Frontal eye field neurons predict "anti-Bayesian" but not Bayesian judgments of visual stability across saccades

Divya Subramanian[1,2,3*], John Pearson[1,2,4,6], Marc A. Sommer[1,2,5,6,]

[1]Department of Neurobiology, Duke School of Medicine, Duke University, Durham, NC, 27710, USA

[2]Center for Cognitive Neuroscience, Duke University, Durham, NC, 27708, USA

[3]Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD, 20892, USA

[4]Department of Biostatistics & Bioinformatics, Duke School of Medicine, Duke University, 27710, Durham, NC, USA

[5]Department of Biomedical Engineering, Pratt School of Engineering, Duke University, Durham, NC, 27708, USA

[6]Department of Psychology & Neuroscience, Trinity College of Arts and Sciences, Duke University, Durham, NC, 27708, USA

*Correspondence: divya.subramanian@nih.gov

Acknowledgments

We would like to thank Jessi Cruger for helping with the care of the non-human primates used in this study. We would also like to thank Steve Lisberger, Jeff Beck, and Greg Field for their insights, feedback, and discussions on the work. This work was supported by a 2020 Duke Institute for Brain Sciences Research Incubator Award to JMP and MAS.

Abstract

Bayesian models, in which priors are used to optimally compensate for sensory uncertainty, have had wide-ranging success in explaining behavior across sensorimotor contexts. We recently reported, however, that humans and monkeys use a combination of Bayesian and non-Bayesian strategies when making categorical judgments of visual stability across saccades. While they used priors to compensate for internal, *movement*-driven sensory uncertainty, consistent with Bayesian predictions, they decreased their use of priors when faced with external, *visual image* uncertainty—an “anti-Bayesian” adjustment consistent with the use of a simple classifier. Here, we tested for neural correlates of these Bayesian and classifier-based strategies in the frontal eye field (FEF), a prefrontal region shown to be important for the perception of visual stability across saccades. We recorded from single FEF neurons while two rhesus macaques performed the internal (motor) and external (image) noise tasks in each session, interleaved trial by trial. FEF activity correlated with and predicted the anti-Bayesian, but not the Bayesian, behavior. These results suggest that the two computational strategies for visual stability are implemented by distinct neural circuits and provide a first step toward an integrated understanding of the computational and neural mechanisms underlying visual perception across saccades.

Introduction

A fundamental challenge for perception is to identify consistent attributes of stimuli based on changing sensory inputs. For example, an apple can be successfully identified under a wide range of viewing angles, luminance levels, and locations in the visual field, and even though repeated, identical viewings of it will evoke slightly different patterns of neural activity each time. How do perceptual systems resolve such sensory uncertainty without conscious effort?

This challenge – was the current pattern of neural activity caused by an apple or a pear? – may be cast as a perceptual decision-making problem[1-4] , and two broad classes of computational strategies have been proposed as solutions: Bayesian models resolve uncertainty by relying on prior beliefs about the natural statistics of the environment and other perceptual regularities, and behavioral performance is consistent with Bayesian predictions in many simple, sensorimotor contexts[5-13]. An alternative solution, common in everyday life, is to use heuristics or cognitive shortcuts rather than optimally weighing the costs and benefits of each option. Heuristic models for higher-order cognitive decisions have been widely demonstrated in both humans and monkeys[14-20], and there is a growing interest in their contributions to perceptual processing[21-25]. We recently showed that for active visual perception, in which primates move their eyes to see the world around them, Bayesian and heuristic strategies may be combined in the same task context[26]. Here, we investigated how the two strategies are integrated in the primate brain during active visual perception.

Our entry point for this investigation is natural viewing. Primates view the world using rapid eye movements called saccades. On the retinas, each saccade causes the projected image to move. This complicates the identification of a fundamental attribute of objects in the world: whether they are stable or moving. The visual system needs to disambiguate the saccade-induced retinal shifts from movements in the visual world. One way the visual system does this is by using a copy of the saccadic command (corollary discharge) to predict the

expected visual inputs after a saccade[27]. Inputs that match the prediction can be attributed to one's own movement and mismatches to external movement. This process is susceptible to two major causes of sensory uncertainty: 1) internally generated noise caused by a degradation of visual processing around the time of saccades, called saccadic suppression[28] (*motor-driven noise*), and 2) external noise in the visual input *(image noise*). In our previous study, we tested, in humans and rhesus macaques, the Bayesian prediction that the use of priors should increase with increasing sensory uncertainty. In that experiment, subjects reported whether they perceived the movement of a visual probe that was displaced either while they made saccades or kept their eyes stationary. Varying the internally generated, motor-driven noise (by either requiring a saccade or not) recruited a Bayesian strategy: prior use increased with increasing noise. By contrast, varying the external image noise (by blurring the probe) resulted in the opposite pattern: prior use decreased with increasing noise in an "anti-Bayesian," manner. The anti-Bayesian results were best explained as resulting from the use of a simpler, heuristic-like model: a sensory classifier that learned to categorize visual inputs into response categories.

A primary candidate region for investigating how the Bayesian and classifier-based strategies may be neurally implemented for perception is the frontal eye field (FEF), a critical prefrontal node where motor and sensory information are combined for the perception of stability across saccades (Fig. 1a). The FEF receives corollary discharge signals that are used by some of its neurons to remap receptive fields to their post-saccadic locations even before the onset of the saccade[29-31]. Since remapping allows neurons to sample the same region of space before and after a saccade, it can serve as the basis for predicting the expected post-saccadic visual input[32-33]. Consistent with this idea, inactivating the corollary discharge input to FEF alters the perception of visual stability[34]. Moreover, the FEF receives retinocentric information[35] which it can evaluate for mismatches with the prediction. Indeed, neurons in FEF signal this mismatch, providing the relevant sensory evidence required for the judgment of stability in our paradigm[36].

We recorded from single neurons in FEF while two rhesus macaques performed the image- and motor-noise tasks, interleaved trial by trial, during each session. After confirming that FEF neurons were selective for both learned priors and sensory evidence, we then evaluated whether the neuronal prior selectivity increased with motor-driven noise (the Bayesian prediction) or decreased with image noise (an “anti-Bayesian” prediction). Surprisingly, we found that FEF activity did not correlate with increasing prior use in the motor noise task but did correlate with and predict decreasing prior use in the image noise task. That is, the activity of FEF neurons deviated from the Bayesian prediction, suggesting that FEF is part of a circuit that selectively instantiates a sensory classifier for judgments of visual stability across saccades.

Results

Behavioral prior use increased with motor-driven noise but decreased with image noise

Two rhesus macaques (Monkey S and Monkey T, both males) performed a Saccadic Suppression of Displacement task[37] in which they were trained to report whether a visual stimulus in the periphery moved during an instructed saccade (Fig. 1b). The task was designed to test the Bayesian hypothesis that learned priors are used more with greater sensory uncertainty, either induced by the saccade itself or by noise in the external image. Subjects first fixated a central square and then, on most trials, made a saccade to a target square. A Gaussian visual stimulus (“probe”), initially presented during fixation, was then displaced by varying amounts *during* the saccade. If the probe moved, subjects were required to report it with a second saccade to the probe. If it did not move, they were required to withhold their second saccade and maintain fixation on the target.

In conditions where the probe moved during a saccade, the instructed “intervening” saccade was always directed away from the recorded neuron’s receptive field (RF) and the probe jumped parallel to the saccade vector to land in the center of the RF after the saccade. The probability of a probe jump (0.2 or 0.8) was indicated by the color of the fixation and target

square. Two noise manipulations defined two experiments conducted on randomly interleaved trials (Fig. 1c; a complete visual depiction of the trial types along with their spatial layout is provided in Fig. S1): For the *image noise experiment*, the width of the Gaussian probe had low, medium, or high noise, and all trials required an intervening saccade to the target. For the *motor-driven noise experiment*, uncertainty was manipulated by either requiring an intervening saccade (high noise) or requiring steady fixation instead (low noise), all while a low image noise Gaussian probe moved. Critically, the two experiments had a condition in common: the low image noise condition with an intervening saccade occurred as a trial type for both.

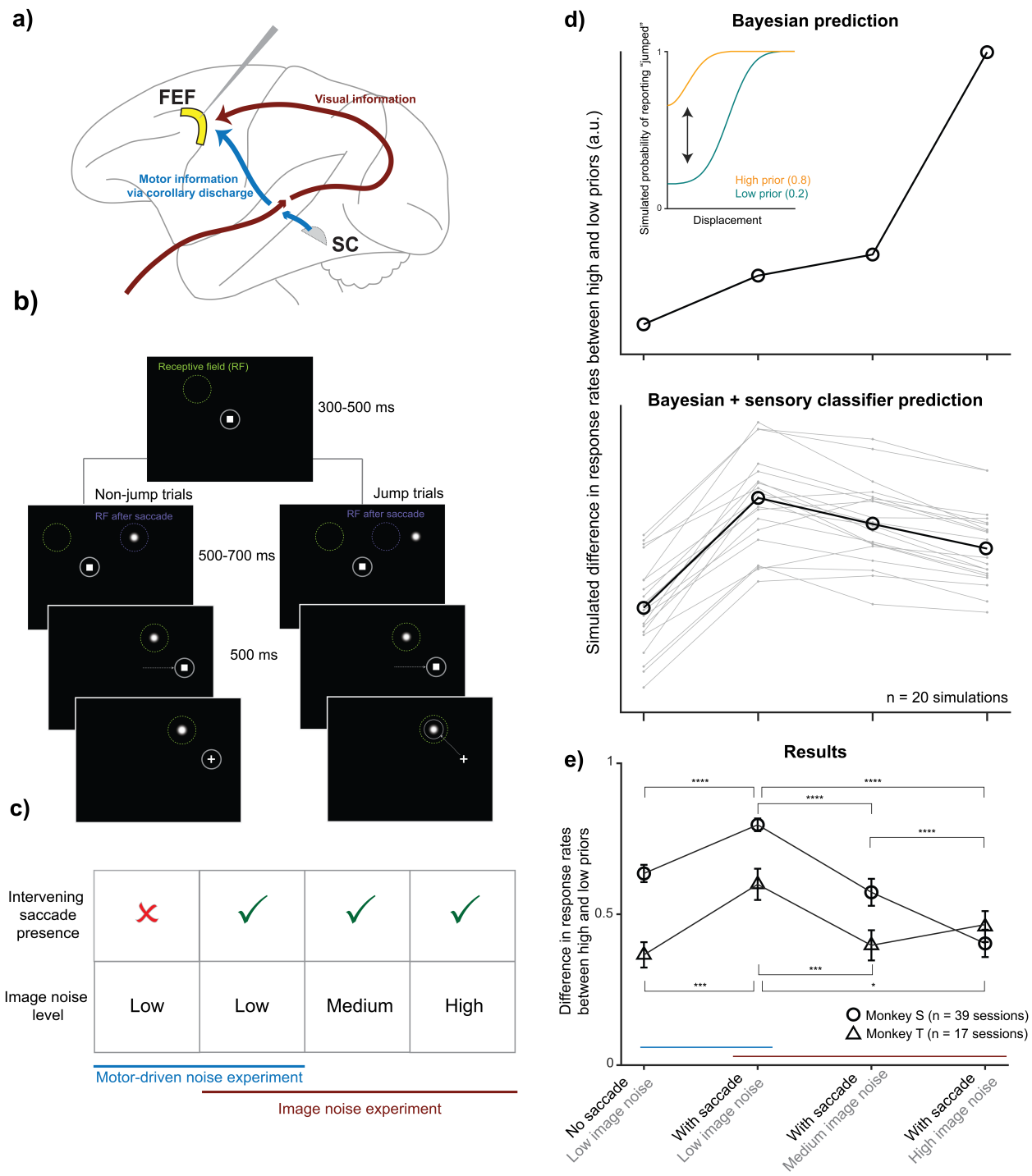

Figure 1. Experimental design and behavior. a) We recorded from single neurons in the frontal eye field (FEF), which receives visual and saccade-related inputs. SC, superior colliculus. b) Typical spatial configuration in the task. Subjects initiate a trial by fixating on a central square. White circle depicts the eye position. A Gaussian visual probe was presented at this time. On jump trials, the probe was presented near the center of the post-saccadic receptive field (purple circle) and for non-jump trials, it was presented at the center. After a brief fixation period, the square moved to the target location and subjects followed it with a saccade (dashed

