Contributions of Bayesian and Discriminative Models to Active Visual Perception across Saccades

by

Divya Subramanian

Department of Neurobiology
Duke University

Date:_____

Approved:

_____
Marc Sommer, Supervisor

_____
Greg Field

_____
Stephen Lisberger

_____
Jeffrey Beck

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Neurobiology in the Graduate School
of Duke University

2022

ABSTRACT

Contributions of Bayesian and Discriminative Models to Active Visual Perception across Saccades

by

Divya Subramanian

Department of Neurobiology
Duke University

Date:_____

Approved:

_____
Marc Sommer, Supervisor

_____
Greg Field

_____
Stephen Lisberger

_____
Jeffrey Beck

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Neurobiology in the Graduate School of Duke University

2022

Abstract

The brain must interpret sensory inputs to guide movement and behavior, but movements themselves disrupt sensory inputs. Maintaining perceptual continuity through these disruptions requires one to resolve whether sensory inputs were externally generated or caused by one's own movements. Understanding the sensory world while moving through it constitutes *active perception*. Saccadic eye movements in primates are a good model system for studying active perception. Eye movements displace the image sensed by the eye, yet the visual system can distinguish movement-induced displacement from external object displacement. How does the brain resolve this uncertainty?

One way is to directly *discriminate* between sensory states and map them onto percepts in a bottom-up manner. Alternatively, the system could develop an internal model of the world it could use to generate predictions for its sensory inputs. If the input is then ambiguous, the system can default to its predictions more for perception. *Bayes rule* formalizes how internally generated predictions may compensate for sensory uncertainty. The goal of this dissertation is to investigate the relative contributions of Discriminative and Bayesian processes to active visual perception across saccades.

We performed a series of psychophysical, computational, and neural recording experiments grounded in variations of a task known as "saccadic suppression of displacement," in which subjects report whether a visual object moved while they made a saccade. First, we found that when humans provided *continuous* estimates of where an object landed across a saccade, they used a Bayesian model. That is, they used internally generated predictions, or *priors*, to compensate for sensory uncertainty. However, when asked to provide a categorical report ("did the object move? Yes or no?") in the same task, they were Anti-Bayesian. They used their priors less with increasing uncertainty. Further investigation in another primate species, rhesus macaques, showed that in the categorical task, priors were used more to compensate for motor-induced uncertainty generated by the saccade. When visual noise was added to the viewed object, however, prior use was Anti-Bayesian, consistent with results from human participants. Decreasing prior use was explained by a Discriminative, neural network model instead. In the macaques, we then recorded single neuron activity during the categorical tasks in a brain region known to signal object displacement across saccades, the Frontal Eye Field (FEF). We compared FEF activity to Bayesian and Discriminative behavior in the motor- and image-noise tasks, respectively. The results showed a clear distinction: the activity of FEF neurons predicted Discriminative but not Bayesian behavior.

In summary, we show that the selection of Bayesian vs. Discriminative models depends on both task requirements and the source of uncertainty. Further, a neural pathway which includes FEF selectively predicts behavior consistent with the use of Discriminative model, implying that the Bayesian model is implemented in a different circuit. These results demonstrate a dissociation between Bayesian and Discriminative models at the computational and neural levels and set the stage for understanding how they interact for perception across saccades.

Dedication

To my parents and grandparents.

Contents

List of Tables

No table of figures entries found.

List of Figures

Acknowledgements

This dissertation is the outcome of a lifetime of guidance, mentorship, friendship, and support from people with whom I am lucky to have crossed paths.

First, I am deeply grateful to my advisor, Marc Sommer. Marc, thank you for granting me the independence and the space to discover and develop my voice as a scientist. At the same time, thank you for sharing your wisdom, enthusiasm, and sense of wonder with me on this journey. I have grown as a scientist and communicator because of you. I hope to approach life the way you do: with optimism, generosity, an open mind, and a sense of humor.

Thank you to my committee, Steve Lisberger, Greg Field, and Jeff Beck for your insights and support. Greg, thank you for going above and beyond to be a mentor and advocate.

Thank you to my collaborator, John Pearson. John, your astute insights, typically provided in the form of a passing remark in a meeting, have helped me get unstuck on problems several times. I am grateful to have worked with and learned from you.

I would like to thank my communities at Duke and in Durham for being a wonderful support system through the years. First, thank you to all the members of the Sommer lab for being brilliant, generous colleagues. In particular, I would like to acknowledge Anthony Alers, who is not only intellectually and technically gifted, but also extraordinarily generous with his time and talents. I would also like to

acknowledge Jessi Cruger for all her behind-the-scenes work that made this dissertation possible, for caring for the animals used in these studies with compassion, and for being a friend. Thank you to all my friends for all the love and support. To the peas, Brenda Yang, Charlotte Moore, Hannah Moshontz, Jane Tandler, and Paula Yust, I have grown as a human due to your friendship. I am so grateful we found each other.

Thank you to my childhood friends whose presence in my life will always remain a source of comfort. To Ridhi and Tanvi, thank you for unconditionally being there for as long as I can remember. Your friendship means the world to me.

To my partner, Vivek, I don't know how I lucked into getting to spend my life with you. Thank you for somehow always knowing the right thing to say or do. Thank you, also, for filling our home with lightness, for having "enjoyed" taking tasks off my plate during crunch times, and for showing me that life really is better when we're willing to laugh at ourselves. Much love to our pup, Bhoomi, who has an uncanny ability to convert treats into pure joy.

Most importantly, no words can do justice to the sincerest love and gratitude I feel towards my family – mom, dad, Sriram, and my grandparents – who have of course known me since before I knew myself. Thank you for making my dreams yours, and for learning with me what goes into following them.

1. Introduction

1.1 Active Perception: Situating the self within the world

Perception is a constructive process in which sensory inputs are mapped onto meaningful percepts. For example, consider vision. Although we "see" the world effortlessly, the only external physical input available to the visual system is light. Patterns of light activate a sheet of light-sensitive photoreceptors in the eye. The visual system, given only the activity of a sheet of sensory receptors, manages to create the experience of a coherent world that can be interpreted and navigated. How are physical inputs processed to construct these useful experiences we call *percepts?*

The computational and neural processes that link sensory inputs to percepts can be split into two theoretical stages. First, a stimulus in the world is detected by receptors and *encoded* into signals that are meaningful to the rest of the system (Barlow, 1961). Photoreceptors convert light into electrical signals, which are packaged by the retina into neural signals, i.e., action potentials. In the second stage, the neural signals must be interpreted or *decoded* to reconstruct the stimulus (Johnson, 2000; Britten et al., 1996) in a form that is useful for guiding the appropriate behavior (Fig. 1).

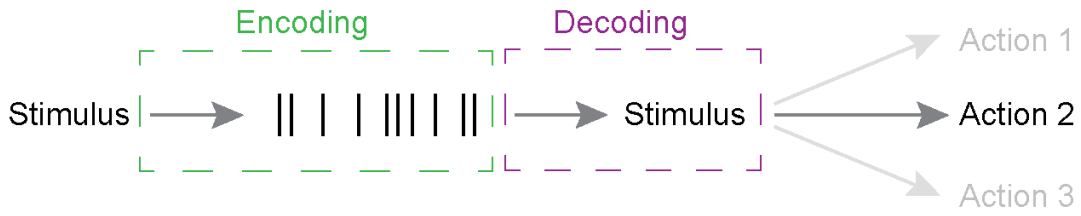

Figure 1. Perception as encoding and decoding for action selection. A stimulus is encoded into neural spike trains (green dashed box). The spike train is used to reverse-engineer or decode the original stimulus (purple box). The decoded representation can then guide the appropriate action.

For example, consider the action of reaching out towards a glass of water. Sensory receptors might encode the distance of the glass from one's hand as a pattern of action potentials or spikes. A decoder can decode the specific pattern and map it onto the action of extending the arm by that distance. The action of reaching out itself, however, continuously changes the distance of the glass from the hand. Further, imagine that in the process of reaching out to the glass, you leaned forward a bit. That would additionally change the location of the glass relative to your eyes and disrupt any existing encoding-decoding-action mappings. In other words, our own behavior is a constant source of disruption and change in our sensory inputs. Generating coherent percepts of the world while we navigate and move through it constitutes *active perception* (Gibson, 1966; Bajcsy, 1988). Critically, accounting for active perception requires updating the encoding-decoding-behavior framework to an encoding-decoding-behavior-encoding loop (Fig. 2).

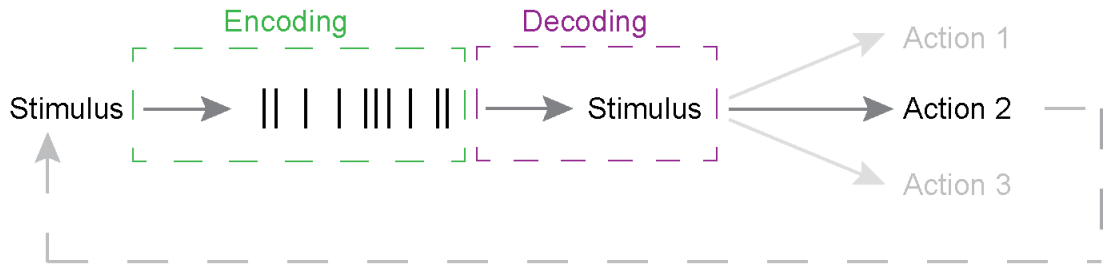

Figure 2. The Active Perception loop. A stimulus is encoded and decoded to guide action, but the action itself changes the stimulus.

This dissertation builds on the assumption that there are computational and neural circuit motifs underlying the active perception loop in the brain, and that understanding these motifs is critical to understanding our moment-to-moment perceptual experience. Specifically, it focuses on a key aspect of active perception, namely, that the brain must distinguish which sensory inputs were caused by oneself and which ones due to external agents or events in the world. For example, imagine you are walking through the woods. Your footsteps cause rustling and snapping sounds, but those sounds might also come from a lurking predator. Assigning sensory inputs to the appropriate source – self-generated vs. external – is a critical component of normal perceptual functioning, and its dysfunction may contribute to hallucinations and delusions that characterize disorders such as psychosis and schizophrenia (Feinberg and Guazzelli, 1999; Ford and Mathalon, 2005; Thakkar and Rolfs, 2019).

1.2 Prediction and Probabilities: Bayesian inference as a candidate computational motif for active perception

1.2.1 Predictive perception

How does the brain distinguish self-generated sensory inputs from external inputs? When a movement is made, a copy of the movement command, called the "corollary discharge", is simultaneously broadcast to other areas in the brain (Sperry, 1950; von Holst and Mittelstaedt, 1950). Forward Models formalize how corollary discharge may be used to distinguish self-generated sensory inputs from external inputs (Jordan and Rumelhart, 1992; Wolpert and Miall, 1996). Corollary discharge, in conjunction with information about the current state of the world and oneself, can be used to *predict* the sensory consequence of the movement. This prediction may then be compared to the sensory input to decide if there is a discrepancy, or prediction error, between the prediction and the input (Fig. 3). A prediction error would indicate that the sensory input might have had an external cause.

