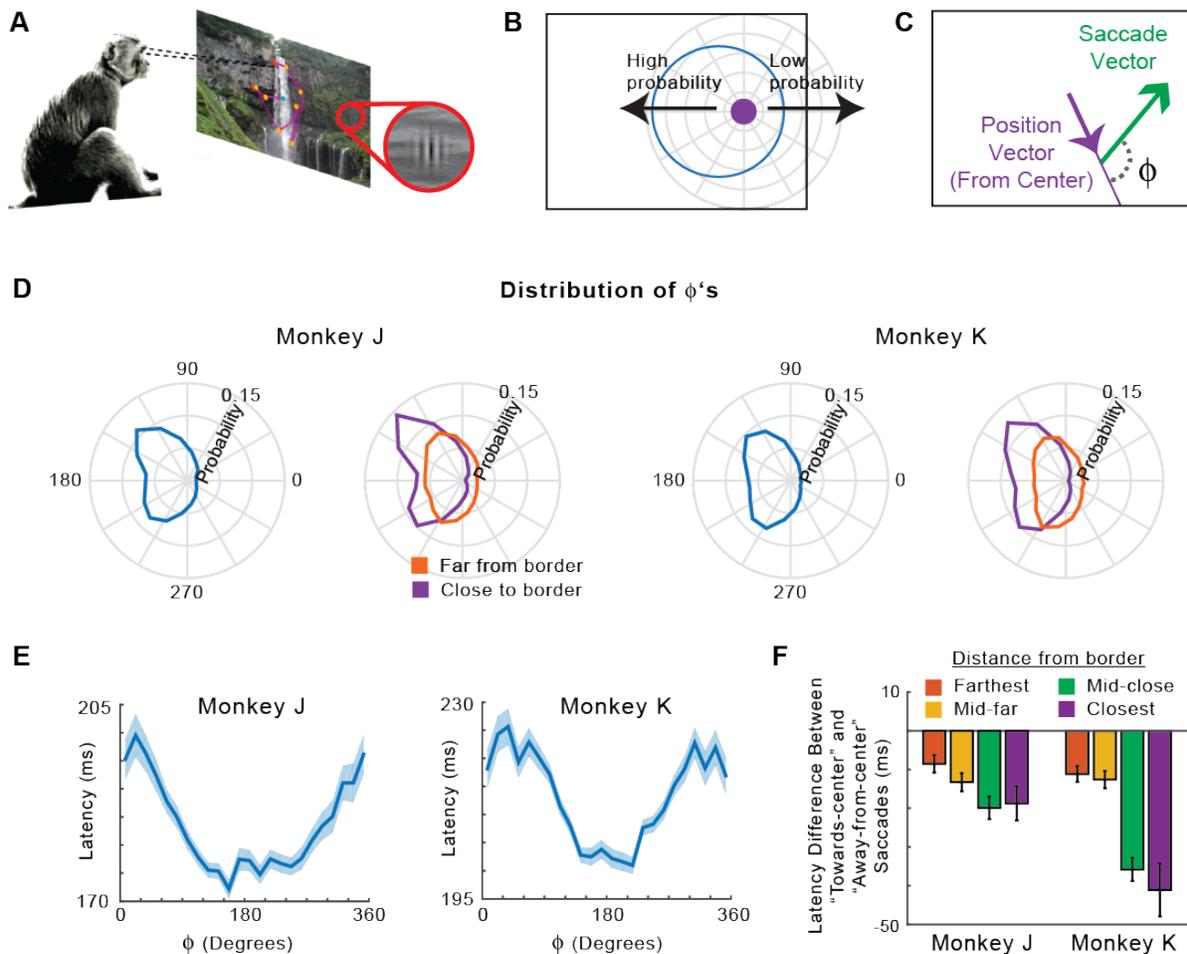


Figure 4.1: **Experiment and behavior.** (A) Monkeys freely searched for a Gabor target embedded in natural scenes. (B) The probability of the direction of the upcoming saccade is dependent on the eye position on the screen. This is an example where the eye position is to the right of the screen. (C) We quantify the relationship between the upcoming saccade direction and position using ϕ , the angle between the position vector (relative to center), and the upcoming saccade vector. (D) The distribution of ϕ 's for all saccades (blue), and split according to whether the starting eye position was close (purple) or far (orange) from the border. The close/far distinction was based on being less/more than the median distance from a border. (E) The mean latency of saccades as a function of ϕ . (F) The latency difference between saccades towards the center ($|\phi - 180^\circ| < 60^\circ$) and saccades away from the center ($|\phi - 180^\circ| > 120^\circ$), as a function of the starting eye position's distance from a border. The distance from the border was divided into quartiles.

saccade prior to fixation. As E-Sel neurons demonstrated predictive activity that preceded detailed visual information (available only after fixation), these neurons seemed to represent preliminary planning of saccades.

If the neurons' activities are involved in saccade planning, then we would expect them to relate to the latency of the upcoming saccade, with higher activity predictive of shorter saccade latency [118, 119, 120, 121]. We found that this is the case for both E-Sel and L-Sel neurons. For saccades into neurons' PDs, shorter saccade latencies were associated with greater neural activity (Figs. 2B and S1B). This provides additional evidence that both E-Sel and L-Sel neurons are involved in planning upcoming saccades, but at different times.

Preliminary planning based on eye position

In the preceding section, we showed that E-Sel neurons became predictive of the upcoming saccade during early stages, which suggested preliminary planning. These preliminary plans could be related to the monkey's eye position on the screen. For instance, when the monkey is fixating on the left side of the screen, it may make preliminary plans to move rightward. Thus, if a neuron with a PD to the right was involved in preliminary planning, we would expect this neuron to have greater activity when the monkey's eye position was on the left side of the screen. Generalizing this example, we would expect neurons involved in preliminary planning to have greater activity when the monkey's eye positions are such that the next saccade is likely to be into the neurons' PDs. This occurs when the eye positions are approximately opposite of the neurons' PDs. Importantly, we

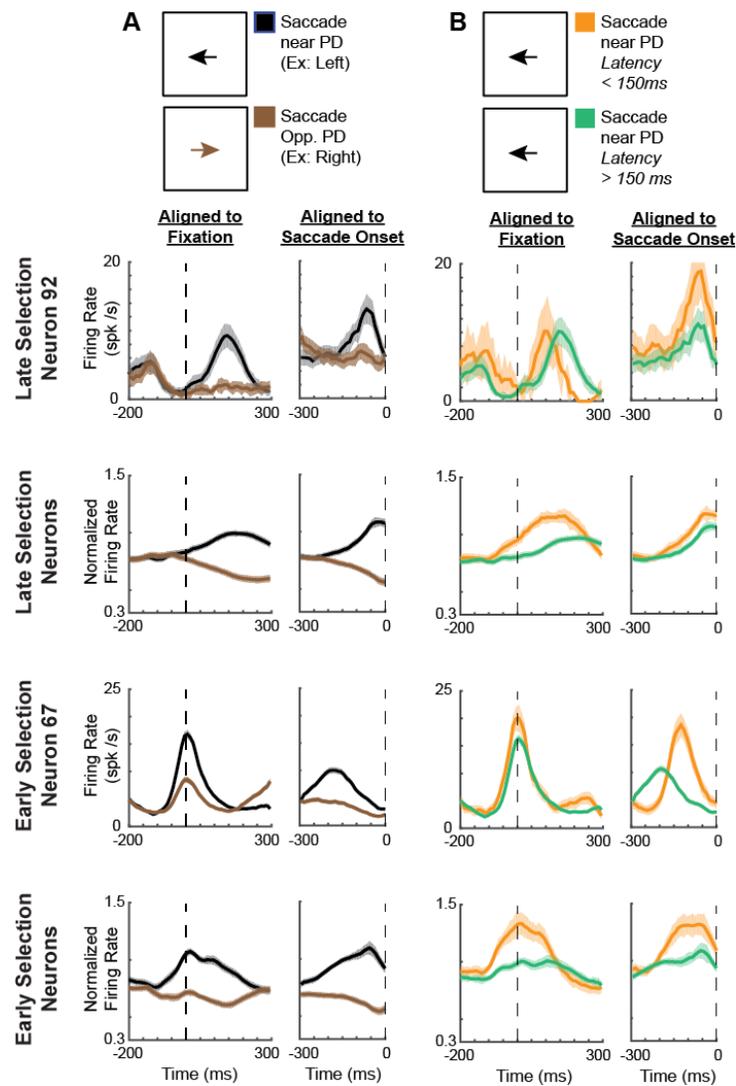


Figure 4.2: **Different times of saccade selectivity for E-Sel and L-Sel neurons, and neural differences related to saccade latencies.** Peri-event time histograms (PETHs), aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). **First Row of PETHs:** An example late selection neuron. **Second Row:** Normalized averages of late selection neurons. **Third Row:** An example early selection neuron. **Bottom Row:** Normalized averages of early selection neurons. **(A)** PETHs of saccades near the preferred direction (PD; black) versus opposite the PD (brown). **(B)** PETHs of saccades near the PD, for saccade latencies less than 150 ms (orange) versus latencies greater than 150 ms (green).

should be able to see evidence of this preliminary planning based on position, regardless of the direction of the actual saccade.

We investigated whether the time course of neural activity (as shown by peri-event time histograms, PETHs) of E-Sel neurons depended on the initial eye position. Indeed, we found that these neurons did have greater activity when the eye position at fixation was opposite of the neurons' PDs (Figs. 3 and S2, PETH rows 3,4; red trace higher than blue trace), i.e. when the upcoming saccade is likely to go into the neurons' PDs. This differentiation of activity based on the initial eye position began prior to fixation, long before the upcoming saccade. This is around the same time these neurons became selective for the upcoming saccade (Figs. 2 and S1). Importantly, even when controlling for the direction of the upcoming saccade, E-Sel neurons still had greater activity when the eye position was opposite the PD (Figs. 3 and S2). Moreover, this activity difference was not due to latency differences (Fig. S3). In other words, neurons had higher activity when the eye was in a position that made a saccade into the PD likely, even if the saccade didn't actually end up going into the PD (Figs. 3B and S2B). This finding, along with the early timing of these neurons' responses, supports the idea that E-Sel neurons are involved in preliminary planning.

Do L-Sel neurons also represent this information about preliminary plans? The PETHs reveal that these neurons did not have higher activity when the initial eye position was opposite the PD (Figs. 3 and S2, PETH rows 1,2; red trace does not rise above the blue trace). In fact, on average, activity was higher for positions near neurons' PDs, likely due to persisting activity from the previous saccade. This finding, along with the late

timing of these neurons' responses, supports the idea that L-Sel neurons are not involved in preliminary planning.

Model-based analysis

The previous PETH-based analyses contained several confounding factors. For instance, activity around the time of fixation could have been related to the previous saccade. Plus, there were correlations between eye position and upcoming saccades that were not fully disentangled. To more rigorously determine how much of the observed neural activity related to the actual upcoming saccades versus the eye position (which can be used in preliminary planning), we used a generalized linear model (GLM) analysis. As input variables to the model, we included the eye position, upcoming saccade vector, upcoming saccade velocity, and previous saccade vector (see *Methods*). This model-based analysis allowed us to estimate the importance of saccade and position variables.

For E-Sel neurons, the upcoming saccade covariate's importance began to increase approximately 100 ms before fixation, and ramped up for about 200 ms (Figs. 4B and S4B). The position covariate's importance was also high at this time, and showed a similar increase in importance over time. This analysis supports that early activity was indeed predictive of the actual saccade that would occur, but the activity was also independently influenced in a similar manner by eye position.

For L-Sel neurons, the saccade covariate's importance began to increase after fixation (Figs. 4A and S4A). While the position covariate had some importance on average, its importance across time was static, rather than ramping up. Additionally, the importance of the saccade parameter was much greater than that of the position parameter. This

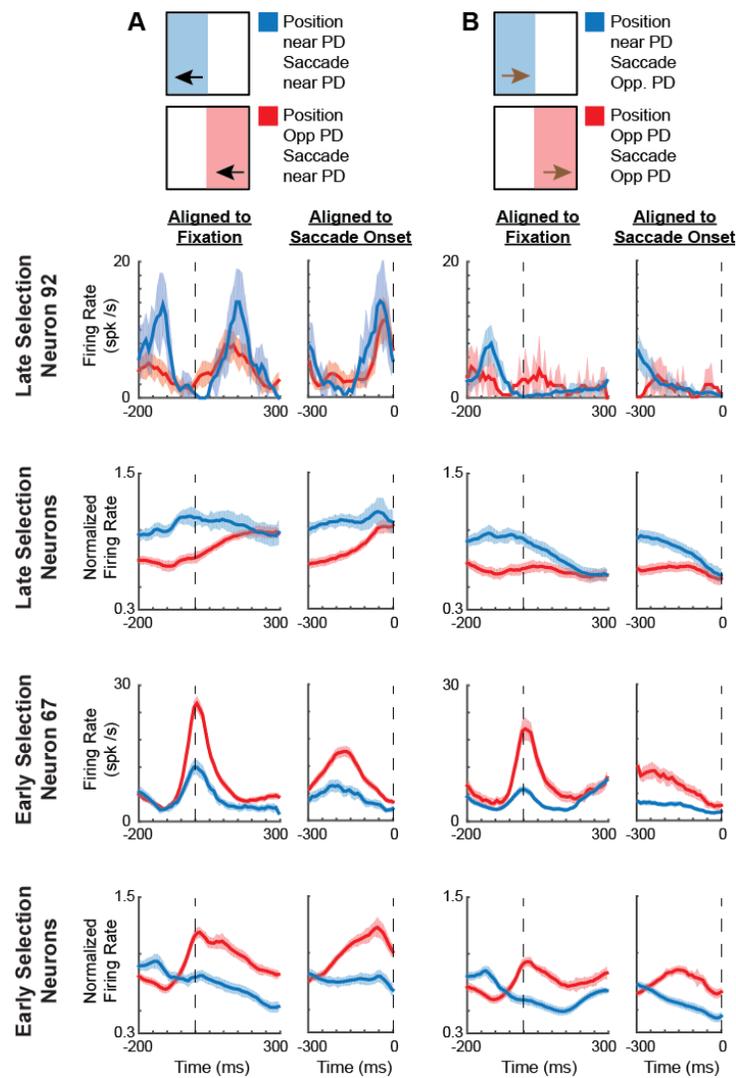


Figure 4.3: **E-Sel, but not L-Sel, neurons have higher activity in positions that are more likely to result in saccades near the PD.** Peri-event time histograms (PETHs), aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). **First Row of PETHs:** An example late selection neuron. **Second Row:** Normalized averages of late selection neurons. **Third Row:** An example early selection neuron. **Bottom Row:** Normalized averages of early selection neurons. **(A)** PETHs of saccades near the PD, with a starting position near the PD (unlikely that upcoming saccade will be near PD; blue) versus a position opposite the PD (likely that upcoming saccade will be near PD; red). **(B)** PETHs of saccades opposite the PD, with a starting position near the PD (blue) versus a position opposite the PD (red).

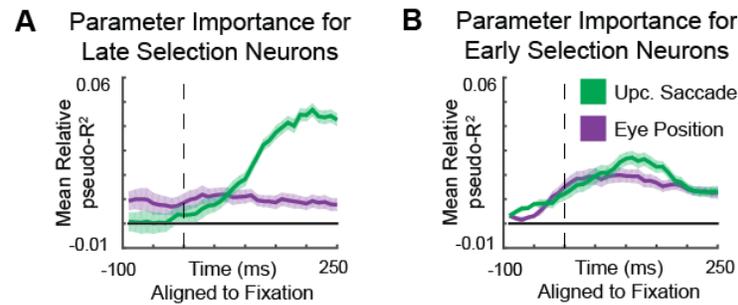


Figure 4.4: **GLM results - importance of upcoming saccade and eye position parameters.** Importance of parameters in the generalized linear model, across time, aligned to fixation. **(A)** The mean relative pseudo- R^2 of the upcoming saccade (green) and eye position (purple) covariates for late selection neurons. **(B)** Same plot for early selection neurons. Shaded areas represent SEMs.

confirms that this later activity was primarily related to the actual saccade that would occur.

Early activity reflects probabilities of upcoming saccades

Above, we observed that E-Sel neurons had greater activity when the eye position made it more likely that the upcoming saccade would be toward neurons' PDs. But does this activity more precisely relate to the task behavior? That is, does the early neural activity reflect the probability distribution of upcoming saccades (Fig. 1d)?

To understand how E-Sel activity relates to the distribution of upcoming saccades, for each neuron we calculated the average activity as a function of the “relative position angle”. The relative position angle is the position relative to the PD of the neuron (Fig. 5a), and is an analog to the behavioral measure ϕ (Fig. 1c, Fig. 5b). We averaged across all neurons and tracked this distribution over the course of saccade planning. Importantly, we only included fixation periods preceding saccades that were made opposite of the

neurons' PDs in order to minimize contamination by neural activity that was related to definitive saccade plans or execution.

In the resulting plot, we observed greater activity when the eye position was approximately opposite the neurons' PDs (Figs. 5c and S5a). This relation was not due to activity from the previous saccade (Fig. S5c). Looking more closely at the 100 ms around fixation, we can see that the greatest neural activity occurred when the eye position was not exactly opposite the PD. It was greatest at a relative position angle of about 135° , similar to the saccade statistics in the task. In fact, the circular means of the neural activity distribution and behavioral distribution were not statistically different ($p > 0.05$ via bootstrapping; Fig. 5d). Thus, the neural population activity reflected the probability distribution of upcoming saccades in the task.

We also examined whether there were differences in FEF activity when the eye position was close versus far from the border. Like before, we only included fixation periods preceding saccades that were made opposite of the neurons' PDs. We found that for eye positions opposite to the PD, there was greater activity when close to the border (Figs. 5e and S5b). This relates to the behavior, in that the probability of saccades opposite the current position is greater when closer to the border (Fig. 5f). For eye positions in the same direction as the PD, there was lower activity when close to the border (Figs. 5e and S5b). This relates to the behavior, in that the probability of saccades in the same direction as the current position is smaller when closer to the border (Fig. 5f). Thus, the early activity of the population of E-Sel neurons closely relates to the probabilities of saccades that will later occur.

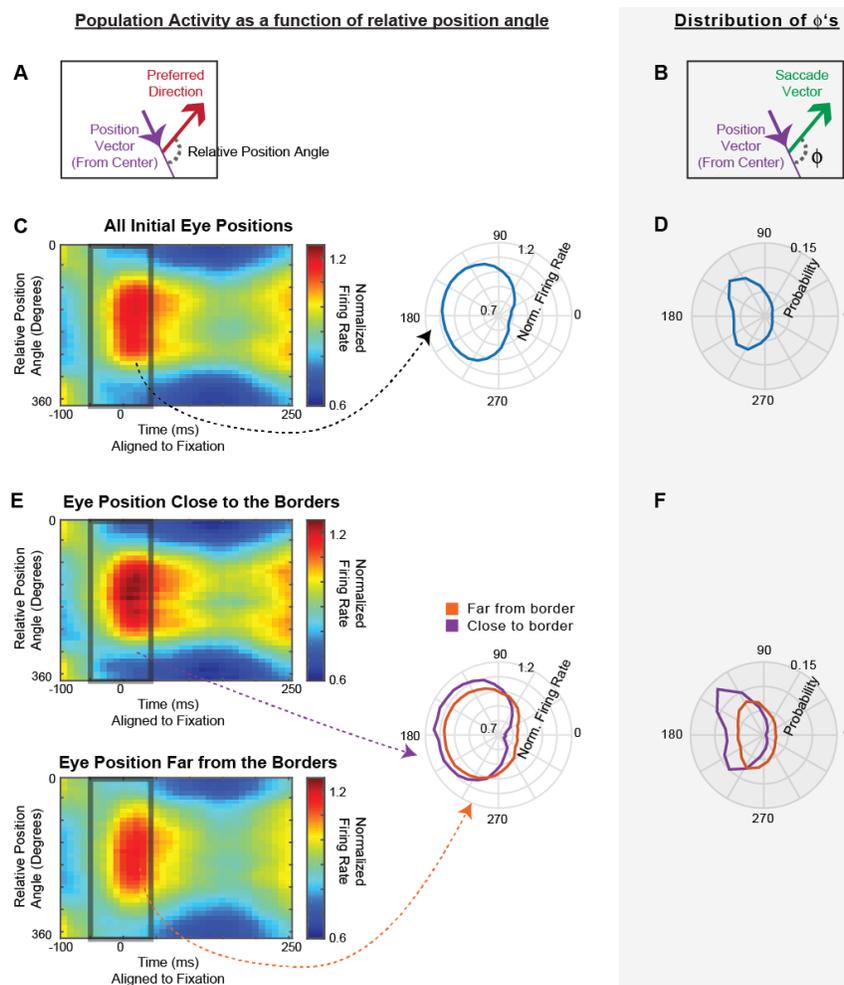


Figure 4.5: **E-Sel population activity reflects the probabilities of upcoming saccades.** On the left, we plot the population activity of E-Sel neurons as a function of the relative position angle. On the right, with a gray background, we show how this relates to the behavioral distribution of ϕ 's. (A) The relative position angle is the difference between a neuron's preferred direction and the eye position vector (from the center). (B) Copied from Fig. 1c, ϕ is the difference between the upcoming saccade and the eye position vector. (C) On the left, a heat map of normalized activity over time, as a function of relative position angle, averaged across neurons. On the right, the normalized average activity in the 100 ms surrounding fixation, plotted as a function of the relative position angle. Only saccades away from the PD are included. (D) The distribution of ϕ 's across all saccades combined across monkeys. (E) Same as panel C, but now separated for initial eye positions close to the borders (top left, right in purple) and far from the borders (bottom left, right in orange). (F) Same as panel D, but now separated for initial eye positions close to the borders (purple) and far from the borders (orange).

While the PETH results suggested that L-Sel neurons are not involved in preliminary planning, we still analyzed L-Sel neurons' activity as a function of eye position to check whether they represent the probabilities of upcoming saccades (Fig. 6). Early on, around the time of fixation, there was the greatest activity when the eye position was in the same direction as the preferred direction. This was due to activity related to the previous saccade, which was more likely to be into the PD when the current eye position was in the direction of the PD. Later during the fixation period (150-250 ms from fixation onset), activity as a function of relative position angle matched the statistics of saccades (Fig. 6a). However, this was only because we included all saccades, and the saccade-related activity captured the correlations with eye positions. When we only included saccades away from the PD (as in Fig. 5), almost all of the activity disappeared (Fig. 6b). Another way we tested whether the apparent position-related activity was actually due to the correlation between position and upcoming saccades was to resample saccades to get a different correlation structure between eye positions and saccades (i.e., we resampled to get a different distribution of ϕ 's). When we did this, L-Sel neurons' activity no longer matched the true distribution of ϕ 's (Fig. 6c). On the other hand, for E-Sel neurons, this resampling still yielded activity that was close to the true distribution of ϕ 's (Fig. 6c). This demonstrated that the activity of L-Sel neurons related to the actual saccades themselves, while E-Sel neurons' activities related to the probabilities that saccades would occur.

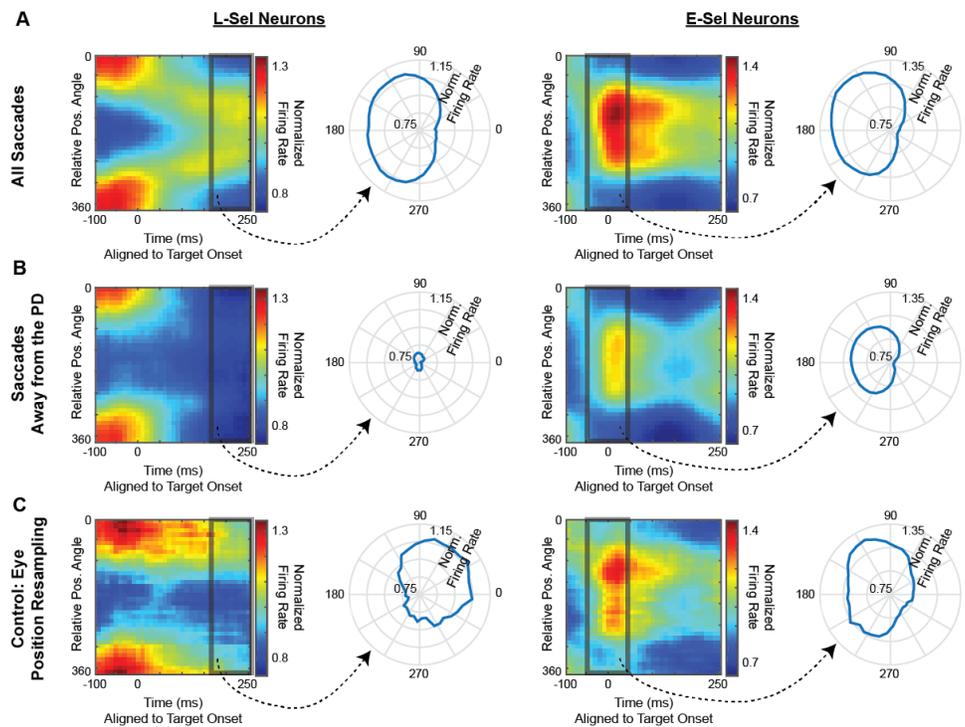


Figure 4.6: **L-Sel population activity does not reflect the probabilities of upcoming saccades.** We plot the population activity of L-Sel neurons (left) and E-Sel neurons (right) as a function of the relative position angle, as in Figs. 5 and S4. For the polar plots on the right, we use activity in the 100 ms around fixation for E-Sel neurons, as in Fig. 5. For L-Sel neurons, we use activity from 150-250 ms following fixation. Note that the color bars for E-Sel neurons differ from those in Fig. 5. **(A)** All saccades are included. **(B)** Only saccades away from the PD are included. **(C)** We control for the correlation between the eye position and upcoming saccade direction. We resampled saccades to create a distribution of saccade directions relative to position angles that were centered on 45° . This plot ensures that our main results are not simply caused by a correlation with the true upcoming saccade (in which case the activity after resampling would become peaked closer to 45°), but rather reflect the distribution of upcoming saccades (in which case the activity after resampling would remain peaked at $\sim 135^\circ$).

Discussion

Here, during a self-guided search task we found separate functional classes of neurons within the frontal eye field (FEF). Early selection (E-Sel) neurons became predictive of

the upcoming saccade direction prior to fixation, before new visual information could be processed, while late selection (L-Sel) neurons became predictive of the upcoming saccade direction following fixation. Along with being predictive of the actual upcoming saccade, E-Sel neurons activity reflected the probability distribution of upcoming saccades, regardless of the actual upcoming saccade direction. This was not the case for L-Sel neurons, whose activity primarily related to the saccade itself. This suggests E-Sel neurons are involved in preliminary planning of possible movements that may or may not be selected, while L-Sel neurons are involved in definitive planning.

Our findings have some overlap with the results of Phillips and Segraves [112], who also studied FEF during a natural scene search task. Like us, they found early saccade predictive activity in many neurons, sometimes prior to fixation. In their study, they also found that many neurons' activities were predictive of future saccades (not just the upcoming saccade), which they called "advanced predictive activity". When they split neurons into two subpopulations depending whether the neurons had advanced predictive activity or not, they found that neurons with advanced predictive activity also became selective for the upcoming saccade significantly earlier. Their classification based on whether neurons had advanced predictive activity (i.e. whether their activity predicted future saccades), may thus overlap with our classification into E-Sel and L-Sel neurons. Importantly, our study goes beyond confirming their finding that neurons have predictive activity for the upcoming saccade at widely varying times (some early and some late). Here, we have shown that E-Sel neurons don't only have activity related to the actual upcoming saccade, but they (unlike L-Sel neurons) have activity related to the probability of the upcoming saccade (based on eye position), regardless of the actual direction of the

saccade that is made. This allowed us to more generally suggest a role for E-Sel neurons in preliminary plans, which may or may not be executed.

While we aimed to record from neurons within FEF, it is possible that some of the recorded neurons were from nearby areas within prefrontal cortex (see *Methods*). Thus, it is possible that both E-Sel and L-Sel may not both have been in FEF; one class may have been in a nearby area. However, we believe this is unlikely as there were many instances in which both E-Sel and L-Sel neurons were recorded from the same electrodes at the same depths. Moreover, in M16, we recorded both E-Sel and L-Sel neurons from electrodes that were confirmed to be in FEF. Thus, while some of the neurons may have been in nearby prefrontal oculomotor structures, at least a portion of both E-Sel and L-Sel neurons were from FEF proper.

We divided neurons into E-Sel neurons and L-Sel neurons based on the time they became selective for the upcoming saccade. Classically, FEF neurons have been categorized as having visual, delay, and/or movement activity using a memory-guided saccade task [80, 10] (although see [122] for recent work revising these classifications). Since both E-Sel and L-Sel neurons are selective for the upcoming movement, one would initially assume that they would both have classical movement activity. However, during movements in natural scenes, neurons with classical visual activity have also been shown to be selective for the upcoming saccade [55, 33]. Moreover, Phillips and Segraves [112] found that neurons with classical visual activity can be selective for the upcoming saccade at widely varying times (both early and late). In the future, it would be beneficial to characterize how the classical neuron types relate to E-Sel and L-Sel neurons.

Here, we assumed that the neural activity of E-Sel neurons related to eye position was used for preliminary planning. This is in line with previous research showing that a lower stimulation threshold in FEF was required to elicit saccades opposite of the current eye position [123], which suggested that eye position biases upcoming saccades. However, it is possible that FEF activity related to position was used for computations other than, or in addition to, saccade planning. For instance, Cassanello and Ferrera [124] found that there was generally greater activity in FEF neurons when the initial eye position was opposite the neurons' PDs. However, they argued that this position-based modulation of activity could allow vector subtraction, with the purpose of keeping a memory of the target location across saccades. It is important to note that a position signal could ultimately be used for multiple purposes. For instance, there could be multiple read-outs of this position signal, one that is used for saccade planning, and another that does vector subtraction for the purpose of stability across saccades. Ultimately, given that E-Sel neurons' activities with respect to position matched the statistics of upcoming saccades, and given that FEF has a known role in saccade planning [12, 125, 15], it is improbable that E-Sel neurons were representing position solely for a purpose other than making saccade decisions.