white arrow), during which the probe was moved on jump trials. For both jump and non-jump trials, the probe is at the center of the receptive field when the subject makes a response. A perceived jump was then reported with a second saccade to the probe. Perceived non-jumps were reported by maintaining fixation. c) The two experiments included 4 noise conditions in total, with one condition in common between them. The motor noise experiment included two conditions, one where the subject made an intervening saccade (as shown in b) and another where they maintained central fixation (shown in Fig. S1) during probe movement. Visual noise was manipulated by the width of the Gaussian probe. d) Top: The Bayesian model for this task (Subramanian et al., 2023[26] and Fig. S2) predicts that prior use, quantified as the difference in response rates between the high and low priors, increases steadily with increasing motor and image noise. Inset depicts model-simulated psychometric curves for the high (orange) and low (teal) priors. Bottom: A model combining Bayesian ideal observer responses for motor noise with sensory classifier responses for image noise predicts that prior use increases with increasing motor noise but decreases with additional image noise. e) Data from the two monkeys is consistent with the predictions of the combined Bayesian and sensory classifier model. Error bars indicate the standard error of the mean. Prior use increased with motor noise, consistent with Bayesian predictions, but decreased with additional image noise, consistent with the use of a sensory classifier. Asterisks at the top for Monkey S and on the bottom for Monkey T. *, $p < 0.05$; ***, $p < 0.001$; ****, $p < 0.0001$.

For this task, a Bayesian ideal observer model predicts an increase in prior use – measured by the difference in response rates across priors (Fig. 1d top, inset) – with increasing noise (Fig. 1d, top; detailed model schematic and predicted psychometric curves for all noise conditions are shown in Fig. S2). As we reported previously[26], subjects relied more on priors with increasing motor noise, consistent with Bayesian predictions, but they relied *less* on priors with increasing image noise (Fig. 1e). The difference in response rates across priors was greater for high compared to low motor-driven noise (Monkey S: $p=1.8\times10^{-6}$ and Monkey T: $p=9.32\times10^{-4}$; paired t-tests). However, it decreased with greater image noise, consistent with the use of a sensory classifier rather than Bayesian inference (Monkey S: $F(2) = 50.5$; $p=1.1\times10^{-14}$ on a repeated-measures ANOVA with $p=5.04\times10^{-6}$ for low vs. medium noise, $p=1.60\times10^{-10}$ for low vs. high noise, and $p=4.74\times10^{-7}$ for medium vs. high noise; Monkey T: $F(2) =8.67$, $p=0.001$ on a repeated-measures ANOVA with $p=5.38\times10^{-4}$ for low vs. medium noise, $p=0.01$ for low vs. high noise, and no significant difference between medium and high noise ($p=0.27$). A model that uses Bayesian ideal observer predictions when faced with motor noise and sensory classification when faced with image noise explained the results (Fig. 1d, bottom)[26]. Detailed

schematic and example simulated psychometric curves using the combined Bayesian and sensory classifier model are shown in Fig. S3.

FEF firing rates reflected the prior, but only when the sensory evidence was available

We first assessed whether the activity of FEF neurons reflected priors at all. Analyses were based on 79 FEF neurons for which there were at least 10 trials completed in each task condition (53 from Monkey S, 26 from Monkey T). Here we focused on the common condition (low image noise with an intervening saccade) in which behavioral prior use was the highest (Fig. 1e) because it serves as the comparison condition between the two experiments.

Since FEF activity is sensitive to the jump size[36], and jump sizes are unequally represented across the two prior conditions, we first confirmed that the results were qualitatively similar even when jump size was restricted to 0 (Fig. S4). Then we pooled across jump sizes to maximize statistical power and ensure approximately equal numbers of trials across the prior conditions. As shown in Figure 2a, population firing rates in the low (0.2, teal) and high (0.8, orange) prior conditions were different. Yet this difference did not emerge when the prior for the trial was first revealed to the animal, but only when the sensory evidence about the displacement entered the RF after the intervening saccade. This type of post-saccadic sensory stimulation of the RF is called *reafference*. The prior-modulated firing rate difference persisted from the start of the reafferent stimulation, past the median response time (611 ms), and then converged back together around 800 ms after the reafferent event.

To quantify this prior-modulated activity, we focused on two task epochs (highlighted in gray in Fig. 2a) during which the spatial locations of the fixation square, target square, and Gaussian probe relative to the RF were the same across all task conditions (Fig. S1). While only 12.6% of the neurons (10/79) showed a statistically significant prior selectivity in the fixation

epoch (0-250 ms from fixation acquisition, Fig. 2b), over half the neurons (58.2%, 46/79) were selective in the post-jump epoch (0-800 ms from the reafferent event, Fig. 2c; significant effect sizes, shaded, were evaluated using Wilcoxon signed-rank tests with α = 0.025 to correct for two comparisons). Distributions for each animal are shown in Fig. S5. We analyzed the time course of the effect in the post-jump epoch using successive 100 ms bins (with 50 ms overlaps). 72.2% of the neurons (57/79) were significantly selective during at least one 100 ms bin (Fig. 2d; we used α = 0.00167 (0.025/15) to further correct for 15 comparisons), and many neurons remained persistently selective rather than exhibiting a clear sequencing of activity across neurons (Fig. 2e). In summary, FEF neuron firing reflected the prior, but only after sensory evidence about the jump was available. We thus focused our subsequent analyses on this reafferent period of the task.

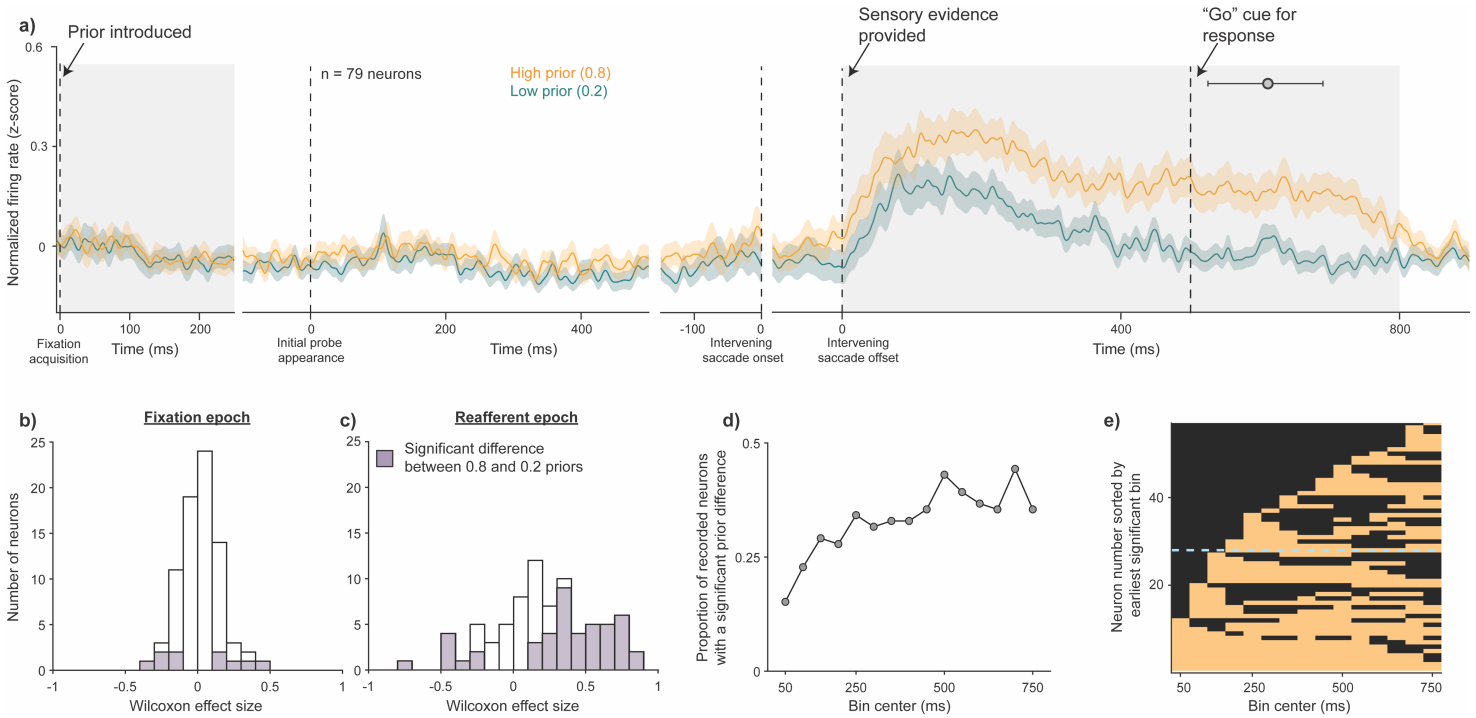

Figure 2. FEF activity reflected the prior but only when the sensory evidence was presented. Data are from the common condition (low image noise with an intervening saccade; Fig. 1c bold outline). a) Normalized peri-stimulus time histograms averaged across all 79 neurons aligned to various task events. A clear difference between the high (orange) and low (teal) prior responses emerged not when the prior was initially introduced at fixation acquisition

but only when the sensory evidence for the task (i.e., the post-saccadic location of the probe) was presented. Gray circle at 611 ms indicates the median response time on trials where a jump was reported with a saccade (error bars range from the 25th to 75th quartiles). b-c) Histograms of Wilcoxon effect sizes between the high and low prior conditions in the fixation (0-250 ms from fixation acquisition) and reafferent (0 – 800 ms from the intervening saccade offset) epochs. Cells with significant differences between the prior conditions ($p < 0.025$ to correct for two comparisons) are shown in purple. d) Proportion of total neurons with significant prior effects during each overlapping 100 ms bin ($p < 0.00167$ to correct for 15 comparisons). e) Neurons ordered by the earliest significant bin. Almost half of the significant cells (dashed blue line) showed a prior effect within 150 ms of the reafferent event.

FEF neurons predicted anti-Bayesian but not Bayesian prior use

Next, we asked how the difference in firing rates between prior cues identified in the common condition (second column of all rows in Fig. 3) was modulated by motor noise or image noise. Possible outcomes were that the prior signal might be modulated 1) only by motor noise but not image noise (i.e., Bayesian, Fig. 3a); 2) only by image noise but not motor noise (anti-Bayesian, Fig. 3b); 3) by neither the image nor the motor noise (Fig. 3c), suggesting that FEF neurons report the prior but do not reflect either computational strategy for visual stability; or 4) by both the image and motor noise (Fig. 3d), suggesting that FEF neurons do not differentiate between the computations but simply reflect the final behavioral output.