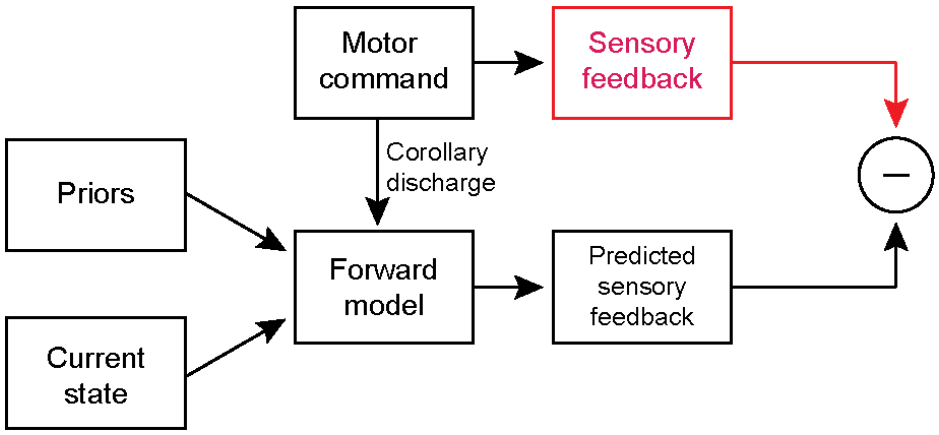

Figure 3. Schematic of Forward models.

While Forward Models formalize the relationship between motor predictions and sensory information, a related idea called Predictive Coding proposes that sensory predictions may arise from later stages in sensory processing itself (Rao and Ballard, 1999). Many sensory systems, vision included, are organized hierarchically in that the earlier stages of the system process the simpler, more elemental components of the world and sequentially combine them to create more complex representations (Felleman and Van Essen, 1991). Such systems are well-suited to make predictions about the sensory input at each level in the hierarchy via feedback connections. The input at each level can then be used to compute a prediction error. Section 1.3.2 addresses how the related notions of Forward Models and Predictive Coding may be combined in the visual system. Overall, active perception requires that internally generated predictions and external sensory inputs combine to create our perceptual experience.

1.2.2 Probabilistic perception

A further consideration is that external and biological signals are often the result of imperfect processes. Thus, these input signals have considerable variability across instances, and it is challenging for a perceptual system to know what's out in the world with certainty. Even for the simplest sensory encoding → perceptual decoding case, the same stimulus presented over and over will result in a slightly different pattern of neural activity each time (Fig. 4). An effective decoder must thus be capable of reconstructing the same stimulus from that variable pattern of activity. Probabilistic theories of perception provide a framework for thinking about how a system might deal with uncertainty (Murphy, 2013; Ng and Jordan, 2002).

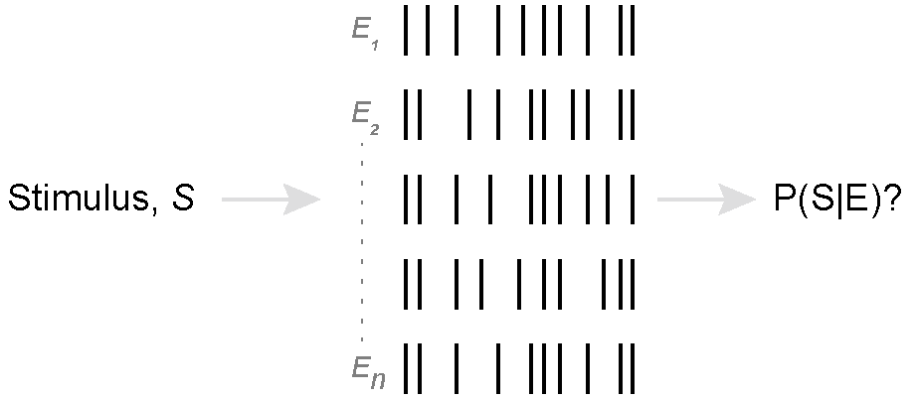

Figure 4. Variable encoding and probabilistic decoding.

Consider the instance of a simple stimulus, *S*. Receptors and neurons early in the visual system encode the stimulus and provide evidence, *E*, for it to the rest of the visual

system. Across instances, E consists of a variable set of signals E_1-E_n, each of which provides an imperfect encoding of S. The goal of a decoder is to reverse-engineer the stimulus given the evidence. Since the mapping between the evidence and the stimulus cannot be one-to-one given the variability in E, the probabilistic casting of this problem is that the decoder must infer the *probability* of having encountered the stimulus, S, given the evidence, E. That is, $P(S|E)$.

Models of decoding can take two broad forms. *Discriminative* models simply learn the boundaries between various evidence states and learn to map stimulus states onto them (Rumelhart et al., 1986; Hinton, 1992). In other words, they directly learn P(S|E) (Fig. 5, top). *Bayesian* (or Generative) models, by contrast, build and use internal models of the world to infer P(S|E) (Fig. 5, bottom) (von Helmholtz, 1924; Rao and Ballard, 1999; Knill and Richards, 1996).

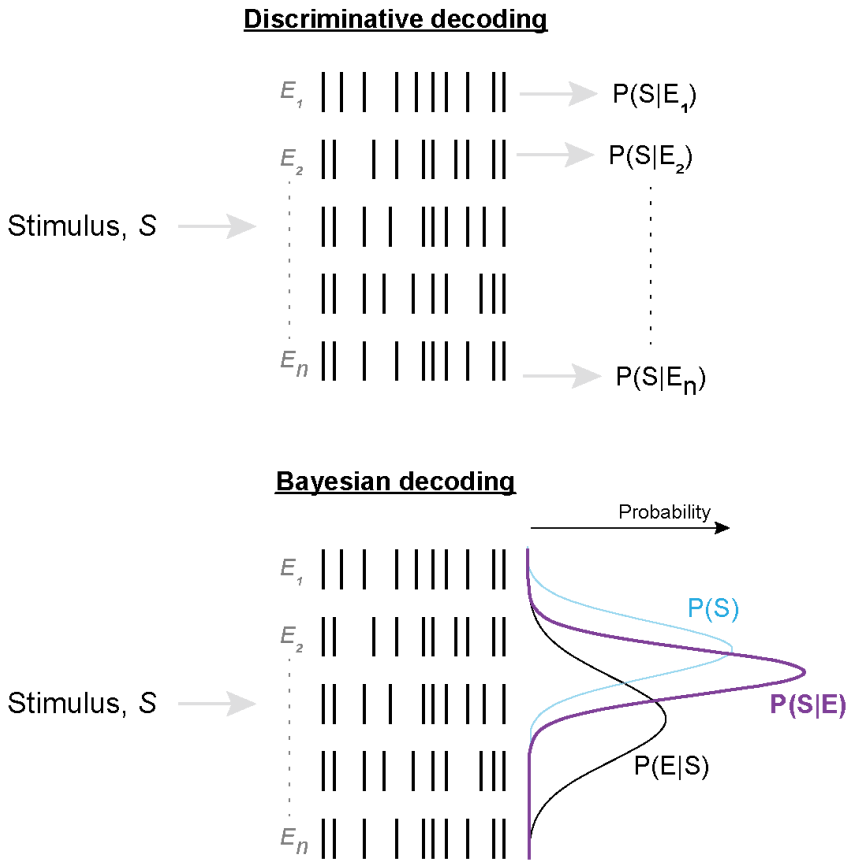

Figure 5. Discriminative (top) and Bayesian (bottom) decoding.

Under a Bayesian framework, the uncertainty in encoding can be captured as the width of the sensory *likelihood* distribution, $P(E|S)$, which is the probability of the encoding evidence, E, given the stimulus, S. P(S|E), the *posterior*, can be inferred from the likelihood, P(E|S), using Bayes' rule if one has an internal model or *prior* of possible stimulus states, P(S):

$$P(S|E) = \frac{P(E|S)P(S)}{P(E)} \tag{1.1}$$

Discriminative and Bayesian models each have distinct strengths and are thought to combine for perception (Gardner, 2019; Sohn and Jazayeri, 2021; DiCarlo et al., 2021). For example, Discriminative models are flexible but vulnerable to sensory input uncertainty that may arise from noise in the environment or receptors. If the input signals are noisy, then they have a higher chance of overlapping with the discrimination boundary, thus blurring the distinction between possible stimulus states. Bayesian models, on the other hand, use prior knowledge to "optimally" resolve input noise. The Bayesian estimate of the stimulus (i.e., the posterior, purple distribution in Fig. 5, bottom) is the optimal combination of information provided by the prior (blue distribution) and the sensory input (i.e., the likelihood, black distribution) because 1) its center (the final estimate) is closer to the less variable of the two sources of information (e.g., the purple distribution is closer to the blue than the black distribution), and 2) the posterior is less variable than each source individually (in Fig. 5, the purple posterior distribution is less wide than the prior and likelihood distributions). That is, assuming a constant prior, Bayesian models can shift their estimates to rely on their priors more when sensory uncertainty is greater.

Due to their robustness to noise, Bayesian models have been successful in explaining behavior in a wide range of sensorimotor contexts including depth perception (Jacobs, 1999), multi-sensory integration (Ernst and Banks, 2002; Battaglia et al., 2003; Fetsch et al, 2012), motion perception (Weiss et al, 2002), sensorimotor

integration (Kording and Wolpert, 2004; Darlington et al., 2017), interval timing (Jazayeri and Shadlen, 2010), and orientation detection (Girschick et al, 2011). Further, Bayesian inference provides a formal structure by which a sensory system may implement probabilistic predictive coding (Aitchison and Lengyel, 2017).

Given its theoretical appeal and empirical success, Bayesian inference has been postulated to be a fundamental computational motif for perception (Knill and Richards, 1996; Knill and Pouget, 2004, Friston, 2012; Pouget et al., 2013). The overall hypothesis of this dissertation is that Bayesian inference is a useful model for active perception. My goal is to understand how Bayesian inference might be computationally and neurally implemented for active perception, and my approach is to investigate this using psychophysics, computational modeling, and neurophysiology in the specific context of seeing while moving the eyes.

1.3 The Saccadic System: A good model system for studying active perception

In primates, saccades are a ballistic eye movement made to focus targets of interest in the environment on a central region of high-resolution vision on the retina called the fovea. Saccadic eye movements serve as a useful, tractable model system for studying how Bayesian Inference is computationally and neurally implemented for active perception. Since eyes themselves are the visual sensor, there is a necessary, principled coupling between incoming visual information and eye movements. Thus,

the visual system in primates is inherently an active perceptual system. In addition, the saccadic system is well-studied both at the level of perceptual effects and its underlying neural circuitry. Indeed, neural circuits for corollary discharge have been identified in the saccadic system, and their interaction with visual processing (Section 1.3.1) is well-suited to implement Forward Models and Predictive Coding (Section 1.3.2). This positions us well to make and test informed hypotheses about where in the brain and how a computational motif like Bayesian inference might be implemented for active perception.

1.3.1 Corollary Discharge and Visual Processing

On average, primates make about 3 saccades a second. The purely visual feed coming into the eye is thus like that of a camera held by unsteady hands. The visual system corrects for the effects of saccades in two major ways. First, it corrects for blur by suppressing visual processing around the time of saccades. This phenomenon is called Saccadic Suppression (Zuber and Stark, 1966; Bridgeman et al, 1975) Second, it constructs a continuous, predictable percept out of the" snapshots" presented to it before and after a saccade (Wurtz, 2008). A specific problem that it must solve while stitching together these snapshots is deciding whether an object in the environment moved relative to its pre-saccadic location. Each saccade causes the visual image to be displaced on the retina, but objects in the world may also move. Thus, across every

saccade, the visual system must distinguish saccade-induced retinal displacement from external retinal displacement (Fig. 6). The process by which the visual system achieves this is the focus of this dissertation.