Previous studies have also suggested that neural activity in superior colliculus (SC) is modulated by eye position in order to bias upcoming saccades. Pare and Munoz [113] found that burst neurons in SC had higher firing rates when the eye position was opposite the neurons' PDs, as we found here for FEF. However, other studies in SC [126, 127] found the opposite result (although in different tasks) - that firing rates were generally higher when the eye position was in the same direction as the neurons' PDs. It is thus possible that a subset of SC neurons use position for preliminary planning. Given the

effect of eye position on saccade latencies (Fig. 1), it makes sense that it would affect neural activity related to saccade planning throughout the oculomotor system.

How does the FEF have access to eye position information to use for saccade planning? Given that E-Sel neurons are modulated by the fixation position prior to the start of fixation, these neurons cannot be using a sensory eye position signal. Rather, we suspect that this information is computed based on a corollary discharge signal of the saccade plan coming from superior colliculus (via mediodorsal thalamus) [128, 129]. In fact, when this corollary discharge pathway is blocked, monkeys are not able to successfully make sequences of saccades [128, 129]. Thus, we would hypothesize that blocking this corollary discharge pathway would interfere with E-Sel neurons' preliminary planning.

When averaging activity across saccades, the activity of E-Sel neurons reflected the full continuous probability distribution of upcoming saccades. This extends previous work showing that neural activity in the oculomotor system reflects the probabilities of upcoming saccades when deciding between a small number of discrete targets [108, 109, 14, 130, 131]. Importantly, because our results were based on averaging across saccades, we do not know whether the FEF population, prior to single saccades, reflects the probability distribution of the upcoming saccade. An alternative explanation is that the population always makes preliminary plans for a single saccade, and when averaged across saccades, these individual plans create a distribution. In the future, it would be beneficial to simultaneously record many FEF neurons and perform a single-trial decoding analysis (as in [4]) to determine whether probability distributions are represented prior to single saccades.

Our findings suggest a link between previous studies showing pre-target preparatory activity in constrained tasks and studies showing advanced saccade planning during self-guided saccades. Everling and Munoz [14] showed that FEF neurons had higher activity prior to target onset when there was a higher probability the target would be shown in the neurons' PDs. This parallels our finding that E-Sel neurons had higher activity when there was a greater probability of the upcoming saccade being in their PDs. Additionally, during self-guided search, researchers have provided evidence for FEF planning more than one saccade in advance [112, 13]. These advanced plans could be reflected by E-Sel neurons, which are predictive of the upcoming saccade before gathering new information. Thus, there may be a common mechanism, where E-Sel neurons are involved in preliminary planning, whether based on saccade probabilities or some saccade sequence planned in advance.

Previous studies have shown that burst neurons in SC and corticotectal neurons in FEF have activity related to the probabilities of upcoming targets prior to target onset [108, 109, 14]. After target onset, the activity of these neurons relates to where the target (and saccade) actually are. In the framing of this paper, the same neurons would be involved in both preliminary planning (prior to target onset) and definitive planning (after target onset). Interestingly, our finding of two separate classes is unique from these past studies.

One possible reason for this difference is that we have simply recorded different neurons than in these previous studies, possibly because we used chronic arrays rather than targeting neurons with single electrodes. For instance, the layer 5 corticotectal neurons

recorded by Everling and Munoz [14] may be more motor-related since they are the output cells of FEF, while we did not focus on this neuron type. Alternatively, it could be due to differences in the experiments, as the previous experiments involved trials with single saccades to displayed targets, while our trials had multiple saccades and no target onset. In previous experiments, the activity after the target was displayed could have specifically been a visually-evoked response, which we did not see because we did not flash a target onscreen. Additionally, it is possible that E-Sel neurons' activity is specific to our continuous and self-guided saccades. For instance, the activity could be triggered by the end of the previous saccade, or based on planning multiple saccades in advance, both of which would not happen in an experiment with a single saccade. In the future, in order to better understand the differences in our findings, it would be useful to conduct both a classical task with targets and a self-guided search task, while recording from the same neurons.

While we have shown separate neurons related to preliminary (E-Sel) and definitive (L-Sel) planning, the neural circuits involved in transforming preliminary to definitive plans remain unclear. One possibility is that E-Sel neurons project to L-Sel neurons within FEF to influence the saccade plan. Another possibility is that E-Sel neurons project to neurons in SC, which then go on to influence the saccade plan. Both possibilities could also happen simultaneously. Our work clearly could be explained by a large number of different circuit models. Future work should aim to elucidate the circuit mechanisms behind the transformation from preliminary to definitive saccade plans.

Methods

Many of the methods here, especially for neural data analysis, are the same as in our other recent manuscripts [2, 4], and are described in the same way.

Behavioral Paradigms

Experiment

Two monkeys (Monkeys J and K; in previous papers referred to as M15 and M16 [2, 1]) freely searched for an embedded Gabor target in a natural scene background, as in [2, 1]. They were rewarded for fixating near the target for 200 ms. If they did not find the target within 20 saccades, the trial ended.

Eye tracking Eye movements were tracked with an infrared eye tracker (ISCAN Inc., Woburn, MA, <http://www.iscaninc.com/>) at 60 Hz.

Saccade detection

The start of saccades was determined by when the velocity of eye movements went above 80 degrees / sec. The end of saccades was marked by when the velocity fell below 100 degrees / sec. Saccades could only be detected after an intersaccadic interval (latency) of 90 ms. To be conservative about saccades, we only included saccades of at least 5 degrees (so that noise in the eye tracker was not classified as a saccade). Saccades longer than 80 degrees or with duration longer than 150 ms were discarded as eye-blinks or other artifacts.

Neural Data Acquisition and Preprocessing

Monkeys J and K were implanted with a 32 channel chronic electrode array (Gray Matter Research, Bozeman, MT, USA) over the frontal eye field (FEF). The depth of each individual tungsten electrode (Alpha-Omega, Alpharetta, GA) could be independently adjusted over a range of 20 mm.

Automatic spike sorting with some manual correction was performed offline using the Plexon Offline Sorter (Plexon, Inc., Dallas, TX, USA). Because any given electrode was often left in place for multiple days, we often recorded from the same neuron across sessions. To make use of this, we combined data from units that persisted across recording sessions on different days. To do this, we manually compared spike waveforms from units recorded at the same site on different days. Generally, we merged units sharing waveform shape (rise/fall characteristics, concavity/convexity, etc.), and time course. Ambiguous cases were not combined. Additionally, we included multi-unit activity in our present analysis.

As stimulation quickly degrades the recording fidelity of the tungsten electrodes in the array, we were unable to stimulate during recording to verify FEF location. Thus (in addition to having the array stereotaxically above FEF), we used functional measures to include neurons that were likely in FEF. We only included neurons that either had visual onset activity or presaccadic activity. To determine whether there was visual onset activity, we compared neural activity in the 100 ms prior to image onset with activity 50 to 150 ms after image onset, to see whether there was a significant difference. To determine whether there was presaccadic activity, we looked at peri-saccadic time histograms aligned to the start of the upcoming saccade, binned into 8 angular directions (according to

saccade direction), with each bin subtending 45 degrees. In any bin, if there was a significant difference between a firing rate and baseline, and the rise started before the saccade onset, then there was presaccadic activity. More details about recording locations and electrical stimulation results can be found in [2]. In sum, while most of the neurons were likely in FEF, it is possible that some neurons were in nearby areas.

We also only included neurons with average firing rates of at least 2 spikes / second in either the early or late period. For the eye movement experiment, the early/late periods were defined as -50 to 100 ms from the start of fixation, and -100 to 0 from saccade onset, respectively. This left us with 104 neurons from Monkey J and 122 neurons from Monkey K.

Behavioral Analysis

We excluded saccades that started or ended off of the screen. Behavioral data was combined across all sessions.

Statistics of movement

We defined the position angle, ϕ_P , as the initial fixation location (prior to saccades) relative to the center of the screen (Fig. 1c). We defined ϕ as the angular difference between the upcoming saccade direction, ϕ_S , and the position angle (Fig. 1). That is, $\phi = \phi_S - \phi_P$.

Latency effects

Latency was defined as the time from fixation to saccade onset. Latencies greater than 400 ms were excluded as outliers, as latencies of this duration could have been due to an undetected saccade.

We computed the mean latency of movements as a function of ϕ . When claiming that latencies were lower when making saccades opposite the eye position (when ϕ is near 180°), we did the following test: We calculated the Pearson's correlation between latency and $|\phi - 180^\circ|$. We then calculated the p-value associated with the correlation (using a 2-sided one-sample t-test).

We also analyzed differences in latencies between saccades that returned towards the center ($|\phi - 180^\circ| < 60^\circ$) and saccades away from the center ($|\phi - 180^\circ| > 120^\circ$), based on the distance of the eye position from the borders of the screen. To test whether the latency difference between towards-center and away-from-center saccades depended on the distance from the border, we used linear regression to fit the latency of saccades as a function of distance from the center. We then did a 2-sided unpaired t-test with unequal sample variances to analyze whether the slope was less (more negative) for towards-center saccades compared to away-from-center saccades.

Neural Data Analysis

As in our behavioral analyses, we only included saccades that remained on the screen.

Smoothed maps of neural activity

For many aspects of the following neural data analysis, we computed smoothed maps of neural activity in relation to some variable (position, previous movement, or the upcoming movement). For instance, we created a map of how neural activity varied over all positions on the screen, and a map of how neural activity varied in response to all upcoming saccade vectors. For our maps, we estimated the average firing rate at each point in space using weighted k -nearest neighbor smoothing. As an example, for the saccade variable (previous or upcoming), for each saccade we found the k nearest saccade vectors (based on Euclidean distance). We then averaged the firing rates associated with each of the k saccades, but with each weighted proportional to its distance from the given saccade to the d power.

The parameters (k and d) we used to generate the smoothed maps of neural activity are as follows. For the smoothed maps used in the generalized linear models (see section below), $k =$ the smaller of 30% of the data points and 500, $d = 0$. These parameters were found using cross-validation on held out data sets, in order to not inflate the number of significant neurons in the GLM analysis. For all other times, $k =$ the smaller of 30% of the data points and 400, $d = -0.5$. These parameters were found using cross-validation on the current data sets in order to create as accurate maps as possible. Importantly, all results were robust to a wide range of smoothing parameters.

For any single variable (e.g. position) we can get the associated estimated firing rate, θ_P , by looking up the firing rate for that position on the smoothed map. If, for instance, we want to get the estimated firing rates due to position in a time interval before every saccade, we would get a vector θ_P , which contains the estimate before each saccade.

The same can be done to estimate the firing rate due to the upcoming saccade, θ_{US} , or previous saccades, θ_{PS} .

Determining Preferred Directions of Neurons

When determining the PD, we used the 100 ms preceding saccade initiation. Let \mathbf{Y} be the vector of firing rates in that interval for every saccade. It is possible that some of the neural activity during these time periods is related to the eye position and/or previous saccade, rather than the upcoming saccade. Thus, we first aimed to remove the effect of these variables that might bias the calculated PD. To do so, we created a smoothed map of how position and previous saccades were related to neural activity, and then subtracted the activity related to these variables from the total activity (leaving activity due to upcoming saccades). That is, if θ_P is the vector of expected firing rates at given positions and θ_{PS} is the vector of expected firing rates with given previous saccades, then we fit the tuning curves to $\mathbf{Y} - \theta_P - \theta_{PS}$. More specifically, we fit a von Mises function to relate the movement directions to the firing rate due to movement:

$$\mathbf{Y} - \theta_P - \theta_{PS} = \alpha \exp(\beta \cos(\phi_S - \phi_S^*))$$

where ϕ_S is the vector of upcoming saccade directions, and α , β , and ϕ_S^* are the parameters that we fit. ϕ_S^* is the PD of the neuron.

PETHs

When plotting PETHs of individual neurons, we plotted the mean firing rate across saccades. The error bars on PETHs are the standard error of the mean (SEM) across saccades. When plotting the PETHs averaged across neurons, we first calculated the mean

firing rate (across saccades of the given condition) over time for each neuron. We then normalized this activity trace for each neuron by dividing by the maximum firing rate of the average trace (across all conditions). We then show the average of these normalized firing rates across neurons. Error bars are the SEM across neurons. All traces are smoothed using a 50 ms sliding window.

PETHs were made for different categories of movements. For the PETHs, saccades near the PD were defined as those that were within 60° of the PD. Saccades opposite the PD were defined as those greater than 120° from the PD. Positions opposite the PD were defined as position angles greater than 120° away from the PD. Positions near the PD were defined as position angles less than 60° away from the PD.

Generalized Linear Model

To determine which variables were reflected in the neural activity, we used a Poisson Generalized Linear model (GLM). Let \mathbf{Y} be a vector containing the number of spikes in the time interval we are considering, for every saccade. It has size $m \times 1$. We aimed to predict \mathbf{Y} based on several factors. We used the eye position, the previous saccade vector, the upcoming saccade vector, the peak velocity of the upcoming saccade, and a baseline term. More specifically, the covariate matrix \mathbf{X} was:

$$\mathbf{X} = \begin{bmatrix} | & | & | & | & | \\ \mathbf{1} & \boldsymbol{\theta}_P & \boldsymbol{\theta}_{US} & \boldsymbol{\theta}_{PS} & \mathbf{v}_{\max} \\ | & | & | & | & | \end{bmatrix}$$

where $\boldsymbol{\theta}_P$, $\boldsymbol{\theta}_{US}$, and $\boldsymbol{\theta}_{PS}$ are generated from the smoothed maps (see *Smoothed maps of neural activity* above). Essentially, these covariates are the expected firing rates from

position, upcoming saccade, and previous saccade (respectively) by themselves. \mathbf{v}_{max} is the vector of peak velocities of movements. The peak velocity was relative to the main sequence [77], to control for the changes of velocity with saccade amplitude (as in [2]). Note that when we run GLMs during different time intervals, we make separate smoothed maps for these time intervals.

Overall, the model that generates the firing rate (λ ; also known as the conditional intensity function) can be written as:

$$\lambda = \exp(\mathbf{X}\beta)$$

where β is a vector of the weights for each covariate that we fit, and \mathbf{X} is the matrix of covariates, which is z-scored before fitting. If there are j covariates, then β has size $j \times 1$. \mathbf{X} has size $m \times j$. Note the use of an exponential nonlinearity to ensure that firing rates are positive. The model assumes that the number of spikes, \mathbf{Y} , is generated from the firing rate, λ , according to a Poisson distribution.

We fit the model weights to the data using maximum likelihood estimation. That is, we found β that was most likely to produce the true spike output (assuming spikes were generated from the firing rate in a Poisson nature). Critically, we used (5-fold) cross-validation, meaning that the model was fit to the data using one set of data (the training set), and model fits were tested with an independent set of data (the testing set). Similarly, when calculating the test set covariates for movement and position (described in *Smoothed maps of neural activity*), we only used k-nearest neighbors from the training set, to avoid overfitting.

To test whether an individual covariate significantly influenced neural activity, we first made sure that a simplified model with only that individual covariate had significant predictive power. To determine the value of a model fit, we used pseudo- R^2 [132, 33], a generalization of R^2 for non-Gaussian variables. The pseudo- R^2 of a model is defined as:

$$R_D^2(\text{model}) = 1 - \frac{\log L(n) - \log L(\hat{\lambda})}{\log L(n) - \log L(\bar{n})}$$

where $\log L(n)$ is the log likelihood of the saturated model (i.e., one that perfectly predicts the number of spikes), $\log L(\hat{\lambda})$ is the log likelihood of the model being evaluated, and $\log L(\bar{n})$ is the log likelihood of a model that uses only the average firing rate.

Then, in order to determine the importance of that covariate to the full model, we test whether the full model predicts neural activity significantly better than a model where that covariate is left out (reduced model). To compare the fits between the reduced model (model 1) and full model (model 2), we used relative pseudo- R^2 , which is defined as:

$$R_D^2(\text{model 1, model 2}) = 1 - \frac{\log L(n) - \log L(\hat{\lambda}_2)}{\log L(n) - \log L(\hat{\lambda}_1)}$$

where $\log L(\hat{\lambda}_2)$ is the log likelihood of the full model and $\log L(\hat{\lambda}_1)$ is the log likelihood of the reduced model.

To determine significance, we bootstrapped the fits to create 95% confidence intervals, and checked whether the lower bounds of these confidence intervals were greater than 0. Note that the pseudo- R^2 and relative pseudo- R^2 values can be less than 0 due to overfitting.

Neuron types

The activities of both Early Selection and Late Selection neurons were significantly modulated by movement in the GLM analysis (in the early and/or late period). Because many neurons were significantly modulated by movement in both the early and late period, we could not split them into Early and Late Selection neurons based on the GLM results. Rather, we compared the PETHs for saccades near the PD and opposite the PD. To determine the time when a neuron became selective for movement, we found the time when the difference between the “near PD” and “opposite PD” PETHs was growing the most (constrained to be between -150 and 200 ms from fixation). This metric essentially found the time when neural activity was separating most between saccades near and opposite the PD. Neurons with a selection time before/after fixation were categorized as Early/Late Selection, respectively.

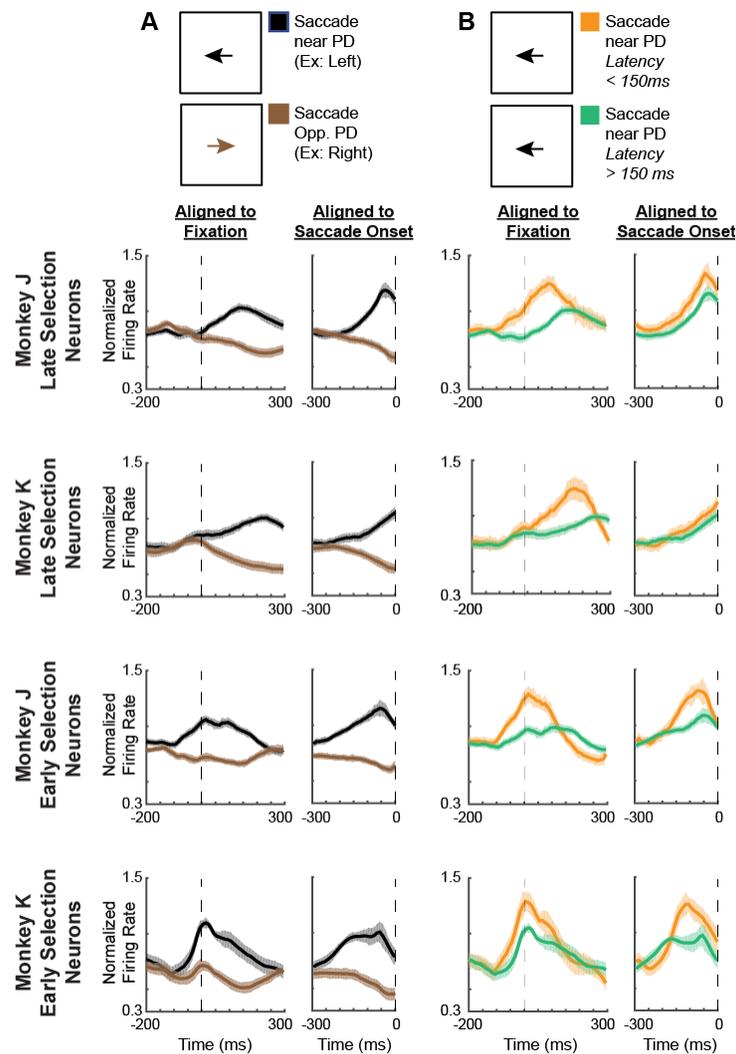
Population activity over time averaged across trials

For each neuron (of the category we were plotting), we calculated the firing rate as a function of the relative position angle (Figs. 5,6). We defined the relative position angle as the difference between a neuron’s PD and the eye position (the PD minus the eye position angle). We then normalized each neuron by dividing by its mean firing rate, and then averaged the normalized activity across neurons. We then smoothed the activity for plotting using the parameters from the smoothed maps.

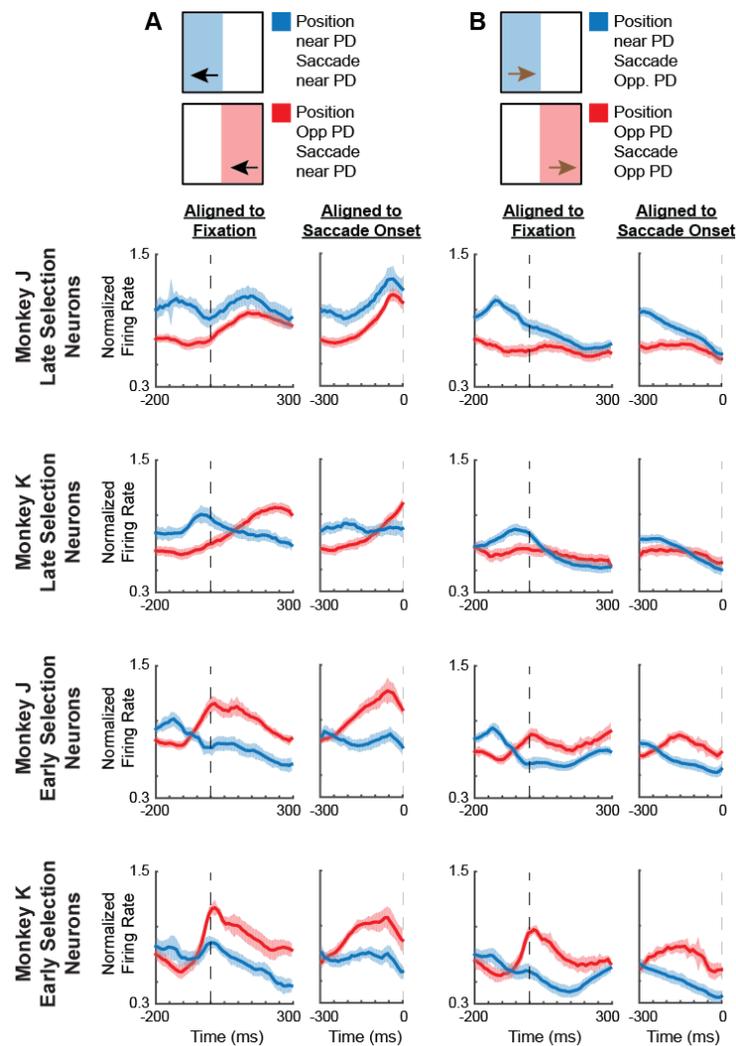
We also made several variants of the above plot (Figs. 5 & 6, Supplementary Fig. 5). We made plots in which only saccades near or far from the border (split based on median distance to the nearest border) were included. To control for the correlation

between the previous and upcoming saccades, we made a plot where saccades were only used if the angle between previous and upcoming saccades was less than 90° . To show that the apparent position-related activity was not just due to the correlation with the actual upcoming saccades, we made a plot in which the saccades were resampled, so that the distribution of ϕ 's was peaked at 45° , rather than being peaked at $\sim 135^\circ$. More specifically, we resampled saccades so that the resampled ϕ 's came from a von Mises distribution: $g(\phi) \propto \exp(\cos(\phi - 45^\circ))$.

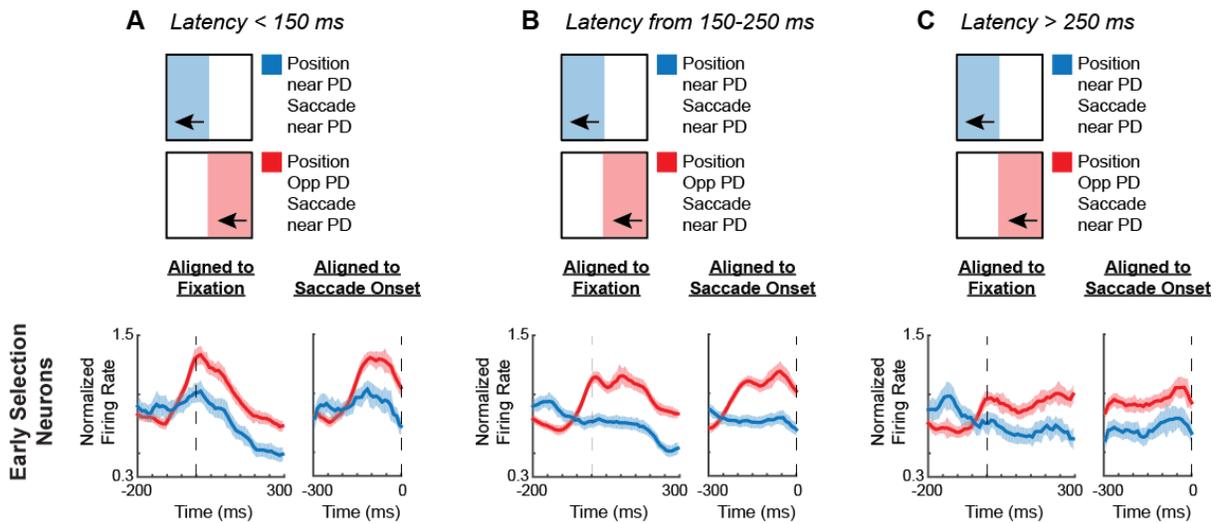
To determine the relative angular position resulting in peak activity (in the 100 ms before target onset), we calculated the activity at 20 relative angular positions (evenly spaced from 0 to 360°), and calculated the circular means of the angles weighted by their activities. We created a 95% confidence interval of peak relative position angles by bootstrapping over the set of neurons., and checked whether this overlapped with the circular mean of the behavioral distribution of ϕ 's.



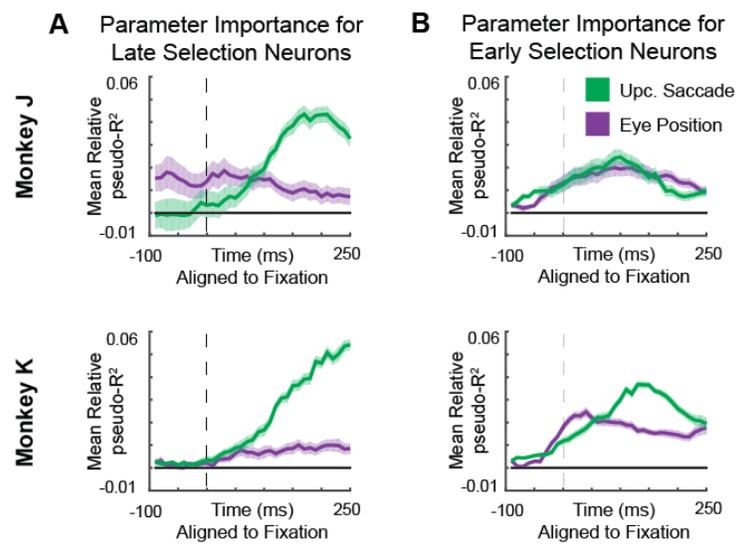
Supplementary Figure 4.1: **PETHs based on saccade direction and latency, for individual monkeys.** Peri-event time histograms (PETHs), aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). **First Row of PETHs:** Normalized averages of late selection neurons from Monkey J. **Second Row:** Normalized averages of late selection neurons from Monkey K. **Third Row:** Normalized averages of early selection neurons from Monkey J. **Bottom Row:** Normalized averages of early selection neurons from Monkey K. **(A)** PETHs of saccades near the preferred direction (PD; black) versus opposite the PD (brown). **(B)** PETHs of saccades near the PD, for saccade latencies less than 150 ms (orange) versus latencies greater than 150 ms (green).



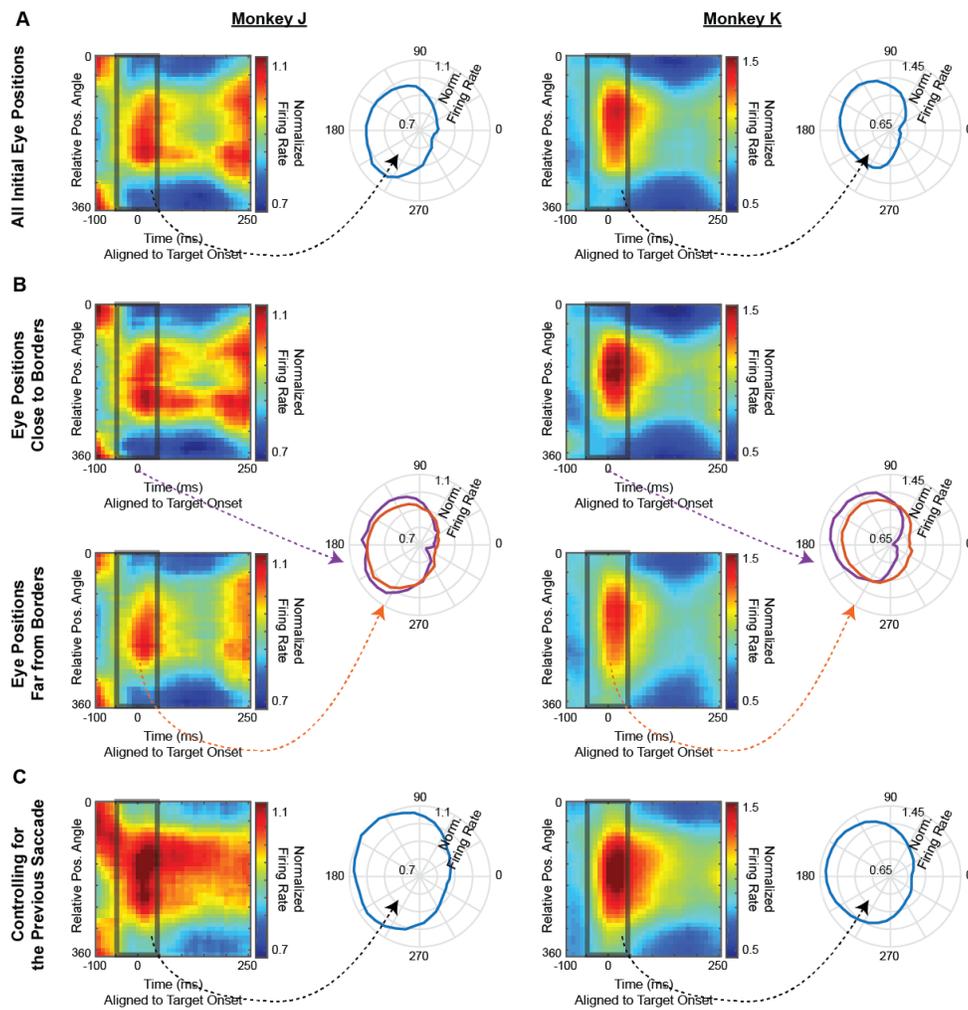
Supplementary Figure 4.2: **PETHs based on eye position, for individual monkeys.** Peri-event time histograms (PETHs), aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). **First Row of PETHs:** Normalized averages of late selection neurons from Monkey J. **Second Row:** Normalized averages of late selection neurons from Monkey K. **Third Row:** Normalized averages of early selection neurons from Monkey J. **Bottom Row:** Normalized averages of early selection neurons from Monkey K. **(A)** PETHs of saccades near the PD, with a starting position near the PD (blue) versus a position opposite the PD (red). **(B)** PETHs of saccades opposite the PD, with a starting position near the PD (blue) versus a position opposite the PD (red).