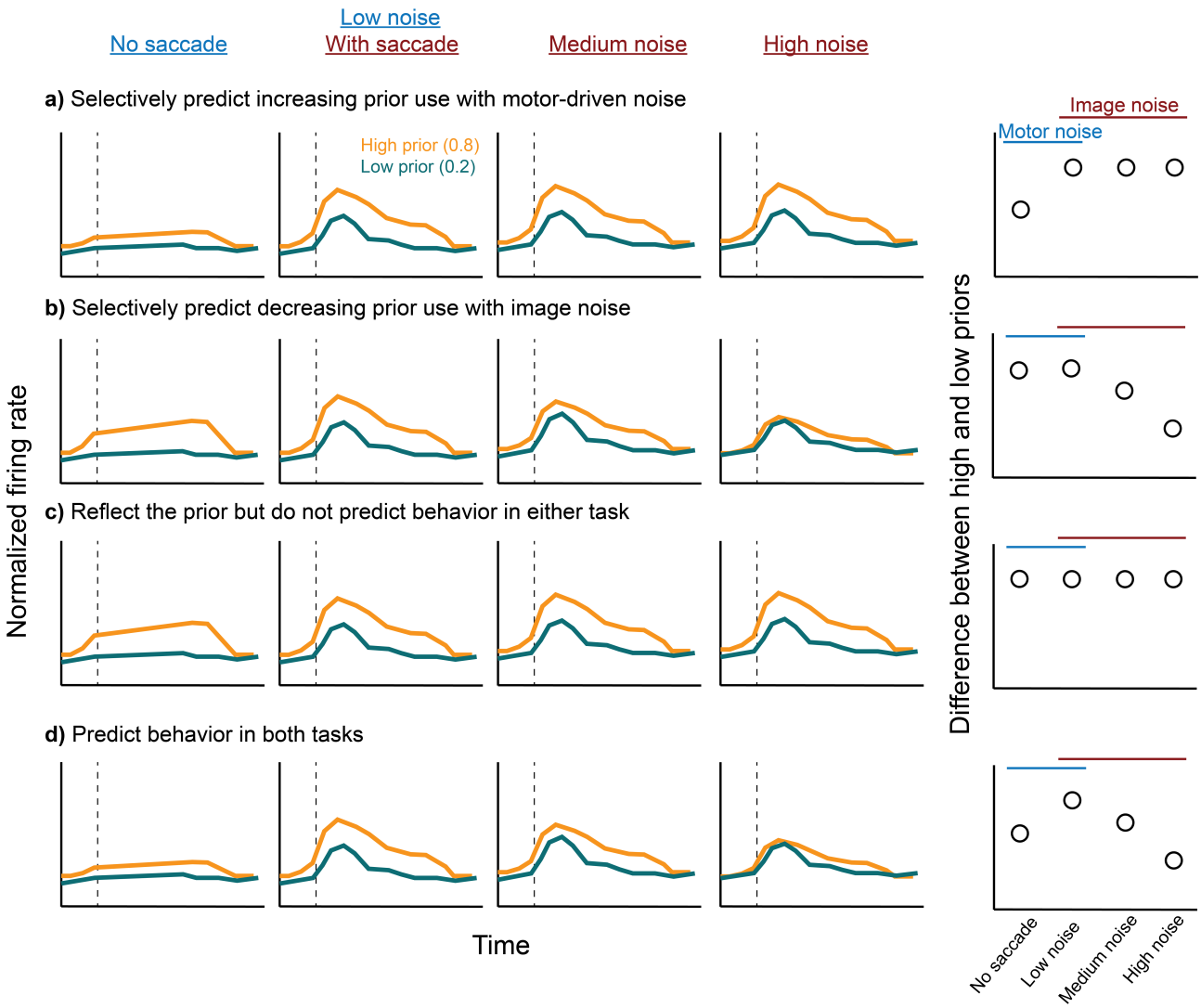

Figure 3. Predicted outcomes for prior use across the four noise conditions. Each row shows a different prediction. Columns 1-4 show cartoons of peristimulus time histograms in the high (orange) and low (teal) prior conditions, aligned to the reafferent event (dashed line), for the four noise conditions across the columns (see label above each column). The second column depicts the common condition (same across all predictions). The fifth column summarizes the prior difference across the four noise conditions.

The results most closely matched the second possible outcome (Fig. 3b): the prior signal was not modulated by motor noise but, instead, only by image noise (Fig. 4). The difference in firing rates between priors did not change between the no-saccade (Fig. 4a) and the saccade condition, which was also the common condition (Fig. 4b). However, the difference in firing rates between priors decreased from the common condition, which used low image noise, to the

medium (Fig. 4c) and high image noise (Fig. 4d) conditions. The same effect was found with the jump size restricted to 0 (Fig. S6). Statistically, the Wilcoxon effect sizes of the difference between priors across all 79 neurons were not significantly different between the no-saccade condition and the common condition ($p = 0.81$, paired t-test), but they were significantly lower than the common condition in the medium ($p = 0.013$) and high ($p = 0.0047$) image noise conditions (Fig. 4e).

Behavioral prior use could vary substantially from session to session, which provided an opportunity to quantify whether the FEF neuronal activity was predictive of behavior. For neurons with significant prior selectivity in the common condition (i.e., $p < 0.025$, corrected for comparisons in the fixation and post-jump epochs), we found that neuronal prior use predicted behavioral prior use session-by-session in the image noise but not motor noise experiment. The correlation between neuronal and behavioral prior use, both divided by the common condition, was not significant for the no-saccade condition (Fig. 4f, $p = 0.21$) but was significant with a positive correlation for the medium (Fig. 4g) and high (Fig. 4h) image noise conditions ($r = 0.41$, $p = 0.0050$ for medium noise and $r = 0.37$, $p = 0.0151$ for high noise; Pearson correlations).

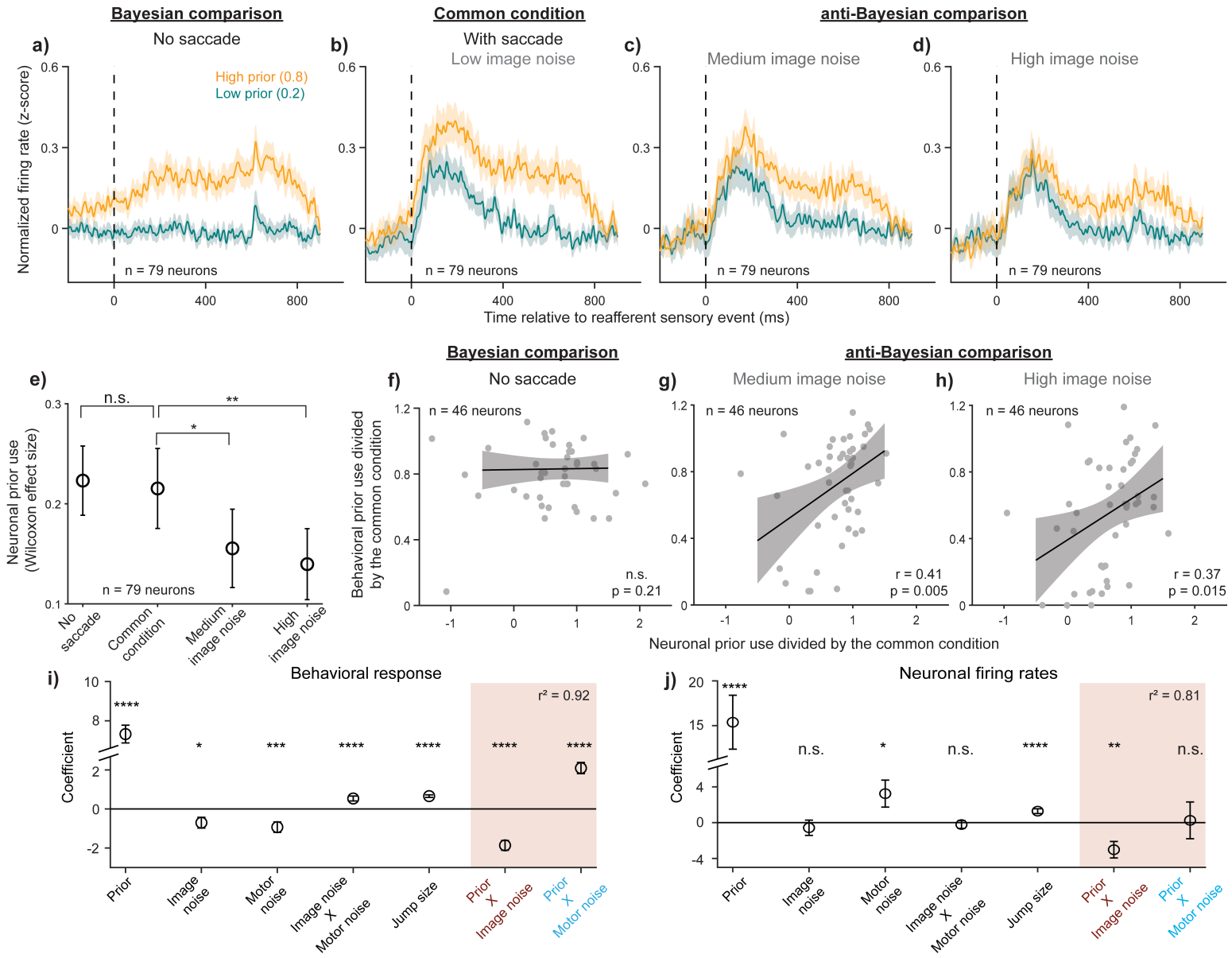

Figure 4. FEF neurons predicted the anti-Bayesian but not Bayesian behavior. a-d) Peristimulus time histograms in the high (orange) and low (teal) prior conditions, aligned to the reafferent event (dashed line), for the no-saccade (a), with saccade and low image noise (b), with saccade and medium image noise (c), and with saccade and high image noise (d) conditions. Comparison between the no-saccade condition and common condition with a saccade but low image noise allows for the evaluation of whether FEF activity reflects Bayesian behavior, while comparison between the common, medium image noise, and high image noise conditions allows for evaluation of whether activity reflects the anti-Bayesian behavior. e) Wilcoxon effect sizes between the high and low prior conditions across the four noise conditions. f-h) Correlation between high vs. low prior Wilcoxon effect sizes (i.e., neuronal prior use) and behavioral prior use for the no-saccade (f), medium image noise (g), and high image noise (h) conditions, when all values were divided by the common condition. i) Coefficients from

a mixed-effects, binomial generalized linear regression predicting the binary behavioral response ("jumped" or "did not jump"). j) Coefficients from a mixed-effects, linear regression predicting firing rates. *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$, ****, $p < 0.0001$. Data in f-j include only the neurons, and their corresponding behavioral sessions, that had significant prior selectivity in the common condition (n = 46 neurons, evaluated using a Wilcoxon signed-rank test).

We additionally considered that for all conditions with an intervening saccade, the probe was initially placed near the center of the cell's RF *after* the saccade, but for the no-saccade condition, it was placed near the pre-saccadic RF center since there was no movement of the RF (Fig. S1, leftmost column). Since the jump sizes we used were typically smaller than 5°, it was likely that the initial location of the probe was already within the RF in this condition[38]. Therefore, it was possible that the prior selectivity and its interaction with behavior aligned better to the initial probe presentation than to the probe jump event (analogous to a reafferent response) for the no-saccade condition. However, while there was clear transient visual activity aligned to probe presentation near the RF center as expected, the neurons showed no prior selectivity during the initial probe presentation epoch, and we did not find a significant correlation between neuronal prior use and behavior (Fig. S7).

Finally, to examine the possibility that other experimental variables (jump size, image noise levels, and the presence of an intervening saccade) may have had confounding effects on the correlation tests of Figure 4f-h, we ran two multiple, mixed-effects regression models. One used the behavioral responses of the animals as the regression model outcome. The other used the firing rates of the neurons as the outcome (detailed description in Methods). All experimental variables, including the interactions between priors and both image (the anti-Bayesian effect) and motor noise (the Bayesian effect), significantly predicted the behavioral response, as expected (Fig. 4i). Prior-noise interaction effects are highlighted in red shading in Fig. 4i – j. When predicting the firing rates, however, only the interaction between priors and image noise

had a significant effect, while the interaction between priors and motor noise did not, even with all other variables accounted for as regressors (Fig. 4j). Three individual variables had a significant effect on firing rates, too, all of which were expected. Priors had a strong effect since only the cells with a significant prior effect were included. Additionally, motor noise (i.e., the presence of an intervening saccade) and jump size had significant coefficients indicating that FEF activity is sensitive both to the presence of an image at the center of the post-saccadic receptive field and the amount by which it moved there[36]. Taken together, although the animals' behavioral prior use was modulated *both* by motor and image noise, the prior signal in FEF was modulated only by image noise. That is, FEF activity selectively correlated with and predicted only the "anti-Bayesian" prior use, consistent with the use of a sensory classifier.