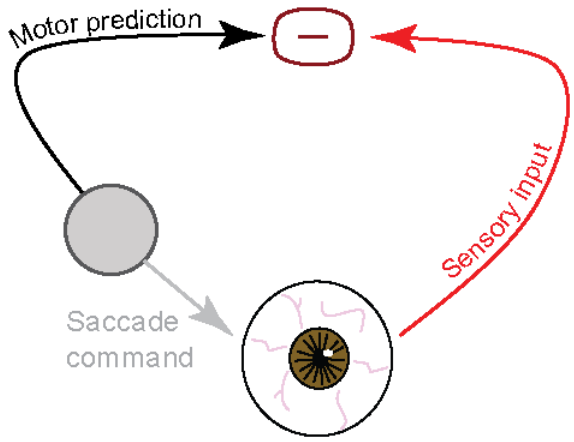

Figure 6. Distinguishing saccade-induced vs. externally caused retinal displacement.

For both compensation functions, pathways from a saccade-generating brainstem region called the Superior Colliculus (SC; represented by the gray circle in Fig. 6) to the neocortex carry information that relates upcoming saccades to visual processing. First, a pathway from visually responsive neurons in superficial SC to the Middle Temporal visual area (MT), which processes motion information, has been shown to carry suppressed visual activity around the time of saccades (Goldberg and Wurtz, 1972; Robinson and Wurtz, 1976; Berman and Wurtz, 2010; Berman et al., 2017). This pathway travels via a thalamic nucleus called the Pulvinar. Saccadic suppression is, however, a

gross sensory suppression and cannot account for how the system maintains a continuous percept or distinguishes self-generated retinal displacement from external movement. A second corollary discharge pathway serves this function. Sommer and Wurtz (2004a, b; 2008) discovered a pathway from saccade-generating neurons in the Superior Colliculus (SC) to the Frontal Eye Field (FEF), via the Mediodorsal thalamus (MD). The impairment of corollary discharge from inactivating this pathway can be dissociated from motor impairment using the double-step task, in which subjects make saccades to the remembered locations of two sequentially flashed targets in the order that they appeared. To make the second saccade correctly, one must monitor where the first saccade landed, information provided by corollary discharge. It was shown that inactivating MD spares the ability to make both saccades but impairs the accuracy and precision of the second one, revealing a selective deficit of corollary discharge (Sommer and Wurtz, 2002; 2008).

The SC-MD-FEF corollary discharge pathway, via its influence on visual processing in FEF, supports the continuity of visual perception across saccades. One such influence is the pre-saccadic remapping of visual receptive fields (Sommer and Wurtz, 2006). Neurons in the visual system, beginning at the retina, respond to a specific region of space. This region of space, the neuron's window to the world, is known as its receptive field. Many neurons across several visual regions including FEF shift their

responsiveness just *before* a saccade to the part of the visual field where their receptive field will land *after* the saccade (the "future field") (Fig. 7).

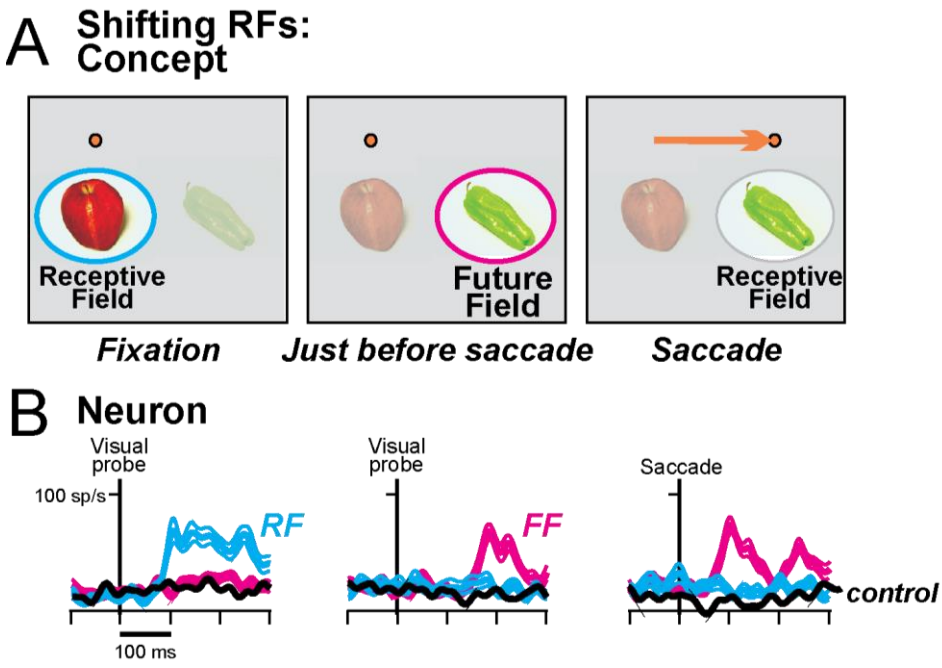

Figure 7. Pre-saccadic remapping of visual receptive fields. During fixation (at the orange dot), the neuron's responsiveness is at its receptive field (blue ellipse). Just before a saccade, its responsiveness shifts to its future field, i.e., its post-saccadic receptive field. Reproduced from Sommer, M. A., & Wurtz, R. H. (2008). Brain circuits for the internal monitoring of movements. ***Annual Review of Neuroscience*****, *31*, 317. https://doi.org/10.1146/annurev.neuro.31.060407.125627, Figure 3**

Since pre-saccadic remapping links visual processing before and after a saccade, it is thought to contribute to the perception of visual continuity across saccades (reviewed in Hall and Colby, 2011; Wurtz, 2008; Wurtz, 2018). Critically, inactivation of

the corollary discharge pathway decreased FEF remapping and disrupts the perception of visual continuity across saccades (Cavanaugh et al., 2016).

1.3.2 Predictive Visuo-saccadic Processing

Does the identified SC-MD-FEF corollary discharge pathway in the saccadic system participate in a Forward Model to support predictive processing? Justification for this claim depends on evidence consistent with three major features of any forward model. First, the corollary discharge would need to generate sensory predictions about the consequence of movement. This capability is demonstrated, as discussed above, by the role of the SC-MD-FEF pathway in presaccadic remapping. Neurons that remap seem to sample, before each saccade, the part of visual space that they will "see" after the saccade. Presaccadic remapping in the FEF therefore provides a signal in visual coordinates that is appropriate for predicting the consequence of each saccade (Crapse and Sommer, 2008; Sommer and Wurtz, 2008; Rao et al., 2016).

Presaccadic remapping also links Forward Models, which have been proposed for motor-sensory processing, with Predictive Coding models that have traditionally been proposed in the sensory domain. Neurons that remap at later stages in visual processing such as in FEF can serve as the source of the feedback predictions postulated by Predictive Coding models (Fig. 8).

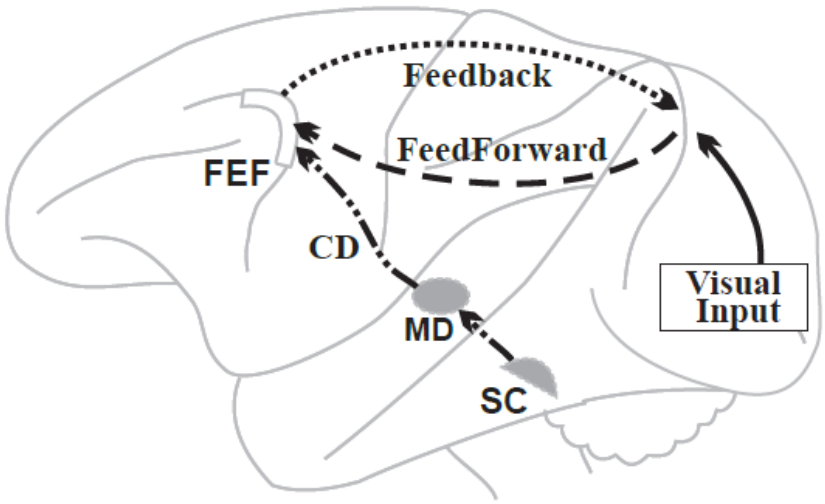

Figure 8. Pathways supporting visuo-saccadic predictive coding. Reproduced from Crapse, T. B., & Sommer, M. A. (2008). The frontal eye field as a prediction map. *Progress in Brain Research*, *171*, 383–390. https://doi.org/10.1016/S0079-6123(08)00656-0, Fig. 4b

Indeed, analysis of microcircuitry in the FEF suggests it could generate the remapping signal locally (Shin and Sommer 2012), and projections from the FEF could relay the signals back to extrastriate cortex for prediction of incoming visual input (Stanton et al., 1995; Schall et al., 1995). Even though the FEF is the only known target of a corollary discharge pathway in the saccadic system, remapping is found in earlier visual areas as well. It is prevalent in the lateral intraparietal cortex (LIP) (Duhamel et al., 1992), V4 (Neupane et al., 2016a, b), and MST (Inaba and Kawano, 2014), but diminishes lower in the visual hierarchy from V3 to V2 and V1 (Nakamura and Colby, 2002; reviewed in Hall and Colby, 2011). This waning of remapping from higher to

lower visual areas is consistent with its possible function as a top-down mechanism for prediction of visual input (Crapse and Sommer, 2008; Rao et al., 2016). Presaccadic remapping has been reported in human visual cortex as well, with a similar decrease in remapping from higher to lower visual areas (Merriam et al., 2003; 2007).

The second feature of a forward model is that the predicted sensory consequence of a movement is compared with the actual sensory consequence, i.e. the "reafferent" response. The difference between these signals, the prediction error, is used to update future predictions. Crapse and Sommer (2012) tested whether FEF neurons make such comparisons about predicted versus reafferent visual responses across saccades. They recorded from FEF visual neurons in monkeys that made saccades to visual targets in the presence of a behaviorally irrelevant visual probe. During each saccade, the probe would stay still or move, but after the saccade it was always at the center of the neuron's postsaccadic receptive field. Because the postsaccadic location of the probe was constant, the null hypothesis was that reafferent responses would be constant. The result, however, was that reafferent responses varied as a function of how much the probe moved during the saccade. The reafferent responses also were sensitive to other changes in the probe (size or color) across saccades. Hence, FEF neurons exhibit a robust prediction error signal about transsaccadic changes of visual stimuli.

A final feature of a forward model is that it must account for the current state of the system, in conjunction with corollary discharge, while making predictions. The

current state of the system includes knowledge, or *priors*, about the effectors or the environment. A prior about effectors could be, for example, information that one of the eyes is injured, limiting the range of saccadic motion. The forward model would need to account for the gain reduction to generate accurate predictions. A prior about the environment could include knowledge about objects in the world, such as their propensity to move. Animals often move, for example, while rocks rarely do. Accounting for the intrinsic nature of objects would improve predictions about their states across saccades. Rao et al. (2016) found that humans take priors about the environment into account while judging whether an object moved across a saccade. Multiple lines of evidence therefore support the hypothesis that corollary discharge in the SC-MD-FEF pathway contributes to predictive processing in the visuo-saccadic system.

1.3.3 Probabilistic Visuo-saccadic Processing

As with other systems, a visuo-saccadic forward model faces the issue that its multiple sources of information are all likely to be variable across instances and, therefore, noisy and uncertain. For example, the motor-driven prediction, the sensory input, and priors are all potential sources of noises (van Opstal, & van Gisbergen, 1989; Collins et al., 2009; Bansal et al., 2015). As noted in Section 1.2.1, the Bayesian prediction is that the judgment of object displacement is closest to the most reliable source of

information about it. Indeed, previous work has shown that saccadic suppression, one main consequence of saccadic corollary discharge, is the result of an optimal Bayesian calculation between motor predictions and sensory inputs in the saccadic system (Niemeier et al., 2003; Crevecoeur and Kording, 2017).