Supplementary Figure 4.3: **PETHs based on eye position, for different saccade latencies.** Peri-event time histograms (PETHs) of E-Sel neuron averages, aligned both to fixation (left part of each column) and the upcoming saccade onset (right part of each column). As in Fig. 3A, we compare PETHs of saccades near the PD, with a starting position near the PD (blue) versus a position opposite the PD (red). Now, we show these plots for (A) saccade latencies less than 150 ms, (B) saccade latencies from 150-250 ms, and (C) saccade latencies greater than 250 ms.



Supplementary Figure 4.4: **GLM results, for individual monkeys.** Importance of parameters in the generalized linear model, across time, aligned to fixation. Results for Monkeys J and K are in the top and bottom rows, respectively. **(A)** The mean relative pseudo- R^2 of the upcoming saccade (green) and eye position (purple) covariates for late selection neurons. **(B)** Same plot for early selection neurons. Shaded areas represent SEMs.



Supplementary Figure 4.5: **E-Sel population activity reflects the probabilities of upcoming saccades, for individual monkeys.** Population activity of E-Sel neurons as a function of relative position angle, for Monkeys J (left) and K (right). **(A)** On the left, a heat map of normalized activity over time, as a function of relative position angle, averaged across neurons. On the right, the normalized average activity in the 100 ms surrounding fixation, as a function of the relative position angle. Only saccades away from the PD are included. **(B)** Same as panel A, but now separated for initial eye positions close to the borders (top left, right in purple) and far from the borders (bottom left, right in orange). **(C)** We control for the correlation between the previous and upcoming saccade directions. We only included saccades in which the previous and upcoming saccade directions were less than 90° apart (while in the actual data, previous and upcoming saccades are more likely to be in opposite directions). Unlike in panels A and B, we do not exclude saccades towards the PD.

CHAPTER 5

Population coding of conditional probability distributions in dorsal premotor cortex**Abstract**

Our bodies and the environment constrain our movements. For example, when our arm is fully outstretched, we cannot extend it further. More generally, the distribution of possible movements is conditioned on the state of our bodies in the environment, which is constantly changing. However, little is known about how the brain represents such distributions, and uses them in movement planning. Here, we recorded from dorsal premotor cortex (PMd) and primary motor cortex (M1) while monkeys reached to randomly placed targets. The hand's position within the workspace created probability distributions of possible upcoming targets, which affected movement trajectories and latencies. PMd, but not M1, neurons had increased activity when the monkey's hand position made it likely the upcoming movement would be in the neurons' preferred directions. Across the population, PMd activity represented probability distributions of individual upcoming reaches, which depended on rapidly changing information about the body's state in the environment.

Introduction

To plan movements, we must incorporate knowledge of the state of our bodies within the current environment. For example, if we are standing in front of a wall, we cannot walk forwards; if our arm is fully outstretched, extending it further is not possible. Considerations like these make some movements more likely than others, resulting in probability distributions over possible movements (Fig. 1a). To understand everyday movement planning, it is essential to understand how the brain represents these probability distributions.

Several studies have investigated whether the brain represents probabilities during movement planning [108, 109, 131, 20]. In most, subjects needed to decide to move in one of a small number of directions, and the probabilities of those choices were manipulated [108, 109, 131]. These studies have shown that neurons in several brain areas have higher firing rates when there is a greater probability of an upcoming movement planned in those neurons' preferred directions (PDs). Recently, we began to study how the brain represents a continuous probability distribution rather than probabilities of discrete movements[20]. We displayed a point cloud representing an uncertain target location for movement. When we increased the uncertainty, there was a broader recruitment of dorsal premotor cortex (PMd) neurons, suggesting that PMd activity can reflect a distribution of possible movements.

Still, there is a large gap between these previous experiments and the real world, which contains dynamically changing *conditional probability distributions*, i.e. probabilities dependent on some background knowledge (here, the current state of the body in the environment). As the body moves, the probability distributions of possible upcoming

movements change. If the brain is to make use of these conditional probability distributions, it must rapidly compute updated probability distributions. Are these conditional probability distributions represented in the motor cortex, and if so where? How does the population of neurons function to represent these rapidly changing probabilities?

Here, we recorded from PMd and primary motor cortex (M1) while macaque monkeys reached to a series of targets that were chosen approximately randomly within the workspace. At all times, the position of the hand relative to the borders of the workspace dictated a conditional probability distribution of possible upcoming target locations. Behaviorally, the latencies and trajectories of the monkeys' movements were affected by this distribution, suggesting that they used this information during movement planning. Critically, neurons in PMd, but not M1, reflected these conditional probability distributions of upcoming movements prior to individual reaches, suggesting that such distributions are incorporated by the planning areas of motor cortex when coordinating movement.

Results

Experiment and behavior

To study conditional probability distributions about upcoming movements, we recorded from three monkeys with electrode arrays chronically implanted in PMd and/or M1 while conducting a random-target reaching experiment. Monkey T had an array in PMd, monkey C had an array in M1, and monkey M had arrays in both areas. In the experiment (Fig. 1b), the monkeys reached sequentially to four targets, before receiving a reward. About 200 ms after the cursor reached a given target, a new target appeared, to which the monkey could reach immediately. Due to the borders of the workspace, upcoming targets

were more likely to be presented approximately opposite of the current hand position (Fig. 1c). That is, if a monkey's hand just landed on a target on the right side of the screen, it was more likely that the next target (and therefore, movement) would be to the left of this current hand position. Therefore, probability distributions in this experiment were conditioned on the hand's current position in the workspace at the time of target presentation.

The statistics of target presentation were not completely random within the workspace; rather, targets were slightly more likely to be selected in a clockwise direction (Fig. 1d; see *Methods*). To summarize the dependence of upcoming target locations on the current hand position, we first found the angular position of the hand relative to the center of the workspace (Fig. 1e). We analyzed the distribution of ϕ 's: the angular differences between the current hand position vector and the upcoming movement vector (the vector that moves the hand to the target; Fig. 1e,f). A ϕ of 180° signifies that the target was exactly opposite of the current angular hand position. Importantly, this distribution had a circular mean of 150° rather than 180° because of the slight clockwise bias in target selection. Additionally, we can see that the farther the hand position was from the center of the workspace, the more likely the upcoming target was to be in the opposite direction (Fig. 1d). When in a position near the center, there is little information about the upcoming target direction (Fig. 1g). We aimed to determine the effect of the conditional probability distributions of upcoming movements on behavior and neural activity.

To determine whether these conditional probability distributions influenced behavior, we first analyzed movement trajectories. The reaches generally did not go straight from

one target to the next; they had some curvature that was influenced by the target probabilities (Fig. 2a,b). Early in the reach, trajectories were biased toward the expected target direction, defined as the most probable direction given the distribution of target presentations (i.e., 150° relative to the angular hand position; Fig. 1f). Additionally, the initial reach directions were more biased toward the expected target direction than simply toward the center of the workspace (Supplementary Fig. 1). Further, when the hand position was farther from the center (and the potential target distribution was more peaked) the magnitude of this bias was larger (Fig. 2b). This supports previous behavioral results showing that movement trajectories reflect uncertainty about the movement goal[133]. Our behavioral results suggest that the monkeys learned and accounted for the conditional probability distributions of possible upcoming targets when planning movements.

We then analyzed how the conditional target probabilities affected movement latencies. We found shorter latencies when the target appeared close to the vector of the expected direction (Fig. 2c; Monkey M, Pearson's $r = 0.26$, $p < 1e - 10$; Monkey T, Pearson's $r = 0.20$, $p < 1e - 10$; Monkey C, $r = 0.041$, $p = 0.0045$). Note that this result is opposite of what we would expect due to momentum from the previous movement, as the expected direction is generally approximately opposite of the previous movement. Moreover, the distance of the hand from the center also affected the latency. For initial hand positions farther from the center (resulting in a tighter probability distribution), there was a larger latency difference between reaches to targets in expected and unexpected directions. (Fig. 2d; Monkey M, $p = 2.5e - 5$; Monkey T, $p = 0.011$; Monkey C, $p = 0.068$). It is important to note that the latency and trajectory results are not independent. Since the latency is defined as the time to reach a velocity threshold (see *Methods*), the monkeys could have

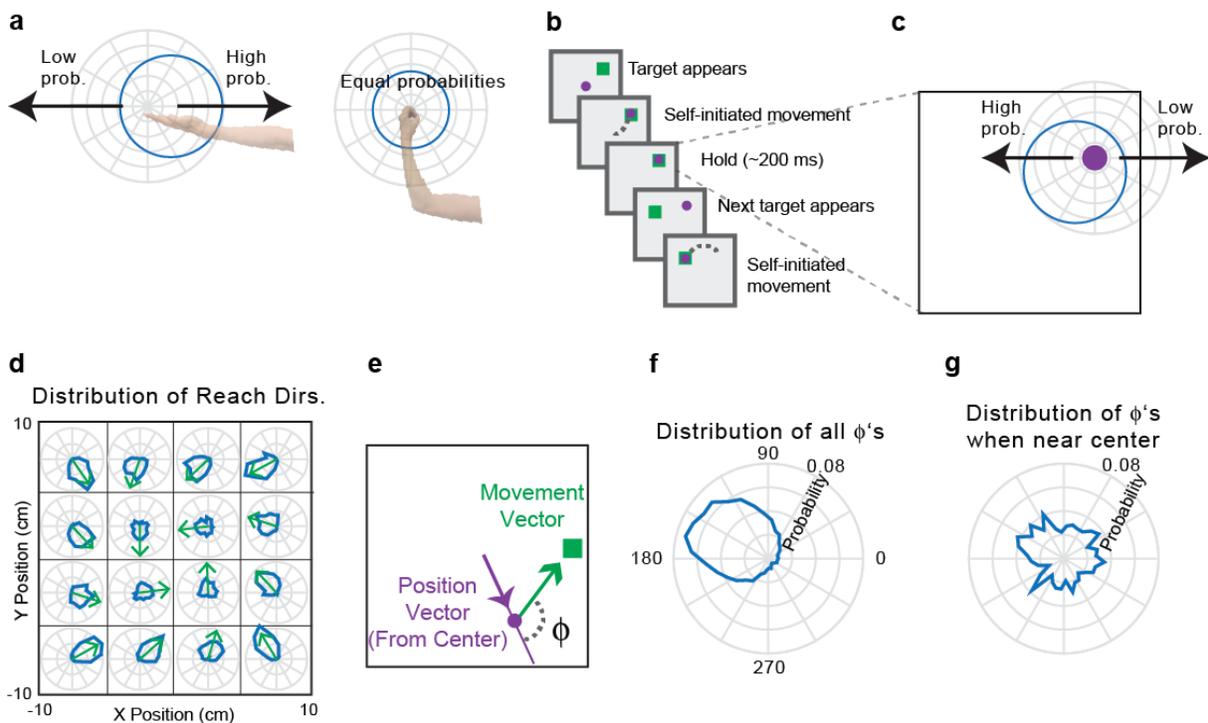


Figure 5.1: **Experimental design and statistics.** (a) If the arm is outstretched, the only possible arm movements are back toward the body (**left**). In other limb postures, it may be possible to move the arm in any direction (**right**). In blue, circular probability distributions are shown for the possible upcoming movements based on the current arm posture. (b) Experimental design. The monkey makes sequences of four reaches, briefly holding within each target box before the next target appears. (c) The current hand position limits the range of possible locations of the next target, due to the borders of the workspace and target presentation algorithm. (d) The probability distributions of upcoming reach directions (blue) from different areas of space (x and y divided into quartiles). Green arrows point toward the circular means of the distributions. (e) ϕ is the angular difference between the upcoming movement vector (the vector that brings the hand to the target) and the current angular hand position (relative to the center of the workspace). (f) The probability distribution of ϕ 's from all hand positions. (g) The probability distribution of ϕ 's from initial hand positions within 2 cm of the center of the workspace.

shorter latencies when the initial trajectory was closer to the direction of the target, since they didn't need to change direction. Overall, the monkeys' behaviors suggest that the

motor system began movement preparation towards highly probable directions prior to target appearance.

Single PMd neurons are modulated by the conditional probabilities of upcoming movements

Given that the conditional probability distributions about the potential upcoming movements affected behavior, we asked whether PMd represented this information in two monkeys. If PMd represents these probabilities, then we would expect neural activity preceding target onset to be modulated based on the anticipated possible target locations.

When observing peristimulus time histograms (PSTHs; Fig. 3a-c, Supplementary Fig. 2a-c for individual monkeys), we found some “potential-response” (PR) neurons (nomenclature as in [18]). As expected for PMd neurons, these neurons’ activity increased when a target was presented near their PDs (Fig. 3a). Crucially, PR neurons’ activity was also modulated prior to target presentation by the range of possible upcoming movements. When the angular hand position was opposite these neurons’ PDs (causing a higher probability that the upcoming target would be near the PD), pre-target activity increased. That is, for PR neurons, the red traces in Fig. 3b and c were elevated prior to target onset. Note that the activity prior to target onset in Fig. 3a for PR neurons is due to the correlation between the upcoming target and the current hand position; monkeys were apparently able to anticipate the upcoming reach direction. We also found “selected-response” (SR) neurons, whose activity was significantly modulated only after target presentation. That is, for SR neurons, the red and blue traces barely differed prior

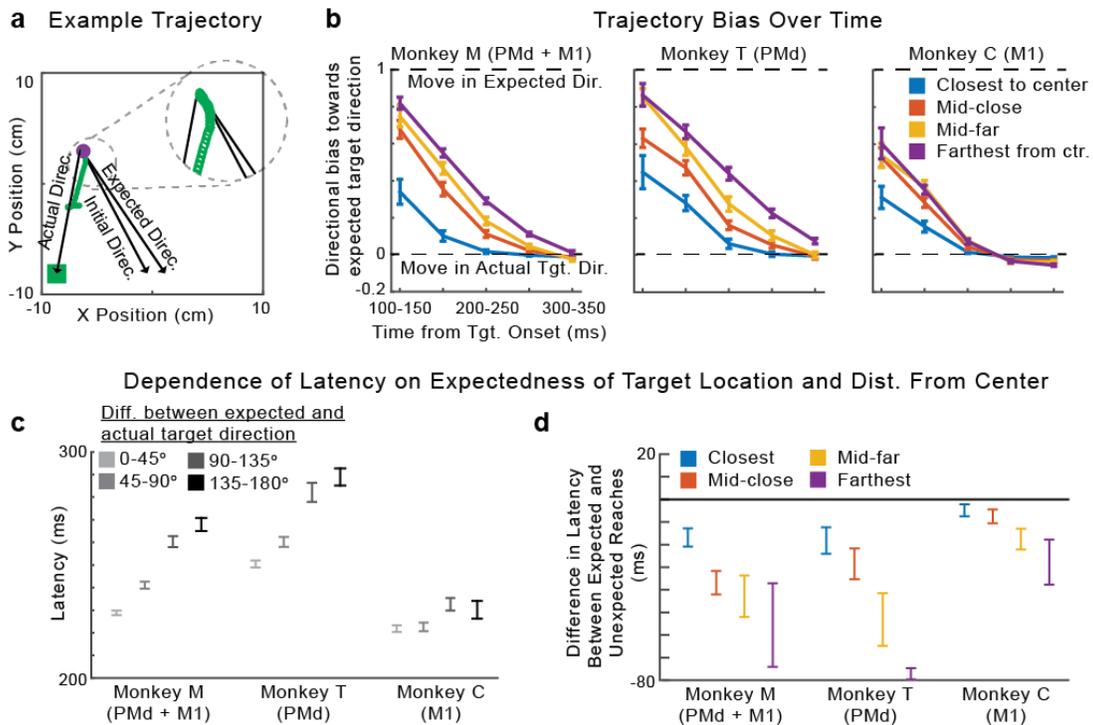


Figure 5.2: **Behavior.** (a) An example trajectory. The initial direction of the reach (green) starts toward the expected direction of the target, given the current hand position. It later moves in the actual direction of the target. The inset shows an enlarged view of the beginning of the reach. (b) The median bias of the trajectory over time. A bias of 1 signifies that the direction of the trajectory is toward the expected target direction, while a bias of 0 signifies that the direction of the trajectory is toward the actual target direction. Negative biases signify movement away from the expected direction. Different traces are shown for hand positions at varying distances from the center of the workspace. Error bars are standard errors of the median. (c) The mean latency of reaches as a function of the angular difference between the actual and expected target directions. (d) The difference in mean latency between expected and unexpected reaches (expected minus unexpected), depending on the hand's distance from the center. "Expected" reaches are those that had an angular difference between the actual and expected target directions of less than 60° . "Unexpected" reaches had an angular difference of more than 120° . In panels c and d, error bars represent SEMs. In panels b and d, distances from the center are divided as follows: "closest" is 0-20% of distances from the center, "mid-close" is 20-40%, "mid-far" is 40-60%, and "farthest" is 60-100%. We used these divisions for plotting, rather than standard quartiles, to ensure that there were "unexpected" reaches in each bin.

to target onset (see Supplementary Fig. 3 for an explanation of the slight difference before target onset in the PSTHs). Thus, PSTHs suggest that a subset of PMd neurons is modulated by the probability of upcoming movements, which seems to form part of the monkey's movement planning.

To analyze the factors contributing to neural activity more rigorously, we used a generalized linear modeling (GLM) approach. This approach can inform us whether the current hand position (and consequent probability distribution of upcoming target locations) significantly modulated neural activity above potential confounds related directly to the previous and upcoming movements. The GLM found that 13% (99/770) of neurons were PR and 42% (322/770) were SR neurons using this conservative classification approach (see *Methods* for classification criteria, and Supplementary Fig. 4 for percentages with a less conservative criteria). For both types of neurons, the upcoming movement covariate began to matter after target onset (green trace in Fig. 3d, Supplementary Fig. 2d for individual monkeys). For PR neurons, but not SR neurons, the importance of the hand position covariate (purple trace) began to increase more than 200 ms prior to target onset, until target onset. The GLM analysis thus supports our PSTH results; prior to target presentation, PR neurons' activities are modulated by hand position, which determines the distribution of potential upcoming targets.

PMd population jointly represents the distribution of upcoming movements

An important question is how the neural population represents the probabilities about upcoming movements. We showed above that when the monkey's hand is in a position that makes an upcoming target more likely to appear near a PR neuron's PD, the neuron

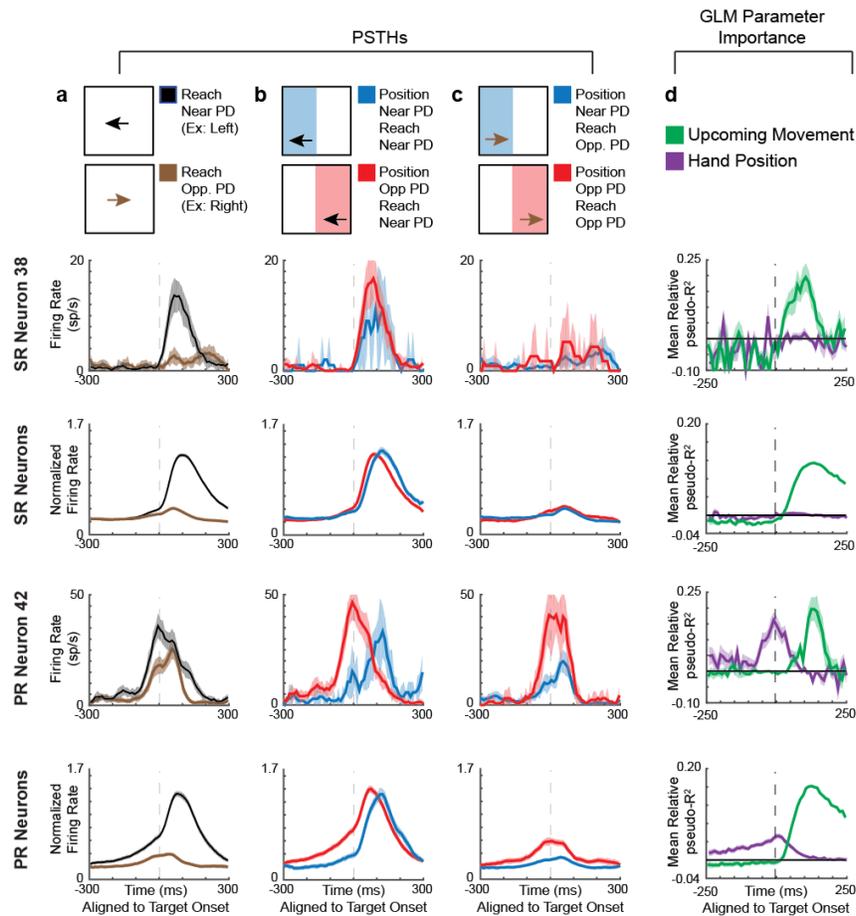


Figure 5.3: **PMd PSTHs and GLM Results.** **First Row:** A selected-response (SR) neuron. **Second Row:** Normalized averages of SR neurons. **Third Row:** A potential-response (PR) neuron. **Bottom Row:** Normalized averages of PR neurons. **(a-c)** Peristimulus time histograms (PSTHs) for PMd neurons, aligned to target onset. Shaded areas represent SEMs. **(a)** PSTHs of reaches near the preferred direction (PD; black) versus opposite the PD (brown). **(b)** PSTHs of reaches near the PD, with a starting hand position near the PD (lower probability of moving near the PD; blue) versus a position opposite the PD (higher probability of moving near the PD; red). **(c)** PSTHs of reaches opposite the PD, with a starting hand position near the PD (blue) versus a position opposite the PD (red). **(d)** We utilized a generalized linear model (GLM) to control for confounds in the PSTHs, including different distributions of starting positions, upcoming movements, and previous movements. Here, we show the importance of parameters in the GLM, across time, for PMd neurons. We show mean relative pseudo- R^2 over time, of the upcoming movement (green) and hand position (purple) covariates. For the 2nd and bottom row, shaded areas represent SEMs across neurons. For individual neurons, shaded areas represent the standard deviation across bootstraps.

will have greater activity prior to target onset than it otherwise would (Fig. 3). This could be because the neural population activity is related to the statistical distribution of possible upcoming movements, conditioned on the current state. Alternatively, the neural population could be using some type of heuristic to determine the likely location of the next target (e.g., assuming the next target will always be toward the center).

To understand how the neural population activity relates to the conditional probability distributions of possible movements, we calculated the average activity (across PR neurons and reaches) as a function of the current angular hand position relative to each neuron's PD (Fig. 4a, Supplementary Fig. 5 for individual monkeys). Neural activity during the 100 ms prior to target appearance closely reflected the statistics of possible target locations (Fig. 4b). The peak angle of the neural activity was not significantly different than 150° , the most likely ϕ determined by the experimental design. This finding was not simply due to the correlation between the previous and upcoming movements or the correlation between the hand positions and upcoming movements (Supplementary Fig. 5). Moreover, when only looking at reaches starting near the center, activity prior to target-onset was clearly diminished (Fig. 4c), reflecting the lower and more uniform probabilities of upcoming reaches (Fig. 4d). Thus, when averaging across reaches and neurons, the population does represent the distribution of upcoming reaches.

How do neurons function together to create this distribution of upcoming reaches? It is possible that individual neurons reflect this distribution, and thus the population does as well. In this scenario, the firing rate of each neuron as a function of position would correspond to the probability of movement into its PD. Alternatively, the distribution

could be created only by many neurons working in concert. In this scenario, not all individual neurons' activities would correspond to the probabilities of upcoming movements into their PDs, but activity across the population would represent upcoming movement probabilities.

To differentiate between these possibilities, we analyzed how neurons' activities as a function of position related to the neurons' PDs. When we look at neurons with an upward preferred direction, we see that many individual PR neurons do not have maximal firing rates at hand positions corresponding to a maximum probability of moving upwards (Fig. 4e,f). Rather, these neurons have different "preferred positions", spanning many different areas at which upward movements are possible. However, when the activity of all these neurons is summed, the activity as a function of position closely matches ($r = .94$) the probability of an upward movement as a function of position. This suggests the movement probabilities are represented across the population rather than by individual neurons. Conversely, we can look at reach PDs relative to preferred angular position (the angular hand position leading to peak activity). When we orient the preferred position to be down, we see that there are a wide range of reach PDs, reflecting possible upcoming movements from an initial downward position (Fig. 4g). The PDs of neurons are distributed approximately in proportion to how likely upcoming movement directions are (Fig. 4h). The population of PR neurons in PMd works together to represent the probability distribution of available upcoming movements given the current hand position.

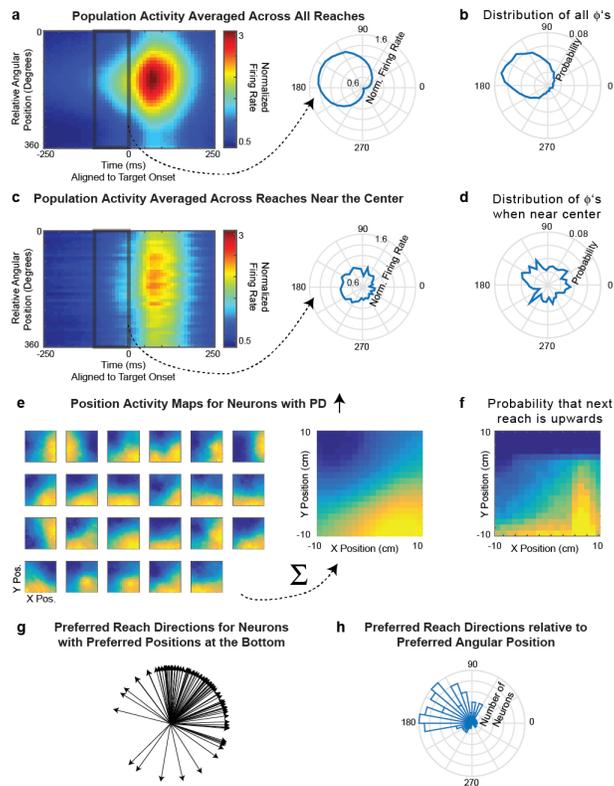


Figure 5.4: **PMd population jointly represents the distribution of upcoming movements.** **a) Left:** The average normalized firing rate of all PR neurons, over time, as a function of relative angular hand position. For each neuron, the relative angular position is the preferred direction of the neuron minus the angular hand position. **Right:** The average normalized firing rate in the 100 ms prior to target onset, as a function of the relative angular hand position. **b)** The distribution of upcoming movement directions relative to position angles (duplicated from Fig. 1f). **c,d)** Same as panels a and b, but for only for reaches starting near (within 2 cm of) the center. **e) Left:** Position activity maps for example PR neurons with preferred movement directions oriented upwards. Position activity maps show the neurons' activity as a function of hand position (blue is low; yellow is high) from -100 to 50 ms surrounding target onset. **Right:** The sum of position maps for all PR neurons, when their preferred directions are oriented upwards. **f)** A map showing the probability that the next movement will be upwards, as a function of initial hand position. **g)** Preferred reach directions for all PR neurons, when space is rotated so that their preferred hand position angle is oriented to be at the bottom (270°). **h)** A histogram of preferred reach directions of all PR neurons relative to their preferred angular hand position (the reach PD minus the preferred angular hand position).

PMd population activity represents the distribution of upcoming movements on single reaches

While we have shown that the PMd population represents conditional probability distributions averaged across trials, we also want to know what is occurring prior to single reaches. Because we recorded many neurons simultaneously, we can decode the monkey's intended movement prior to each reach. To do this, we first trained a naive Bayes decoder to predict the reach direction (see *Methods*) during the time period 50-200 ms after target presentation. We then used this decoder (with firing rate rescaling due to differing firing rates before and after target presentation; see *Methods*) to estimate what movement the neural population was planning in the 100 ms prior to target presentation. Note that this method assumes that the PDs of neurons stayed the same between these two time periods. While PMd neurons are known to have different PDs during preparation and movement[134, 135, 136], both our time periods were during preparation. Thus, we believe it is reasonable to use knowledge about neurons' PDs after target onset to decode planning prior to target onset.

As expected, the planned reaches decoded prior to target onset were usually approximately opposite of the current angular hand position (Fig. 5a,b). This can be seen in example trials (Fig. 5a), where the pre-target decoded reach direction was to the left when the hand position was on the right, while the post-target decoded reach direction was toward the target. Moreover, the distribution (across reaches) of pre-target decoded reach directions relative to the angular hand position approximately represented the experimentally-defined distribution of target presentations determined by the current hand position (Fig. 5b,c; compare to Fig. 1f,g).

There are two explanations for PMd’s apparent representation of the distribution of potential upcoming movement directions (Fig. 4a, Fig. 5b). One hypothesis is that PMd consistently preplans a specific reach prior to target presentation. That is, the PMd population does not actually represent a distribution of reaches on single trials, but averaging across trials yields the observed distribution. The alternative hypothesis is that PMd represents a distribution of possible movements prior to single reaches.

To distinguish between these two hypotheses, we looked at how single reach decoding depended on possible upcoming movements. The output of our probabilistic decoding method is a probability distribution reflecting the animal’s movement intention encoded by the population for single reaches. Uncertainty in the population about the upcoming movement will make this distribution wider. We used the width of the decoded distribution to distinguish between the two hypotheses. If a single movement was being preplanned every reach, then the width of the decoded distributions should be approximately the same for every reach (Fig. 5d; Hypothesis 1). However, if PMd represents a distribution of possible movements, when there are fewer possibilities for upcoming target locations (i.e., when the hand position is farther from the center), the distribution should be narrower (Fig. 5d; Hypothesis 2). The data show that the distributions are narrower prior to target onset when the hand position is farther from the center (Fig. 5e, f, g; Monkey M, $p < 1e - 10$; Monkey T, $p = 1.8e - 6$). Thus, the decoded distributions suggest that PMd does represent a distribution of possible movements prior to single reaches.