Discussion

While we previously judgments of visual stability across saccades exhibit both Bayesian and anti-Bayesian effects of noise from motor and visual sources, respectively, here we found that FEF neural activity is only modulated by the latter. This suggests that the two behavioral strategies are implemented by distinct circuits in the brain and that FEF is part of a pathway that selectively implements the anti-Bayesian behavior. Indeed, the timing of the observed prior signal in FEF is consistent with the classifier model we proposed in Subramanian et al. (2023)[26] and recapitulated in Fig. S3. We proposed that decreasing prior use could be explained by a perceptron-like, two-layer neural network that associates the continuous sensory input (displacement amounts) with categorical "jumped' or "did not jump" response outputs. A critical assumption for explaining the data was that the association (continuous to categorical) was dependent on the context provided by the prior information, leading to a difference across prior conditions in the response rates at each displacement. When external noise is added to the sensory evidence, it blurs the input-output associations across prior contexts such that they appear to collapse together, both in behavior and, as shown here, neuronal activity. In this

sense, our findings that prior and image noise effects emerge at the reafferent event (rather than when the prior is introduced) and span the time frame until the response is made are both consistent with a representation of the sensory input about displacements being associated with the response.

These results raise the complementary question of where in the brain the Bayesian compensation for saccadic motor noise may be implemented. One possibility is the superior colliculus (SC). Neurons in the intermediate layers of the SC, an origin of saccadic corollary discharge signals[27], have also been shown to signal probabilistic priors for binary sensory decisions[39-40] and neurons in superficial SC are modulated by saccadic suppression[41]. The SC thus signals both necessary components of the Bayesian computation for our task. However, the neural signatures of saccadic suppression have also been observed in a variety of regions in the visual system including the Medial Temporal region (MT), the Medial Superior Temporal region (MST), the Ventral Intraparietal region (VIP) and the Lateral Intraparietal region (LIP)[42-43], area V4[44], and even in the retina[45]. The possible contributions of these regions, along with the SC and FEF, to a distributed circuit for visual stability remain to be clarified. Our results, showing the use of two distinct computational strategies as well as their dissociation in the brain, provide a first step towards an integrated understanding of the behavioral, computational, and neural circuit basis for perception across saccades.

We consider a few additional points in interpreting the results. First, what is the scope of the dissociation between Bayesian and heuristic-like processing in FEF? That is, is there a context in which the activity of FEF neurons may be consistent with the implementation of Bayesian inference? We and others have found that humans are Bayesian when asked to report the *continuous* displacement of a probe rather than providing a categorical report of whether it jumped or not[26, 46]. Visual receptive fields in the FEF, both before and after remapping, maintain their spatially continuous retinotopic organization, and FEF neurons are

continuously tuned to the magnitude of image displacements across saccades[36]. It is possible, therefore, that the FEF supports the implementation of a continuous Bayesian model for visual stability across saccades, as it does for smooth pursuit eye movements[47]. A direct test of this possibility is feasible but has not been done. If indeed FEF activity conformed with the Bayesian prediction in this case, this would suggest that FEF supports Bayes optimal stimulus estimates when the stimuli match how visual location information is intrinsically organized within the area, i.e., in continuous rather than categorical coordinates.

Second, we consider the implications of our results in the context of the apparent ubiquity of Bayesian inference in sensory systems. The success of the Bayesian framework in explaining behavior has spurred a search for how Bayesian computations may be implemented by the brain. Leading theoretical proposals include the ideas of "Probabilistic Population Codes," in which sensory variability automatically represents the likelihood and prior distributions required for Bayesian integration[48-49]; of Bayesian inference via efficient coding, in which priors are inherently represented by the matched allocation of neural infrastructure to the natural statistics of the environment[50-52]; and of predictive coding models, in which feedback projections convey prior information that is integrated with feedforward sensory likelihoods[53]. These proposals have in common the intuition that Bayesian inference in simple sensory contexts results from the natural, evolutionarily shaped functioning of sensory circuitry rather than reflecting an explicit, reward-maximizing strategy that needs to be effortfully generated. If this were true, Bayesian inference stands to serve as a general computational principle of sensory function[54-55]. Our results showing that the reafferent visual responses in FEF are not consistent with Bayesian ideal observer predictions, even in a context where the behavior *is*, caution that Bayesian inference is not the default mode of sensory circuits, but that the precise scope and limits of the Bayesian framework for sensory processing should be examined in a context-dependent manner.

Finally, our results addressing the neural basis for how Bayesian and heuristic models may be implemented in the same sensory task context could help to serve as a bridge between our understanding of lower-level sensory processes and decision-making in higher order cognitive contexts.

References

1. Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (pp. xi, 455). John Wiley.
2. Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420
3. Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, *30*(Volume 30, 2007), 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038
4. Hanks, T. D., & Summerfield, C. (2017). Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron*, *93*(1), 15–31. https://doi.org/10.1016/j.neuron.2016.12.003
5. Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision Research*, *39*(21), 3621–3629. https://doi.org/10.1016/S0042-6989(99)00088-7
6. Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. https://doi.org/10.1038/415429a
7. Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*(6), 598–604. https://doi.org/10.1038/nn0602-858
8. Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*, *43*(24), 2539–2558. https://doi.org/10.1016/s0042-6989(03)00458-9
9. Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*(6971), 244–247. https://doi.org/10.1038/nature02169
10. Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, *14*(7), 926–932. https://doi.org/10.1038/nn.2831
11. Fetsch, C. R., Pouget, A., DeAngelis, G. C., & Angelaki, D. E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nature Neuroscience*, *15*(1), 146–154. https://doi.org/10.1038/nn.2983
12. Darlington, T. R., Tokiyama, S., & Lisberger, S. G. (2017). Control of the strength of visual-motor transmission as the mechanism of rapid adaptation of priors for Bayesian inference in smooth pursuit eye movements. *Journal of Neurophysiology*, *118*(2), 1173–1189. https://doi.org/10.1152/jn.00282.2017
13. Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), 1020–1026. https://doi.org/10.1038/nn.2590

14. Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124
15. Watzek, J., & Brosnan, S. F. (2018). (Ir)rational choices of humans, rhesus macaques, and capuchin monkeys in dynamic stochastic environments. *Cognition*, *178*, 109–117. https://doi.org/10.1016/j.cognition.2018.05.019
16. Toader, A. C., Rao, H. M., Ryoo, M., Bohlen, M. O., Cruger, J. S., Oh-Descher, H., Ferrari, S., Egner, T., Beck, J., & Sommer, M. A. (2019). Probabilistic inferential decision-making under time pressure in rhesus macaques (Macaca mulatta). *Journal of Comparative Psychology (Washington, D.C.: 1983)*, *133*(3), 380–396. https://doi.org/10.1037/com0000168
17. Yang, Q., Lin, Z., Zhang, W., Li, J., Chen, X., Zhang, J., & Yang, T. (2022). Monkey plays Pac-Man with compositional strategies and hierarchical decision-making. *eLife*, *11*, e74500. https://doi.org/10.7554/eLife.74500
18. Santos, L. R., & Platt, M. L. (2014). Chapter 7 - Evolutionary Anthropological Insights into Neuroeconomics: What Non-Human Primates can Tell us About Human Decision-Making Strategies. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics (Second Edition)* (pp. 109–122). Academic Press. https://doi.org/10.1016/B978-0-12-416008-8.00007-3
19. Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, *62*(Volume 62, 2011), 451–482. https://doi.org/10.1146/annurev-psych-120709-145346
20. O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, Reward, and Decision Making. *Annual Review of Psychology*, *68*(Volume 68, 2017), 73–100. https://doi.org/10.1146/annurev-psych-010416-044216
21. Carandini, M. (2024). Sensory choices as logistic classification. *Neuron*, *112*(17), 2854-2868.e1. https://doi.org/10.1016/j.neuron.2024.06.016
22. Gardner, J. L. (2019). Optimality and heuristics in perceptual neuroscience. *Nature Neuroscience*, *22*(4), 514–523. https://doi.org/10.1038/s41593-019-0340-4
23. Laquitaine, S., & Gardner, J. L. (2018). A Switching Observer for Human Perceptual Estimation. *Neuron*, *97*(2), 462-474.e6. https://doi.org/10.1016/j.neuron.2017.12.011
24. Rahnev, D., & Denison, R. N. (2018). Suboptimality in perceptual decision making. *The Behavioral and Brain Sciences*, *41*, e223. https://doi.org/10.1017/S0140525X18000936
25. Sohn, H., & Jazayeri, M. (2021). Validating model-based Bayesian integration using prior-cost metamers. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(25), e2021531118. https://doi.org/10.1073/pnas.2021531118

26. Subramanian, D., Pearson, J. M., & Sommer, M. A. (2023). Bayesian and Discriminative Models for Active Visual Perception across Saccades. *eNeuro*, *10*(7), ENEURO.0403-22.2023. https://doi.org/10.1523/ENEURO.0403-22.2023
27. Sommer, M. A., & Wurtz, R. H. (2008). Brain circuits for the internal monitoring of movements. *Annual Review of Neuroscience*, *31*, 317–338. https://doi.org/10.1146/annurev.neuro.31.060407.125627
28. Zuber, B. L., & Stark, L. (1966). Saccadic suppression: Elevation of visual threshold associated with saccadic eye movements. *Experimental Neurology*, *16*(1), 65–79. https://doi.org/10.1016/0014-4886(66)90087-2
29. Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *Journal of Neurophysiology*, *78*(3), 1373–1383. https://doi.org/10.1152/jn.1997.78.3.1373
30. Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, *444*(7117), 374–377. https://doi.org/10.1038/nature05279
31. Wang, X., Zhang, C., Yang, L., Jin, M., Goldberg, M. E., Zhang, M., & Qian, N. (2024). Perisaccadic and attentional remapping of receptive fields in lateral intraparietal area and frontal eye fields. *Cell Reports*, *43*(3), 113820. https://doi.org/10.1016/j.celrep.2024.113820
32. Hall, N. J., & Colby, C. L. (2011). Remapping for visual stability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1564), 528–539. https://doi.org/10.1098/rstb.2010.0248
33. Wurtz, R. H. (2018). Corollary Discharge Contributions to Perceptual Continuity Across Saccades. *Annual Review of Vision Science*, *4*, 215–237. https://doi.org/10.1146/annurev-vision-102016-061207
34. Cavanaugh, J., Berman, R. A., Joiner, W. M., & Wurtz, R. H. (2016). Saccadic Corollary Discharge Underlies Stable Visual Perception. *The Journal of Neuroscience*, *36*(1), 31–42. https://doi.org/10.1523/JNEUROSCI.2054-15.2016
35. Bruce, C. J., & Goldberg, M. E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. *Journal of Neurophysiology*, *53*(3), 603–635. https://doi.org/10.1152/jn.1985.53.3.603
36. Crapse, T. B., & Sommer, M. A. (2012). Frontal Eye Field Neurons Assess Visual Stability Across Saccades. *Journal of Neuroscience*, *32*(8), 2835–2845.
37. Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*(6), 719–722. https://doi.org/10.1016/0042-6989(75)90290-4