The current work focuses on testing the Bayesian hypothesis for *visual continuity*, the other function of saccadic corollary discharge, across saccades. To evaluate whether an object in the environment moved, the saccadic system must make a corollary discharge-driven prediction about the sensory consequences of the saccade and compare it with the post-saccadic sensory input. Under a probabilistic framework, the goal of the decoder is to infer the probability that an object moved given the evidence for a prediction error (detailed model in Chapter 2). Rao et al. (2016) demonstrated that expectation-based priors are used for this process. The studies in this Dissertation build on this finding and test the Bayesian hypothesis directly. We tested whether expectation-based priors are used more when there is sensory uncertainty in distinguishing self-generated retinal displacement from external object movement.

1.4 Dissertation roadmap

The experiments reported in Chapter 2 tested whether human participants used prior expectations in a Bayesian manner for the judgment of object displacement across saccades. Contrary to our expectations, we found that participants were *Anti-Bayesian*:

they used their priors less with increasing sensory uncertainty. We replicated this finding in two rhesus macaques trained on a comparable task. We next demonstrated that the Anti-Bayesian prior use was specific to categorical judgments, because when human participants were required to provide continuous rather than categorical reports, their behavior became Bayesian. The results of Chapter 2 showed that the saccadic system is Bayesian under some circumstances but not others, implying that it sometimes operates under a Discriminative regime.

In Chapter 3, we first examined the conditions under which the system selects Bayesian vs. Discriminative models. We focused on the categorical task in Chapter 2 to understand the origins of the unexpected Anti-Bayesian, putatively Discriminative behavior. We tested the Bayesian hypothesis separately for sensory noise and motor noise to see if one of them selectively drives Anti-Bayesian behavior. We leveraged the precise eye movement monitoring and extended within-subject testing possible in rhesus macaques for this. We found that monkeys were Bayesian in using priors to compensate for motor noise, but not for visual noise. We then identified a simple neural network model that explains how a Discriminative model boundary may be drawn between categories to account for the Anti-Bayesian results. The outcomes of Chapter 3 were to identify the relative contributions of Bayesian and Discriminative models to active visual perception, and to make predictions for what the neuronal correlates of each would look like.

In Chapter 4, we examined the activity of neurons in the Frontal Eye Field (FEF) to test whether they supported the Bayesian model, Discriminative model, or both. We chose the FEF as a candidate brain region since, as detailed in Sections 1.3.1 and 1.3.2, it is well-positioned to use corollary discharge information to make sensory predictions and evaluate incoming visual information for a discrepancy. The first question was thus whether FEF neurons also reported an expectation-based prior. If so, are the priors modulated in accordance with Bayesian behavior or Discriminative behavior? That is, does prior modulation in FEF neurons increase with increasing uncertainty (Bayesian) or does it decrease (Discriminative)? We found that FEF neurons do encode expectations. Additionally, the modulation of prior-driven activity predicted Discriminative behavior but not Bayesian behavior. Thus, neuronal activity in FEF better correlates with Discriminative, categorical behavior than Bayesian categorical behavior in the context of sustaining visual continuity across saccades.

Finally, Chapter 5 synthesizes the results from Chapters 2, 3, and 4 in the context of our current understanding and outlines future directions.

2. Is the perception of object displacement across saccades Bayesian?

2.1 Introduction

The neural processes that mediate the mapping of sensory information onto meaningful percepts can be split into two theoretical stages. First, a stimulus must be detected by sensory receptors and encoded into neural signals (Barlow, 1961). Thus, sensory receptors provide evidence for the stimulus to the rest of the system. In the second stage, the evidence must be decoded to reconstruct the stimulus and guide behavior (Johnson, 2000; Britten et al., 1996). However, both sensory inputs and early sensory processing can be variable across instances leading to uncertainty in interpreting them. Bayesian models have formalized how sensorimotor systems may use priors to compensate for sensory uncertainty (Knill and Richards, 1996; Knill and Pouget, 2004, Friston, 2012; Pouget et al., 2013).

In addition to input noise, one's own movements also disrupt sensory inputs to cause uncertainty in the evidence sent to the brain. Saccadic eye movements, for example, introduce uncertainty to visual signals. In primates, each saccade causes the image projected on the retinas to be displaced, but this displacement must be discerned from true object motion in the world. The visual system can do this by comparing the predicted sensory consequence of a saccade (Fig. 9, black arrow) with the incoming sensory input (Fig. 9, red arrow). Rao et al. (2016) showed that humans use priors about

the probability of an object moving or not in judging its stability across a saccade (Fig. 9, top). They used a modified Saccadic Suppression of Displacement (SSD) task (Bridgeman et al., 1975). Participants fixated a central location and, upon being cued, they made a saccade to a target. On some trials, the target was displaced or "jumped" during the saccade. Participants reported whether the target jumped or not. The color of the target corresponded with the probability that it would jump. When informed of this color-probability association, participants behaved as if they used the colors as expectations or "priors". They were more likely to respond that the target jumped if the prior indicated high probability of jumping and vice versa if the prior indicated low probability of jumping. Although Rao et al. (2016) established that priors can be used for the SSD task, whether that use was Bayesian remained an open question since sensory uncertainty was not systematically varied.

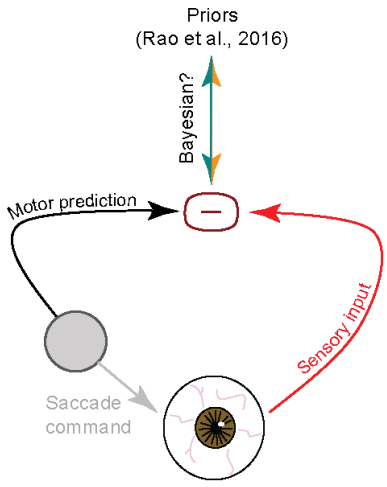

Figure 9. Components for judging object displacement across saccades.

The overall design of the current study extended the SSD paradigm of Rao et al. (2016) to test the Bayesian hypothesis that priors are used more when sensory uncertainty about object displacement is higher. In Experiment 1, we identified a stimulus manipulation that reliably decreased perceptual sensitivity, thus inducing sensory uncertainty. Experiment 2 tested the Bayesian hypothesis in human participants who provided a categorical (binary) report of whether the peripheral stimulus moved or not during the saccade. To our surprise, we found that participants were Anti-Bayesian: they used their priors less with increasing noise. We replicated this finding in two rhesus macaques trained on a comparable categorical task. Experiment 3 demonstrated that this Anti-Bayesian prior use was specific to categorical judgments, because if human participants were required to provide continuous rather than categorical reports, their behavior became Bayesian.

2.2 General Methods: Human Psychophysics

2.2.1 Materials and Paradigm

45 adult volunteers with normal or corrected-to-normal vision participated in the experiments. All procedures were explained verbally to participants beforehand and written, informed consent was obtained. Participants were paid $12/hr. and informed

that participation was completely voluntary. All procedures were performed in accordance with protocols approved by the Duke University Institutional Review Board.

Participants sat alone in a darkened room in front of a monitor with their head stabilized using a chin- and forehead-rest. The monitor was positioned at 60cm from the center of the head. Experiments 1 and 2 were displayed on a CRT monitor (Accusync 120) at 120Hz. Experiment 3 was displayed on a Dell LCD monitor at 60Hz. The experiment was written in and displayed using Presentation (Neurobehavioral systems). Monocular eye position was recorded with an eye-tracking software developed by Matsuda et al. (2014).

Participants performed a modified Saccadic Suppression of Displacement task (Bridgeman et al., 1975) (Fig. 10).

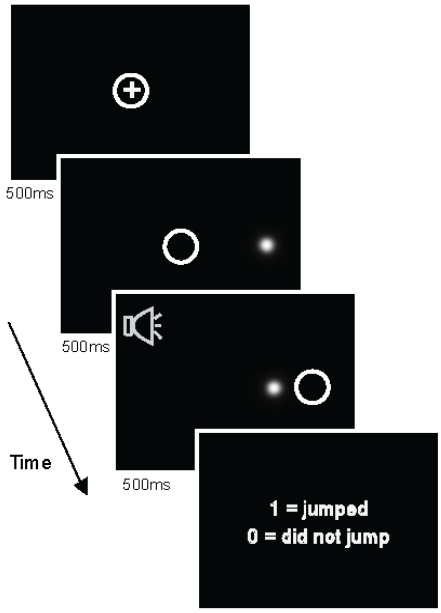

Figure 10. Modified Saccadic Suppression of Displacement task.

On each trial, a fixation cross first appeared near the center of the screen. Once participants had acquired and maintained fixation for 500ms, a saccade target appeared at one of two average positions relative to the center of the screen: 10° or -10°. A target at 10° appeared in the right half of the screen while a target at -10° appeared in the left half of the screen. Additionally, on every trial, the position of the target and fixation cross were both jittered by -0.5 to 0.5 degrees relative to the average position to mitigate the confounding effects of adaptation to either a constant saccade amplitude (Hopp and Fuchs, 2004) or a constant distance between the target and the edge of the screen. The fixation cross then disappeared for 500ms, and an auditory cue was presented to signal

to participants that it was legal to make a saccade to the target. If fixation was broken before the auditory cue, the trial was aborted and a new one began immediately. On some trials, saccade initiation (defined as the time the eye left a window of 2°. relative to the fixation cross) triggered target displacement. In Experiments 1 and 2, participants provided a binary, categorical report on whether they had perceived the target as having jumped or not. The target remained on the screen for 500ms after it was displaced, after which it was replaced by a response prompt screen (5 = moved, 0 = remained stationary).

In these experiments, target displacement was drawn from overlapping Gaussian distributions designated as the "jump" and "non-jump" distributions (Fig. 11).

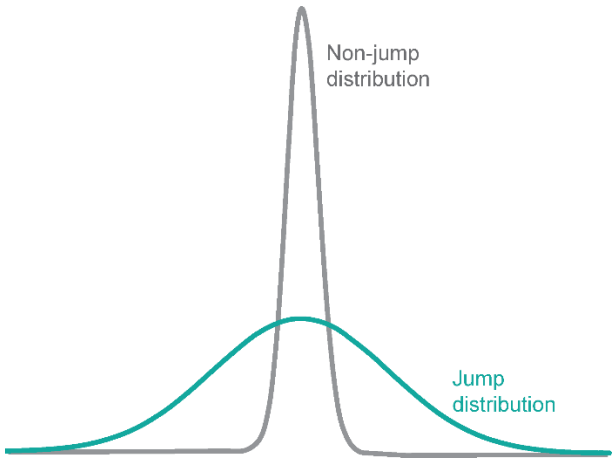

Figure 11. Schematic of jump and non-jump distributions (not to scale).

Displacements were drawn from overlapping, rather than distinct, distributions to ensure that the solution to the task was probabilistic. On trials where the target

moved, the displacement was drawn from a relatively broad Gaussian distribution centered around 0 ($\mu = 0°$, $\sigma = 1.5°$). On "no movement" trials, the displacement was drawn from a very narrow Gaussian distribution centered around 0 ($\mu = 0°$, $\sigma = 0.017°$). A positive displacement meant that the target moved rightward, and a negative displacement meant it moved leftward.

In Experiment 3, participants provided a continuous report of the target's postsaccadic location. For this study, the target stayed visible for 100ms after displacement and was then replaced by a screen where the mouse cursor (shaped "+") was placed at the center of the screen and restricted to the horizontal meridian. Participants could then move the mouse cursor to where they perceived the target as having landed.

2.2.2 Data preparation

Data from individual trials were analyzed offline to confirm that the visual probe landed in its displaced location before the end of the saccade. The saccade end time was defined as the time at which the eye velocity dropped below 40°/s. For human participants, the time at which the target jump command was sent was recorded for each trial. Trials with a recorded jump time greater than 1 whole frame (8.33ms for Experiments 1 and 2, and 16.7ms for Experiment 3) before the detected end of the saccade were excluded from analysis. Participants for whom at least 90% of all trials did

not meet this criterion were excluded from analyses entirely. No participants were excluded in Experiment 1, three participants were excluded from Experiment 2, and three participants were excluded from Experiment 3.