Our decoding results provide insight into how neural activity prior to target onset influences individual upcoming reaches. The decoded pre-target reach directions were predictive of the monkeys’ subsequent behavior. First, when the decoded direction was closer

to the true target direction, reach latencies were shorter (Fig. 5h; Monkey M, $p < 1e - 10$; Monkey T, $p = 2.1e - 5$). Second, the initial direction of many reaches was initially biased toward the pre-target decoded reach direction (Fig. 5i; both monkeys $p < 0.05$ using bootstrapping). Additionally, the uncertainty of the decoded distributions influenced upcoming reaches. For targets in an expected upcoming direction, latencies were shorter when the width of the decoded distribution was narrower (Monkey M, $p < 1e - 10$; Monkey T, $p = 3.3e - 4$). These decoding results could provide a neural explanation for our observed latency and trajectory behavioral effects (Fig. 2). Overall, our decoding results provide insight into the expectations represented by the neural population prior to target onset.

Control: Visuomotor Rotation

In our task, the probabilities of upcoming target locations were determined by the current hand position. Thus, it is theoretically possible that the neural activity could only be modulated by position for some purpose other than representing upcoming movements[**137**, **138**]. As a control, we used data where the monkeys performed a visuomotor rotation (VR) learning task, which changed the probabilities of upcoming movements for the same cursor position in the workspace. In this task, cursor feedback on the screen was rotated by 30° counterclockwise relative to the hand movement. That is, if a target were directly upwards on the screen, the monkey would now need to reach up and right to get there (Fig. 6a). The monkeys performed a block of random-target reaches with normal feedback (baseline), followed by a block of movements with the VR. In this task, the statistics of reach directions relative to the hand position on the screen rotate by 30° (Fig. 6b).

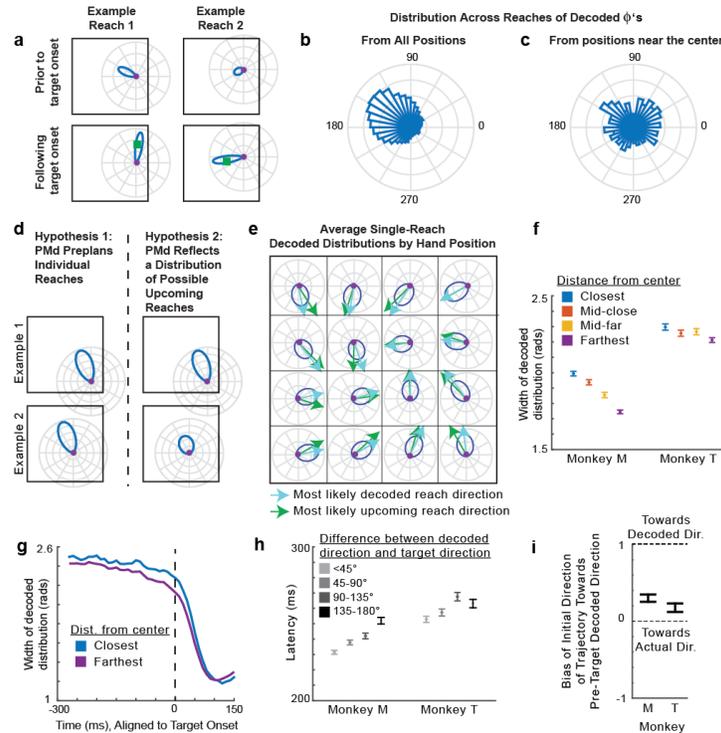


Figure 5.5: PMd population activity represents the distribution of upcoming movements- single reach decoding. (a) The distribution of decoded reach directions (blue) from the population of PMd neurons for two example reaches (left and right), before and after target onset (top and bottom). The purple circle is the current hand position, and the green square is the target location. (b) The distribution across all reaches of pre-target decoded reach directions relative to the hand's angular position (this is ϕ from Fig. 1, except with decoded reach directions instead of actual reach directions). (c) The distribution across reaches starting within 2 cm of the center, of pre-target decoded reach directions relative to the hand's angular position. (d) The predictions of two hypotheses (left and right), shown for two different hand positions (example 1 vs. example 2). (e) Average pre-target decoded reach direction distributions as a function of hand position. These distributions are constructed to have the average width and peak angle of all decoded distributions from the hand positions within the grid square. (f) The full width at half maximum of the pre-target decoded distribution as a function of hand distance from the center of the workspace. Distances from the center are binned as in Fig. 2. Error bars represent SEMs. (g) The width of the decoded distributions over time, for starting positions that are the closest (blue) and farthest (purple) from the center. (h) The latency of the reach as a function of the angular difference between the pre-target decoded direction and the actual target direction. Error bars represent SEMs. (i) The bias of the initial trajectory of the reach (100-150ms from target onset) toward the pre-target decoded direction. 95% confidence intervals, computed via bootstrapping, are shown.

Thus, for the same cursor position in the workspace, the monkeys should plan movements 30° more clockwise than in the baseline condition. In fact, the monkeys did mostly learn to adapt their expectations of the upcoming movement directions, judging by their initial reach directions (Fig. 6c). We can therefore determine whether PMd activity changed to reflect the new movement probabilities at the same workspace positions.

Did PMd activity prior to target onset change to reflect the modified expected movement directions? Looking at individual neurons, the majority of PR neurons' preferred positions shifted counterclockwise in the VR task (Fig. 6d,e). This was expected, as the hand position most likely to result in a movement into a neuron's PD will rotate counterclockwise in the VR task. There was a wide range of changes across neurons, again demonstrating that distributions are represented across the population.

To look at population level changes, we decoded reach direction using activity prior to target onset, and compared the distributions of these decoded reaches in the baseline and rotation periods of the task (Fig. 6f,g). As the rotation was small, we looked only at hand positions not near the center, where monkeys had more information about the upcoming movement directions. In every session for both monkeys, the reach directions shifted counterclockwise relative to the workspace position. On average, there was a small, but significant, clockwise shift in decoding of about 10° for each monkey (Fig. 6d,e; $p < 0.05$ for both monkeys; for all decoded reaches see Supplementary Fig. 6a,b). Additionally, when we analyzed the data by looking at the average activity of PR neurons prior to target onset (as in Fig. 4a) rather than decoding, we found a similar shift in activity in every experimental session (Supplementary Fig. 6c,d). Thus, PMd activity prior to target onset (see Supplementary Fig. 7 for effects after target onset) is modulated by

the probabilities of upcoming movements, not simply by hand position for some other purpose.

M1 does not represent the conditional probability distribution of upcoming movements

To determine whether primary motor cortex (M1) also represents conditional probability distributions of upcoming movements, we ran the same set of analyses for M1 as we did for PMd. When we did a GLM analysis, we found that 28% (176/618) of M1 neurons met the criteria of PR neurons, meaning they had significant modulation with hand position prior to target onset and movement after target onset. However, these neurons did not respond to hand position in the same manner as the PR neurons in PMd. These M1 neurons had increased activity prior to target onset when the hand position was in the same direction as the neurons' PDs, rather than in the opposite direction of the neurons' PDs (Fig. 7, Supplementary Fig. 8 for individual monkeys). This activity could be explained by the end of the previous movement, since previous movements into the PD (which correspond to higher M1 activity), often result in angular hand positions in the same direction as the PD (Supplemental Fig. 9). Moreover, the effect of the hand position didn't ramp up as it did in PMd; rather, it appears to be a decreasing effect from the previous movement (Fig. 7, Supplementary Fig. 8). Thus, while M1 activity varies according to position (as in [139, 140]), it likely does so in a way that reflects movement execution, rather than information about the upcoming movement.

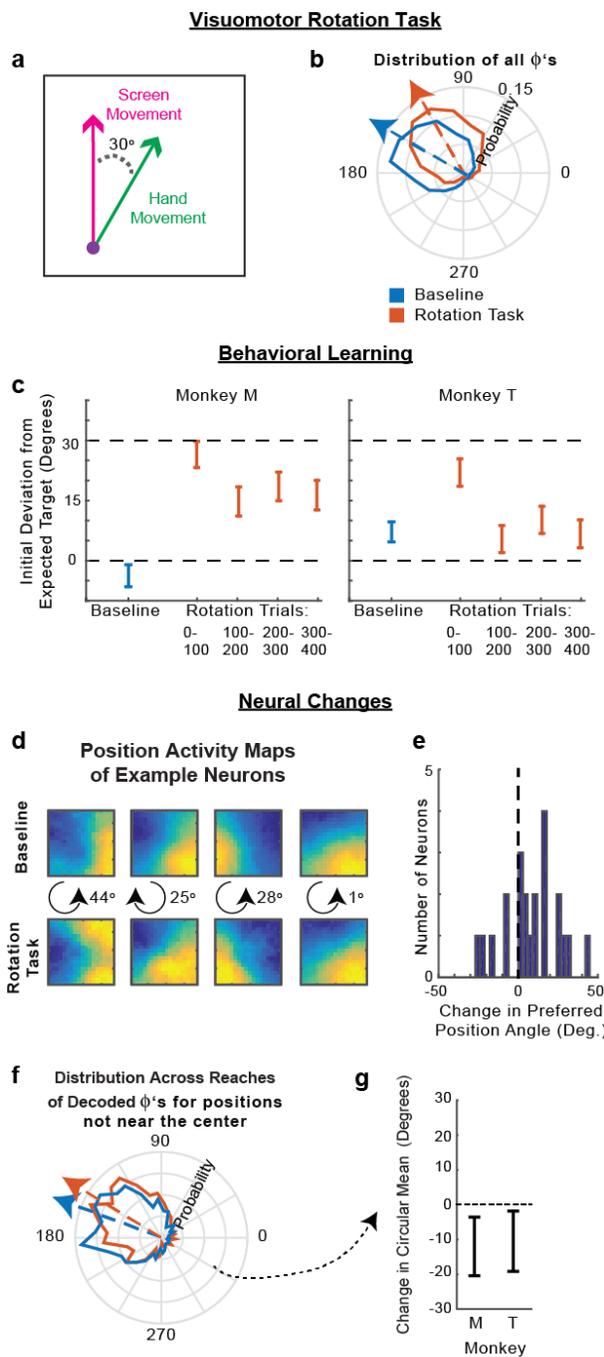


Figure 5.6: Visuomotor Rotation Control Task. (a) The visuomotor rotation (VR) task. Movements on the screen (in the workspace) are rotated 30° counterclockwise relative to the hand movement. (b) The distribution of hand movements relative to the position angle in the workspace, i.e. ϕ 's, for the baseline (blue) and VR (orange) tasks. Arrows point toward the distributions' circular means. (c) The difference between the initial reach direction (100-150ms from target onset) and the expected reach direction during the baseline task (blue) and different periods during the VR task (orange). The expected reach direction was the most likely upcoming reach direction given the current workspace position and movement statistics (panel b). Positive values mean the initial reach direction was counterclockwise of the expected target direction, meaning the monkey had not adapted. (d) Position activity maps (activity as a function of position in the workspace) of example PR neurons in the baseline (top) and VR task in the second 2/3 of trials (bottom), as in Fig. 4c. In the middle, we show the direction (clockwise or counterclockwise) and magnitude of change of the preferred angular position. (e) The change in preferred angular position of all PR neurons (VR minus baseline). Positive means a counterclockwise shift. (f) The distribution of pre-target decoded reach directions relative to the hand's angular position (decoded ϕ 's) for positions not near the center (greater than the median distance). Decoding from the VR task used the second 2/3 of trials. (g) The difference between the circular mean of the distributions of decoded ϕ 's in panel f, between the baseline and VR tasks (VR minus baseline). Error bars represent 95% confidence intervals from bootstrapping.

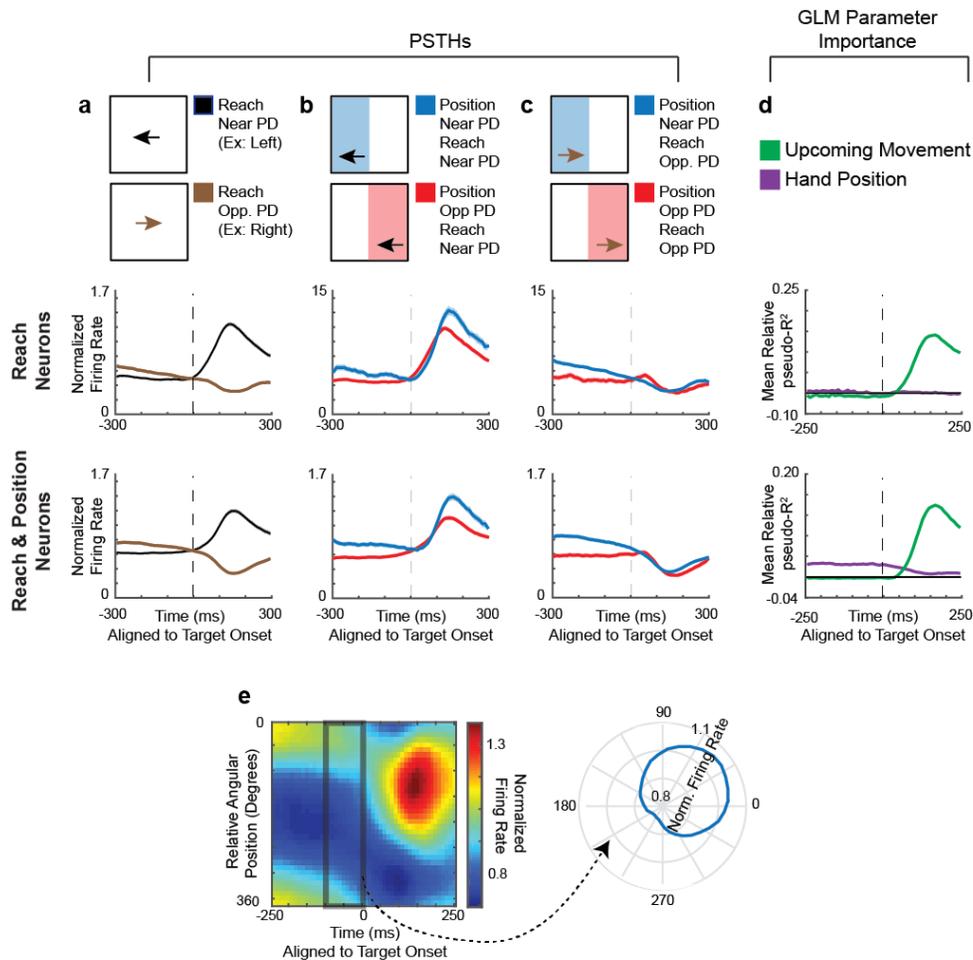


Figure 5.7: **M1 does not reflect the probability of upcoming movements.** (a-d) PSTHs and GLM results for M1 neurons. Columns have the same schematics as Fig. 3. **First Row of PSTHs:** Normalized averages of “reach” neurons, defined as those neurons significant for movement during the late period, but not position in the early period of the GLM. This was the same criteria as for SR neurons in PMd. **Second Row:** Normalized averages of “reach and position” neurons, defined as those neurons significant for movement during the late period, and position in the early period of the GLM. This was the same criteria as for PR neurons in PMd. Note that we did not use the same “SR/PR” nomenclature as PMd, because there was no evidence in the PSTHs of M1 neurons that position was used to represent potential upcoming movements. (e) Same schematic as Fig. 4a, but for M1 reach and position neurons. **Left:** The normalized average firing rate, as a function of time and relative angular position. Activity is averaged across all reach and position neurons. **Right:** The normalized average firing rate in the 100 ms prior to target onset, plotted as a function of the relative angular hand position.

Discussion

In order to plan everyday movements, we take into account the probability distributions of possible movements determined by the state of our body in the environment. Here, we have demonstrated that these conditional probability distributions influence behavior, specifically movement trajectories and latencies. A subpopulation of neurons in PMd, but not M1, function together to represent these probabilities, even prior to individual reaches. We used a visuomotor rotation task to show that the effect was not simply a position dependent component of the firing rate.

Information that shapes conditional probability distributions of upcoming movements

Probability distributions can be conditioned on many sources of information. Here, we focused on probability distributions that were determined by the body's state in the task environment. When hand position within the environment changed, the probability distributions of upcoming movements changed. Likewise, changes in the task environment caused by the visuomotor rotation altered the probability distributions. Although not the focus of our study, another source of probability distributions is biomechanics. In the extreme, biomechanics limit the possible movements. Softer constraints may arise from biomechanical costs, e.g., the ease of movement, which can affect both choice of arm movements [141, 142] and PMd activity[142]. While biomechanics may influence the representation within PMd of possible upcoming movements, the fact that PMd's probability distributions also changed when the task environment changed in the VR task, suggests our findings are not solely due to biomechanical constraints.

In order to represent the probability distributions of upcoming movements, PMd needed to have information about the hand position in the workspace. Previous studies have shown that PMd activity is modulated by hand position, either to make the neurons' PDs compatible with the orientation of the arm[137], or to encode the relative position between the hand and eye[138]. Thus, there is evidence that PMd neurons have access to information about hand position. In our task, PMd began to represent the possible upcoming movement while the current movement was ongoing and hand position was changing. Thus, if PMd was using proprioceptive information to determine the possible upcoming movements, it would have been using a changing position estimate. Alternatively, PMd could have made use of visual information about the current target, or an efference copy of the current movement command, to determine the probability distribution of the next movement. In everyday life, PMd likely uses a mixture of sensory, proprioceptive, and movement information to determine possible upcoming movements.

Representation of probability distributions in the brain

There is much debate on how probabilities are represented in the brain. Some argue for a temporal coding of uncertainty [143], while others argue that probability distributions are represented across populations of neurons[144, 145, 146] (e.g., probabilistic population coding[145]). Several previous studies have proposed models of movement in which distributed neural populations represent a probability density function across movement directions[147, 148, 149, 150, 151]. We showed that neurons were more strongly active in workspace locations from which movements into their preferred direction were more likely. Only when looking across neurons did the activity as a function of location closely

match the task's movement probabilities. Also, when looking at neurons that were active at a nearby location, the distribution of PDs was proportional to the probability of upcoming movement directions from that location. Our results are consistent with the interpretation that coding of probability distributions across populations of neurons plays a central role in the movement decision process.

Still, our data is consistent with several possible interpretations for how the brain represents probabilities prior to movement. The fact that the decoded distributions are narrower when there are fewer possible upcoming movements could mean that PMd represents a continuous probability distribution at any given time. Alternatively, the brain could be discretely sampling multiple possibilities from the probability distribution before a reach. For example, when there are few movement possibilities, the monkey could be simultaneously pre-planning two movements on average, and when there are many movement possibilities, the monkey might typically be pre-planning three movements. Another alternative interpretation is that the monkeys are rapidly sampling (and pre-planning) individual reaches from the probability distribution at a rate much faster than 50 ms (the bin size used to calculate the distribution width in Fig. 5g). Future experiments with many more recorded neurons, resulting in more precise decoding, could help resolve these questions of how the probability distribution is represented.

Dissociation between visual and motor responses in the visuomotor rotation task

When analyzing the visuomotor rotation task (Fig. 6), we found changes in PMd activity that corresponded to changes in the probability distribution of the upcoming movements,

even though the visual distribution of targets did not change. On first glance, this would appear to contradict previous studies, which have suggested that PMd tracks the visual spatial parameters more than the actual movement direction[152, 153, 154]. However, we were analyzing PMd responses prior to target onset, while other studies have looked at PMd activity following target onset. In fact, when we analyzed PMd results after the target was displayed (Fig. S5), we found that the visuomotor rotation did not change PMd's representation of the target, consistent with previous findings. It is only during the time prior to the target being visually displayed, that the activity of PR neurons in PMd changes to reflect the changing probabilities of the upcoming movements themselves.

PMd's role in representing movement possibilities

Our study builds on much research about the role of PMd in planning upcoming movements [20, 18, 21, 155, 156, 19, 157]. Previous work has demonstrated that monkeys represent possible movements when selecting between alternatives[18, 155, 156, 19, 157] and when estimating the likely target location from visual cues[20]. These studies suggested that PMd can represent a probability distribution. Our work extends these findings by showing that PMd also represents dynamically changing probability distributions that are dependent on interactions between the body and the environment. Moreover, our work shows that PMd does not only represent probability distributions that are explicitly manipulated. Here, PMd represented probability distributions even in a standard reaching task that has been used in a variety of motor studies[158, 159], where probabilities are usually considered to be irrelevant. The representation of conditional probability distributions of possible movements in PMd appears to be ubiquitous.

Methods

Behavioral Paradigms

Random-target Experiment

Three monkeys (Monkeys M, T, and C) performed a random-target reaching task (similar to the experiments in [158, 159]) in which they controlled a computer cursor using arm movements (Fig. 1). Monkeys were seated in a primate chair while they operated a two-link planar manipulandum. Arm movements were constrained to a horizontal plane within a workspace of 20 cm x 20 cm. On each trial, the monkey consecutively reached to 4 targets (2 cm x 2 cm squares), with each new target appearing once the monkey reached the previous target. More precisely, once a target was reached, a new target was triggered 100 ms later, as long as the cursor remained on the target. The target appeared on-screen 96 ms after this trigger on average, due to delays from graphics processing and the monitor refresh rate. In accordance, the monkey was required to keep the cursor on the target for an additional 100 ms after a new target was triggered. Thus, in total there was a 200 ms hold period after landing on the target. This brief hold period forced the monkeys to decelerate as they approached the target, but was not so long that the monkeys completely stopped on the target. After a successful trial (4 successful reaches), the monkey received a liquid reward. The next trial started after a delay of one second with a new random target presentation. Target locations were chosen to be 5 - 15 cm from the current target. Specifically, they were chosen as follows. 1) Randomly choose a distance between 5 and 15 cm, and an angle between 0° and 360° for the new target (relative to the current target). 2) If the new target falls outside of the workspace, add

90° to the angle and set the distance to be 5 cm. 3) Repeat step 2 until the target is in the workspace.

Many of the analyses are aligned to target onset. These experiments did not use a photodiode to determine the exact moment the target was displayed. Thus, in all analyses, the target onset time we used was the time the computer sent the target command plus the average delay time (96 ms).

In total, we recorded 8 sessions for monkey M, 6 sessions for monkey T, and 5 sessions for monkey C.

Visuomotor Rotation Experiment

Monkeys M and T each performed three sessions in which a visuomotor rotation (VR) task followed the baseline random-target task. The VR task was equivalent to the random-target task, with the exception that the movement vectors displayed on the screen (workspace) were rotated 30° counterclockwise relative to the hand movement vectors (as in Fig. 6a).

Neural Data Acquisition and Preprocessing

Monkeys M and T were implanted with 100-electrode Utah arrays (Blackrock Microsystems, Salt Lake City, UT) in dorsal premotor cortex (PMd). Monkeys M and C were implanted with Utah arrays in primary motor cortex (M1). See [20] for the location of the arrays in Monkeys M and T. Units were manually sorted with Offline Sorter (Plexon, Inc, Dallas, TX, USA). Only well-isolated individual units were included. Since we used chronically implanted arrays, it is likely that some neurons were recorded on multiple sessions and thus were not unique. We only included neurons with firing rates of at least

2 spikes / second in either the early or late period. The early/late periods were defined as -100 to 50 and 50 to 200 ms from target onset, respectively. In PMd, this left us with 520 neurons from Monkey M and 250 neurons from monkey T. In M1, this left us with 352 neurons from Monkey M and 266 neurons from Monkey C.

Behavioral Analysis

Each trial consisted of 4 reaches. We did not include the first reach in any of our analyses, as this was preceded by a reward period without movement (rather than being in the midst of a continuous movement). Reaches were also excluded if the monkey did not hold on the previous target for 200 ms, or if it took greater than 1.4 seconds to reach the target. These “error” reaches were rare, and occurred 2.3%, 4.9%, and 0.8% of the time in monkeys M, T, and C, respectively. Behavioral data was combined across all sessions for each monkey.

Statistics of target presentation

We defined the angular position, ϕ_P , as the hand position (prior to movement) relative to the center of the workspace (Fig. 1). We defined ϕ as the angular difference between the upcoming movement direction (also the direction to the target), ϕ_T , and the angular position. That is, $\phi = \phi_T - \phi_P$.

Trajectory bias

We calculated whether the movement trajectory within a given time interval was biased toward the expected target direction, ϕ_E . The expected direction was the most likely direction of the next target given the current hand position, based on the distribution of ϕ 's. So if ϕ^* is the value corresponding to the circular mean of the distribution of ϕ 's,

$\phi_E = \phi_P + \phi^*$. The bias of the movement trajectory within a given time interval was defined as follows. First, a movement direction, ϕ_M , was determined within that time interval based on the start and end hand position in that time interval. We calculated the bias, $B = \frac{\phi_M - \phi_T}{\phi_E - \phi_T}$, where the numerator and denominator were made to be in the interval of $[-180^\circ \ 180^\circ]$ prior to dividing. When the current movement direction is toward the expected direction, B will be near 1, and when the movement direction is toward the actual target direction, B will be near 0. B can also be negative when the movement direction is away from the expected direction. For the summary statistics of B , we used the median and standard error of the median, as B has outliers when dealing with circular variables. To calculate the standard error of the median, we used bootstrapping. Note that in Supplementary Fig. 1, we also calculated the bias toward the center, rather than the expected direction. This has the exact formulation as above, except $\phi^* = 180^\circ$.

Latency effects

The latency of a reach was defined as the time from target onset until the movement surpassed a velocity of 8 cm / sec. Latencies greater than 6 standard deviations from the mean were excluded as outliers.

We computed the mean latency of movements as a function of the expectedness of the target location, which was defined as the difference between the target direction and expected direction: $|\phi_T - \phi_E|$. We calculated the Pearson's correlation between latency and the expectedness of the movement, and determined significance based on the p-value associated with the correlation (2-sided one-sample t-test).

We also analyzed differences in latencies between expected reaches (expected direction $< 60^\circ$ from target direction) and unexpected reaches (expected direction $> 120^\circ$

from target direction), based on the distance of the hand position from the center of the workspace. To test whether the latency of expected reaches decreased as a function of distance from the center more than unexpected reaches, we used linear regression to fit the latency of reaches as a function of distance from the center. We then did a 2-sided unpaired t-test with unequal sample variances to analyze whether the slope was less (more negative) for expected reaches.

Neural Data Analysis

As with the behavioral analyses, we only included successful reaches, and did not include the first reach of each trial.

Smoothed maps of neural activity

For many aspects of the following neural data analysis, we computed smoothed maps of neural activity in relation to some variable (hand position, previous movement, or the upcoming movement). For instance, we created a map of how neural activity varied over all hand positions in the workspace, and a map of how neural activity varied in response to all upcoming movement vectors. For our maps, rather than assuming a parametric form, we non-parametrically estimated the average firing rate at each point in space using weighted k-nearest neighbor smoothing. The parameters were the number of nearest neighbors, k , and a decay parameter, d . As an example, for the movement variable (previous or upcoming), for each movement we found the k nearest movement vectors (based on Euclidean distance). We then averaged the firing rates associated with each of the k movements, but with each weighted proportional to its distance from the given movement vector to the d power.

The parameters we used for the generalized linear models (GLMs) were $k = 20\%$ of the data points, $d = 0$. The parameters were found using cross-validation on a held out data set, in order to not inflate the number of significant neurons in the GLM analysis. The parameters we used at other times (including in plots) were $k = 30\%$ of the data points, $d = -1$. These parameters were found using cross-validation on the current data sets in order to create as accurate maps as possible. Importantly, all results were robust to a wide range of smoothing parameters.

For visualizing the position maps in Fig. 4c, they were rotated either 90° , 180° , or 270° so that the PD of that neuron (after the same rotation) was always upward (between 45° and 135° relative to horizontal), which facilitated the interpretation and comparison of the results.

To get the estimated firing rate due to a single variable (e.g. position), we could use these smoothed maps. For any position we can get the associated estimated firing rate, θ_P , by looking up the firing rate for that position on the smoothed map. If, for instance, we want to get the estimated firing rates due to position in a time interval prior to every reach, we would get a vector θ_P , which contains the estimate prior to each reach. The same can be done to estimate the firing rate due to the upcoming movement, θ_{UM} , or previous movements, θ_{PM} .

Determining PDs of Neurons

We determined the preferred movement direction (PD) of each neuron from 50 to 200 ms following target onset. Let \mathbf{Y} be the vector of firing rates in that interval for every movement. It is possible that some of the neural activity during these time periods was related to the hand position, rather than the upcoming movement. Thus, we first aimed

to remove the effect of any position-related signal that might bias the calculated PD. Let θ_P be the vector of the estimated firing rates due to hand position in the same time interval (see Smoothed Map section above for how we estimate θ_P). We fit the tuning curves to $Y - \theta_P$, i.e., we subtracted out the position-related signal to get a “firing rate due to movement.” More specifically, we fit a von Mises function to relate the movement directions to this “firing rate due to movement”:

$$Y - \theta_P = \alpha \exp(\beta \cos(\phi_M - \phi_M^*))$$

where ϕ_M is the vector of movement directions, and α , β , and ϕ_M^* are the parameters that we fit. ϕ_M^* is the PD of the neuron.

Preferred angular positions

We determined the preferred angular (hand) position of each neuron from -100 to 50 ms following target onset. We first aimed to remove the effect of any movement-related signal that might bias the calculated preferred angular position. To do so, we subtracted the movement-related activity from the total activity (as in the previous section). We then fit a von Mises function to relate the angular positions to this residual. This fitting is identical to the fitting of PDs, except replacing movement directions with angular hand positions.

PSTHs

When plotting PSTHs of individual neurons, we plotted the mean firing rate across movements. The error bars on PSTHs are the standard error of the mean (SEM) across movements. When plotting the PSTHs averaged across neurons, we first normalized the mean

firing rate (across time) for each neuron by dividing by the maximum firing rate of the average trace (across all conditions). We then plotted the average of these normalized firing rates across neurons. Error bars are the SEM across neurons. All traces were smoothed using a 50 ms sliding window.