38. Mayo, J. P., DiTomasso, A. R., Sommer, M. A., & Smith, M. A. (2015). Dynamics of visual receptive fields in the macaque frontal eye field. *Journal of Neurophysiology*, *114*(6), 3201–3210. https://doi.org/10.1152/jn.00746.2015
39. Crapse, T. B., Lau, H., & Basso, M. A. (2018). A Role for the Superior Colliculus in Decision Criteria. *Neuron*, *97*(1), 181-194.e6. https://doi.org/10.1016/j.neuron.2017.12.006
40. Jun, E. J., Bautista, A. R., Nunez, M. D., Allen, D. C., Tak, J. H., Alvarez, E., & Basso, M. A. (2021). Causal role for the primate superior colliculus in the computation of evidence for perceptual decisions. *Nature Neuroscience*, *24*(8), 1121–1131. https://doi.org/10.1038/s41593-021-00878-6
41. Walker, M. F., Fitzgibbon, E. J., & Goldberg, M. E. (1995). Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *Journal of Neurophysiology*, *73*(5), 1988–2003. https://doi.org/10.1152/jn.1995.73.5.1988
42. Bremmer, F., Kubischik, M., Hoffmann, K.-P., & Krekelberg, B. (2009). Neural Dynamics of Saccadic Suppression. *The Journal of Neuroscience*, *29*(40), 12374. https://doi.org/10.1523/JNEUROSCI.2908-09.2009
43. Thiele, A., Henning, P., Kubischik, M., & Hoffmann, K.-P. (2002). Neural Mechanisms of Saccadic Suppression. *Science*, *295*(5564), 2460–2462. https://doi.org/10.1126/science.1068788
44. Zanos, T. P., Mineault, P. J., Guitton, D., & Pack, C. C. (2016). Mechanisms of Saccadic Suppression in Primate Cortical Area V4. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(35), 9227–9239. https://doi.org/10.1523/JNEUROSCI.1015-16.2016
45. Idrees, S., Baumann, M. P., Franke, F., Münch, T. A., & Hafed, Z. M. (2020). Perceptual saccadic suppression starts in the retina. *Nature Communications*, *11*(1), 1977. https://doi.org/10.1038/s41467-020-15890-w
46. Niemeier, M., Crawford, J. D., & Tweed, D. B. (2003). Optimal transsaccadic integration explains distorted spatial perception. *Nature*, *422*(6927), 76–80. https://doi.org/10.1038/nature01439
47. Darlington, T. R., Beck, J. M., & Lisberger, S. G. (2018). Neural implementation of Bayesian inference in a sensorimotor behavior. *Nature Neuroscience*, *21*(10), 1442–1451. https://doi.org/10.1038/s41593-018-0233-y
48. Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, *9*(11), 1432–1438. https://doi.org/10.1038/nn1790
49. Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., & Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron*, *60*(6), 1142–1152. https://doi.org/10.1016/j.neuron.2008.09.021

50. Wei, X.-X., & Stocker, A. A. (2012). Bayesian Inference with Efficient Neural Population Codes. In A. E. P. Villa, W. Duch, P. Érdi, F. Masulli, & G. Palm (Eds.), *Artificial Neural Networks and Machine Learning – ICANN 2012* (pp. 523–530). Springer. https://doi.org/10.1007/978-3-642-33269-2_66
51. Ganguli, D., & Simoncelli, E. P. (2014). Efficient sensory encoding and Bayesian inference with heterogeneous neural populations. *Neural Computation*, *26*(10), 2103–2134. Neural Computation. https://doi.org/10.1162/NECO_a_00638
52. Park, I. M., & Pillow, J. W. (2024). *Bayesian Efficient Coding* (p. 178418). bioRxiv. https://doi.org/10.1101/178418
53. Aitchison, L., & Lengyel, M. (2017). With or without you: Predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology*, *46*, 219–227. https://doi.org/10.1016/j.conb.2017.08.010
54. Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719. https://doi.org/10.1016/j.tins.2004.10.007
55. Friston, K. (2012). The history of the future of the Bayesian brain. *Neuroimage*, *62–248*(2), 1230–1233. https://doi.org/10.1016/j.neuroimage.2011.10.004
56. Robinson, D. A. (1963). A Method of Measuring Eye Movement Using a Scieral Search Coil in a Magnetic Field. *IEEE Transactions on Bio-Medical Electronics*, *10*(4), 137–145. IEEE Transactions on Bio-medical Electronics. https://doi.org/10.1109/TBMEL.1963.4322822
57. Judge, S. J., Richmond, B. J., & Chu, F. C. (1980). Implantation of magnetic search coils for measurement of eye position: An improved method. *Vision Research*, *20*(6), 535–538. https://doi.org/10.1016/0042-6989(80)90128-5

Methods

Surgical preparation and electrophysiological recording

Two rhesus macaques (Monkey S and Monkey T, both males) were used for procedures described in this manuscript. Animals were brought into the lab in custom-made chairs (Crist Instruments, Hagerstown, MD) and their heads were stabilized using a head-post that attached to both the chair and a surgically implanted socket (Crist Instruments, Hagerstown, MD) on the skull. The socket, a stereotaxically positioned recording chamber over the frontal eye field (FEF), and monocular scleral search coils[56-57] were implanted in an aseptic surgical procedure. All surgical and experimental procedures were performed in accordance with protocols approved by the Duke Institutional Animal Care and Use Committee.

Neurons in FEF were recorded by placing a 1 mm x 1 mm grid within the recording chamber and advancing sterilized tungsten microelectrodes with impedances ranging from 750-950 kΩ (FHC, Bowdoin, ME) through a guide tube. The FEF was identified anatomically in each animal using structural MRI and by the visual and presaccadic response properties of recorded neurons. Neurons were extracellularly isolated by identifying the auditory and visual signatures of individual waveforms using an audio speaker and oscilloscope respectively. Continuous data for all recorded signals (neural signals and eye position) were stored. Online, the neural activity of one or more putative single neurons was recorded, aligned to behavioral trial events, and monitored using rasters and spike density functions using Spike2 (Cambridge Electronic Design, Cambridge, UK) and custom Matlab code (Mathworks). This online monitoring aided in ensuring sustained isolation of specific action potential waveforms throughout the session, which could last 1-3 hours. Offline, action potentials were sorted to confirm single neuron origins using a Principal Components Analysis based clustering algorithm in Spike2. All other analyses were performed in Matlab (Mathworks).

Receptive field mapping

After a neuron was isolated in each session, we first mapped its visual, motor, or visuomotor spatial receptive field in two delayed saccade tasks. On every trial, the animal first fixated a white, central square (1° x 1°) for 300-500ms. A target square (also white and 1° x 1°) was then presented concurrently with the fixation square for a delay of 500-700ms. After the delay, the fixation square was extinguished cuing the animal to make a saccade to the target. Upon target acquisition, the animal was further required to fixate on the target for an additional 500 ms after which, fluid reward was delivered. Since the first occurrence of the target, i.e., its onset, is separated by 500-700ms from the time that the animal makes a saccade to it, we can distinguish the neuron's visual response (typically 50-100ms after target onset) from its motor response (typically 150-100ms before the saccade). We first mapped the neuron's responsiveness to the target placed in one of 8 potential directions: 4 cardinal directions and 4 along the diagonal. The targets were placed at a retinal eccentricity that was most likely to match neurons' preferences in the subregion of FEF. Typically, this was at 10°. Next, based on the neuron's direction preference, we then ran the amplitude task, where the target was restricted to the preferred direction but presented at one of four amplitudes: 2°, 5°, 10°, and 20°. The putative center of the receptive field was inferred based its response properties in the direction and amplitude tasks and used to guide the placement of the probe in the Saccadic Suppression of Displacement (SSD) task.

Saccadic Suppression of Displacement (SSD) task

The overall goal of the current manuscript was to record the activity of FEF neurons while we tested the use of learned priors to compensate for internal, motor-driven or external, image

noise in the context of a saccadic suppression of displacement task. Animals made saccades and reported whether a visual stimulus in the environment moved during the saccade.

On each trial, a fixation square (1ºx1º) first appeared at the center of the screen. After fixation had been acquired and maintained for a randomized duration of 300-500ms, the visual probe appeared at one of the 4 locations on the screen for 500-700ms. The monkey was required to maintain fixation on the central fixation square for that duration, after which the fixation square was replaced by the saccade target (1ºx1º) indicating to the animal they could make a saccade. Saccade initiation (defined as the time the eye crossed a threshold set at 20% of the saccade length) triggered the displacement of the probe on some trials. Animals were further required to maintain postsaccadic fixation for 500ms after which the saccade target was replaced by a white cross in the same location. To report that the probe had moved during the saccade, the monkey was required to make a saccade to the probe within 500ms and then fixate on it for 400ms. To report that it had remained stationary during the saccade, the monkey had to remain fixated on the cross for 1000ms. The precise timing of stimulus presentation was verified with a photodiode taped to the top left corner of the monitor, where a white square (invisible to the monkey) was flashed within the same frame as the measured stimulus.
The probe was initially placed at a location determined by the jump size and was always displaced parallel to the intervening saccade to land at the center of the cell's post-saccadic receptive field. The intervening saccade vector was chosen on a session-by-session basis to be *away* from the RF as depicted in Fig. 50, right. The saccade amplitude was jittered by 1° trial-by-trial, so that an absolute target location did not correspond with a jump or non-jump response. For trials in which no saccade was made (described below), the probe landed in the center of the receptive field. Displacements were drawn from relatively broad and narrow Gaussian distributions in the movement (μ = 0º, σ = 2.5º) and non-movement (μ = 0º, σ = 0.2º) conditions, respectively. Positive displacements were either rightward or upward, and negative displacements were leftward or downward. The continuous nature of these distributions was

maintained across sessions. For each session, four to six discretized displacements, which maintained the probabilistic characteristics of the Gaussian distributions, were chosen to allow for a comparison of neural responses across prior and noise conditions with variability minimized across all other parameters.

Priors (0.2, 0.5, and 0.8) were cued by the color of the fixation and target squares. For monkey S, green squares meant that the probe had a 0.2 probability of being displaced while magenta squares indicated a 0.8 probability of displacement. For monkey T, blue squares were associated with a 0.2 probability of displacement while orange squares were associated with a 0.8 probability of displacement. Animals were trained on priors over multiple sessions using performance-based feedback. External, image noise was manipulated by the width of the Gaussian probe, which had one of three possible standard deviations: 0.5º ("low noise"), 1.25º ("medium noise"), and 2º ("high noise") for Monkey S and 0.5º ("low noise"), 1.25º ("medium noise"), and 1.75º ("high noise"), for Monkey T. To manipulate internal, motor-driven uncertainty, we added a condition to the experiment where monkeys did not make a saccade. The purpose was to eliminate a major form of saccade-driven sensory uncertainty, the saccadic suppression of visual signals. The monkeys remained fixated in the center while the Gaussian, visual probe was displaced. This no-saccade condition served as the "low motor noise" condition and was compared to a "high motor noise" condition where animals made a saccade. The Gaussian probe had low image noise in these conditions. The temporal structure of the no-saccade trials was identical to the trials with a saccade.

The low image noise condition with a saccade served as the common condition between the motor and image noise experiments. Thus, there were a total of four noise conditions at three prior levels. All trial types were equally likely and randomly interleaved.

Data inclusion

We recorded from 91 FEF neurons (55 from Monkey S across 40 sessions, and 36 from Monkey T across 24 sessions). For each session, a small minority of trials in which the probe did not land in its final position before the end of the intervening saccade were excluded from analysis. As in Subramanian et al. (2023)[26], we measured the timing of the probe landing using the response of a photodiode to a white square, invisible to the monkey, presented in the top left corner of the screen simultaneously as the probe. We verified the maximum duration of a frame as being 7ms from top left to bottom right using a second photodiode. Since the probe was presented at various locations on the screen, we set the most conservative criterion such that the photodiode timestamp had to be at least 7ms before the detected end of the saccade. We limited our analyses to neurons recorded on behavioral sessions in which the animal completed at least 10 valid trials in each condition of the experiment. This resulted in a total of 79 included neurons (53 from Monkey S across 39 sessions and 26 from Monkey T across 17 sessions). For control analyses performed only on trials where the displacement was 0, we used neurons for which the animal completed at least 5 valid trials with displacement = 0 in all conditions. This resulted in the inclusion of 64 total neurons.

Behavioral prior use and comparison to model predictions

We quantified animals' "prior use" as being the difference in the proportion of trials in which they reported "jumped" across prior conditions. We compared the pattern of their prior use data across noise conditions to the simulated predictions from two models evaluated in in Subramanian et al. (2023): a Bayesian ideal observer model and a model combining the outputs of a Bayesian model and a perceptron-based classifier model.