All data were analyzed using MATLAB (Mathworks, Inc.). Psychometric curves were fit to binary responses using a 4-parameter logistic regression model:

$$y = max + \frac{min - max}{1 + \left(\frac{x}{thresh}\right)^{slope}} \tag{2.1}$$

where x is the absolute value of the presented displacement, y is the value of the psychometric function, min is the minimum value of the function (i.e., y at $x = 0$), max is the maximum value, $thresh$ is the inflection point, and $slope$ is the slope of the psychometric function. min, max, $thresh$, and $slope$ terms were fit to binary data by minimizing mean squared error.

For Experiment 1, we quantified sensitivity to target jumps using d', a measure from Signal Detection Theory:

$$d' = Z(Hit\ rate) + Z(False\ alarm\ rate) \tag{2.2}$$

The hit rate is the proportion of "jumped" responses to the total number of trials where the target jumped. The false alarm rate is the proportion of "jumped" responses to the number of trials in which the target did *not* jump.

For Experiment 2, we quantified the upward and downward shifts of the psychometric curve using its intercept, i.e., its minimum value. The intercept is at jump size zero, so reports of "jumped" for those data are expected to directly indicate prior use. The vertical shifts were quantified to measure prior use for statistical tests and to compare the data with predictions of the Bayesian model. We repeated all the analyses of prior use presented in the manuscript using a measure from Signal Detection Theory, the Criterion. Criterion provides an alternative measure of bias in responses (i.e., a translational shift in psychometric curves). It is given by:

$$C = -0.5 \times \big(Z(Hit\ rate) + Z(False\ alarm\ rate)\big) \tag{2.3}$$

The results using Criterion replicated the findings using the intercept. These results are not shown.

For all statistical comparisons, the assumption of normality was first tested for each sample using a Kolmogorov-Smirnov (KS) test. If met, then we used a parametric comparison such as an ANOVA or a t-test. Otherwise, the appropriate non-parametric test was used.

2.3 Experiment 1: Choosing a sensory uncertainty manipulation

2.3.1 Methods

The goal of Experiment 2 was to test the use of prior expectations about object movement across saccades when the sensory is relatively uncertain. For this, we would need to experimentally manipulate participants' prior expectations and the sensory uncertainty of the visual probe. We planned to train participants on priors by associating the color of the fixation cross with a given probability that the target would move and reinforcing this association with performance-based feedback on each trial. We were relatively confident that this prior manipulation would be successful since it was only a slight variation from previous work in the lab demonstrated color of an object can be used to indicate the probability of movement across saccades (Rao et al., 2016). As a first step, therefore, we sought to confirm that our experimental manipulation for inducing sensory uncertainty reliably decreased discriminability in the task. Nine human participants completed at least 100 trials each in 8 experimental conditions: 4 candidate stimulus manipulations at two uncertainty levels each. The probability of target movement across all stimulus conditions was 0.5.

The 4 possible manipulations were:

1) Arrows (1° long with 0.5° width) that either pointed in the direction of the jump (congruent) or in the opposite direction (incongruent). We predicted that incongruent jumps might induce greater uncertainty and decrease discriminability.

2) A Gaussian cloud of 20 white squares (0.25° x 0.25°) where the uncertainty corresponded to the standard deviation of the cloud (low uncertainty = 0.0625° and high uncertainty = 0.25°).

3) Squares (0.5° x 0.5°) at two levels of contrast (low uncertainty = 78.4% and high uncertainty = 29.4%).

4) Gaussian "blobs" (isoluminant, Gaussian distributions of light) uncertainty corresponded to the standard deviation of the blob (low uncertainty = 0.19° and high uncertainty = 0.47°).

The results of this experiment would determine how we manipulated the sensory uncertainty of the stimulus in all other experiments.

2.3.2 Results

We assessed participants' (n = 9) sensitivity to target jumps under conditions of low and high uncertainty for 4 different target manipulations. For the arrow (Fig. 12, first column from the left) and Gaussian cloud (Fig. 12, second column from the left) manipulations, we found no significant difference between the low noise (mean = 2.02, SE = 0.16 for arrow; mean = 1.98, SE = 0.23 for Gaussian cloud) and high noise (mean = 2.24, SE = 0.16 for arrow; mean = 1.99, SE = 0.20 for Gaussian cloud) conditions ($p = 0.091$ and $p = 0.97$ for Gaussian cloud on a paired t-test). For stimuli with different contrasts (Fig. 12, third column from the left), sensitivity in the low noise condition (mean = 2.12,

SE = 0.16) trended higher than in the high noise condition (mean = 1.61, SE = 0.16), but the difference was not statistically significant at n = 9 participants (p = 0.055 on a paired t-test). For the Gaussian blob condition, however, d' in the low noise condition (mean = 2.08, SE = 0.15) was significantly higher than in the high noise condition (mean = 1.62, SE = 0.14; p = 0.0023; tested against α = 0.0125 Bonferroni-corrected for 4 multiple comparisons). Given the confirmation that it reliably induced sensory uncertainty, we chose the Gaussian blob (Fig. 12, right-most column, highlighted with the black box) as the target stimulus for following experiments.

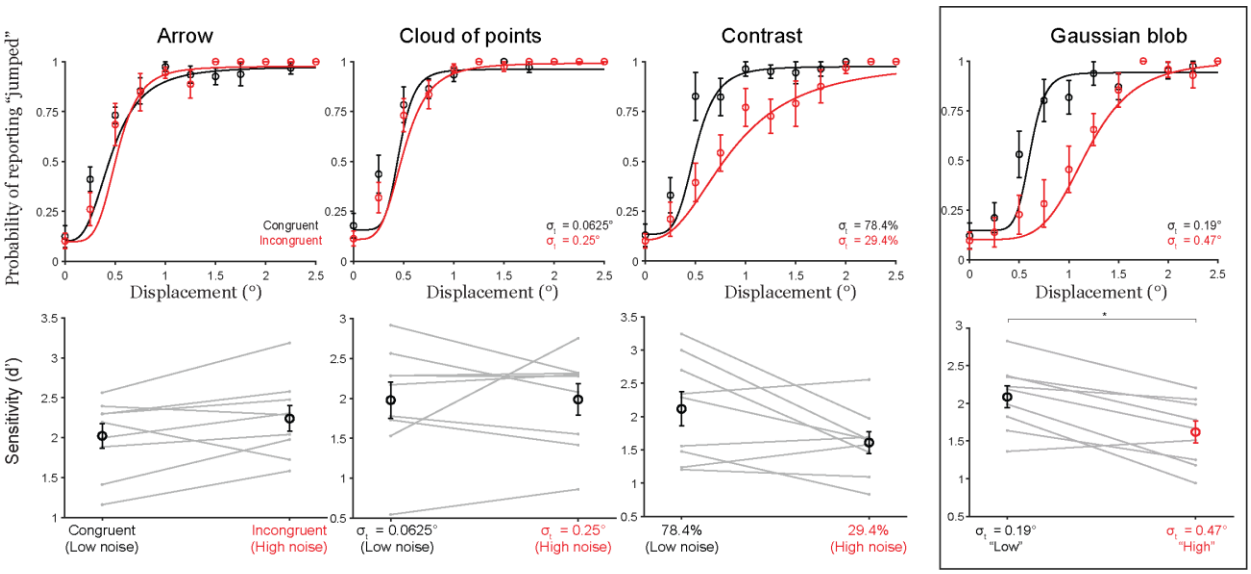

Figure 12. Results in the sensory uncertainty manipulation experiment.

2.4 Experiment 2: Testing the Bayesian hypothesis using a categorical prior

2.4.1 Experimental Methods

The overall goal of Experiment 2 was to test the hypothesis that participants use their learned priors more with increasing sensory uncertainty. We trained participants to develop expectations about the probability of target displacement (i.e., the prior) associated with the color of the fixation cross on each trial (Fig. 13, top dashed box). Participants were trained on this association using performance-based feedback. That is, they were told whether their responses were correct or incorrect on each trial using an image of a smiling or frowning face, respectively. Based on the results from Experiment 1, we chose the isoluminant Gaussian blob as sensory uncertainty manipulation. The target was white on every trial; only its width given by the standard deviation changed. The target had one of three possible standard deviations for the whole experiment: 0.1 deg. ("low noise"), 0.25 deg. ("medium noise) or 0.5 deg. ("high noise") (Fig. 13 left, bottom dashed box).

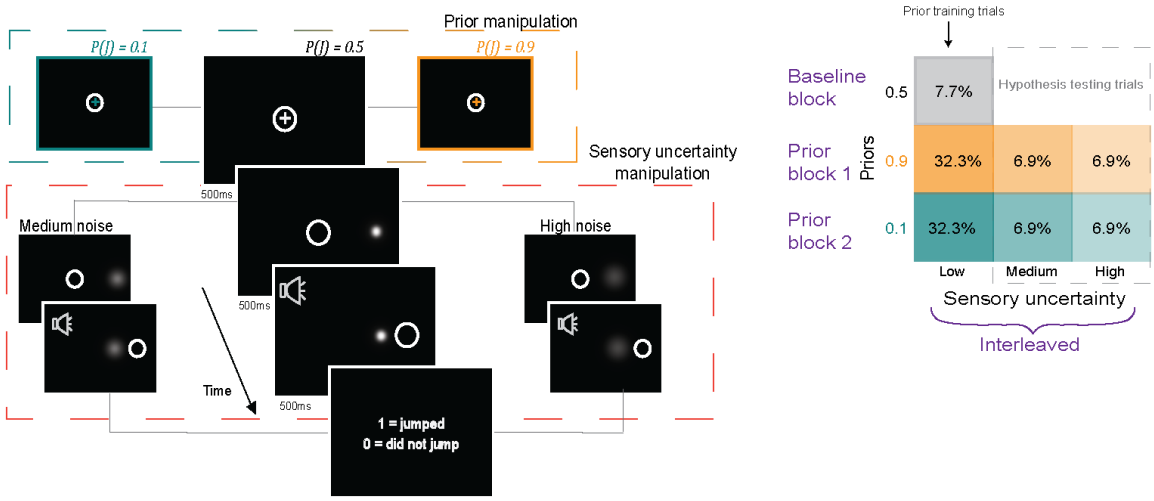

Figure 13. Categorical Bayesian task for human participants. Prior and sensory uncertainty manipulations shown on the left, and task structure shown on the right. Numbers in the boxes on the right indicate the proportion of the given trial type.

Twenty participants completed a total of 1300 trials each. Trials were presented in 100-trial blocks. For all participants, the first block was a baseline block where the color of the fixation cross was white, and the target moved on 50% of the trials. In the next 6 blocks, the fixation cross was either green or red, and vice versa for last 6. Each of these fixation colors was associated with one of two probabilities of target movement (0.9 or 0.1). The order of the two prior conditions and color-probability associations were counterbalanced across participants (task structure schematized in Fig. 13, right).

As in Experiment 1, displacements were drawn from a relatively broad Gaussian distribution ($\mu = 0°$, $\sigma = 1.5°$) on "movement" trials and from a narrow Gaussian distribution ($\mu = 0°$, $\sigma = 0.017°$) on "non-movement" trials to ensure that the solution to

the task was probabilistic. Thus, the optimal solution to the task was to learn the probability that any given displacement was drawn from the "movement" distribution relative to the "non-movement" distribution. In conditions with a biased prior (0.9 or 0.1), the optimal solution would be to weight this relative probability by the appropriate prior. In other words, the optimal solution to this task is the Bayesian solution (details in Section 2.4.2).