PSTHs were made for different categories of movements. For the PSTHs, movements near the PD were those that were within 60° of the PD. Movements opposite the PD were those greater than 120° from the PD. Hand positions opposite the PD were angular positions greater than 120° away from the PD. Hand positions near the PD were angular positions less than 60° away from the PD.

Generalized Linear Model

To determine which variables were reflected in the neural activity, we used a Poisson Generalized Linear model (GLM). Let \mathbf{Y} be a vector containing the number of spikes in the time interval we are considering, for every movement. It has size $m \times 1$, where m is the number of movements. We aimed to predict \mathbf{Y} based on several factors. We used the hand position, the previous movement vector, the upcoming movement vector, the peak velocity of the upcoming movement, and a baseline term. More specifically, the covariate matrix \mathbf{X} was:

$$\mathbf{X} = \begin{bmatrix} | & | & | & | & | \\ \mathbf{1} & \boldsymbol{\theta}_P & \boldsymbol{\theta}_{UM} & \boldsymbol{\theta}_{PM} & \mathbf{v}_{\max} \\ | & | & | & | & | \end{bmatrix}$$

where \mathbf{v}_{\max} is the vector of peak velocities of movements, and $\boldsymbol{\theta}_P$, $\boldsymbol{\theta}_{UM}$, and, $\boldsymbol{\theta}_{PM}$ are generated from the smoothed maps (see *Smoothed Maps* above). Essentially, these

covariates are the expected firing rates from position, upcoming movement, and previous movement (respectively) by themselves. Note that the previous and upcoming movement covariates were fit separately and do not need to have the same smoothed map (as PDs can be different during planning and movement [134, 135, 136]). Also note that when we run GLMs during different time intervals, we make separate smoothed maps for these time intervals.

Overall, the model that generates the firing rate (λ ; also known as the conditional intensity function) can be written as:

$$\lambda = \exp(\mathbf{X}\boldsymbol{\beta})$$

where $\boldsymbol{\beta}$ is a vector of the weights for each covariate that we fit, and \mathbf{X} is the matrix of covariates, which is z-scored before fitting. If there are j covariates, then $\boldsymbol{\beta}$ has size $j \times 1$. \mathbf{X} has size $m \times j$. Note the use of an exponential nonlinearity to ensure that firing rates are positive. The model assumes that the number of spikes, \mathbf{Y} , is generated from the firing rate, λ , according to a Poisson distribution.

We fit the model weights to the data using maximum likelihood estimation. That is, we found $\boldsymbol{\beta}$ that was most likely to produce the true spike output (assuming spikes were generated from the firing rate in a Poisson nature). Critically, we used (5-fold) cross-validation, meaning that the model was fit to the data using one set of data (the training set), and model fits were tested with an independent set of data (the testing set). Similarly, when calculating the test set covariates for movement and position (described in *Smoothed maps of neural activity*), we only used k-nearest neighbors from the training set, to avoid overfitting.

To test whether an individual covariate significantly influenced neural activity, we first made sure that a simplified model with only that individual covariate had significant predictive power. To determine the value of a model fit, we used pseudo- R^2 [132, 33], a generalization of R^2 for non-Gaussian variables. The pseudo- R^2 of a model is defined as:

$$R_D^2(\text{model}) = 1 - \frac{\log L(n) - \log L(\hat{\lambda})}{\log L(n) - \log L(\bar{n})}$$

where $\log L(n)$ is the log likelihood of the saturated model (i.e., one that perfectly predicts the number of spikes), $\log L(\hat{\lambda})$ is the log likelihood of the model being evaluated, and $\log L(\bar{n})$ is the log likelihood of a model that uses only the average firing rate.

Then, in order to determine the importance of that covariate to the full model, we test whether the full model predicts neural activity significantly better than a model where that covariate is left out (reduced model). To compare the fits between the reduced model (model 1) and full model (model 2), we used relative pseudo- R^2 , which is defined as:

$$R_D^2(\text{model 1, model 2}) = 1 - \frac{\log L(n) - \log L(\hat{\lambda}_2)}{\log L(n) - \log L(\hat{\lambda}_1)}$$

where $\log L(\hat{\lambda}_2)$ is the log likelihood of the full model and $\log L(\hat{\lambda}_1)$ is the log likelihood of the reduced model.

To determine significance, we bootstrapped the fits to create 95% confidence intervals, and checked whether the lower bounds of these confidence intervals were greater than 0. Note that the pseudo- R^2 and relative pseudo- R^2 values can be less than 0 due to overfitting.

Neuron types

In PMd, we defined selected-response (SR) neurons as those that were significantly modulated by upcoming movement in the late period in the GLM, but were not significantly modulated by hand position in the early period. Potential-response (PR) neurons were significantly modulated by upcoming movement in the late period and by hand position in the early period. Using a more relaxed criterion for defining neurons (as described in *Decoding* below) greatly increases the number of PR neurons. While our criteria for determining SR and PR neurons was different from [18], we used the same terminology due to the same perceived function. In the VR task, PR neurons were those that were significant during both the baseline and VR periods.

Population activity over time averaged across trials

For each neuron, we calculated the firing rate as a function of the relative angular position (Fig. 4). We defined the relative angular position as the difference between a neuron's PD and the hand's angular position (the PD minus the angular hand position). In the VR task (Supplementary Fig. 6c), the relative angular position was calculated relative to the neuron's PD in the baseline task. We then normalized each neuron by dividing by its mean firing rate, and then averaged the normalized activity across neurons. We then smoothed the activity for plotting using the parameters from the smoothed maps.

We also made several variants of the above plot (Supplementary Fig. 5). We made a plot where movements were only used if the angle between previous and upcoming movements was less than 90° . We made a plot in which the movements were resampled,

so that the distribution of ϕ 's was centered at 180° , rather than being off center. More specifically, we resampled from a von Mises distribution: $g(\phi) \propto \exp(\cos(\phi - 180^\circ))$.

To determine the relative angular position resulting in peak activity (in the 100 ms before target onset), we calculated the activity at 20 relative angular positions (evenly spaced from 0 to 360°), and calculated the circular means of the angles weighted by their activities. We determined whether the activity prior to target onset was related to the distribution of upcoming movements by testing whether the relative angular position resulting in peak activity was significantly different from 150° (the circular mean of the distribution of ϕ 's). We created a 95% confidence interval of peak relative angular positions by bootstrapping over the set of neurons, and checked whether this overlapped with 150° .

Decoding

We aimed to determine the movement intention of the neural population in the 100 ms prior to target onset. As only including PR neurons would give us a small number of neurons per session for decoding, we expanded our criteria. While PR neurons were significant for hand position and upcoming movement with 95% confidence, here we included neurons that were significant at a level of 50% (the median pseudo- R^2 and relative pseudo- R^2 values were greater than 0). For comparing between the VR task and baseline, we required that neurons had positive median pseudo- R^2 values in both conditions. Additionally, as neurons changed from session to session, separate decoders were trained for each session.

We first fit tuning curves to each neuron during 50 to 200 ms after target onset, when the neurons' preparatory responses to different target directions was known. This was

done using a von Mises function, as in *Determining PDs of Neurons*. We wanted to use these tuning curves to decode during the 100 ms prior to target onset. However, as firing rates were greatly different during these two time periods, we needed to rescale the tuning curves. To do so, we fit tuning curves to the future movement in the 100 ms prior to target onset. We then modified this tuning curve by giving it the preferred direction calculated after target onset. This essentially gives us rescaled versions of the tuning curves determined when the target is known.

Note that for decoding in the VR task, we still fit the initial tuning curve using activity after target onset in the baseline task. We then decoded using activity from before target onset in the VR task, using the rescaling described above. We use this procedure to make the comparison meaningful.

Next, for each neuron, we found the likelihood of the number of spikes given all possible movement directions (in 1° increments). This was done by assuming the number of spikes during the time period is a Poisson random variable with a mean rate determined by the value of the tuning curve at the direction being tested. If r_i is the number of spikes during the interval for neuron i , s is the direction, and $f_i(s)$ is the value of the tuning curve (the expected number of spikes) for neuron i at direction s :

$$P(r_i|s) = \frac{\exp(-f_i(s)) f_i(s)^{r_i}}{r_i!}$$

We assumed that neurons' activities were conditionally independent given the direction (a naive Bayes decoder), and thus multiplied their probability distributions:

$$P(\mathbf{r}|s) \propto \prod_i P(r_i|s)$$

We can use Bayes rule to determine the likelihood of all the movement directions given the number of spikes of all neurons. Assuming a uniform prior, by Bayes rule:

$$P(s|\mathbf{r}) \propto P(\mathbf{r}|s)$$

Finally, we normalized $P(s|\mathbf{r})$ (so it was a probability distribution), and this term was the decoded distribution.

The decoded direction was the direction corresponding to the peak of the distribution (the maximum likelihood decoded direction). The width of the decoded distribution was the full width half maximum (FWHM) of the decoded distribution.

To calculate the width of the decoded distribution over time (Fig. 5g), we decoded using a 50 ms sliding window of neural activity. All methods were the same as above, just replacing the 100ms of activity prior to target onset with the given 50 ms of activity. This choice allowed us a better temporal resolution.

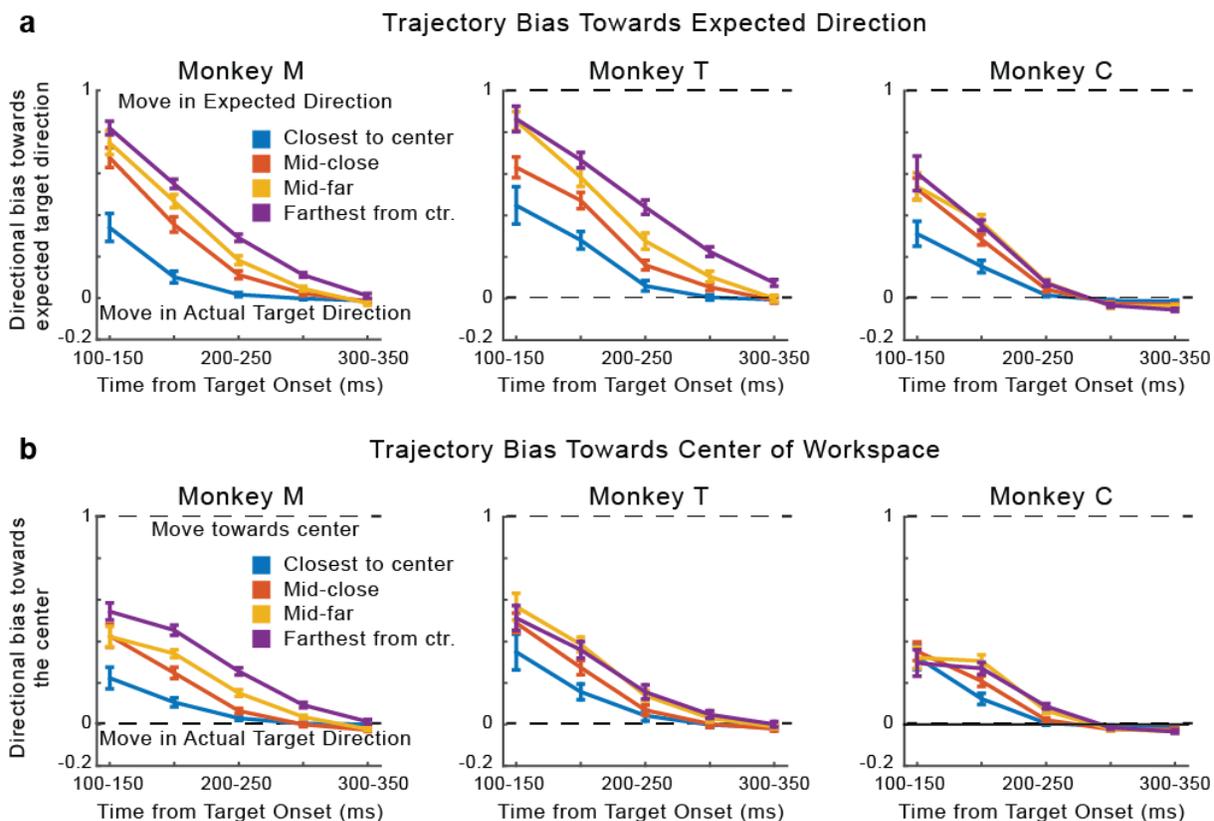
As we did decoding separately for each session, to do significance testing, we used a simple multilevel model analysis - specifically, a random intercepts model, where the baseline (intercept) can be different for every session. Thus, if there were 4 sessions, when looking at the width of the decoded distribution (w) as a function of the distance from center (d), we wrote the model as $w = \beta_1 \mathbf{I}_1 + \beta_2 \mathbf{I}_2 + \beta_3 \mathbf{I}_3 + \beta_4 \mathbf{I}_4 + \beta_5 d$, where I_1 through I_4 are indicator variables for whether the values are from a given session. We looked at whether β_5 was significantly different from 0 using a 2-sided one-sample t-test. We used

an equivalent approach to determine the significance of the relationship between latency and the difference between the decoded angle and the target angle. We also used an equivalent approach to determine the significance of the relationship between latency and the width of the distribution for expected reaches (expected reaches are those where the difference between the actual and expected target direction is less than 60°).

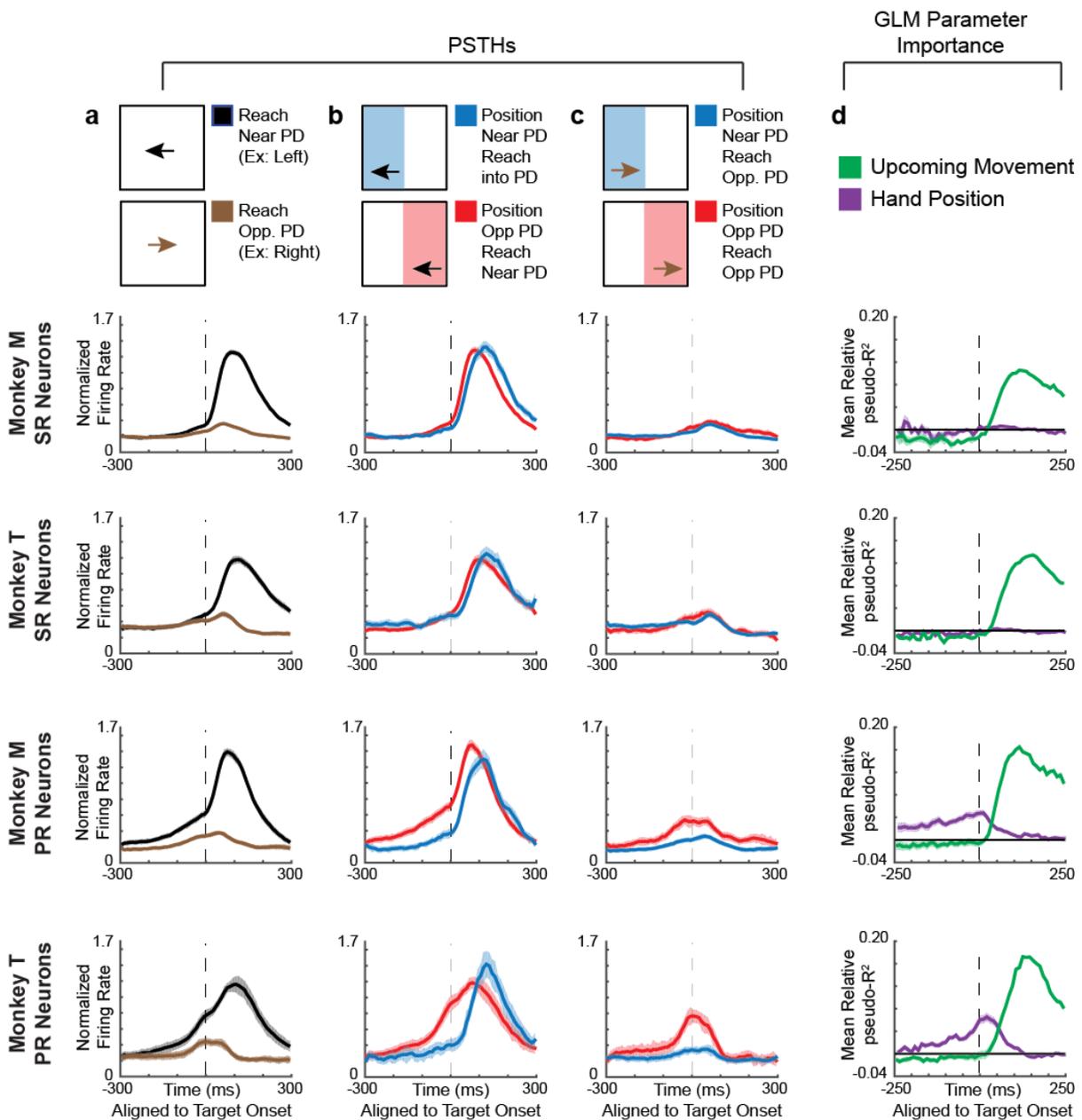
We calculated the bias of the initial trajectory toward the decoded direction, ϕ_D , equivalently to how we calculated the behavioral bias toward the expected target direction:

$$B = \frac{\phi_M - \phi_T}{\phi_D - \phi_T}.$$

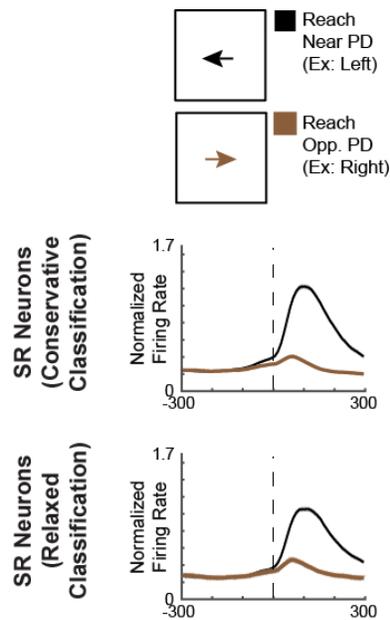
To determine whether the distribution of decoded reach directions shifted from the baseline task to the VR task, we calculated the difference between the circular means of the distributions of “decoded ϕ ’s” (the decoded reach direction relative to the angular position). For significance testing, we bootstrapped this difference in circular means. More specifically, 1000 times, we resampled decoded reaches within the baseline task and calculated the baseline circular mean, and did the same thing for the VR task. This led to 1000 differences in circular means. We looked at the 95% confidence interval of this difference.



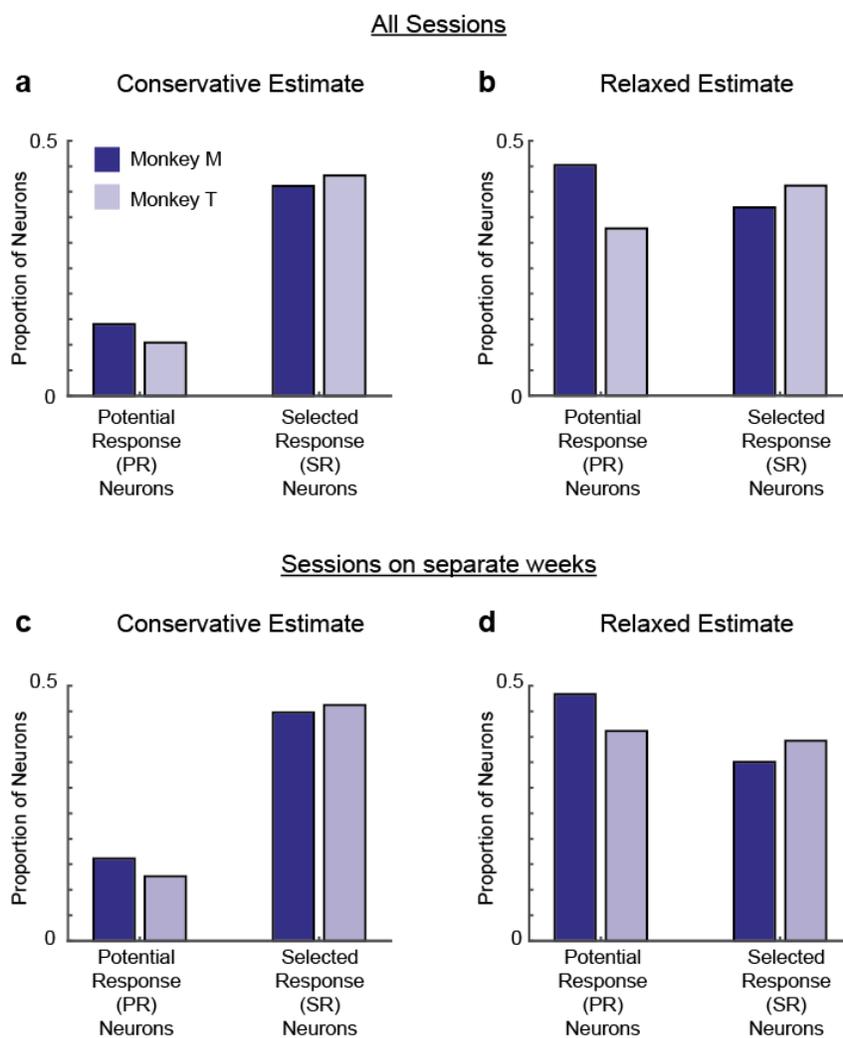
Supplementary Figure 5.1: **Trajectory Biases**. We compare (a) the biases in the reach direction towards the expected direction of the reach given the statistics of target presentation (identical to Fig. 2b) versus (b) the biases in the reach direction towards the center of the workspace.



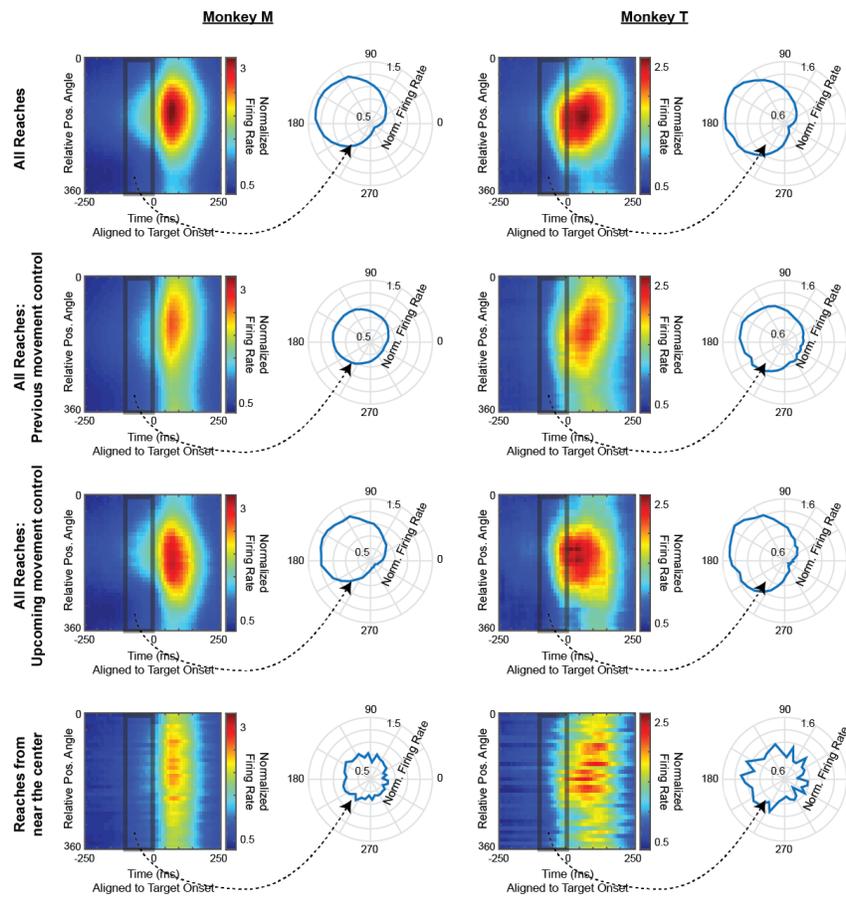
Supplementary Figure 5.2: **PMd PSTHs and GLM results, for individual monkeys.** Peristimulus time histograms (PSTHs) and GLM results for PMd neurons, for individual monkeys. All columns are the same as in Fig. 3. **First and Second PSTH Rows:** Normalized averages of selected-response (SR) neurons from Monkey M and Monkey T, respectively. **Third and Bottom Rows:** Normalized averages of potential-response (PR) neurons from Monkey M and Monkey T, respectively.



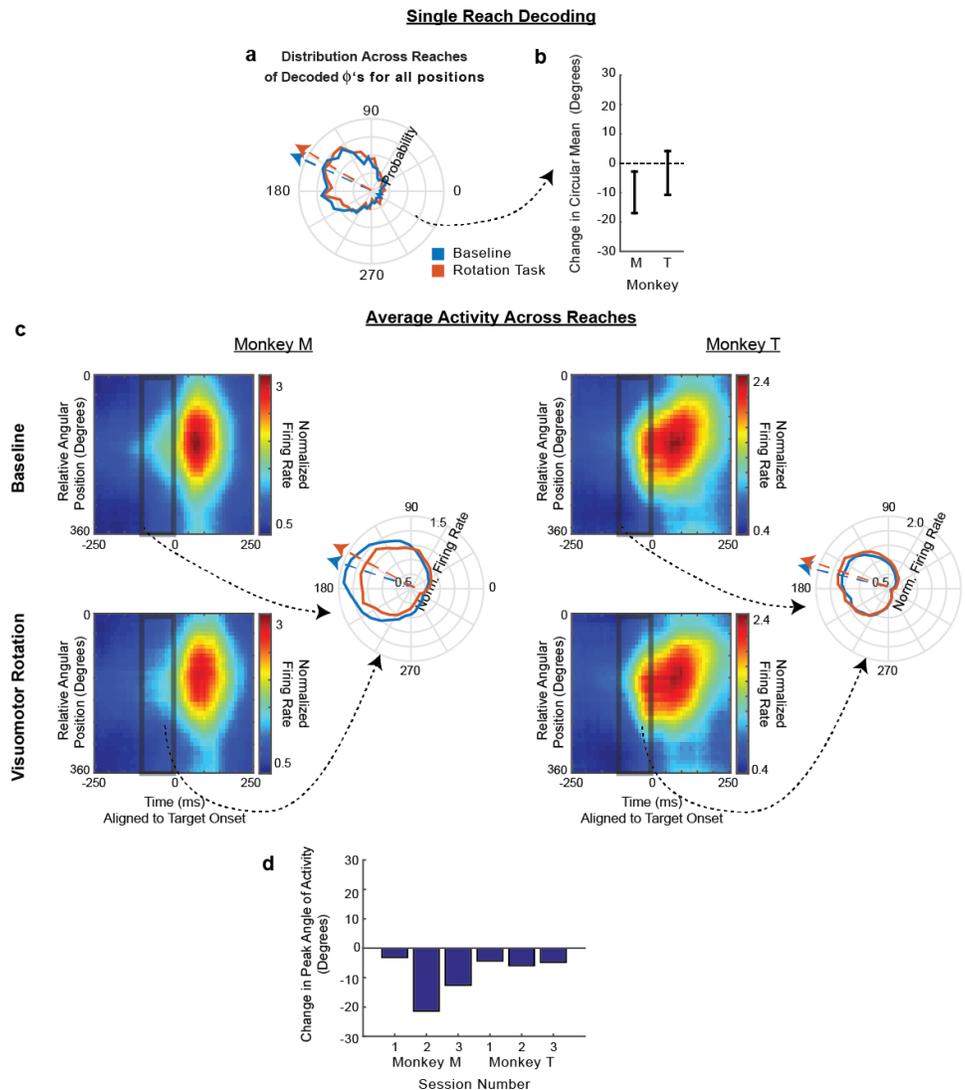
Supplementary Figure 5.3: **Explaining pre-target activity for SR neurons.** In Fig. 3, for SR neurons, prior to target onset there began to be a slight separation between activity traces depending on whether the reach will be near vs. opposite the PD. Given that SR neurons are supposed to only respond after target onset, this is initially surprising. However, there are two likely reasons for SR neurons' apparent pre-target activity. The first reason is our classification criteria of SR and PR neurons. PR neurons, unlike SR neurons, were significantly modulated by hand position prior to target onset (see Methods for details). That is, if a neuron was modulated by hand position with $> 95\%$ (e.g. 96%) confidence, then it was a PR neuron, but if it was modulated by hand position with $< 95\%$ (e.g. 94%) confidence, it would be an SR neuron. Thus, using this "conservative classification" (as we do in Fig. 3), we are likely including some PR neurons in the SR category. A PSTH using this conservative classification, copied from Fig. 3, is shown in the top row. Instead, if we use a "relaxed classification" that includes neurons as PR neurons if they are modulated by hand position with $> 50\%$ confidence, then SR neurons should not include any true PR neurons. When we plot SR neurons using this relaxed classification (bottom row), the differential activity prior to target onset diminishes, demonstrating that some PR neurons being included as SR neurons was a cause of the differential activity. Note that Supplementary Fig. 4 gives more details about different "conservative" and "relaxed" classification types. A second reason for the pre-target-onset differentiation of SR neurons in Fig. 3 is jitter in the time of target onset. While we subtracted the average delay for the target to be displayed on screen, there was some jitter in this timing (see Methods). Thus, some activity aligned to target onset could appear slightly earlier than it occurred.