Bayesian ideal observer model

The ideal observer makes a probabilistic decision about binary variable, J, indicating whether the target jumped or not. $\neg J$ indicates that the target did not jump. Since the true

displacement is experimentally drawn but not available to the observer, they make this decision given the *perceived* displacement, $\hat{x}$. The decision is based on the relative probabilities of the target having jumped or not jumped given the perceived displacement,

$$D(\hat{x}) = \mathbb{I}\{P(J|\,\hat{x}) > P(\neg J|\,\hat{x})\,\} \tag{3}$$

where $D(\hat{x})$ is the decision given the perceived displacement, $\hat{x}$, and is determined by a binary indicator function, $\mathbb{I}$. $\mathbb{I}$ = 0 (no jump) if the condition in braces is not met. Otherwise, $\mathbb{I}$ = 1 (jumped); $P(J|\,\hat{x})$ is the probability that the probe jumped given $\hat{x}$; $P(\neg J|\,\hat{x})$ is the probability that the probe did not jump given $\hat{x}$.

Using Bayes' rule for the condition within braces in Eq. 3,

$$D(\hat{x}) = \mathbb{I}\left\{\frac{P(\hat{x}|J)P(J)}{P(\hat{x}|\neg J)\,(1 - P(J))} > 1\right\} \tag{4}$$

The simulated decision of the ideal observer, however, must be compared to the responses of participants. We do not have access to participants' perceived displacement, but instead can only infer their decision given the true experimental displacement, x. We assume that the perceived displacement is a Gaussian random variable where the mean is the true displacement, and its variance given by the width of the blob on that trial:

$$\hat{x} \sim N(x, \sigma_t^2) \tag{5}$$

where σ_t is the variance of the target.

The decision given the true displacement can thus be modeled as,

$$D(x) = \int \mathbb{I}\left\{\frac{P(\hat{x}|J)\,P(J)}{P(\hat{x}|\neg J)\,(P\,(\neg J))} > 1\right\} P(\hat{x}|\,x)dx \tag{6}$$

That is, the decision value given the true displacement, $D(x)$, is the integral of the perceived displacement distribution that falls above the point at which the indicator function, $\mathbb{I}$, is non-zero.

Based on the distributions used in the experiment,

$$x|J \sim N\left(0, \sigma_J^2\right) \tag{7}$$

and

$$x|\neg J \sim N\left(0, \sigma_{\neg J}^2\right) \tag{8}$$

Since $P(\hat{x}|\, x)$, $P(x|J)$, and $P(x|\neg J)$ are Gaussian distributions, we integrate over x such that,

$$\hat{x}|J \sim N\left(\mu_J, \sigma_J^2 + \sigma_t^2\right) \tag{9}$$

and

$$\hat{x}|\neg J \sim N\left(\mu_{\neg J}, \sigma_{\neg J}^2 + \sigma_t^2\right) \tag{10}$$

Thus, the expression inside the indicator function in Eq. 6, when replaced with the appropriate Gaussian probability density functions, equals,

$$\frac{\frac{1}{\sqrt{(\sigma_J^2 + \sigma_t^2)2\pi}} e^{-\frac{1}{2}\left(\frac{(\hat{x}-\mu_J)^2}{\sigma_J^2+\sigma_t^2}\right)} P(J)}{\frac{1}{\sqrt{(\sigma_{\neg J}^2 + \sigma_t^2)2\pi}} e^{-\frac{1}{2}\left(\frac{(\hat{x}-\mu_{\neg J})^2}{\sigma_{\neg J}^2+\sigma_t^2}\right)} (P\,(\neg J))} > 1 \tag{11}$$

Taking the log on both sides provides the condition under which the indicator function is > 0,

$$\log\left(\frac{\frac{1}{\sqrt{(\sigma_J^2 + \sigma_t^2)2\pi}} e^{-\frac{1}{2}\left(\frac{(\hat{x}-\mu_J)^2}{\sigma_J^2+\sigma_t^2}\right)} P(J)}{\frac{1}{\sqrt{(\sigma_{\neg J}^2 + \sigma_t^2)2\pi}} e^{-\frac{1}{2}\left(\frac{(\hat{x}-\mu_{\neg J})^2}{\sigma_J^2+\sigma_t^2}\right)} (P\,(\neg J))}\right) > \log(1) = 0 \tag{12}$$

That is,

$$\log\left(\frac{1}{\sqrt{(\sigma_J^2 + \sigma_t^2)2\pi}} e^{-\frac{1}{2}\left(\frac{(\hat{x}-\mu_J)^2}{\sigma_J^2+\sigma_t^2}\right)} P(J)\right) - \log\left(\frac{1}{\sqrt{(\sigma_{\neg J}^2 + \sigma_t^2)2\pi}} e^{-\frac{1}{2}\left(\frac{(\hat{x}-\mu_{\neg J})^2}{\sigma_J^2+\sigma_t^2}\right)} P\,(\neg J)\right) > 0 \tag{13}$$

Rearranging terms, the indicator function is > 0 when

$$\frac{-(\hat{x}-\mu_J)^2}{\sigma_J^2+\sigma_t^2}+\frac{(\hat{x}-\mu_{\neg J})^2}{\sigma_{\neg J}^2+\sigma_t^2}>\log\left(\frac{\sigma_J^2+\sigma_t^2}{\sigma_{\neg J}^2+\sigma_t^2}\right)+\ 2\log\frac{P(\neg J)}{P(J)} \tag{14}$$

When $\mu_J = \mu_{\neg J} = 0$, the indicator function is > 0 when $\hat{x}^2$ is greater than a criterion value, $\hat{x}_C^2$, defined as,

$$\hat{x}^2\ >\hat{x}_C^2=\frac{\log\left(\frac{\sigma_J^2+\sigma_t^2}{\sigma_{\neg J}^2+\sigma_t^2}\right)+\ 2\log\frac{P(\neg J)}{P(J)}}{\frac{1}{\sigma_{\neg J}^2+\sigma_t^2}-\frac{1}{\sigma_J^2+\sigma_t^2}} \tag{15}$$

Since $\hat{x}$ is a Gaussian random variable, $\hat{x}\sim \mathrm{N}(x,\sigma_t^2)$, $\frac{\hat{x}^2}{\sigma_t^2}$ is a non-central chi-square random variable, $\frac{\hat{x}^2}{\sigma_t^2}\sim X^2\left(1,\frac{x^2}{\sigma_t^2}\right)$. Thus, the decision, $D(x)$, can be modeled as the integral of a non-central chi-square distribution that lies above the criterion, $\hat{x}_C^2$. That is,

$$D(x)=1-\ F_{X^2}(\frac{\hat{x}_C^2}{\sigma_t^2},df\ =1,\lambda=\ \frac{x^2}{\sigma_t^2}) \tag{16}$$

where F_{X^2} is the cumulative distribution function of X^2 with degrees of freedom, $df\ =1$ and gamma, $\lambda=\ \frac{x^2}{\sigma_t^2}$ up to $\frac{\hat{x}_C^2}{\sigma_t^2}$.

Based on the results in Subramanian et al. (2023), we simulated the three levels of image noise by varying the values of σ_t and set them at 0.4º, 0.65º, and 0.9º for low, medium, and high noise respectively. Motor-driven noise was simulated by varying the width of the non-jump distribution, $\sigma_{\neg J}$. We set $\sigma_{\neg J}$ = 0.1º for the no-saccade condition, and $\sigma_{\neg J}$ = 0.75º for the conditions with an intervening saccade. The width of the "jump" distribution, σ_J, was set to 5º for all conditions. Simulation parameters for both models were chosen to best illustrate the predicted pattern of prior use across the four noise levels, and simulated results are expressed in arbitrary units.

Classifier model

We set up a two-layer neural network to classify displacements into categories of "jumped" and "did not jump" and simulated its performance under experimental conditions. The input layer consisted of units representing continuous displacements and the output layer had two units: "jump (J)" and "no jump (NJ)." For ease of computation, continuous input displacements were discretized into bins of 0.1°. Displacements ranged from 0 - 7.5°. That is, there were 75 input units for each network. Sensory noise was simulated as a Gaussian distribution of input unit activation, truncated at the two ends of the input range (0 and 7.5), such that the total activation of input units was always 1. On each trial, the distribution was centered on the true displacement for the trial and the width of activation was determined by the sensory noise level. Each input unit was connected to both output units. The activation of each output unit was the *weighted* sum of inputs, i.e.,

$$a_j = \sum_{i=1}^{n} a_i \, w_{ij} \tag{17}$$

where a_j is the activation of the output unit, j; a_i is the activation of the output unit, i; and w_{ij} is the weight of the connection between input unit, i, and output unit, j. The final output on each trial was the normalized activation of the "jump" and "no jump" output units such that the output for each unit was bounded between 0 and 1:

$$o_{jump} = \frac{a_{jump}}{a_{jump} + a_{no\ jump}} \tag{18}$$

and

$$o_{no\ jump} = \frac{a_{no\ jump}}{a_{jump} + a_{no\ jump}} \tag{19}$$

where o_{jump} is the final output of the "jump" unit, $o_{no\ jump}$ is the final output of "no jump" unit, and a_{jump} and $a_{no\ jump}$ are activations of the "jump" and "no jump" output units, respectively. There were separate sets of inputs for each prior. During the training phase, weights between inputs and outputs were updated using an error-based learning rule:

$$\Delta w_{ij} = \beta a_i (d - o_j) \tag{20}$$

where Δw_{ij} is the change in weights between input unit, i, and output unit, j; β is the learning rate, a_i is the activation of the input unit, i; o_j is the final output of unit j, and d is the desired state of output unit, j.

Combined model

The final output of the combined Bayesian and classifier model was obtained using a weighted combination of the output of each model individually:

$$O_C = w_B O_B + w_P O_P \tag{21}$$

where O_C is the output of the combined model, O_B is the output of the Bayesian model, O_P is the output of the perceptron classifier model, and w_B and w_P are the weights assigned to the Bayesian and classifier models, respectively. Further, the weights of the two component models added up to 1:

$$w_b + w_P = 1 \tag{22}$$

We simulated behavior in the motor and image noise experiments by modeling motor noise within the Bayesian ideal observer model and image noise in the classifier model as described in Subramanian et al. (2023). Each simulation consisted of 2000 trials. The width of the "jump" distribution, σ_J, was set to 2.5º. We set $\sigma_{\neg J}$ = 0.5º for the no-saccade condition, and $\sigma_{\neg J}$ = 1º for the conditions with an intervening saccade. Image noise (σ_t) was set to a constant 0.1º for all conditions in the Bayesian model. For the classifier model, low, medium, and high image noise

were simulated at 0.1º, 1º, and 2º respectively. For the first 1000 trials, the model was trained under conditions similar to those experienced by the subjects. The weights between the inputs and outputs were trained in the low image noise, with saccade condition using a learning rate of 0.5. The remaining 1000 trials formed the test phase with the 4 noise conditions described in this manuscript. For these trials, the weights remained constant, and motor and image noise were simulated as described above. w_b and w_P were set to 0.1 and 0.9, respectively.

Neuronal firing rate analyses

Normalization

To visualize population peri-stimulus time histograms aligned to specific task events, we first normalized the activity of each neuron and then averaged their firing rates. Firing rates were normalized by binning spikes across all task conditions into non-overlapping 20ms bins, fitting a Gaussian distribution to the binned firing rates to obtain a mean and standard deviation, and z-scoring firing rates across the various prior and noise conditions using the fit mean and standard deviations. Firing rates were additionally convolved with 20ms Gaussian kernels.