For 70% of the trials in blocks 2-13, the target had "low noise" (standard deviation of 0.1 deg.) and the probability of target displacement conformed to the experimental prior, i.e. 0.9 or 0.1. These 70% of the trials were considered "training trials" where the intended prior was reinforced and maintained. The other 30% of trials were "testing trials", where we tested the Bayesian hypothesis that participants would use their learned prior more when the evidence was relatively uncertain. Thus, on these trials, the target had either medium or high sensory noise. Additionally, both to isolate the effects of a learned, color-associated expectation on performance and to mitigate the possibility that our sensory manipulation affected participants' representation of the prior, the testing trials comprised a neutral condition where the target had a 0.5 probability of moving, but the fixation color cuing the prior was the same as the rest of the block. Training and testing trials were randomly interleaved. To preserve a sense of experiential continuity across the experiment, 5% of the targets in block 1 had a standard

deviation of 0.25° ("medium-noise") and the remaining 5% had a standard deviation of 0.5° ("high-noise"). Data from those trials are not analyzed.

2.4.2 Model Simulation

To test the Bayesian hypothesis, we compared participants' reports with the performance of a Bayesian ideal observer. This section provides a detailed mathematical description of the Bayesian model. An illustrated comparison of the structure and performance of the model with empirical results in Experiment 2 is provided in Section 2.4.3.

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:

$$D(\hat{x}) = \mathbb{I}\{P(J \mid \hat{x}) > P(\neg J \mid \hat{x})\} \quad (2.4)$$

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); $P(J \mid \hat{x})$ is the probability that the probe jumped given $\hat{x}$; $P(\neg J \mid \hat{x})$ is the probability that the probe did not jump given $\hat{x}$.

Using Bayes' rule for the condition within braces in Eq. 2.4,

$$D(\hat{x}) = \left\{ \frac{P(\hat{x} \mid J)P(J)}{P(\hat{x} \mid \neg J)\ (1 - P(J))} > 1 \right\} \tag{2.5}$$

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:

$$\hat{x} \sim N(x, \sigma_t^2) \tag{2.6}$$

where σ_t is the variance of the target.

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

$$D(x) = \int \mathbb{I} \left\{ \frac{P(\hat{x} \mid J)\ P(J)}{P(\hat{x} \mid \neg J)\ (P\ (\neg J))} > 1 \right\} P(\hat{x} \mid x) dx \tag{2.7}$$

That is, the decision 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 (Fig.11),

$$x|J \sim N\left(0, \sigma_J^2\right) \tag{2.8}$$

and

$$x|\neg J \sim N\left(0, \sigma_{\neg J}^2\right) \tag{2.9}$$

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

$$\hat{x}|J \sim N\left(0, \sigma_J^2 + \sigma_t^2\right) \tag{2.10}$$

and

$$\hat{x}|\neg J \sim N\left(0, \sigma_{\neg J}^2 + \sigma_t^2\right) \tag{2.11}$$

Thus, the expression inside the indicator function in Eq. 2.7, when replaced with the appropriate Gaussian probability density function, equals:

$$\frac{\frac{1}{\sqrt{(\sigma_J^2 + \sigma_t^2)2\pi}} e^{-\frac{1}{2}\left(\frac{(\hat{x}-0)^2}{\sigma_J^2+\sigma_t^2}\right)} P(J)}{\frac{1}{\sqrt{(\sigma_{\neg J}^2 + \sigma_t^2)2\pi}} e^{-\frac{1}{2}\left(\frac{(\hat{x}-0)^2}{\sigma_{\neg J}^2+\sigma_t^2}\right)} (P\,(\neg J))} > 1 \tag{2.12}$$

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

$$log\left(\frac{\frac{1}{\sqrt{(\sigma_J^2+\sigma_t^2)2\pi}}e^{-\frac{1}{2}\left(\frac{(\hat{x}-0)^2}{\sigma_J^2+\sigma_t^2}\right)}P(J)}{\frac{1}{\sqrt{(\sigma_{\neg J}^2+\sigma_t^2)2\pi}}e^{-\frac{1}{2}\left(\frac{(\hat{x}-0)^2}{\sigma_J^2+\sigma_t^2}\right)}(P\,(\neg J))}\right) > \log(1) = 0 \qquad (2.13)$$

That is,

$$log\left(\frac{1}{\sqrt{(\sigma_J^2+\sigma_t^2)2\pi}}e^{-\frac{1}{2}\left(\frac{(\hat{x}-0)^2}{\sigma_J^2+\sigma_t^2}\right)}P(J)\right) - log\left(\frac{1}{\sqrt{(\sigma_{\neg J}^2+\sigma_t^2)2\pi}}e^{-\frac{1}{2}\left(\frac{(\hat{x}-0)^2}{\sigma_J^2+\sigma_t^2}\right)}P\,(\neg J)\right) > 0 \quad (2.14)$$

Rearranging terms to solve for $\hat{x}^2$, we find that the indicator function is > 0 when $\hat{x}^2$ is greater than a criterion value defined as:

$$\hat{x}^2 > \hat{x}_C^2 = \frac{log\left(\frac{\sigma_J^2+\sigma_t^2}{\sigma_{\neg J}^2+\sigma_t^2}\right) + 2\log\frac{P(\neg J)}{P(J)}}{\frac{1}{\sigma_{\neg J}^2+\sigma_t^2} - \frac{1}{\sigma_J^2+\sigma_t^2}} \qquad (2.15)$$

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,

$$D(x) = 1 - F_{X^2}(\frac{\hat{x}_C^2}{\sigma_t^2}, df = 1, \lambda = \frac{x^2}{\sigma_t^2}) \quad (2.16)$$

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 $\frac{\hat{x}_C^2}{\sigma_t^2}$.

2.4.3 Results

In Experiment 2, we first assessed prior learning across the 3 prior training conditions with the predicted performance of the ideal observer model detailed in the previous section (Fig. 14). Briefly, the ideal observer must, on every trial, make a binary choice about whether they perceived the probe as having jumped or not. The ideal observer would decide that it jumped if the probability of jumping given a perceived displacement were greater than the probability that it did not jump. That is,

$$D(\hat{x}) = \mathbb{I}\{P(J \mid \hat{x}) > P(\neg J \mid \hat{x})\} \quad (2.17)$$

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

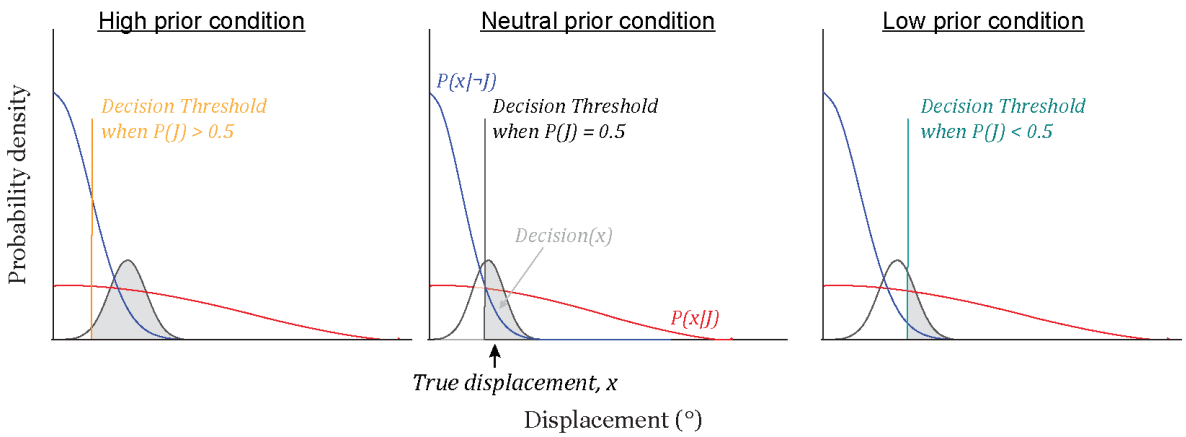

Figure 14. Schematic of the Bayesian ideal observer model.

Using Bayes' rule,

$$D(\hat{x}) = \{P(\hat{x} \mid J)P(J) > P(\hat{x} \mid \neg J)\,(1 - P(J))\} \tag{2.18}$$

If we assume that the ideal observer has fixed representations of the (experimentally defined) likelihood distributions, $P(\hat{x} \mid J)$ and $P(\hat{x} \mid \neg J)$, then for each prior condition, $P(J)$, there is a threshold at which it is equally likely that the probe jumped vs. did not jump. For $P(J) = 0.5$, that threshold jump size is where the two likelihood distributions intersect (Fig. 14 center, black vertical line). If the observer had no sensory

uncertainty, i.e., if $\hat{x} = x$, where x is the true displacement, then for all displacements less than the threshold, the ideal observer would report "did not jump", and for all displacements above the threshold, "jumped", and their decision curve as a function of displacement would be a step function at the threshold. However, we assume that both participants and the ideal observer have some sensory uncertainty. Since the visual probe was a Gaussian distribution of light on the screen, we assume Gaussian uncertainty such that:

$$\hat{x} \sim \mathrm{N}(x, \sigma t) \tag{2.19}$$

where x is the true displacement and σ_t is the standard deviation of the target (Fig. 14 center, black distribution). Therefore, the value of the decision given the true displacement, $D(x)$, can be estimated as:

$$D(x) = \int P(\hat{x} \mid x) \left\{ \frac{P(\hat{x}|J)\, P(J)}{P(\hat{x}|\neg J)\ (1 - P\ (J))} > 1 \right\} d\hat{x} \tag{2.20}$$

or the integral over all possible values of x greater than the decision threshold (Fig. 14 center, shaded region). For a higher prior, e.g., $P(J) = 0.9$, the threshold would move to the left since $P(\hat{x} \mid J)$ is weighted higher than $P(\hat{x}|\neg J)$ (Fig. 14, left) and vice versa for a lower prior, e.g., $P(J) = 0.1$ (Fig. 14, right). Critically, for the same perceived

displacement, the ideal observer is more likely to report that the probe jumped for a higher prior than for a lower prior.

Fig. 15 (left) shows the predicted decision curve of an ideal observer with likelihood distributions, $P(\hat{x}|J) \sim N(0, 2)$ and $P(\hat{x}|\neg J) \sim N(0, 0.017)$, prior P(J) = 0.22 (teal), P(J) = 0.5 (black), and P(J) = 0.78 (orange), and sensory noise, $\hat{x} \sim \mathrm{N}(x, 0.1°)$. We chose P(J) = 0.22 and P(J) = 0.78 to match the most accurate probabilities that the ideal observer would have encountered in the experiment with 70% true-statistic trials and 30% neutral. The key point, however, is that the high prior is greater than 0.5 and vice versa for the low prior. The panel on the right in Fig. 16 shows the value of the curve at displacement = 0 (the intercept) for comparison to participants' data.