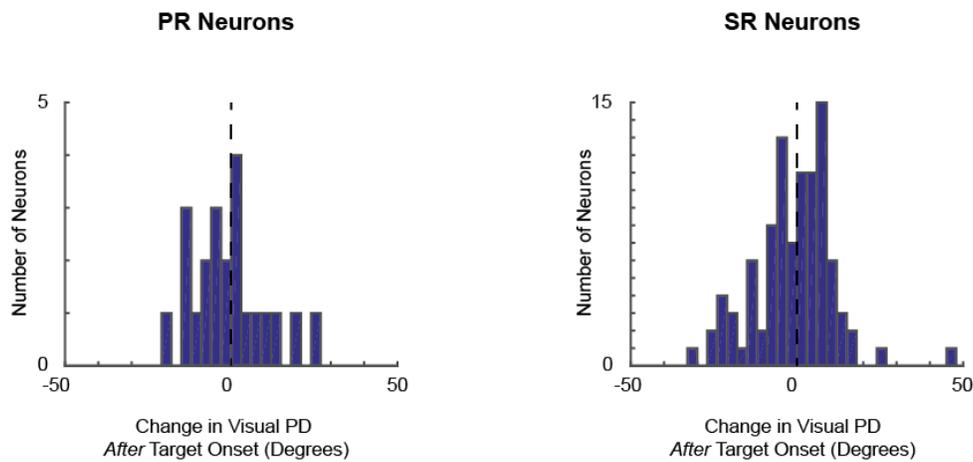
Supplementary Figure 5.4: **Neuron Classification.** The proportion of potential-response (PR) and selected-response (SR) neurons using different classification criteria. **a,b)** We include all sessions (as in the main text) **a)** We defined SR neurons as those significantly modulated by upcoming movement in the late period in the GLM, but were not significantly modulated by hand position in the early period. PR neurons were significantly modulated by upcoming movement in the late period and by hand position in the early period. Significantly modified means that the 95% confidence intervals of pseudo- R^2 and relative pseudo- R^2 values were greater than 0 (see Methods). This was the criteria used for all parts of the main text with the exception of decoding. **b)** A more relaxed criteria is to look at all neurons that had a median pseudo- R^2 and relative pseudo- $R^2 > 0$. In the main text, we used this relaxed criteria to include PR neurons for decoding. **c,d)** These are the same as panels a and b, respectively, except we now only include sessions that were recorded in separate weeks, to decrease the number of “repeat” neurons that were recorded in multiple sessions.



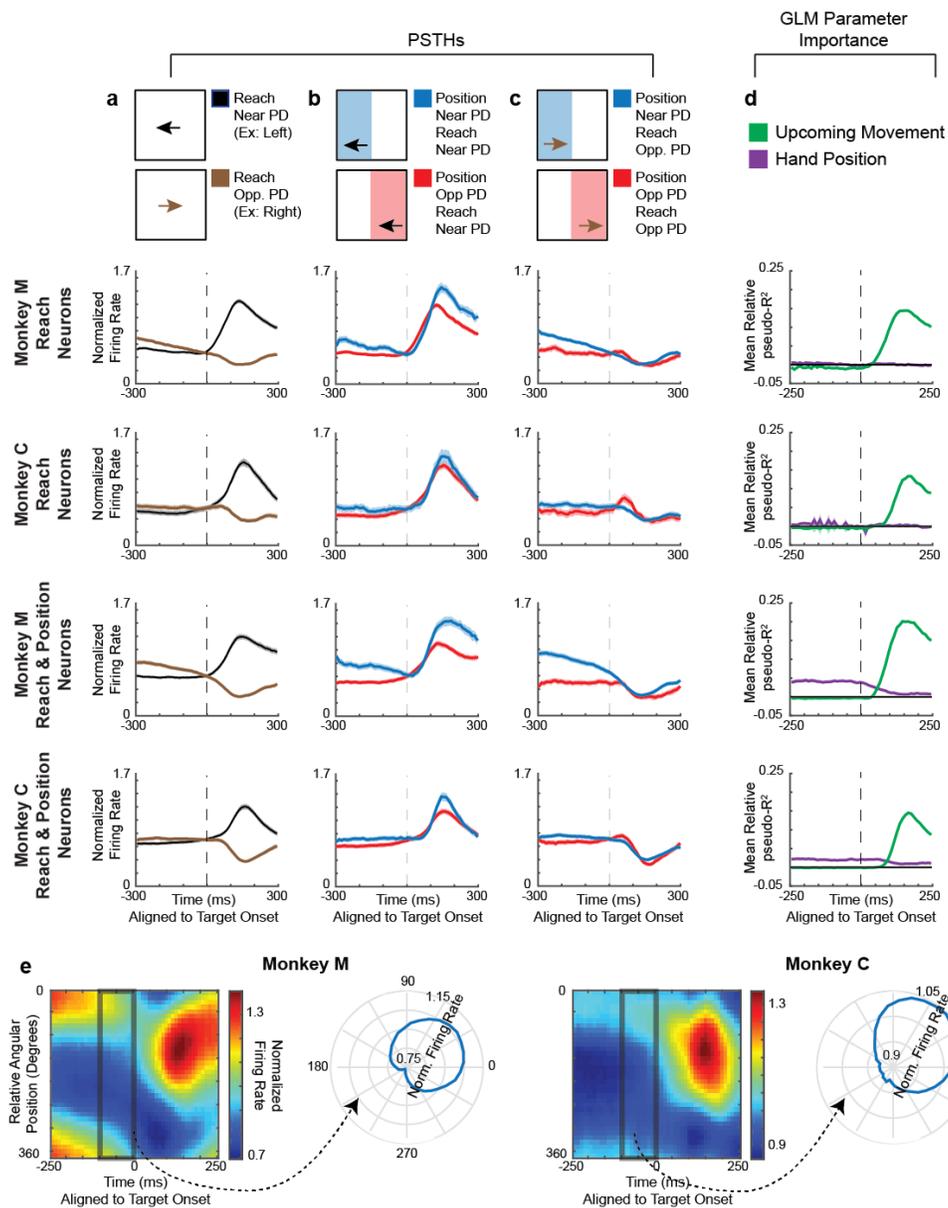
Supplementary Figure 5.5: **Pmd population activity represents the distribution of upcoming movements- accumulated across reaches: individual monkeys and controls.** As in Fig. 4a, all heat maps show normalized average activity over time as a function of relative angular position (the angular hand position relative to neurons' PDs). To the right of each heat map, normalized activity in the 100 ms prior to target onset, as a function of the relative angular position. Each row calculates activity from a different set of reaches. **Top row:** From all reaches (as in Fig. 4a). **Second row:** Control for the correlation between the previous and upcoming movement directions, which tended to be in opposite directions. We examined neural activity in the infrequent cases when pairs of sequential reaches were in similar directions (less than 90° away from each other). **Third row:** We resampled reaches to create a distribution of reach directions relative to angular hand position that were centered on 180° (rather than 150°). This plot ensures that our main results were not simply caused by a correlation between the angular hand position and the true upcoming movement (in which case the activity after resampling would become centered on 180°), but rather reflected the distribution of upcoming movements (in which case the activity after resampling would remain centered at 150°). **Bottom row:** From reaches starting near the center (as in Fig. 4c).



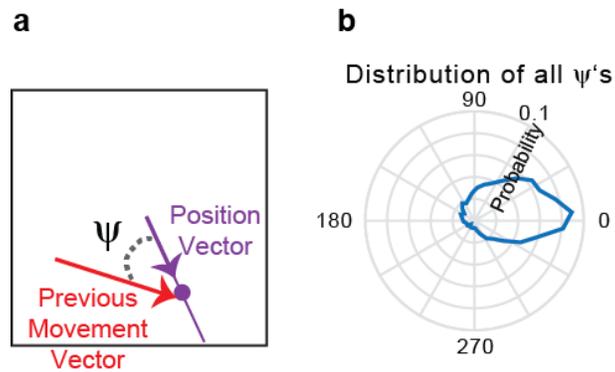
Supplementary Figure 5.6: **Visuomotor Rotation Control Task.** (a) The distribution of pre-target decoded reach directions relative to the hand's angular position (decoded ϕ 's) for all positions. Arrows point toward the circular means of the distributions. (b) The difference between the circular means of the distributions of decoded ϕ 's in panel a (rotation minus baseline). Error bars are 95% confidence intervals from bootstrapping. (c) In both the baseline task and visuomotor rotation task, we display on the **left**: The average normalized firing rate of all PR neurons, over time, as a function of relative angular hand position (as in Fig. 4a). For each monkey, we display on the **right**: The average normalized firing rate in the 100 ms prior to target onset, as a function of the relative angular position (baseline in blue; rotation task in orange). Arrows point toward the peak angles of activity. (d) For each session, the difference between the angles corresponding to peak activity in panel c, between the baseline and rotation tasks (rotation minus baseline). In all panels, results from the VR task used the second 2/3 of trials.



Supplementary Figure 5.7: **Visuomotor Rotation Task - changes in PDs after target onset.** On the **left**, we plot a histogram of changes in the visual PD for PR neurons, in the time period 50-200 ms after target onset. If the neurons were representing the movement rather than the target, then the neurons would be most active when the target is +30 degrees (counterclockwise) in the VR task. If they were representing the visual location of the target, then we would see a change of 0. The median change across PR neurons is -2.1 degrees. On the **right**, we have the same plot for SR neurons, with a median change of 0.4 degrees. Thus, the visuomotor rotation does not change PMd's representation of the target after target onset.



Supplementary Figure 5.8: **M1 does not reflect the probability of upcoming movements, for individual monkeys.** (a-d) PSTHs and GLM results for M1 neurons. Columns have the same schematics as Fig. 3. **First and Second Rows of PSTHs:** Normalized averages of “reach” neurons. **Third and Fourth Rows:** Normalized averages of “reach and position” neurons. (e) Same schematic as Fig. 4a, but for M1 neurons. **Left:** The normalized average firing rate, as a function of time and relative angular hand position. Activity is averaged across all reach and position neurons. **Right:** The normalized average firing rate in the 100 ms prior to target onset, as a function of the relative angular position.



Supplementary Figure 5.9: **Relationship between previous movement and hand position.** **(a)** We define the angle between the previous movement vector and the hand position vector (relative to the center) as ψ . More specifically, ψ is the previous movement vector direction minus the position vector direction. **(b)** The distribution of ψ 's, across all reaches. Note the slight counter-clockwise bias from 0° (it is peaked at about 10° and has a circular mean at 23°), that may explain a portion of the counter-clockwise bias in M1 activity in Fig. 7e.

CHAPTER 6

Machine learning for neural decoding**Abstract**

While machine learning tools have been rapidly advancing, the majority of neural decoding approaches still use last century's methods. Improving the performance of neural decoding algorithms allows us to better understand what information is contained in the brain, and can help advance engineering applications such as brain machine interfaces. Here, we apply modern machine learning techniques, including neural networks and gradient boosting, to decode from spiking activity in 1) motor cortex, 2) somatosensory cortex, and 3) hippocampus. We compare the predictive ability of these modern methods with traditional decoding methods such as Wiener and Kalman filters. Modern methods, in particular neural networks and ensembles, significantly outperformed the traditional approaches. For instance, for all of the three brain areas, an LSTM decoder explained over 40% of the unexplained variance from a Wiener filter. These results suggest that modern machine learning techniques should become the standard methodology for neural decoding. We provide code to facilitate wider implementation of these methods.

Introduction

Decoding is a critical tool for understanding how neural signals relate to the outside world. It can be used to determine how much information the brain contains about an external variable (e.g. sensation or movement) [160, 161, 162], and how this information

differs across brain areas [163, 164, 165], experimental conditions [20, 4], disease states [166], and more. It is also useful in engineering contexts, such as for brain machine interfaces (BMIs), where signals from motor cortex are used to control computer cursors [158], robotic arms [167], and muscles [168]. Decoding is a central tool for neural data analysis.

Because decoding is simply a regression or classification problem, many methods can be used for neural decoding. Despite the many recent advances in machine learning techniques, it is still very common to use traditional methods such as linear regression. Using modern machine learning tools for neural decoding would likely significantly boost performance, and might allow deeper insights into neural function.

Here, we compare many different machine learning methods to decode information from neural spiking activity. We predict movement velocities from macaque motor cortex and sensorimotor cortex, and locations in space from rat hippocampus. In all brain regions, modern methods, in particular neural networks and ensembles, led to the highest accuracy decoding, even for limited amounts of data. We provide code so that others can easily use all the decoding methods we tested.

Methods

Tasks and brain regions

Decoding movement velocity from the motor cortex and somatosensory cortex

In our random-target experiment [4], monkeys moved a planar manipulandum that controlled a cursor on the screen (Fig. 1a). The monkeys continuously reached to new targets that were presented with a brief hold period between reaches. After training,

the monkeys were surgically implanted with 96-channel Utah electrode arrays (Blackrock Microsystems, Salt Lake City, UT) to record the extracellular activity of cortical neurons. In one experiment [4], we recorded from both primary motor cortex (M1) and dorsal premotor cortex (PMd) and combined neurons from both areas. In another experiment we recorded from area 2 of primary somatosensory cortex (S1) [169]. From both brain regions, we aimed to predict the x and y components of movement velocity. The recording from motor cortex was 21 minutes, and contained 164 neurons. The mean and median firing rates, respectively, were 6.7 and 3.4 spikes / sec. The recording from S1 was 51 minutes, and contained 52 neurons. The mean and median firing rates, respectively, were 9.3 and 6.3 spikes / sec.

Decoding position from the hippocampus

We used a dataset from CRCNS, in which rats chased rewards on a square platform (Fig. 1b) [7, 8]. Extracellular recordings were made from layer CA1 of dorsal hippocampus (HC). We aimed to predict the x and y position of the rat. The recording from HC was 93 minutes, and contained 58 neurons. We excluded neurons with fewer than 100 spikes over the duration of the experiment, resulting in 46 neurons. These neurons had mean and median firing rates, respectively, of 1.7 and 0.2 spikes / sec.

General Decoding methods

Decoding movement velocity from the motor cortex and somatosensory cortex

We predicted the average velocity (x and y components) in 50 ms bins. Neural spike trains used for decoding were also put into 50 ms bins. In motor cortex, we used 700

ms of neural activity (13 bins before and the concurrent bin) to predict the current movement velocities, as the primary interest is in investigating how motor cortex causally affects movement. In somatosensory cortex, we used 650 ms surrounding the movement (6 bins before, the concurrent bin, and 6 bins after), as neural activity has been shown both preceding and following movements [170]. Decoding position from the hippocampus: We aimed to predict the position (x and y coordinates) of the rat in 200 ms bins. Neural spike trains used for decoding were also put into 200 ms bins. We used 2 seconds of surrounding neural activity (4 bins before, the concurrent bin, and 5 bins after) to predict the current position.

Scoring Metric

To determine the goodness of fit, we used $R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (y_i - \bar{y})^2}$, where \hat{y}_i are the predicted values, y_i are the true values and \bar{y} is the mean value. This formulation of R^2 (which is the fraction of variance accounted for, rather than the Pearson's correlation coefficient squared [171]) can be negative on the test set due to overfitting on the training set. The reported R^2 values are the average across the x and y components of velocity or position.

Cross-validation

When determining the R^2 for every method (Fig. 3), we used 10 fold cross-validation. For each fold, we split the data into a training set (80% of data), a contiguous validation set (10% of data), and a contiguous testing set (10% of data). For each fold, decoders were trained to minimize the mean squared error between the predicted and true velocities/positions of the training data. We found the algorithm hyperparameters that led to the highest on the validation set using Bayesian optimization [172]. That is, we fit many

models on the training set with different hyperparameters and calculated the R^2 on the validation set. Then, using the hyperparameters that led to the highest validation set R^2 , we calculated the R^2 value on the testing set. Error bars on the test set R^2 values were computed across cross-validation folds.

Bootstrapping

When determining how performance scaled as function of data size (Fig. 5), we used a single test set and validation set, and varying amounts of training data that directly preceded the validation set. The test and validation sets were 5 minutes long for motor and somatosensory cortices, and 7.5 minutes for hippocampus. To get error bars, we resampled from the test set. Because of the high correlation between temporally adjacent samples, we didn't resample randomly from all examples (which would create highly correlated resamples). Instead, we separated the test set into 20 temporally distinct subsets, S_1 - S_{20} (i.e., S_1 is from $t = 1$ to $t = T/20$, S_2 is from $t = T/20$ to $t = 2T/20$, etc., where T is the end time), that were more nearly independent of each other. We then resampled combinations of these 20 subsets (e.g. S_5, S_{13}, S_2) 1000 times to get confidence intervals of R^2 values.

Preprocessing

The training input was normalized (z-scored). The training output was zero-centered (mean subtracted), except in support vector regression, where the output was z-scored. The validation/testing inputs and outputs were preprocessed using the preprocessing parameters from the training set.

Specific Decoders

Kalman Filter

In the Kalman filter, the hidden state at time t is a linear function of the hidden state at time $t - 1$, plus a matrix characterizing the uncertainty. The observation (measurement) at time t is a linear function of the hidden state at time t (plus noise). At every time point, to update the estimated hidden state, the updates derived from the current measurement and the previous hidden states are combined. During this combination, the noise matrices give a higher weight to the less uncertain information. We used a Kalman filter similar to that implemented in [9]. In the Kalman filter, the measurement was the neural spike trains, and the hidden state was the kinematics (x and y components of position, velocity, and acceleration). We had one hyperparameter which differed from the implementation in [9]. This parameter determined the noise matrix associated with the transition in kinematic states (\mathbf{Q} in [9]). We divided the empirical noise matrix of training data (used in [9]) by the hyperparameter scalar C . The rationale for this addition is that neurons have temporal correlations, which make it desirable to have a parameter that allows changing the weight of the new neural evidence. Interestingly, the introduction of this parameter made a big difference for the hippocampus dataset (Fig. S1). We also allowed for a lag between the neural data and predicted kinematics. The lag and hyperparameter were determined based on validation set performance.

Wiener Filter

The Wiener filter uses multiple linear regression to predict the output from multiple time bins of every neurons' spikes. That is, the output is assumed to be a linear mapping of the

number of spikes in the relevant time bins from every neuron (Fig. 1c,d). Here, separate models were used to predict the x and y components of the kinematics.

Wiener Cascade

The Wiener cascade (also known as a linear nonlinear model) fits a linear regression (the Wiener filter) followed by a fitted static nonlinearity (e.g. [173]). This allows for a nonlinear relationship between the input and the output, and assumes that this nonlinearity is purely a function of the linear output. Here, as in the Wiener Filter, the input was neurons' spike rates over relevant time bins. The nonlinear component was a polynomial with degree determined on the validation set. Separate models were used to predict the x and y components of the kinematics.

Support Vector Regression

In support vector machine regression (SVR) [174], the inputs are projected into a higher dimensional space using a nonlinear kernel, and then linearly mapped from this space to the output to minimize an objective function [174]. Here, we used standard support vector regression (SVR) with a radial basis function kernel to predict the kinematics from the neurons' spike rates in each bin. We set hyperparameters for the penalty of the error term and the maximum number of iterations. Separate models were used to predict the x and y components of the kinematics.

XGBoost

XGBoost (Extreme Gradient Boosting) [175] is an implementation of gradient boosted trees. For the regression problem, gradient boosting fits many regression trees. Each subsequent regression tree is fit to the residuals of the previous fit. Regression trees

create nonlinear mappings from the input to output. Here, we used XGBoost to predict the kinematics from the neurons' spike rates in each bin. We set hyperparameters for the maximum depth of the tree, number of trees, and learning rate. Separate models were used to predict the x and y components of the kinematics.

Feedforward Neural Network

A feedforward neural net connects the inputs to sequential layers of hidden units via linear mappings followed by output nonlinearities. This can allow for mapping complex nonlinear functions from input to output. Here, using the Keras library [176], we created a fully connected (dense) feedforward neural network with 2 hidden layers and rectified linear unit activations after each hidden layer. We required the number of hidden units in each layer to be the same. We set hyperparameters for the number of hidden units in the layers, amount of dropout [177], and number of training epochs. We used the Adam algorithm [178] as the optimization routine. This neural network, and all neural networks below had 2 output units. That is, the same network predicted the x and y components rather than there being 2 separate networks. The input was still the number of spikes in each bin from every neuron. Note that we refer to feedforward neural networks as a modern technique, despite their having been around for many decades, due to their current resurgence and the modern methods for training the networks.

Simple RNN

In a standard recurrent neural network (RNN), the hidden state is a linear combination of the inputs and the previous hidden state. This hidden state is then run through an output nonlinearity, and linearly mapped to the output. RNNs, unlike feedforward neural

networks, allow temporal changes in the system to be modeled explicitly. Here, using the Keras library [176], we created a neural network architecture where the spiking input from all neurons was fed into a standard recurrent neural network (Fig. 1e). The units from this recurrent layer were fed through a rectified linear unit nonlinearities, and fully connected to an output layer with 2 units (x and y velocity or position components). We set hyperparameters for the number of units, amount of dropout, and number of training epochs. We used RMSprop [179] as the optimization routine.

Gated Recurrent Unit

Gated recurrent units (GRUs) [180] are a more complex type of recurrent neural network. It has gated units, which in practice allow for better learning of long-term dependencies. For implementation, all methods were the same as for the Simple RNN, except Gated Recurrent Units were used rather than a traditional RNN.

Long Short Term Memory Network

Like the GRU, the long short term memory (LSTM) network [181] is a more complex recurrent neural network with gated units that allow long-term dependencies to be captured better. The LSTM has more parameters than the GRU. For implementation, all methods were the same as for the Simple RNN, except LSTMs were used.

Ensemble

Ensemble methods combine the predictions from several other methods, and thus have the potential to leverage the different benefits of the methods contained within the ensemble. Here, using the predictions from all decoders except the Kalman filter (which had a

different format) as inputs, we predicted the outputs using the feedforward neural network described above.

Code

Python code for all methods is available at https://github.com/KordingLab/Neural_Decoding

Results

We investigated how the choice of machine learning technique affects decoding performance (Fig. 1) using a plethora of common machine learning methods. These ranged from historical linear techniques (e.g., the Wiener filter) to modern machine learning techniques (e.g., neural networks and ensembles of techniques). We tested the performance of all these techniques across datasets from motor cortex, somatosensory cortex, and hippocampus.

We aimed to understand the performance of the methods when fit to neural data. First, in order to get a qualitative impression of the performance, we plotted the output of each decoding method for each of the three datasets (Fig. 2). In these examples, the modern methods, such as the LSTM and ensemble, appeared to outperform traditional methods, such as the Wiener and Kalman filters, as the predictions were slightly closer to the true output. Next, we quantitatively compared the methods. In all three brain areas, modern machine learning methods outperformed traditional decoding methods (Fig. 3). In particular, neural networks and the ensemble led to the best performance, while the Wiener or Kalman Filter led to the worst performance. In fact, the LSTM decoder explained over 40% of the unexplained variance from a Wiener filter (R^2 's of 0.88, 0.86,

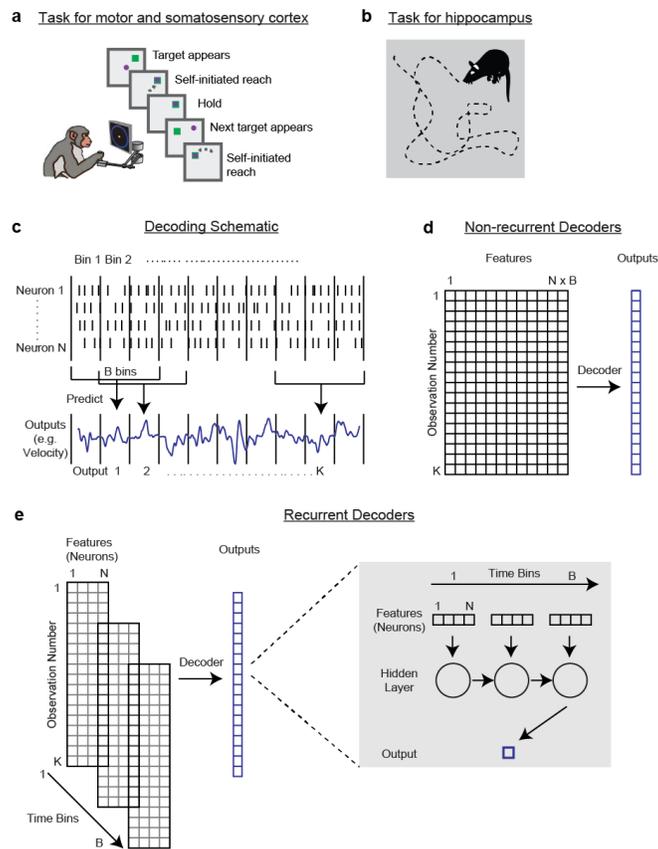


Figure 6.1: **Tasks and Decoding Schematic.** **a)** In the task for decoding from motor and somatosensory cortices, monkeys continuously reached to new targets that were presented, with a brief hold period between reaches [4]. **b)** In the task for decoding from hippocampus, rats chased rewards on a square platform [7, 8]. **c)** To decode (predict) the output in a given time bin, we used the firing rates of all N neurons in B time bins. In this schematic, $N = 4$ and $B = 3$ (one bin preceding the output, one concurrent bin, and one following bin). In our data, we predicted two outputs from each brain region (x and y components of velocity from motor and somatosensory cortex, and x and y components of position from hippocampus). For each region, the number of neurons and time bins used for decoding are described in *Methods*. Also, note that this schematic does not apply for the Kalman Filter decoder. **d)** For the non-recurrent decoders (Wiener Filter, Wiener Cascade, SVR, XGBoost, and Feedforward Neural Network), this is a standard regression problem where $N \times B$ features (the firing rates of each neuron in each relevant time bin) are used to predict the output. **e)** For the recurrent decoders (simple RNN, GRUs, LSTMs), to predict an output, we used N features, with temporal connections across B bins. A schematic of an RNN predicting a single output is on the right.

0.62 vs. 0.78, 0.75, 0.35). Additionally, the feedforward neural network did almost as well as the LSTM in all brain areas. Across cases, the ensemble method added a reliable, but small increase to the explained variance. Modern machine learning methods led to significant increases in predictive power.

While modern machine learning methods yielded the best performance on our full datasets, it is possible, because of their greater complexity, that they would not work well with less data. Thus, we tested the feedforward neural network and LSTM (two modern methods that worked particularly well), along with the Wiener and Kalman filters, on varying amounts of data. Even with limited data, the modern methods worked very well. With only 2 minutes of training data for motor and somatosensory cortices, and 15 minutes of hippocampus data, both modern methods outperformed both traditional methods (Fig. 4,5). When decreasing the amount of training data further, to only 1 minute for motor and somatosensory cortices and 7.5 minutes for hippocampus data, the Kalman filter sometimes performed comparably to the modern methods. Still, the modern methods significantly outperformed the Wiener Filter (Fig. 5). Thus, even for limited data, modern machine learning methods can yield significant gains in decoding performance.

Discussion

Here we tested the performance of a large number of decoding techniques on three different neural decoding problems. We found that, across datasets, neural networks outperformed traditional methods. An ensemble method provided only minor additional

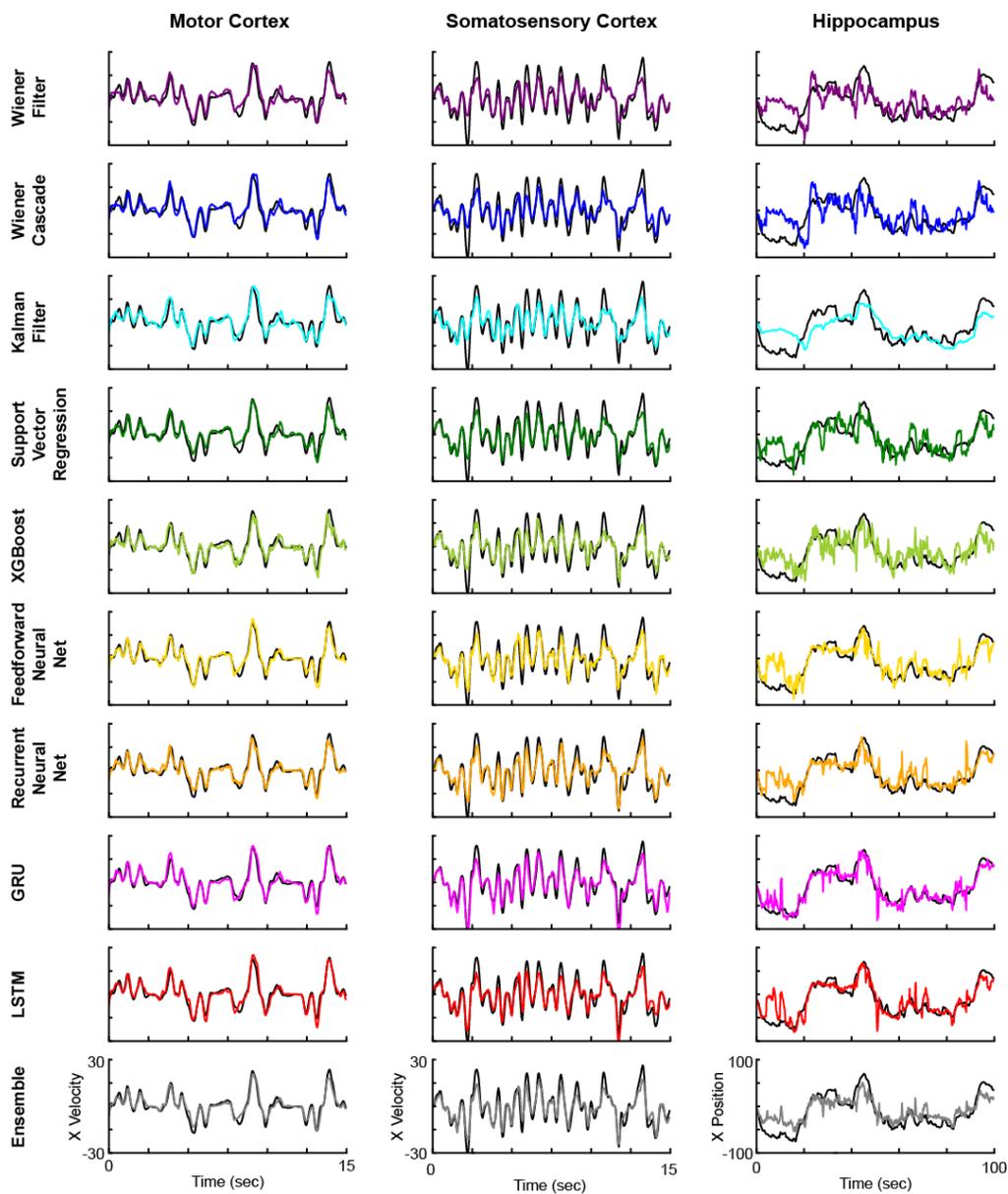


Figure 6.2: **Example Decoder Results.** Example decoding results from motor cortex (left), somatosensory cortex (middle), and hippocampus (right), for all ten methods (top to bottom). Ground truth traces are in black, while decoder results are in various colors.