Non-parametric statistical significance test and effect size calculation

Firing rates in the high and low prior conditions were evaluated for statistical significance using the Wilcoxon ranksum test, and the effect size of the difference was quantified using the Wilcoxon effect size. The Wilcoxon effect size is given by ($\frac{z}{\sqrt{n}}$), where z is the z-statistic used by the rank sum test and n is the total number of trials (low prior + high prior). A positive effect size indicated that firing rates in the high prior condition were greater than in the low prior condition, and vice versa for a negative effect size. Note that the Wilcoxon rank sum test and the Wilcoxon effect size are the non-parametric equivalents of a two-sample t-test and Cohen's d, respectively.

Mixed-effects regression model

To compare neuronal prior use with behavioral prior use while simultaneously accounting for all other experimental variables, we ran two mixed-effects regression models on neurons with significant prior effects in the common condition: one with the neuronal firing rates as the outcome variable and one with the behavioral responses ("jumped" or "did not jump") in the same sessions as the outcome. The models were run on the data used to analyze neuron-behavior correlations in Figure. 4f-h and Figure S7, i.e., the neurons for which there was a significant prior effect in the reafferent epoch. The model to explain the neuronal firing rates was a linear, mixed-effects model while the one describing the binary, behavioral responses in those sessions was a binomial, generalized linear mixed-effects model. Both models were fit using Matlab's *fitlme* function.

For both models, the fixed effects included the prior, motor noise, image noise, displacement, and interaction terms between image and motor noise, image noise and the prior, and motor noise and the prior. For the linear model predicting neuronal firing rates, we also included as random effects distinct intercepts and slopes for each of these predictors for each neuron. For the binomial model predicting behavioral responses, we included the same random effects but grouped by session rather than neuron identity.

Supplementary figures

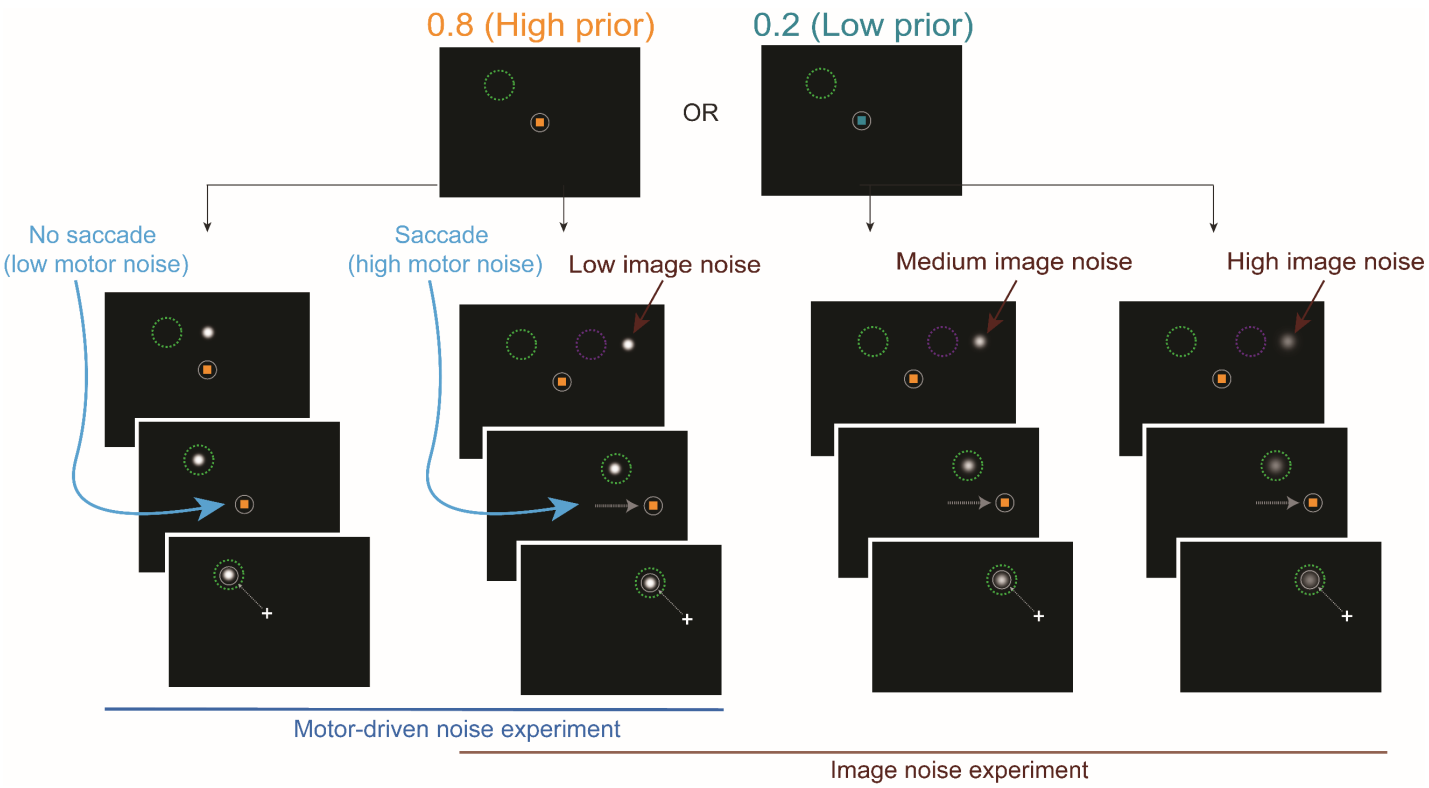

Figure S1. Spatial configuration of the task. Priors were indicated to the subject on every trial by the color of the fixation square. Left-most column: no-saccade condition with low motor-driven noise. Left-middle, right-middle, and right-most columns: conditions with an intervening saccade (with saccade) but with low, medium, and high image noise respectively. The condition with a saccade and low image noise serves as a common condition between the two experiments. Green dashed circle: receptive field of the neuron. Purple dashed circle: receptive field after the intervening saccade. Grey solid circle: eye position. Dashed arrows: saccades. Target, probe, and receptive field sizes are all schematized for clarity and not drawn to scale.

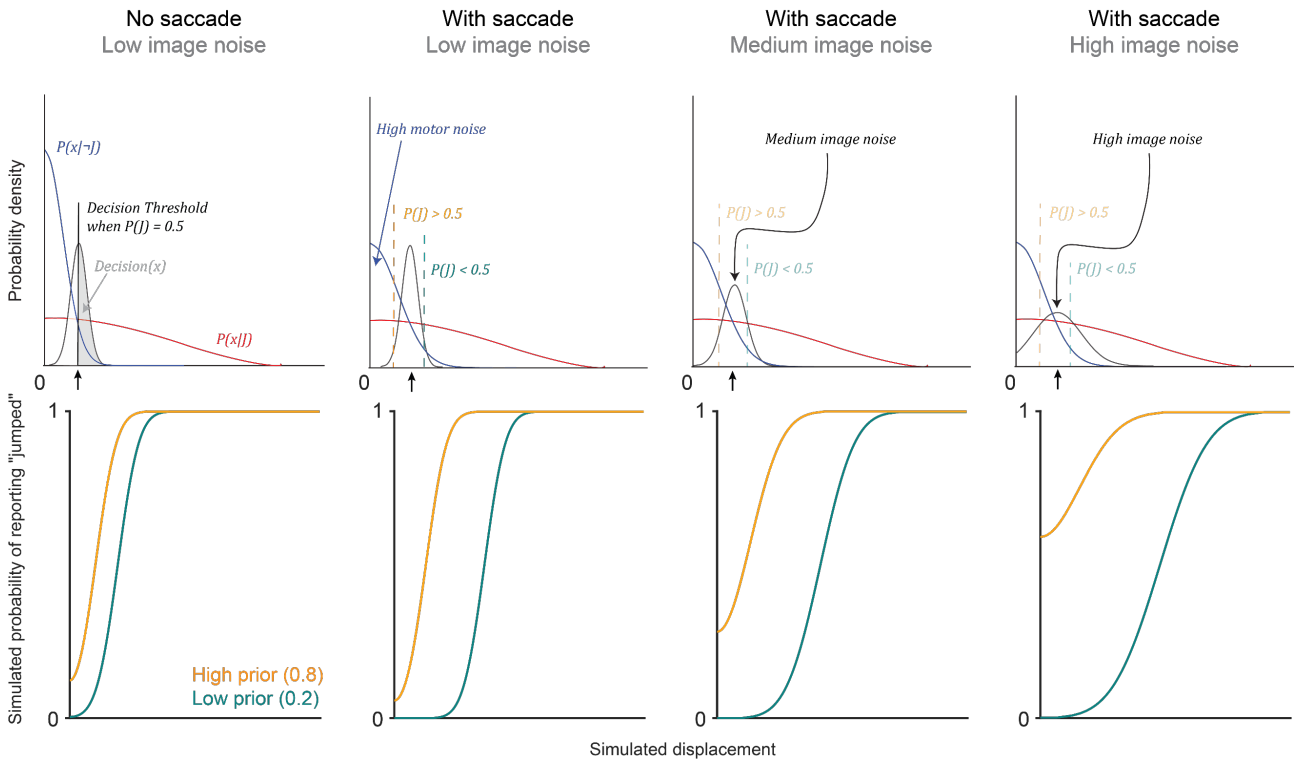

Figure S2. Model schematic (top row) and predicted psychometric curves (bottom row) for the Bayesian ideal observer model used for this task in Subramanian et al. (2023)[26]. Left-most column: no-saccade condition with low motor-driven noise. Left-middle, right-middle, and right-most columns: conditions with an intervening saccade (with saccade) but with low, medium, and high image noise respectively. Top row: Overlapping blue and red distributions depict the Gaussian non-jump and jump distributions, respectively, both centered on displacement = 0. High motor-driven noise analogous to saccadic suppression is modeled by increasing the width of the non-jump distribution. The black distribution depicts the Gaussian distribution associated with probe. The probe distribution is centered on the true displacement (black arrows) with its width determined by the amount of image noise. Black vertical line in the left-most panel depicts the decision threshold, or the point at which a given displacement was likely drawn from the non-jump vs. jump distributions, when the relative probabilities of being drawn from either are equal (i.e., the prior = 0.5). When the prior is less than 0.5, the threshold moves to the right and when greater than 0.5, to the left (orange and teal dashed lines in subsequent columns). The value of the model output, i.e., the “decision”, for each displacement and prior condition is determined by the integral of the probe distribution over displacement values greater than the decision threshold (gray shaded area in left-most panel). Bottom row: simulated decision values for the high and low prior conditions across the four noise levels. Prior use in each condition, quantified as the difference in decision values between the high and low prior conditions, is shown in Fig. 1d (top). Detailed model equations and simulation parameters used are described in the Methods section.