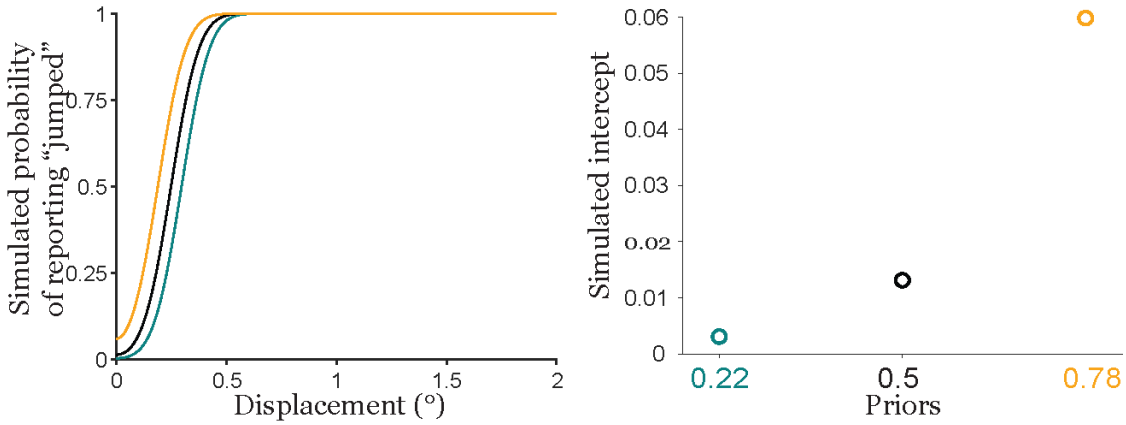

Figure 15. Bayesian predictions for prior learning.

Psychometric curves in the prior training condition from n = 17 participants are shown in Fig. 16 (left) and quantified intercepts for all participants are shown in Fig. 16

(right). As for Experiment 1, psychometric curves were fit to pooled data and the circles show responses binned individually and averaged across participants.

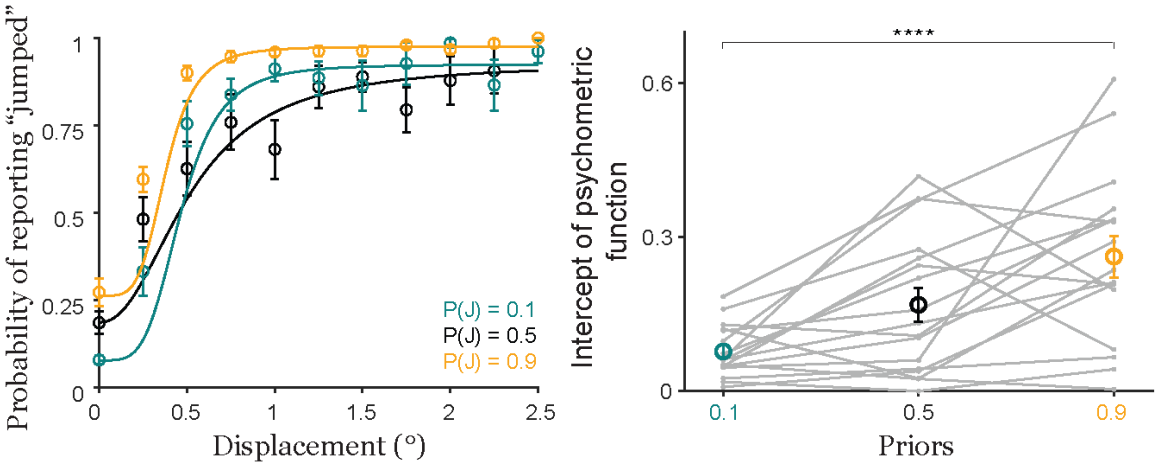

Figure 16. Human participants learned the priors as expected.

Note that although the ideal observer psychometric curves in Fig. 15 are parallel, the teal, low prior psychometric curve rises above the black, baseline curve. This pattern is discussed in Chapter 3. Nevertheless, a lower intercept in the low prior condition and higher intercept in the high prior condition match the performance of the ideal observer during prior acquisition. A repeated-measures ANOVA on the intercepts with prior as the within-conditions factor yielded a significant main effect of priors ($F(2) = 11.82$; $p = 0.0001$). We performed a post-hoc comparison of intercepts in the $P(J) = 0.9$ and $P(J) = 0.1$ conditions, the two priors tested in the medium- and high-noise conditions, using a Tukey HSD test. High-prior intercepts (mean = 0.26, SE = 0.04) were significantly higher

than both low-prior intercepts (mean = 0.08, SE = 0.01; p = 8.61 x 10^{-5}). These results indicated that participants acquired the prior as expected.

We then tested the main hypothesis of Experiment 2, that priors are used more with increasing uncertainty. Fig. 17 shows the predicted decision curves for the Bayesian ideal observer in medium- (left) and high-noise (center) conditions.

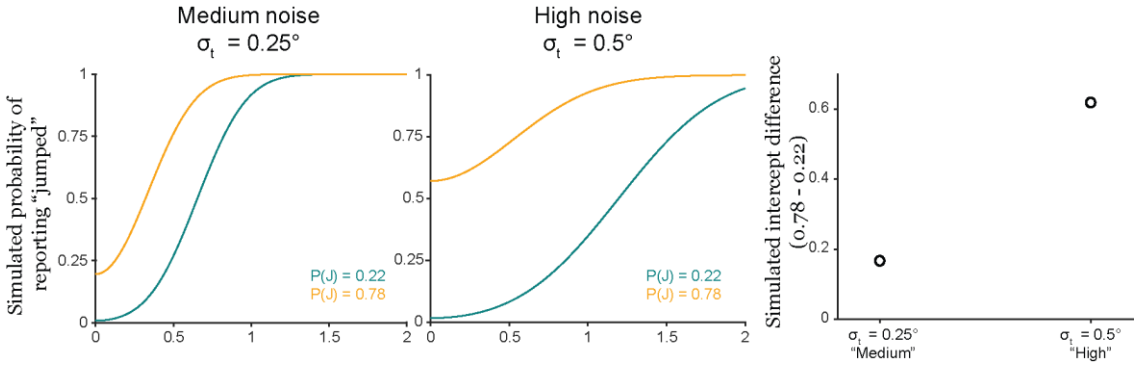

Figure 17. Bayesian predictions for medium- and high-noise hypothesis-testing trials.

The data were simulated using likelihood ratios, $P(\hat{x}|J) \sim N(0,2)$ and $P(\hat{x}|\neg J) \sim N(0,0.017)$ and priors 0.22, and 0.78 (both same as in Fig. 15) but with sensory noise, $\hat{x} \sim \mathrm{N}(x, 0.25°)$ and $\hat{x} \sim \mathrm{N}(x, 0.5°)$, respectively, to match the target widths used in Experiment 2. Qualitatively, there is a greater separation between the low- and high-prior decision curves in the higher noise condition than in the medium noise condition. Quantitatively, "prior use" or the high – low (0.78 – 0.22) prior intercept difference is higher with high sensory uncertainty than with medium uncertainty (Fig. 17, right). In

other words, the Bayesian ideal observer reverts more to the prior when the sensory input is less reliable.

Participant behavior was opposite of this modeled result.

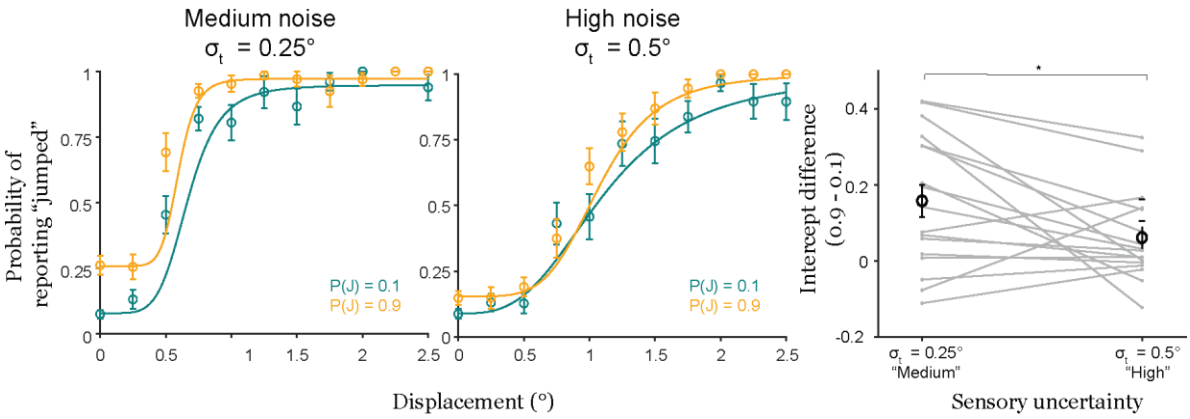

Figure 18. Prior use was Anti-Bayesian.

Fig. 19 (left) shows psychometric curves for human participants for priors 0.1 (blue) and 0.9 (orange) in the medium-noise ($\sigma_t = 0.25°$) condition. Fig. 18 (center) depicts the same results in the high-noise ($\sigma_t = 0.5°$) condition. The shifts in the psychometric curves clearly violated Bayesian predictions in that the separation between the prior curves *decreased* with greater sensory noise. The difference in intercepts was significantly lower in the high-noise condition (mean = 0.07, SE = 0.03) than in the medium-noise condition (mean = 0.16, SE = 0.04; p = 0.0326 on a paired t-test) (Fig. 18, right). Overall, the results in Experiment 2 suggested that human participants were *Anti-Bayesian.*

We considered the possibility that participants weren't Anti-Bayesian, but somehow learned that the medium- and high-noise trials had a neutral jump probability (50%), despite our efforts to keep those trials in the minority and never present them in a block. In this case, their "prior" for these conditions would be P(J) = 0.5, and the Bayesian ideal observer would use that prior, regardless of the color of the fixation cross, more with increasing noise. That is, psychometric curves in the P(J) = 0.9 and P(J) = 0.1 color conditions would get closer together with increasing noise as in Fig. 3d-f. Note that if they *only* learned the prior statistic as a function of target type (i.e., low-noise targets = prior probability, medium and high noise targets = 50% jump probability), then there would be no separation between the orange and teal psychometric curves at all. Therefore, participants learned the color-associated prior. Nevertheless, it was possible that the color-and-noise combination statistic overrode the simpler color-associated prior. To account for this possibility, we analyzed results from a control experiment in monkeys (2 rhesus macaques) where the jump probability matched the color-associated prior for all noise levels. This control condition was drawn from the neuronal recording sessions that are detailed in Chapter 4.

Two monkeys performed the same SSD task as human participants, and the prior and sensory uncertainty manipulations were the same except that the underlying statistics of probe jumps equaled the learned priors. Using monkeys for this control provided the benefit of being able to test them in many more sessions than we could

with individual human participants. A detailed description of the experiments in monkeys are provided in Chapters 3 and 4. In this control experiment, all 7 trial types (3 priors with low sensory noise, 2 each with medium and high noise) were randomly interleaved and had the same relative frequencies (Fig. 19).

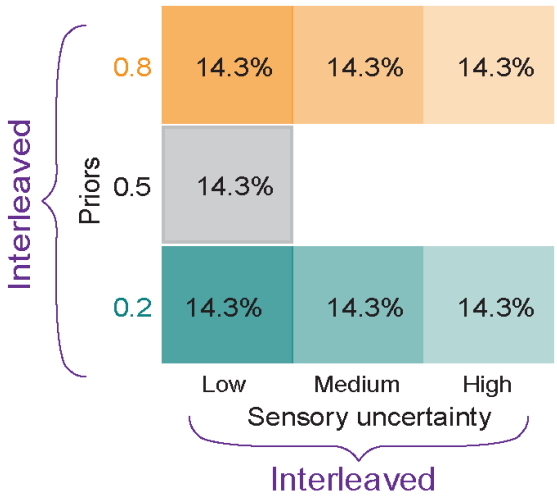

Figure 19. Task structure of the control experiment for the neutral statistic of hypothesis-testing trials.

Consistent with the results in Experiment 2, intercept differences between the P(J) = 0.8 and P(J) = 0.2 conditions *decreased* with increasing sensory noise (Fig. 20) for both animals.