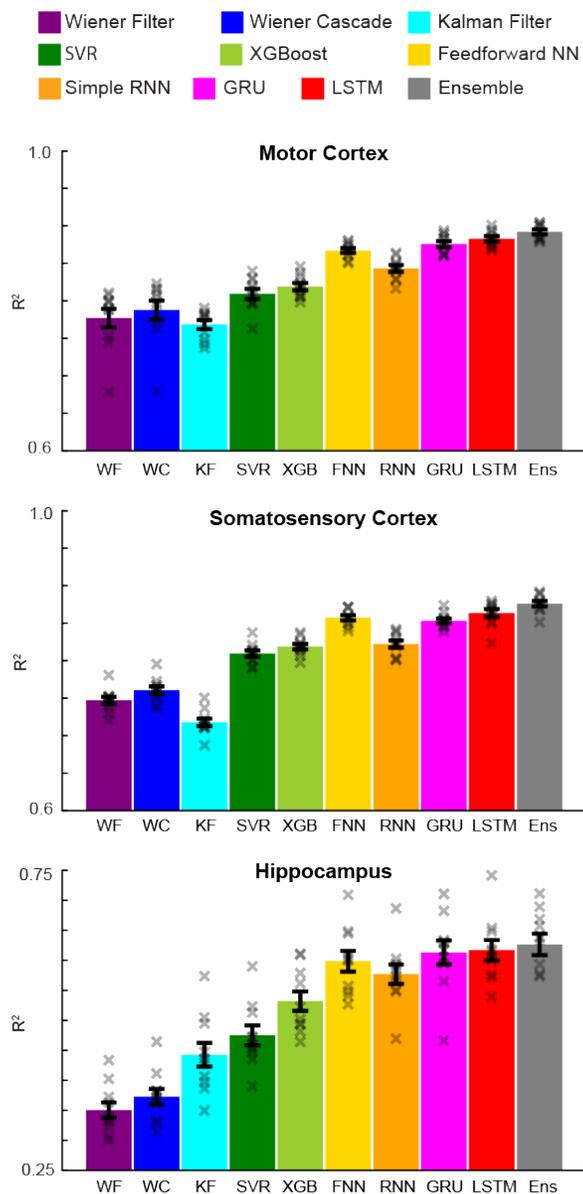


Figure 6.3: **Decoder Result Summary.** R^2 values are reported for all decoders (different colors) for each brain area (top to bottom). Error bars represent the mean \pm SEM across cross-validation folds. X's represent the R^2 values of each cross-validation fold. Note the different y-axis limits for the hippocampus dataset.

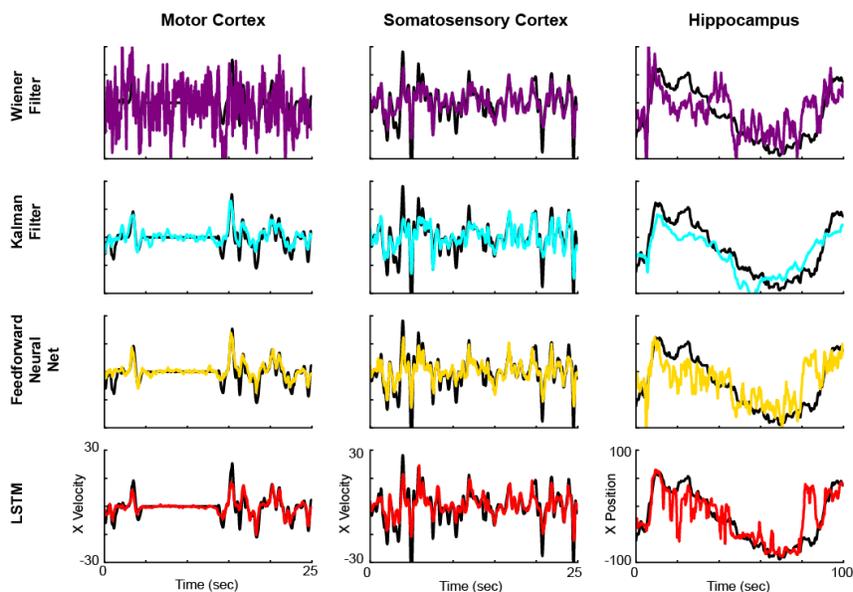


Figure 6.4: **Example results with limited training data.** Using only 2 minutes of training data for motor cortex and somatosensory cortex, and 15 minutes of training data for hippocampus, we trained two traditional methods (Wiener filter and Kalman filter), and two modern methods (feedforward neural network and LSTM). Example decoding results are shown from motor cortex (left), somatosensory cortex (middle), and hippocampus (right), for these methods (top to bottom). Ground truth traces are in black, while decoder results are in the same colors as previous figures.

predictive power. The strong performance of neural networks even persisted for small datasets with as little as one minute of training data.

We find it particularly interesting that the neural network methods worked so well with limited data, which is counter to the common perception. We believe the explanation is simply the size of networks. For instance, our networks have on the order of 100 thousand parameters, while common networks for image classification (e.g. [182]) can have on the order of 100 million parameters. Thus, the reasonable size of our networks (hundreds of hidden units) likely allowed for excellent prediction with limited data [183].

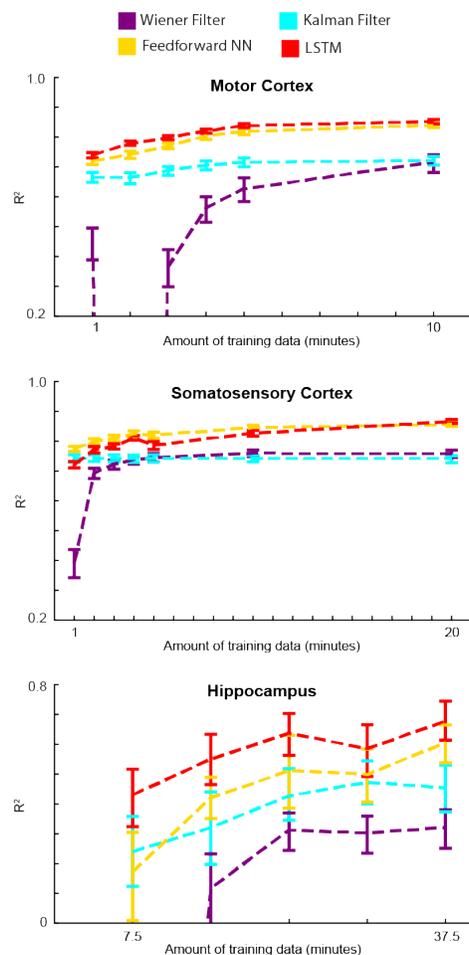


Figure 6.5: **Decoder results with varying amounts of training data.** Using varying amounts of training data, we trained two traditional methods (Wiener filter and Kalman filter), and two modern methods (feedforward neural network and LSTM). R^2 values are reported for these decoders (different colors) for each brain area (top to bottom). Error bars are 68% confidence intervals (meant to approximate the SEM) produced via bootstrapping, as we used a single test set. Values with negative R^2 s were not shown. Also note the different y-axis limits for the hippocampus dataset.

It is also intriguing that the feedforward neural network did almost as well as the LSTM and better than the standard RNN, considering the recent attention to treating the brain as a dynamical system [184]. For the motor and somatosensory cortex decoding,

it is possible that the highly trained monkeys yielded a stereotyped temporal relationship between neural activity and movement that a feedforward neural network could effectively capture. It would be interesting to compare the performance of feedforward and recurrent neural networks on less constrained behavior.

In order to find the best hyperparameters for the decoding algorithms, we used a Bayesian optimization routine [172] to search the hyperparameter space (see Methods). Still, it is possible that some of the decoding algorithms did not use the optimal hyperparameters, which could have lowered overall accuracy. Moreover, for several methods we did not fit all available hyperparameters. We did this in order to simplify the use of these methods, in order to decrease computational runtime during hyperparameter optimization, and because adding additional hyperparameters did not appear to improve accuracy. For example, for the neural nets we used dropout but not L1 or L2 regularization, and for XGBoost we used less than half the available hyperparameters for avoiding overfitting. While our preliminary testing with additional hyperparameters did not appear to significantly change the results, it is possible that we have not achieved optimal performance of our methods.

While we have tested standard algorithms on three different datasets, it is possible that the relative performance of algorithms differs on other datasets. However, many datasets in neuroscience share basic properties with those we used. Most are similar in length (tens of minutes to a couple hours), simply because the length of a recording session is usually limited by both the patience of the animal and the experimentalist. Moreover, most variables of interest have similar relevant timescales, where movement, speech, vision,

and many other phenomena unfold on a timescale of hundreds of milliseconds to seconds. We thus expect that similar results would be obtained for other spiking datasets.

We have decoded from spiking data, but it is possible that the problem of decoding from other data modalities is different. One main driver of a difference may be the distinct levels of noise. For example, fMRI signals have far higher noise levels than spikes. As the noise level goes up, linear techniques become more appropriate, which may ultimately lead to a situation where the traditional linear techniques become superior. Applying the same analyses we did here across different data modalities is an important next step.

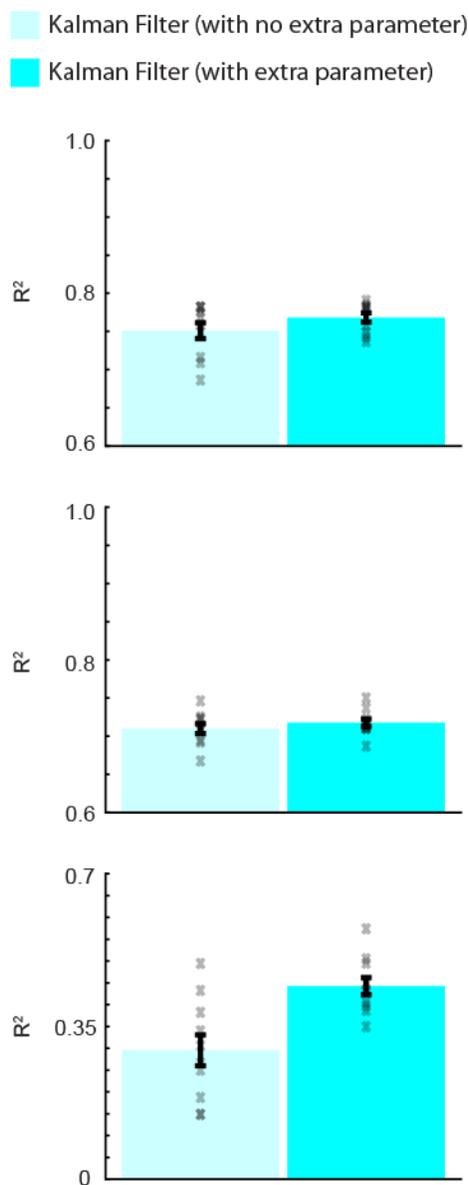
All our decoding was done offline, meaning that the decoding occurred after the recording, and was not part of a control loop. This type of decoding is useful for determining how information in a particular brain area relates to an external variable. However, for engineering applications such as BMIs [185, 186], the goal is to decode information (e.g., predict movements) in real time. Our results here may not apply as directly to online decoding situations, since the subject is ultimately able to adapt to imperfections in the decoder. In that case, even relatively large decoder performance differences may be irrelevant. An additional challenge for online applications is computational runtime, which we have not addressed here. In the future, it would be valuable to test modern machine learning techniques for decoding in online applications (as in [187]).

While modern machine learning methods provide an increase in decoding accuracy, it is important to be careful with the scientific interpretation of decoding results. Decoding can tell us how much information a neural population has about a variable X . However, high decoding accuracy does not mean that a brain area is directly involved in processing X , or that X is the purpose of the brain area. For example, with a powerful decoder,

it could be possible to accurately classify images based on recordings from the retina, since the retina has information about all visual space. However, this does not mean that the primary purpose of the retina is image classification. Moreover, even if the neural signal comes before the external variable, it does not mean that it is causally involved. For example, information could be in somatosensory cortex prior to movement due to an efference copy from M1. Thus, researchers should constrain interpretations to being about the information in neural populations, and how it may vary across brain regions, experimental conditions, or time intervals.

We decoded continuous valued variables. The same methods can be used for classification tasks, which often use classic decoders such as logistic regression and support vector machines. While here we have not demonstrated the benefit of modern machine learning methods for classification, our available code can easily be modified to allow users to do classification.

Neural engineering has a history of developing specialized algorithms meant to increase the performance of decoders [188, 189, 190]. However, these algorithms are not typically tested against state of the art machine learning algorithms. Along with this manuscript, we have released a package to do neural decoding using all the described methods, making it is easy to compare with any new algorithm. Our hunch is that it will be hard for specialized algorithms to compete with the standard algorithms developed by a massive community in machine learning.



Supplementary Figure 6.1: **Kalman Filter Versions**. R^2 values are reported for different versions of the Kalman Filter for each brain area (top to bottom). On the left (in more transparent cyan), the Kalman Filter is implemented as in [9]. On the right (in more opaque cyan), the Kalman Filter is implemented with an extra parameter that scales the noise matrix associated with the transition in kinematic states (see *Methods*). This version with the extra parameter is the one used in the main text. Error bars represent the mean \pm SEM across cross-validation folds. X's represent the R^2 values of each cross-validation fold. Note the different y-axis limits for the hippocampus dataset.

CHAPTER 7

Discussion

Across the chapters of this thesis, I aimed to better understand how the brain plans movements in more naturalistic settings, and I often used machine learning techniques to do so. In Chapter 2, I demonstrated that during a natural scene search task, search-relevant visual features had no significant influence on FEF activity. Rather, FEF activity was primarily modulated by the upcoming eye movement. In Chapter 3, I demonstrated that FEF activity is greater when making a saccade to the target (an exploitative saccade) rather than somewhere else in the scene (an exploratory saccade). In Chapter 4, I found two separate classes of FEF neurons, related to preliminary and definitive planning. Those related to preliminary planning represented the probabilities of potential upcoming saccades, independent of the saccade that would occur. In Chapter 5, I demonstrated that PMd neurons represented the probability distribution of possible upcoming movements that are based on the position of the monkey's hand within the environment. In Chapter 6, I demonstrated that modern machine learning methods such as neural networks and ensemble methods significantly improve predictive performance for neural decoding, even with limited amounts of data.

Multiple chapters of this thesis focused on factors that modulate FEF activity during natural scene search, and I would like to relate the different findings across chapters. In Chapter 2, I did not find that task-relevant visual features modulated FEF activity. This differed from previous findings using artificial stimuli, in which FEF activity was greater

when stimuli similar to the target were displayed in neurons' RFs. Findings from chapter 3 may explain this difference in results. I showed that purposeful, but not accidental, saccades towards a target in neurons' RFs increased their firing rates. It thus may be the case that awareness of a target is necessary for enhanced representation of this target. Artificial stimulus paradigms with few objects most likely consistently lead to awareness of the target and target-similar objects, while in natural scenes, a subject may not be aware of target-similar objects. This difference in awareness could lead to my negative result about the representation of target-similar features in natural scenes.

Chapters 2 and 4 both investigated different environmental factors that could affect saccade planning, and thus FEF activity. In Chapter 2, I analyzed target-related visual features, which had weak predictive ability about the upcoming saccade. In Chapter 4, I analyzed the eye position on the screen, which had a stronger predictive ability about the upcoming saccade direction. In parallel to these behavioral results, I found neural activity was modulated by eye position (which had a strong behavioral effect) but not by the target-related visual features (which had a weak behavioral effect). While I did a power analysis in Chapter 2 to demonstrate that our negative result about visual features was likely not because the effect size was undetectable, it still could be possible that the behavioral effect was just too weak for the model to detect the neural correlate. In general, this shows the importance of having strong behavioral effects when searching for neural correlates.

While all FEF-related chapters found a relation between FEF activity and saccades, some aspects of the activity may have to do with saccade planning, while some activity may have to do with the saccade itself, or saccade execution. In Chapter 2, I did not

analyze this distinction in any way. In Chapter 3, I found that FEF activity related to expected reward (i.e., whether making a saccade to the target or elsewhere), saccade peak velocity, and saccade latency. Expected reward is a cognitive factor, and could relate to planning, as higher firing rates when expecting a reward in a given direction could increase the likelihood of a saccade in that direction. Neural activity related to saccade peak velocity may seem to be related to saccade execution, but it could also be related to a higher confidence about the saccade, which is just correlated with the saccade velocity. Neural activity being correlated with saccade latency is likely related to the saccade decision-making (planning) process, at least based on many models of decision-making [191, 192]. In chapter 4, I separated neurons into two subpopulations that do “preliminary planning” and “definitive planning.” This demonstrated that there is FEF activity related to both planning (regardless of the saccade that actually occurs), along with the final decision (the saccade that actually occurs). Still, it is unknown whether this activity related to the final saccade plan is actually responsible for the execution of the upcoming saccade. In general, it is a great challenge to differentiate movement planning and execution in naturalistic tasks.

Another theme shared across chapters was probabilistic representations for movement planning. In Chapter 4, I found probabilistic representations of upcoming saccade directions in a subset of FEF neurons when averaging across many saccades. In Chapter 5, I found probabilistic representations of upcoming reach directions in a subset of PMd neurons. Because of the many simultaneously recorded neurons in PMd, I was able to do single-trial decoding to demonstrate that probability distributions were represented by the population of neurons preceding individual reaches. In the future, it would be

valuable to record more FEF neurons simultaneously to allow testing whether probability distributions are also represented in FEF populations prior to individual saccades. The findings of these two chapters suggest the generality of probabilistic representations for movements for both the oculomotor and skeletomotor systems.

In chapters 4 and 5, I also was interested in how potential movements evolve into definitive movement plans. In PMd, PR neurons represented the possible movements prior to target onset and represented the final selected movement (i.e. the target location) after target onset. In FEF, around the time of fixation, E-Sel neurons had activity related to the probabilities of upcoming saccades (independent of the actual saccade) and also had activity predictive of the true upcoming saccade. Later on, closer to the time of the saccade, activity of E-Sel neurons did not become more related to the actual saccade that would occur. Because I used different experiments for arm movements and eye movements, I am unable to directly compare how movement plans evolve in PMd and FEF. It would be interesting to examine PMd responses using a more natural task where the monkeys could choose their own movements (as in the eye movement experiment). Would the PR neurons have early “preliminary planning” activity predictive of the actual upcoming reach, like I saw for FEF E-Sel neurons? Additionally, it would be interesting to investigate FEF responses in a task similar to the reaching task, where targets are continuously presented. Would the E-Sel neurons reflect the available options prior to target onset? And would the E-Sel neurons no longer select the upcoming movement, since that experiment would not allow explicitly planning saccades in advance? Future experiments to more directly compare movement planning in FEF and PMd would be valuable.

There are many tradeoffs involved in conducting more naturalistic experiments. Ultimately, researchers want to understand how the brain operates in real world scenarios. Moreover, neural responses in more naturalistic settings can differ from those in more constrained, artificial, paradigms [193, 194], which demonstrates the necessity of more naturalistic experiments. However, it can be challenging to understand the reason why neural responses are different between naturalistic environments and constrained paradigms when there are many differences between the two setups. Returning to the example from chapter 2, was the negative result about the representation of target-similar visual features because of the visual structure of natural scenes? Or was it because the natural scene was not flashed on the screen? In general, it would be beneficial to do experiments that titrate “naturalness”, i.e., make changes that systematically increase the amount of naturalness in one dimension, so that the changes in neural activity can be more clearly understood. Moreover, it would be beneficial to do these experiments with the same sets of neurons, to explicitly determine how responses differ between experimental setups. In general, while clever analyses can help to understand more naturalistic experiments, post-hoc analyses can’t replace a good initial experimental design to understand how naturalness changes responses.

Machine learning served multiple roles in the chapters of this thesis. In chapter 6, I demonstrated how modern machine learning methods can improve predictive performance. Improving predictive performance is generally valuable for engineering applications. In the case of neural decoding, it is useful to better understand what information a population of neurons contains about an external variable. In the case of neural encoding [169], it is useful to better understand how well a neuron’s activity can be predicted based on

multiple external variables. In chapters 2-5, I used machine learning, in particular GLMs, to better understand how several variables modulated neural activity. It is interesting to consider whether I might have gotten any different results if I had used a modern machine learning method, such as neural networks or XGBoost, rather than GLMs. While I believe the general results would be the same, since the PSTH results had the same trends as the GLMs, it is likely that there would be minor differences. For example, some neurons that encode variables in a weak, nonlinear, manner could become significant for those variables. Finally, I want to mention one other valuable use for machine learning in neuroscience - benchmarking. Since modern machine learning methods lead to great predictive ability, we should use them as performance benchmarks for simpler, more understandable, models. In general, machine learning techniques can be very useful within neuroscience.

References

- [1] P. Ramkumar, P. N. Lawlor, J. I. Glaser, D. K. Wood, A. N. Phillips, M. A. Segraves, and K. P. Kording, “Feature-based attention and spatial selection in frontal eye fields during natural scene search,” *Journal of Neurophysiology*, pp. jn–01044, 2016.
- [2] J. I. Glaser, D. K. Wood, P. N. Lawlor, P. Ramkumar, K. P. Kording, and M. A. Segraves, “Role of expected reward in frontal eye field during natural scene search,” *Journal of neurophysiology*, vol. 116, no. 2, pp. 645–657, 2016.
- [3] J. I. Glaser, D. K. Wood, P. N. Lawlor, M. A. Segraves, and K. P. Kording, “From preliminary to definitive plans: two classes of neurons in frontal eye field,” *bioRxiv*, p. 137026, 2018.
- [4] J. I. Glaser, M. G. Perich, P. Ramkumar, L. E. Miller, and K. P. Kording, “Population coding of conditional probability distributions in dorsal premotor cortex,” *bioRxiv*, p. 137026, 2017.
- [5] J. I. Glaser, R. H. Chowdhury, M. G. Perich, L. E. Miller, and K. P. Kording, “Machine learning for neural decoding,” *arXiv preprint arXiv:1708.00909*, 2017.
- [6] P. Ramkumar, H. Fernandes, K. Kording, and M. Segraves, “Modeling peripheral visual acuity enables discovery of gaze strategies at multiple time scales during natural scene search,” *Journal of Vision*, vol. 15, no. 3, pp. 19–19, 2015.
- [7] K. Mizuseki, A. Sirota, E. Pastalkova, and G. Buzski, “Theta oscillations provide temporal windows for local circuit computation in the entorhinal-hippocampal loop,” *Neuron*, vol. 64, no. 2, pp. 267–280, 2009.
- [8] K. Mizuseki, A. Sirota, E. Pastalkova, and G. Buzski, “Multi-unit recordings from the rat hippocampus made during open field foraging.” 2009.
- [9] W. Wu, M. J. Black, Y. Gao, M. Serruya, A. Shaikhouni, J. Donoghue, and E. Bienenstock, “Neural decoding of cursor motion using a kalman filter,” in *Advances in neural information processing systems*, pp. 133–140.

- [10] C. J. Bruce and M. E. Goldberg, "Primate frontal eye fields. I. Single neurons discharging before saccades.," *Journal of neurophysiology*, vol. 53, pp. 603–35, mar 1985.
- [11] N. P. Bichot, J. D. Schall, and K. G. Thompson, "Visual feature selectivity in frontal eye fields induced by experience in mature macaques.," *Nature*, vol. 381, pp. 697–9, jun 1996.
- [12] N. P. Bichot and J. D. Schall, "Effects of similarity and history on neural mechanisms of visual selection.," *Nature neuroscience*, vol. 2, pp. 549–554, 1999.
- [13] H. Zhou and R. Desimone, "Feature-based attention in the frontal eye field and area V4 during visual search.," *Neuron*, vol. 70, pp. 1205–17, jun 2011.
- [14] S. Everling and D. P. Munoz, "Neuronal correlates for preparatory set associated with pro-saccades and anti-saccades in the primate frontal eye field," *Journal of Neuroscience*, vol. 20, no. 1, pp. 387–400, 2000.
- [15] L. Ding and J. I. Gold, "Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field," *Cerebral Cortex*, vol. 22, no. 5, pp. 1052–1067, 2011.
- [16] M. Weinrich, S. Wise, and K.-H. Mauritz, "A neurophysiological study of the premotor cortex in the rhesus monkey," *Brain*, vol. 107, no. 2, pp. 385–414, 1984.
- [17] S. Wise and K.-H. Mauritz, "Set-related neuronal activity in the premotor cortex of rhesus monkeys: effects of changes in motor set," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 223, no. 1232, pp. 331–354, 1985.
- [18] P. Cisek and J. F. Kalaska, "Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action," *Neuron*, vol. 45, no. 5, pp. 801–814, 2005.
- [19] A. Bastian, A. Riehle, W. Erlhagen, and G. Schöner, "Prior information preshapes the population representation of movement direction in motor cortex," *NeuroReport*, vol. 9, no. 2, pp. 315–319, 1998.
- [20] B. M. Dekleva, P. Ramkumar, P. A. Wanda, K. P. Kording, and L. E. Miller, "Uncertainty leads to persistent effects on reach representations in dorsal premotor cortex," *eLife*, vol. 5, p. e14316, 2016.

- [21] D. Thura and P. Cisek, “Deliberation and commitment in the premotor and primary motor cortex during dynamic decision making,” *Neuron*, vol. 81, no. 6, pp. 1401–1416, 2014.
- [22] J. M. Henderson, G. L. Malcolm, and C. Schandl, “Searching in the dark: Cognitive relevance drives attention in real-world scenes,” *Psychonomic Bulletin & Review*, vol. 16, no. 5, pp. 850–856, 2009.
- [23] J. M. Wolfe, “Guided search 2.0 a revised model of visual search,” *Psychonomic bulletin & review*, vol. 1, no. 2, pp. 202–238, 1994.
- [24] D. Ferrier, “The Croonian Lecture: Experiments on the Brain of Monkeys (Second Series),” *Philosophical Transactions of the Royal Society of London*, vol. 165, pp. 433–488, jan 1875.
- [25] C. J. Bruce, M. E. Goldberg, M. C. Bushnell, and G. B. Stanton, “Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements.,” *Journal of neurophysiology*, vol. 54, pp. 714–34, sep 1985.
- [26] K. G. Thompson and N. P. Bichot, “A visual salience map in the primate frontal eye field.,” *Progress in brain research*, vol. 147, pp. 251–62, jan 2005.
- [27] T. Moore and M. 2004, “Microstimulation of the frontal eye field and its effects on covert spatial attention,” *Journal of neurophysiology*, vol. 91, no. 1, pp. 152–162, 2004.
- [28] J. D. Schall, “On the role of frontal eye field in guiding attention and saccades,” *Vision research*, vol. 44, no. 12, pp. 1453–1467, 2004.
- [29] R. J. Schafer and T. Moore, “Attention governs action in the primate frontal eye field,” *Neuron*, vol. 56, no. 3, pp. 541–551, 2007.
- [30] T. J. Buschman and E. K. Miller, “Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations,” *Neuron*, vol. 63, no. 3, pp. 386–396, 2009.
- [31] K. G. Thompson and N. P. Bichot, “A visual salience map in the primate frontal eye field.,” *Progress in brain research*, vol. 147, pp. 251–62, jan 2005.
- [32] J. D. Schall, D. P. Hanes, K. G. Thompson, and D. J. King, “Saccade target selection in frontal eye field of macaque. I. Visual and premovement activation.,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 15, pp. 6905–18, oct 1995.

- [33] H. L. Fernandes, I. H. Stevenson, A. N. Phillips, M. a. Segraves, and K. P. Kording, “Saliency and Saccade Encoding in the Frontal Eye Field During Natural Scene Search.,” *Cerebral cortex (New York, N.Y. : 1991)*, pp. 1–14, jul 2013.
- [34] A. N. Phillips and M. a. Segraves, “Predictive activity in macaque frontal eye field neurons during natural scene searching.,” *Journal of neurophysiology*, vol. 103, pp. 1238–52, mar 2010.
- [35] J. Najemnik and W. S. Geisler, “Optimal eye movement strategies in visual search,” *Nature*, vol. 434, pp. 387–91, mar 2005.
- [36] D. Walther and C. Koch, “Modeling attention to salient proto-objects.,” *Neural networks : the official journal of the International Neural Network Society*, vol. 19, pp. 1395–407, nov 2006.
- [37] C. Kanan, M. H. Tong, L. Zhang, and G. W. Cottrell, “Sun: Top-down saliency using natural statistics,” *Visual Cognition*, vol. 17, no. 6-7, pp. 979–1003, 2009.
- [38] P.-H. Tseng, R. Carmi, I. G. Cameron, D. P. Munoz, and L. Itti, “Quantifying center bias of observers in free viewing of dynamic natural scenes,” *Journal of vision*, vol. 9, no. 7, pp. 4–4, 2009.
- [39] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer series in statistics Springer, Berlin, 2001.
- [40] J. Qian, T. Hastie, J. Friedman, R. Tibshirani, and N. Simon, “Glmnet for matlab,” 2013.
- [41] J. Friedman, T. Hastie, and R. Tibshirani, “Regularization paths for generalized linear models via coordinate descent,” *Journal of statistical software*, vol. 33, no. 1, p. 1, 2010.
- [42] D. McFadden *et al.*, “Conditional logit analysis of qualitative choice behavior,” 1973.
- [43] A. Borji, D. N. Sihite, and L. Itti, “Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study,” *Image Processing, IEEE Transactions on*, vol. 22, no. 1, pp. 55–69, 2013.
- [44] N. P. Bichot, M. T. Heard, E. M. DeGennaro, and R. Desimone, “A Source for Feature-Based Attention in the Prefrontal Cortex.,” *Neuron*, vol. 88, pp. 832–44, nov 2015.