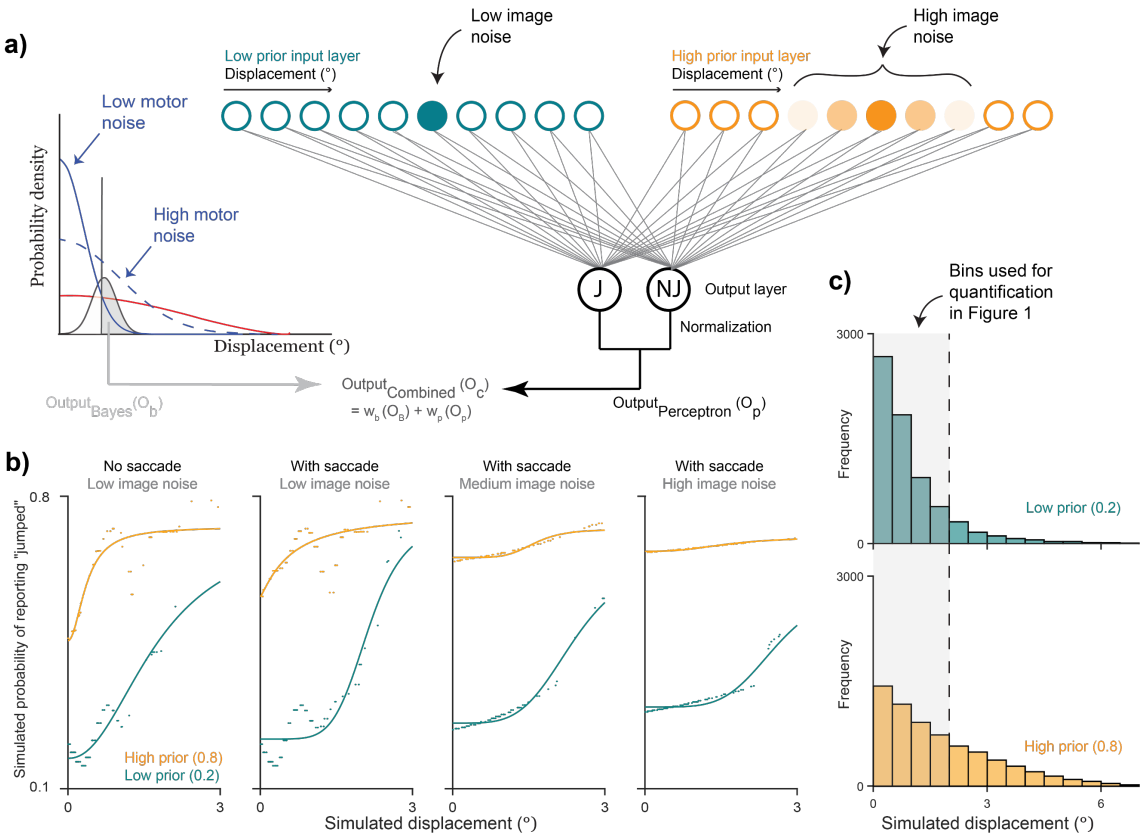

Figure S3. Combined Bayesian and perceptron classifier proposed in Subramanian et al. (2023)[26]. Detailed model equations and simulation parameters used are described in the Methods section. a) Model schematic. Left hand side depicts the Bayesian observer model from Fig. S2. High motor noise, i.e., in conditions with an intervening saccade, is modeled by increasing the standard deviation of the non-jump distribution relative to the jump distribution. The image noise level in the Bayesian model was held constant at the lowest level. Medium and high image noise were applied instead to the input layer of a two-layer sensory classifier network (right hand side). The input layer consisted of displacement units with one set each per prior. Input activation was centered on a given displacement and the width of activation was determined by the image noise level. Example low image noise activation shown on the left and high image noise on the right (intensity of the color reflects the width of the activation). The total input activation on each trial always summed to 1. Activation of the "jump" and "no jump" units in the output layer, which were determined by a weighted sum of activated inputs, were normalized for a perceptron model output. The final, combined model output was a weighted sum of Bayesian and perceptron outputs. b) Simulated outputs and fit psychometric curves for the high and low prior conditions across the four noise levels. c) Histogram of simulated displacements. We quantified prior use (shown in Figure 1d, bottom) as the mean difference in response rates between priors for displacements ranging from 0-2°, depicted by the gray shaded region (in 0.5° bins). This was done to restrict the quantification to displacements that were prevalent across both the low and high prior conditions.

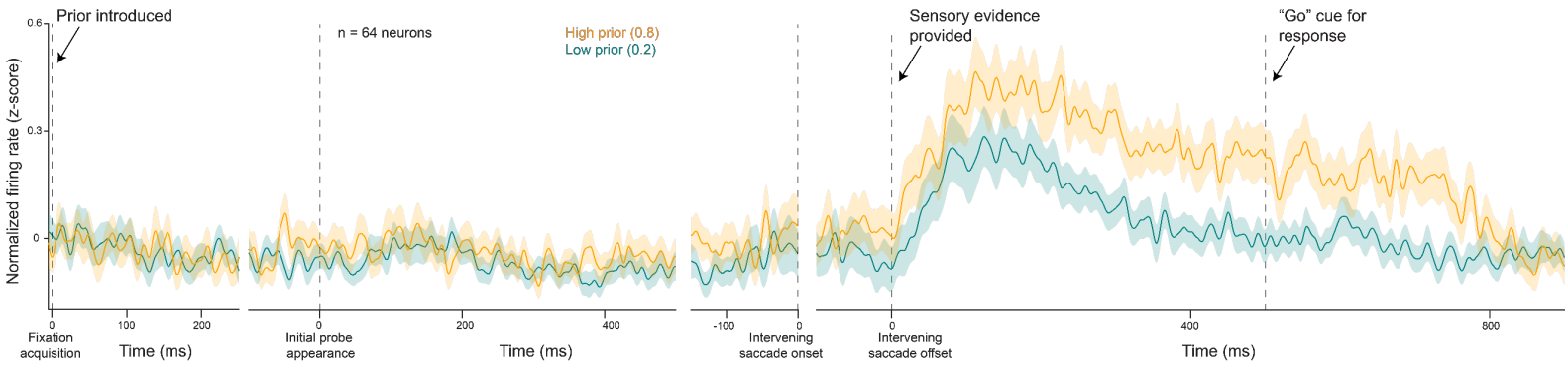

Figure S4. Same as Fig. 2a but restricted to trials where the jump size was 0.

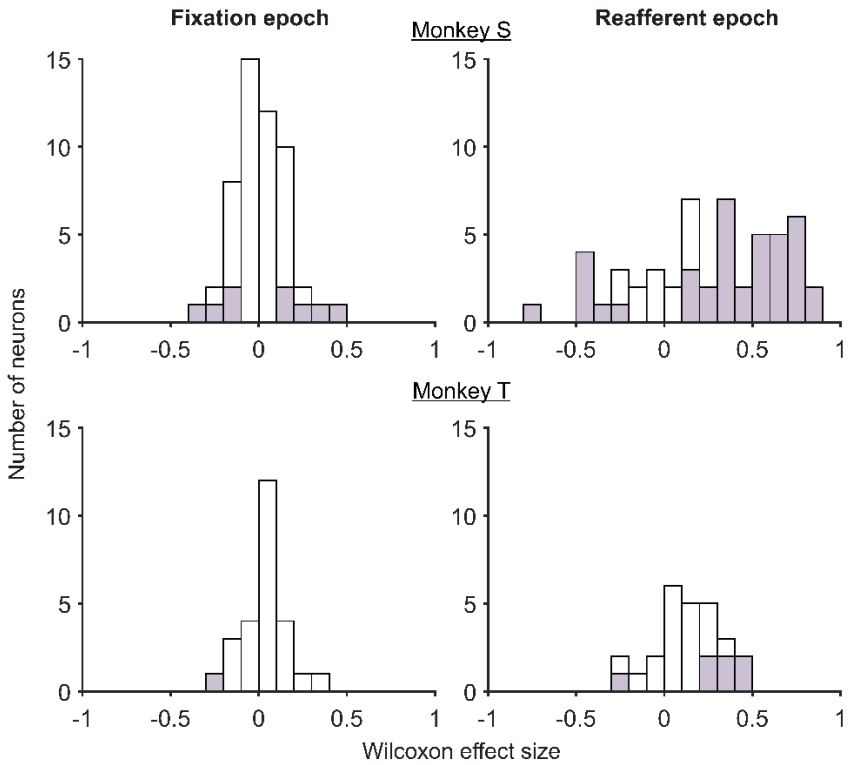

Figure S5. Same as Fig. 2b – c but for each animal individually. Top row: Monkey S, bottom row: Monkey T.

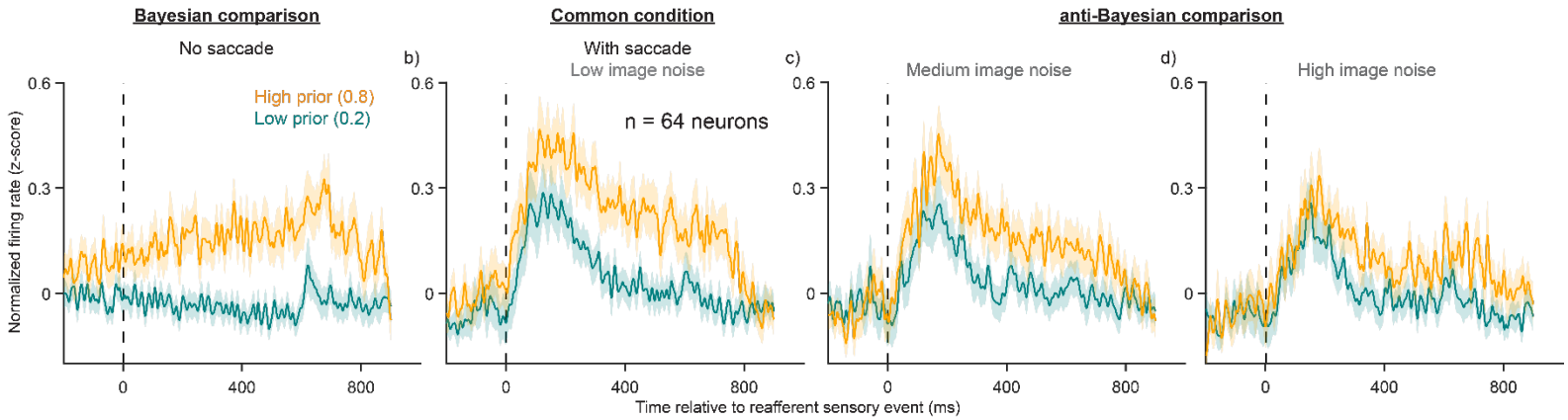

Figure S6. Same as Fig. 4a-d but restricted to trials where the jump size was 0.

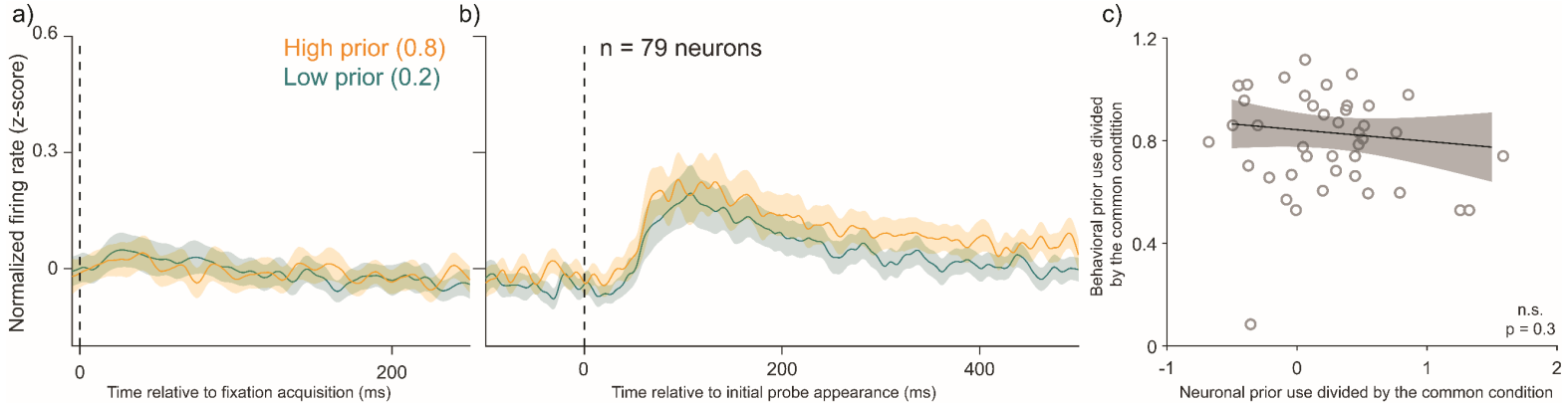

Figure S7. a-b) Peri-stimulus time histograms in the fixation (a) and probe epochs (b) for the no-saccade conditions show no difference in firing rates across the prior conditions. A clear difference emerged only around the time of the sensory event (probe jump) in this condition as well (shown in Fig. 4a). c) There was additionally no correlation between neuronal effect sizes in the probe epoch (0-500 ms from probe appearance) normalized to the common condition and normalized behavioral prior use in the no-saccade, Bayesian comparison condition.