- [45] M. E. Goldberg and C. J. Bruce, "Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal.," *Journal of neurophysiology*, vol. 64, pp. 489–508, aug 1990.
- [46] M. Schlag-Rey, J. Schlag, and P. Dassonville, "How the frontal eye field can impose a saccade goal on superior colliculus neurons," *Journal of Neurophysiology*, vol. 67, no. 4, pp. 1003–1005, 1992.
- [47] K. G. Thompson, N. P. Bichot, and J. D. Schall, "Dissociation of visual discrimination from saccade programming in macaque frontal eye field," *Journal of neurophysiology*, vol. 77, no. 2, pp. 1046–1050, 1997.
- [48] A. Murthy, K. G. Thompson, and J. D. Schall, "Dynamic dissociation of visual selection from saccade programming in frontal eye field.," *Journal of neurophysiology*, vol. 86, pp. 2634–7, nov 2001.
- [49] J. D. Schall and K. G. Thompson, "Neural selection and control of visually guided eye movements," *Annual review of neuroscience*, vol. 22, no. 1, pp. 241–259, 1999.
- [50] J. D. Schall and D. P. Hanes, "Neural basis of saccade target selection in frontal eye field during visual search.," *Nature*, vol. 366, pp. 467–9, dec 1993.
- [51] J. M. Findlay, "Global visual processing for saccadic eye movements," *Vision research*, vol. 22, no. 8, pp. 1033–1045, 1982.
- [52] G. J. Zelinsky, "A theory of eye movements during target acquisition.," *Psychological review*, vol. 115, no. 4, p. 787, 2008.
- [53] G. J. Zelinsky, R. P. Rao, M. M. Hayhoe, and D. H. Ballard, "Eye movements reveal the spatiotemporal dynamics of visual search," *Psychological Science*, pp. 448–453, 1997.
- [54] J. Gottlieb, P. Y. Oudeyer, M. Lopes, and A. Baranes, "Information-seeking, curiosity, and attention: Computational and neural mechanisms," *Trends in Cognitive Sciences*, vol. 17, no. 11, pp. 585–593, 2013.
- [55] D. D. Burman and M. A. Segraves, "Primate frontal eye field activity during natural scanning eye movements.," *Journal of neurophysiology*, vol. 71, pp. 1266–71, mar 1994.
- [56] T. R. Sato and J. D. Schall, "Effects of stimulus-response compatibility on neural selection in frontal eye field.," *Neuron*, vol. 38, pp. 637–48, may 2003.

- [57] T. Ikeda and O. Hikosaka, “Reward-Dependent Gain and Bias of Visual Responses in Primate Superior Colliculus,” *Neuron*, vol. 39, pp. 693–700, aug 2003.
- [58] R. Kawagoe, Y. Takikawa, and O. Hikosaka, “Expectation of reward modulates cognitive signals in the basal ganglia,” *Nature neuroscience*, vol. 1, no. 5, 1998.
- [59] Y. Tachibana and O. Hikosaka, “The primate ventral pallidum encodes expected reward value and regulates motor action,” *Neuron*, vol. 76, no. 4, pp. 826–837, 2012.
- [60] K. Nakamura, M. Matsumoto, and O. Hikosaka, “Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus,” *Journal of Neuroscience*, vol. 28, no. 20, pp. 5331–5343, 2008.
- [61] J. E. Choi, P. A. Vaswani, and R. Shadmehr, “Vigor of movements and the cost of time in decision making,” *Journal of neuroscience*, vol. 34, no. 4, pp. 1212–1223, 2014.
- [62] L. L. Chen, L. Y. Hung, J. Quinet, and K. Kosek, “Cognitive regulation of saccadic velocity by reward prospect,” *European Journal of Neuroscience*, vol. 38, no. 3, pp. 2434–2444, 2013.
- [63] T. R. Reppert, K. M. Lempert, P. W. Glimcher, and R. Shadmehr, “Modulation of saccade vigor during value-based decision making,” *Journal of Neuroscience*, vol. 35, no. 46, pp. 15369–15378, 2015.
- [64] J. H. Maunsell, “Neuronal representations of cognitive state: reward or attention?,” *Trends in cognitive sciences*, vol. 8, no. 6, pp. 261–265, 2004.
- [65] J. W. Bisley and M. E. Goldberg, “Attention, intention, and priority in the parietal lobe,” *Annual review of neuroscience*, vol. 33, pp. 1–21, jan 2010.
- [66] J. Gottlieb, M. Hayhoe, O. Hikosaka, and A. Rangel, “Attention, reward, and information seeking,” *Journal of Neuroscience*, vol. 34, no. 46, pp. 15497–15504, 2014.
- [67] O. Hikosaka, “Basal ganglia mechanisms of reward-oriented eye movement,” *Annals of the New York Academy of Sciences*, vol. 1104, no. 1, pp. 229–249, 2007.
- [68] P. H. Schiller, S. D. True, and J. L. Conway, “Effects of frontal eye field and superior colliculus ablations on eye movements,” *Science*, vol. 206, no. 4418, pp. 590–592, 1979.

- [69] L. Ding and O. Hikosaka, “Comparison of reward modulation in the frontal eye field and caudate of the macaque,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 26, pp. 6695–703, jun 2006.
- [70] M. R. Roesch and C. R. Olson, “Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex,” *Journal of Neurophysiology*, vol. 90, no. 3, pp. 1766–1789, 2003.
- [71] A. Hays Jr, B. Richmond, and L. Optican, “Unix-based multiple-process system, for real-time data acquisition and control,” 1982.
- [72] S. J. Judge, B. J. Richmond, and F. C. Chu, “Implantation of magnetic search coils for measurement of eye position: an improved method,” *Vision research*, vol. 20, no. 6, pp. 535–538, 1980.
- [73] D. A. Robinson, “A method of measuring eye movement using a scleral search coil in a magnetic field,” *IEEE Transactions on bio-medical electronics*, vol. 10, no. 4, pp. 137–145, 1963.
- [74] B. Amirikian and A. P. Georgopoulos, “Directional tuning profiles of motor cortical cells,” *Neuroscience research*, vol. 36, no. 1, pp. 73–79, 2000.
- [75] A. M. Haith, T. R. Reppert, and R. Shadmehr, “Evidence for hyperbolic temporal discounting of reward in control of movements,” *Journal of neuroscience*, vol. 32, no. 34, pp. 11727–11736, 2012.
- [76] U. Beierholm, M. Guitart-Masip, M. Economides, R. Chowdhury, E. Düzel, R. Dolan, and P. Dayan, “Dopamine modulates reward-related vigor,” *Neuropsychopharmacology*, vol. 38, no. 8, p. 1495, 2013.
- [77] A. T. Bahill, M. R. Clark, and L. Stark, “The main sequence, a tool for studying human eye movements,” *Mathematical Biosciences*, vol. 24, no. 3-4, pp. 191–204, 1975.
- [78] D. A. Butts and M. S. Goldman, “Tuning curves, neuronal variability, and sensory coding,” *PLoS biology*, vol. 4, no. 4, p. e92, 2006.
- [79] I. E. Monosov, D. L. Sheinberg, and K. G. Thompson, “Paired neuron recordings in the prefrontal and inferotemporal cortices reveal that spatial selection precedes object identification during visual search,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 29, pp. 13105–13110, 2010.

- [80] M. A. Sommer and R. H. Wurtz, "Composition and topographic organization of signals sent from the frontal eye field to the superior colliculus," *Journal of Neurophysiology*, vol. 83, no. 4, pp. 1979–2001, 2000.
- [81] M. A. Segraves and M. E. Goldberg, "Functional properties of corticotectal neurons in the monkeys frontal eye field," *Journal of Neurophysiology*, vol. 58, no. 6, pp. 1387–1419, 1987.
- [82] M. A. Segraves, "Activity of monkey frontal eye field neurons projecting to oculomotor regions of the pons," *Journal of Neurophysiology*, vol. 68, no. 6, pp. 1967–1985, 1992.
- [83] B. Noudoost, K. L. Clark, and T. Moore, "A distinct contribution of the frontal eye field to the visual representation of saccadic targets," *Journal of Neuroscience*, vol. 34, no. 10, pp. 3687–3698, 2014.
- [84] M. S. Bendiksy and M. L. Platt, "Neural correlates of reward and attention in macaque area lip," *Neuropsychologia*, vol. 44, no. 12, pp. 2411–2420, 2006.
- [85] C. J. Peck, D. C. Jangraw, M. Suzuki, R. Efem, and J. Gottlieb, "Reward modulates attention independently of action value in posterior parietal cortex," *Journal of Neuroscience*, vol. 29, no. 36, pp. 11182–11191, 2009.
- [86] N.-Y. So and V. Stuphorn, "Supplementary eye field encodes option and action value for saccades with variable reward," *Journal of Neurophysiology*, vol. 104, no. 5, pp. 2634–2653, 2010.
- [87] M. I. Leon and M. N. Shadlen, "Effect of Expected Reward Magnitude on the Response of Neurons in the Dorsolateral Prefrontal Cortex of the Macaque," vol. 24, pp. 415–425, 1999.
- [88] S. W. Kennerley and J. D. Wallis, "Reward-dependent modulation of working memory in lateral prefrontal cortex," *Journal of Neuroscience*, vol. 29, no. 10, pp. 3259–3270, 2009.
- [89] W. Schultz, L. Tremblay, and J. R. Hollerman, "Reward processing in primate orbitofrontal cortex and basal ganglia," *Cerebral cortex*, vol. 10, no. 3, pp. 272–283, 2000.
- [90] M. R. Roesch and C. R. Olson, "Neuronal activity related to reward value and motivation in primate frontal cortex," *Science*, vol. 304, no. 5668, pp. 307–310, 2004.

- [91] T. Ikeda and O. Hikosaka, “Positive and negative modulation of motor response in primate superior colliculus by reward expectation,” *Journal of neurophysiology*, vol. 98, no. 6, pp. 3163–3170, 2007.
- [92] O. Hikosaka, H. F. Kim, M. Yasuda, and S. Yamamoto, “Basal ganglia circuits for reward value-guided behavior,” *Annual review of neuroscience*, vol. 37, pp. 289–306, 2014.
- [93] M. Matsumoto and O. Hikosaka, “Two types of dopamine neuron distinctly convey positive and negative motivational signals,” *Nature*, vol. 459, no. 7248, pp. 837–841, 2009.
- [94] M. A. Segraves and K. Park, “The relationship of monkey frontal eye field activity to saccade dynamics,” *Journal of Neurophysiology*, vol. 69, no. 6, pp. 1880–1889, 1993.
- [95] P. Cisek, G. A. Puskas, and S. El-Murr, “Decisions in changing conditions: the urgency-gating model,” *Journal of Neuroscience*, vol. 29, no. 37, pp. 11560–11571, 2009.
- [96] A. K. Churchland, R. Kiani, and M. N. Shadlen, “Decision-making with multiple alternatives,” *Nature neuroscience*, vol. 11, no. 6, pp. 693–702, 2008.
- [97] R. P. Rao, “Decision making under uncertainty: a neural model based on partially observable markov decision processes,” *Frontiers in computational neuroscience*, vol. 4, 2010.
- [98] H. Itoh, H. Nakahara, O. Hikosaka, R. Kawagoe, Y. Takikawa, and K. Aihara, “Correlation of primate caudate neural activity and saccade parameters in reward-oriented behavior,” *Journal of Neurophysiology*, vol. 89, no. 4, pp. 1774–1783, 2003.
- [99] J. D. Wallis and E. L. Rich, “Challenges of interpreting frontal neurons during value-based decision-making,” *Frontiers in neuroscience*, vol. 5, 2011.
- [100] K. Louie, L. E. Grattan, and P. W. Glimcher, “Reward value-based gain control: divisive normalization in parietal cortex,” *Journal of Neuroscience*, vol. 31, no. 29, pp. 10627–10639, 2011.
- [101] A. Litt, H. Plassmann, B. Shiv, and A. Rangel, “Dissociating valuation and saliency signals during decision-making,” *Cerebral cortex*, vol. 21, no. 1, pp. 95–102, 2010.

- [102] M. R. Roesch and C. R. Olson, "Neuronal activity related to anticipated reward in frontal cortex," *Annals of the New York Academy of Sciences*, vol. 1121, no. 1, pp. 431–446, 2007.
- [103] J. L. Gallant, C. E. Connor, and D. C. Van Essen, "Neural activity in areas V1, V2 and V4 during free viewing of natural scenes compared to controlled viewing.," *Neuroreport*, vol. 9, pp. 2153–8, jun 1998.
- [104] S. P. MacEvoy, T. D. Hanks, and M. a. Paradiso, "Macaque V1 activity during natural vision: effects of natural scenes and saccades.," *Journal of neurophysiology*, vol. 99, pp. 460–72, feb 2008.
- [105] W. E. Vinje, "Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision," *Science*, vol. 287, pp. 1273–1276, feb 2000.
- [106] E. T. Rolls and M. J. Tovee, "Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex," *Journal of neurophysiology*, vol. 73, no. 2, pp. 713–726, 1995.
- [107] D. L. Sheinberg and N. K. Logothetis, "Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision," *Journal of Neuroscience*, vol. 21, no. 4, pp. 1340–1350, 2001.
- [108] M. A. Basso and R. H. Wurtz, "Modulation of neuronal activity in superior colliculus by changes in target probability," *The Journal of Neuroscience*, vol. 18, no. 18, pp. 7519–7534, 1998.
- [109] M. C. Dorris and D. P. Munoz, "Saccadic probability influences motor preparation signals and time to saccadic initiation," *The Journal of Neuroscience*, vol. 18, no. 17, pp. 7015–7026, 1998.
- [110] K. Mirpour, F. Arcizet, W. S. Ong, and J. W. Bisley, "Been there, seen that: a neural mechanism for performing efficient visual search," *Journal of Neurophysiology*, vol. 102, no. 6, pp. 3481–3491, 2009.
- [111] N. P. Bichot and J. D. Schall, "Priming in Macaque Frontal Cortex during Popout Visual Search : Feature-Based Facilitation and Location-Based Inhibition of Return," vol. 22, no. 11, pp. 4675–4685, 2002.
- [112] A. N. Phillips and M. a. Segraves, "Predictive activity in macaque frontal eye field neurons during natural scene searching.," *Journal of neurophysiology*, vol. 103, pp. 1238–52, mar 2010.

- [113] M. Par and D. P. Munoz, "Expression of a re-centering bias in saccade regulation by superior colliculus neurons," *Experimental Brain Research*, vol. 137, no. 3-4, pp. 354–368, 2001.
- [114] M. Bindemann, "Scene and screen center bias early eye movements in scene viewing," *Vision research*, vol. 50, no. 23, pp. 2577–2587, 2010.
- [115] G. T. Buswell, "How people look at pictures: a study of the psychology and perception in art," 1935.
- [116] D. Zambarbieri, G. Beltrami, and M. Versino, "Saccade latency toward auditory targets depends on the relative position of the sound source with respect to the eyes," *Vision research*, vol. 35, no. 23, pp. 3305–3312, 1995.
- [117] J. H. Fuller, "Eye position and target amplitude effects on human visual saccadic latencies," *Experimental Brain Research*, vol. 109, no. 3, pp. 457–466, 1996.
- [118] R. M. McPeck and E. L. Keller, "Deficits in saccade target selection after inactivation of superior colliculus," *Nature neuroscience*, vol. 7, no. 7, pp. 757–763, 2004.
- [119] O. Hikosaka and R. Wurtz, "Saccadic eye movements following injection of lidocaine into the superior colliculus," *Experimental Brain Research*, vol. 61, no. 3, pp. 531–539, 1986.
- [120] M. C. Dorris, M. Pare, and D. P. Munoz, "Neuronal activity in monkey superior colliculus related to the initiation of saccadic eye movements," *Journal of Neuroscience*, vol. 17, no. 21, pp. 8566–8579, 1997.
- [121] S. Rivaud, R. Mri, B. Gaymard, A. Vermersch, and C. Pierrot-Deseilligny, "Eye movement disorders after frontal eye field lesions in humans," *Experimental Brain Research*, vol. 102, no. 1, pp. 110–120, 1994.
- [122] K. Lowe and J. D. Schall, "Functional categories in macaque frontal eye field," *bioRxiv*, p. 212589, 2017.
- [123] G. S. Russo and C. J. Bruce, "Effect of eye position within the orbit on electrically elicited saccadic eye movements: a comparison of the macaque monkey's frontal and supplementary eye fields," *Journal of neurophysiology*, vol. 69, no. 3, pp. 800–818, 1993.
- [124] C. R. Cassanello and V. P. Ferrera, "Computing vector differences using a gain fieldlike mechanism in monkey frontal eye field," *The Journal of physiology*, vol. 582, no. 2, pp. 647–664, 2007.

- [125] J. I. Gold and M. N. Shadlen, “Representation of a perceptual decision in developing oculomotor commands,” *Nature*, vol. 404, no. 6776, pp. 390–394, 2000.
- [126] M. Campos, A. Cherian, and M. A. Segraves, “Effects of eye position upon activity of neurons in macaque superior colliculus,” *Journal of neurophysiology*, vol. 95, no. 1, pp. 505–526, 2006.
- [127] A. Van Opstal, K. Hepp, Y. Suzuki, and V. Henn, “Influence of eye position on activity in monkey superior colliculus,” *Journal of Neurophysiology*, vol. 74, no. 4, pp. 1593–1610, 1995.
- [128] M. A. Sommer and R. H. Wurtz, “A pathway in primate brain for internal monitoring of movements,” *Science*, vol. 296, no. 5572, pp. 1480–1482, 2002.
- [129] M. A. Sommer and R. H. Wurtz, “What the brain stem tells the frontal cortex. ii. role of the sc-md-fef pathway in corollary discharge,” *Journal of neurophysiology*, vol. 91, no. 3, pp. 1403–1423, 2004.
- [130] M. a. Basso and R. H. Wurtz, “Modulation of neuronal activity by target uncertainty,” *Nature*, vol. 389, no. 6646, pp. 66–69, 1997.
- [131] T. Yang and M. N. Shadlen, “Probabilistic reasoning by neurons,” *Nature*, vol. 447, no. 7148, pp. 1075–1080, 2007.
- [132] T. Waldhr, G. Haidinger, and E. Schober, “Comparison of r2 measures for poisson regression by simulation,” *Journal of Epidemiology and Biostatistics*, vol. 3, no. 2, pp. 209–215, 1998.
- [133] A. M. Haith, D. M. Huberdeau, and J. W. Krakauer, “Hedging your bets: Intermediate movements as optimal behavior in the context of an incomplete decision,” *PLOS Computational Biology*, vol. 11, no. 3, p. e1004171, 2015.
- [134] D. J. Crammond and J. F. Kalaska, “Prior information in motor and premotor cortex: activity during the delay period and effect on pre-movement activity,” *Journal of Neurophysiology*, vol. 84, no. 2, pp. 986–1005, 2000.
- [135] M. Johnson, J. Coltz, M. Hagen, and T. Ebner, “Visuomotor processing as reflected in the directional discharge of premotor and primary motor cortex neurons,” *Journal of neurophysiology*, vol. 81, no. 2, pp. 875–894, 1999.
- [136] M. M. Churchland, J. P. Cunningham, M. T. Kaufman, S. I. Ryu, and K. V. Shenoy, “Cortical preparatory activity: representation of movement or first cog in a dynamical machine?,” *Neuron*, vol. 68, no. 3, pp. 387–400, 2010.

- [137] R. Caminiti, P. B. Johnson, C. Galli, S. Ferraina, and Y. Burnod, "Making arm movements within different parts of space: the premotor and motor cortical representation of a coordinate system for reaching to visual targets," *The Journal of neuroscience*, vol. 11, no. 5, pp. 1182–1197, 1991.
- [138] B. Pesaran, M. J. Nelson, and R. A. Andersen, "Dorsal premotor neurons encode the relative position of the hand, eye, and goal during reach planning," *Neuron*, vol. 51, no. 1, pp. 125–134, 2006.
- [139] A. Georgopoulos, R. Caminiti, and J. Kalaska, "Static spatial effects in motor cortex and area 5: quantitative relations in a two-dimensional space," *Experimental Brain Research*, vol. 54, no. 3, pp. 446–454, 1984.
- [140] W. Wang, S. S. Chan, D. A. Heldman, and D. W. Moran, "Motor cortical representation of position and velocity during reaching," *Journal of neurophysiology*, vol. 97, no. 6, pp. 4258–4270, 2007.
- [141] I. Cos, N. Blanger, and P. Cisek, "The influence of predicted arm biomechanics on decision making," *Journal of neurophysiology*, vol. 105, no. 6, pp. 3022–3033, 2011.
- [142] D. K. Wood, P. A. Chouinard, A. J. Major, and M. A. Goodale, "Sensitivity to biomechanical limitations during postural decision-making depends on the integrity of posterior superior parietal cortex," *Cortex*.
- [143] S. Deneve, "Bayesian spiking neurons i: inference," *Neural computation*, vol. 20, no. 1, pp. 91–117, 2008.
- [144] B. Kim and M. A. Basso, "A probabilistic strategy for understanding action selection," *The Journal of Neuroscience*, vol. 30, no. 6, pp. 2340–2355, 2010.
- [145] J. M. Beck, W. J. Ma, R. Kiani, T. Hanks, A. K. Churchland, J. Roitman, M. N. Shadlen, P. E. Latham, and A. Pouget, "Probabilistic population codes for bayesian decision making," *Neuron*, vol. 60, no. 6, pp. 1142–1152, 2008.
- [146] X.-J. Wang, "Decision making in recurrent neuronal circuits," *Neuron*, vol. 60, no. 2, pp. 215–234, 2008.
- [147] S. P. Tipper, L. A. Howard, and G. Houghton, "Actionbased mechanisms of attention," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 353, no. 1373, pp. 1385–1393, 1998.
- [148] W. Erlhagen and G. Schner, "Dynamic field theory of movement preparation," *Psychological review*, vol. 109, no. 3, p. 545, 2002.

- [149] P. Cisek, “Integrated neural processes for defining potential actions and deciding between them: a computational model,” *Journal of Neuroscience*, vol. 26, no. 38, pp. 9761–9770, 2006.
- [150] M. Furman and X.-J. Wang, “Similarity effect and optimal control of multiple-choice decision making,” *Neuron*, vol. 60, no. 6, pp. 1153–1168, 2008.
- [151] C. Klaes, S. Schneegans, G. Schnier, and A. Gail, “Sensorimotor learning biases choice behavior: a learning neural field model for decision making,” *PLoS computational biology*, vol. 8, no. 11, p. e1002774, 2012.
- [152] L. Shen and G. E. Alexander, “Preferential representation of instructed target location versus limb trajectory in dorsal premotor area,” *Journal of Neurophysiology*, vol. 77, no. 3, pp. 1195–1212, 1997.
- [153] A. B. Schwartz, D. W. Moran, and G. A. Reina, “Differential representation of perception and action in the frontal cortex,” *Science*, vol. 303, no. 5656, pp. 380–383, 2004.
- [154] P. Cisek, D. J. Crammond, and J. F. Kalaska, “Neural activity in primary motor and dorsal premotor cortex in reaching tasks with the contralateral versus ipsilateral arm,” *Journal of neurophysiology*, vol. 89, no. 2, pp. 922–942, 2003.
- [155] A. Riehle and J. Requin, “Monkey primary motor and premotor cortex: single-cell activity related to prior information about direction and extent of an intended movement,” *Journal of Neurophysiology*, vol. 61, no. 3, pp. 534–549, 1989.
- [156] C. Klaes, S. Westendorff, S. Chakrabarti, and A. Gail, “Choosing goals, not rules: deciding among rule-based action plans,” *Neuron*, vol. 70, no. 3, pp. 536–548, 2011.
- [157] J. Rickert, A. Riehle, A. Aertsen, S. Rotter, and M. P. Nawrot, “Dynamic encoding of movement direction in motor cortical neurons,” *The Journal of Neuroscience*, vol. 29, no. 44, pp. 13870–13882, 2009.
- [158] M. D. Serruya, N. G. Hatsopoulos, L. Paninski, M. R. Fellows, and J. P. Donoghue, “Brain-machine interface: Instant neural control of a movement signal,” *Nature*, vol. 416, no. 6877, pp. 141–142, 2002.
- [159] R. D. Flint, E. W. Lindberg, L. R. Jordan, L. E. Miller, and M. W. Slutzky, “Accurate decoding of reaching movements from field potentials in the absence of spikes,” *Journal of neural engineering*, vol. 9, no. 4, p. 046006, 2012.

- [160] D. Raposo, M. T. Kaufman, and A. K. Churchland, “A category-free neural population supports evolving demands during decision-making,” *Nature neuroscience*, vol. 17, no. 12, pp. 1784–1792, 2014.
- [161] E. L. Rich and J. D. Wallis, “Decoding subjective decisions from orbitofrontal cortex,” *Nature neuroscience*, vol. 19, no. 7, pp. 973–980, 2016.
- [162] C. P. Hung, G. Kreiman, T. Poggio, and J. J. DiCarlo, “Fast readout of object identity from macaque inferior temporal cortex,” *Science*, vol. 310, no. 5749, pp. 863–866, 2005.
- [163] R. Q. Quiroga, L. H. Snyder, A. P. Batista, H. Cui, and R. A. Andersen, “Movement intention is better predicted than attention in the posterior parietal cortex,” *Journal of neuroscience*, vol. 26, no. 13, pp. 3615–3620, 2006.
- [164] A. Hernandez, V. Ncher, R. Luna, A. Zainos, L. Lemus, M. Alvarez, Y. Vzquez, L. Camarillo, and R. Romo, “Decoding a perceptual decision process across cortex,” *Neuron*, vol. 66, no. 2, pp. 300–314, 2010.
- [165] M. A. van der Meer, A. Johnson, N. C. Schmitzer-Torbert, and A. D. Redish, “Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task,” *Neuron*, vol. 67, no. 1, pp. 25–32, 2010.
- [166] M. Weygandt, C. R. Blecker, A. Schfer, K. Hackmack, J.-D. Haynes, D. Vaitl, R. Stark, and A. Schienle, “fmri pattern recognition in obsessivecompulsive disorder,” *Neuroimage*, vol. 60, no. 2, pp. 1186–1193, 2012.
- [167] J. L. Collinger, B. Wodlinger, J. E. Downey, W. Wang, E. C. Tyler-Kabara, D. J. Weber, A. J. McMorland, M. Velliste, M. L. Boninger, and A. B. Schwartz, “High-performance neuroprosthetic control by an individual with tetraplegia,” *The Lancet*, vol. 381, no. 9866, pp. 557–564, 2013.
- [168] C. Ethier, E. R. Oby, M. J. Bauman, and L. E. Miller, “Restoration of grasp following paralysis through brain-controlled stimulation of muscles,” *Nature*, vol. 485, no. 7398, pp. 368–71, 2012.
- [169] A. S. Benjamin, H. L. Fernandes, T. Tomlinson, P. Ramkumar, C. VerSteeg, L. Miller, and K. P. Kording, “Modern machine learning far outperforms glms at predicting spikes,” *bioRxiv*, p. 111450, 2017.

- [170] B. M. London and L. E. Miller, “Responses of somatosensory area 2 neurons to actively and passively generated limb movements,” *Journal of neurophysiology*, vol. 109, no. 6, pp. 1505–1513, 2013.
- [171] A. H. Fagg, G. W. Ojakangas, L. E. Miller, and N. G. Hatsopoulos, “Kinetic trajectory decoding using motor cortical ensembles,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 17, no. 5, pp. 487–496, 2009.
- [172] J. Snoek, H. Larochelle, and R. P. Adams, “Practical bayesian optimization of machine learning algorithms,” in *Advances in neural information processing systems*, pp. 2951–2959.
- [173] E. A. Pohlmeyer, S. A. Solla, E. J. Perreault, and L. E. Miller, “Prediction of upper limb muscle activity from motor cortical discharge during reaching,” *Journal of neural engineering*, vol. 4, no. 4, p. 369, 2007.
- [174] C.-C. Chang and C.-J. Lin, “Libsvm: a library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.
- [175] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, ACM.
- [176] F. Chollet *et al.*, “Keras,” 2015.
- [177] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [178] D. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [179] T. Tieleman and G. Hinton, “Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude.,” *COURSERA: Neural Networks for Machine Learning*, 2012.
- [180] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” *arXiv preprint arXiv:1406.1078*, 2014.
- [181] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

- [182] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, pp. 1097–1105.
- [183] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, “Understanding deep learning requires rethinking generalization,” *arXiv preprint arXiv:1611.03530*, 2016.
- [184] K. V. Shenoy, M. Sahani, and M. M. Churchland, “Cortical control of arm movements: a dynamical systems perspective,” *Annu Rev Neurosci*, vol. 36, pp. 337–59, 2013.
- [185] J. C. Kao, S. D. Stavisky, D. Sussillo, P. Nuyujukian, and K. V. Shenoy, “Information systems opportunities in brainmachine interface decoders,” *Proceedings of the IEEE*, vol. 102, no. 5, pp. 666–682, 2014.
- [186] L. F. Nicolas-Alonso and J. Gomez-Gil, “Brain computer interfaces, a review,” *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.
- [187] D. Sussillo, P. Nuyujukian, J. M. Fan, J. C. Kao, S. D. Stavisky, S. Ryu, and K. Shenoy, “A recurrent neural network for closed-loop intracortical brainmachine interface decoders,” *Journal of neural engineering*, vol. 9, no. 2, p. 026027, 2012.
- [188] J. C. Kao, P. Nuyujukian, S. I. Ryu, M. M. Churchland, J. P. Cunningham, and K. V. Shenoy, “Single-trial dynamics of motor cortex and their applications to brain-machine interfaces,” *Nature communications*, vol. 6, 2015.
- [189] E. Corbett, E. Perreault, and K. Koording, “Mixture of time-warped trajectory models for movement decoding,” in *Advances in Neural Information Processing Systems*, pp. 433–441.
- [190] J. C. Kao, P. Nuyujukian, S. I. Ryu, and K. V. Shenoy, “A high-performance neural prosthesis incorporating discrete state selection with hidden markov models,” *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 4, pp. 935–945, 2017.
- [191] R. Ratcliff, “Modeling response signal and response time data,” *Cognitive psychology*, vol. 53, no. 3, pp. 195–237, 2006.
- [192] M. Usher and J. L. McClelland, “The time course of perceptual choice: the leaky, competing accumulator model,” *Psychological review*, vol. 108, no. 3, p. 550, 2001.
- [193] F. E. Theunissen, K. Sen, and A. J. Doupe, “Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds,” *Journal of Neuroscience*, vol. 20, no. 6, pp. 2315–2331, 2000.

- [194] W. E. Vinje and J. L. Gallant, “Natural stimulation of the nonclassical receptive field increases information transmission efficiency in v1,” *Journal of Neuroscience*, vol. 22, no. 7, pp. 2904–2915, 2002.