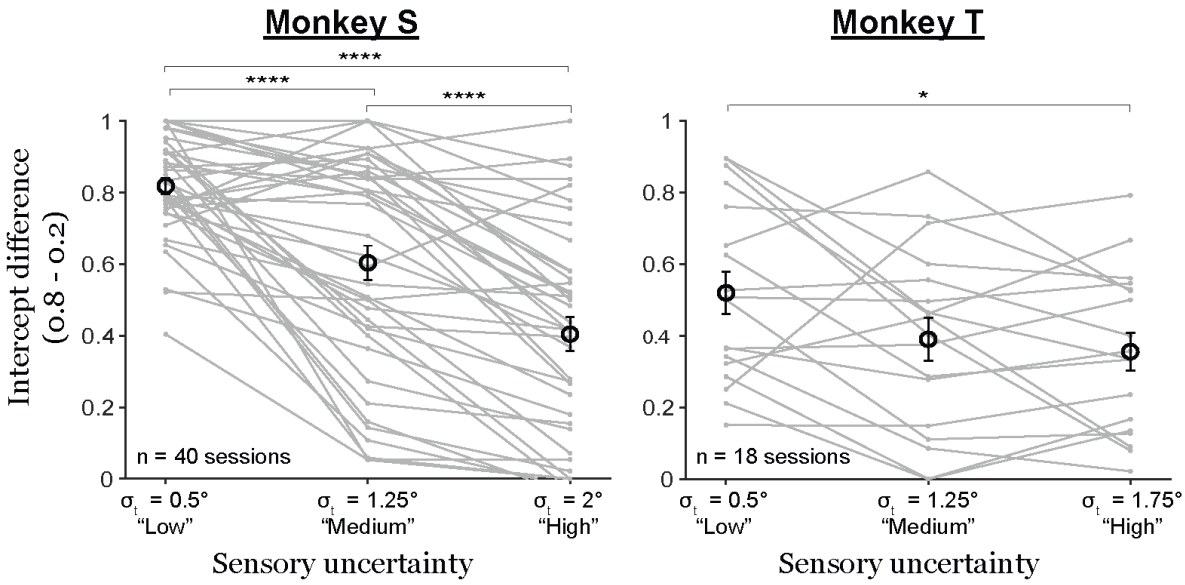

Figure 20. Prior use decreases with increasing noise even when the prior statistic is matched at all noise levels.

A repeated-measures ANOVA with intercept difference as the main within-subjects factor yielded significant effects (Monkey S: $F(2) = 51.75$, $p = 4.97 \times 10^{-15}$; Monkey T: $F(2) = 4.56$, $p = 0.0176$). For monkey S (n = 40 sessions), post-hoc comparisons using a Tukey HSD test showed that intercept differences in the low noise condition ($\sigma_t = 0.5°$; mean = 0.82, SE = 0.02) were significantly higher than in the medium noise ($\sigma_t = 1.25°$; mean = 0.60, SE = 0.05; $p = 3.49 \times 10^{-6}$) and high noise ($\sigma_t = 2°$; mean = 0.40, SE = 0.05; $p = 0$) conditions. Intercept differences in the medium noise condition were also higher than in the high noise condition ($p = 1.51 \times 10^{-5}$). For Monkey T (n = 18 sessions), there was a significant difference between the low ($\sigma_t = 0.5°$; mean = 0.52, SE = 0.06) and high noise ($\sigma_t = 1.75°$; mean = 0.36, SE = 0.05; $p = 0.0190$) conditions. Intercept differences in the

medium noise condition (σ_t = 1.25°; mean = 0.39, SE = 0.06) fell in between the low and high noise conditions but were not statistically different from either. Overall, the results replicated our findings in Experiment 2, showed that the Anti-Bayesian effect was consistent across primate species, and confirmed that the Anti-Bayesian effect persisted even when the task statistic was matched across all sensory noise levels.

2.5 Experiment 3: Testing the Bayesian hypothesis using a continuous prior

2.5.1 Experimental Methods

One possible explanation for the results of Experiment 2 (Section 2.4) is that expectation-based priors are simply not used for Bayesian computations across saccades. Alternatively, since Bayesian inference for sensorimotor stimulus estimation is typically tested in the context of *continuous* stimulus values (e.g., as schematized by the distributions in Chapter 1, Fig. 5 (bottom)), Anti-Bayesian behavior may be selective to binary categorical judgments as used in Experiment 2. To investigate this, in Experiment 3 we tested the Bayesian hypothesis for continuous estimates of displacement across saccades. We ran human participants in the same SSD paradigm, but instead of providing a binary report of "jumped" or "did not jump" at the end, participants provided a continuous report using a mouse cursor (Fig. 21, left). The target jumps were all along the horizontal meridian, and the mouse cursor was restricted to that

dimension. Formulating the task as a unidimensional, continuous problem allowed us to cast it in a form that has been tested across many sensorimotor and cognitive domains (Fig. 21, right) (e.g., Battaglia et al., 2003; Ernst and Banks, 2002; Kording and Wolpert, 2004 amongst others).

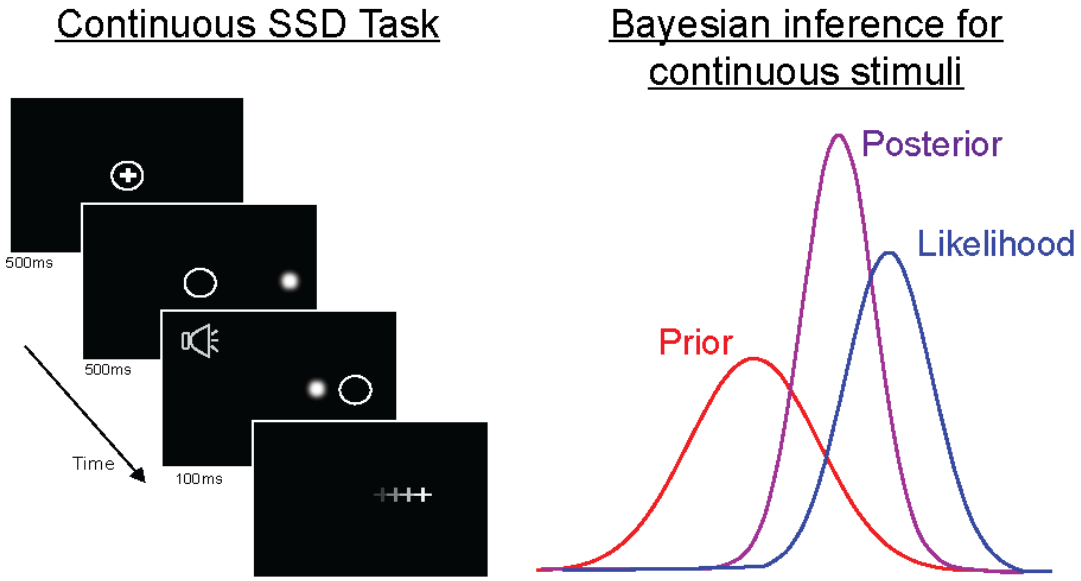

Figure 21. Continuous Bayesian Experiment. Modified SSD task (left) and overall concept for Bayesian Inference (right).

Fourteen human volunteers participated in Experiment 3. The overall paradigm was similar to Experiments 1 and 2. The critical difference was that the target was displayed post-saccadically for a limited period (100ms) and participants provided a continuous report of where they had perceived it as having landed (Fig. 21, left). Participants fixated a central cross and upon being cued, made a saccade to a target

located at either 10° or -10°. The target was displaced horizontally during the saccade, displayed in its new location for 50ms, and then replaced by a response screen. The response screen consisted of a mouse cursor that started out in the center and was restricted to the horizontal meridian to ensure that participants were solving a one-dimensional problem. Participants were required to drag the mouse cursor to the location where the target had landed and click to submit their response.

Participants completed a total of 1000 trials each. We first trained participants on the prior for 600 trials and then tested the use of this prior with increasing sensory noise. As in Experiments 1 and 2, the target was a Gaussian blob and we manipulated sensory uncertainty by varying its width. The prior was a *continuous* Gaussian distribution of displacements, rather than a categorical prior indicating the probability of object displacement. Throughout the experiment, displacements were drawn from a Gaussian distribution with mean 0° and standard deviation 1°. Participants were trained on this prior in the first 600 trials with performance-based feedback. After they submitted their response, the target appeared in its correct postsaccadic location for 500ms. To indicate their degree of correctness, the color of this feedback target ranged continuously from green (correct) to red (incorrect by more than 2°). Targets in this training phase trials had a standard deviation of 0.1° ("low noise").

In the remaining 400 trials, participants underwent a "testing" trials phase during which they were provided no feedback. These trials had one of 3 noise levels:

0.1° (“low noise”), 0.5° (“medium noise”), and 1° (“high noise”). Further, throughout the experiment, in 20% of the trials, the target did not appear post-saccadically. We call these “infinite noise” trials. All 4 noise levels were randomly interleaved throughout the testing phase. All data shown comes from the testing phase of the experiment. We used participants’ performance in the infinite noise condition to evaluate how well they learned the prior in the training phase.

2.5.2 Model simulation and fitting

As for Experiment 2, we evaluated the results for evidence of an underlying Bayesian computation by comparing participants’ responses with the predictions of a Bayesian Ideal Observer Model. A brief mathematical description of the model procedures is provided in this section. A comparison of model predictions to results are in Section 2.5.3.

We first simulated the responses of a Bayesian ideal observer in Experiment 3. The ideal observer infers the perceived displacement as a reliability-weighted combination of the sensory likelihood and prior distributions:

$$D_{perceived} = w_{prior} D_{prior} + w_{likelihood} D_{likelihood} \quad (2.21)$$

where $D_{perceived}$ is the mean of the inferred posterior distribution, w_{prior} is the weight assigned to the prior, D_{prior} is the mean of the prior distribution, $w_{likelihood}$ is the weight assigned to the likelihood, and $D_{likelihood}$ the likelihood distribution. When both D_{prior} and $D_{likelihood}$ are Gaussian distributions, the weight terms are given by:

$$w_{prior} = \frac{\sigma^2_{likelihood}}{\sigma^2_{likelihood} + \sigma^2_{prior}} \tag{2.23}$$

and

$$w_{likelihood} = \frac{\sigma^2_{prior}}{\sigma^2_{likelihood} + \sigma^2_{prior}} \tag{2.24}$$

That is, the more reliable (i.e., less variable) estimate is weighted higher. This reliability-weighted inference is additionally the Bayes optimal estimate because the variance of the estimate, $\sigma^2_{perceived}$, is lower than the variance of both the prior and the likelihood distributions:

$$\sigma^2_{perceived} = \frac{1}{\frac{1}{\sigma^2_{likelihood}} + \frac{1}{\sigma^2_{prior}}} \tag{2.25}$$

We simulated the final response as the mean of the posterior distribution, i.e., its maximum value. The values of the parameters used to simulate ideal observer responses were the same as the ones used in the experiment.

We then minimized squared error between each participant's responses and those of the Bayesian ideal observer model to identify the best-fit values for their internal prior and likelihood distributions. Best-fit parameters were identified on a participant-by-participant basis.

2.5.3 Results

If the uncertainty about the true stimulus is modeled as the sensory likelihood, then the mean of the posterior – its maximum value and thus, our approximation of the inferred response -- would be a reliability-weighted combination of the sensory likelihood and prior distributions. Specifically, the mean of the likelihood and prior distributions would be weighted by the reliabilities of the *other* distribution, that is, the prior and likelihood respectively:

$$\mu_{posterior} = \frac{\sigma^2_{likelihood}\mu_{prior} + \sigma^2_{prior}\mu_{likelihood}}{\sigma^2_{likelihood} + \sigma^2_{prior}} \quad (2.26)$$

In other words, for a given prior with fixed uncertainty, the response should get closer to the prior with greater sensory uncertainty. In this experiment, the prior was a Gaussian statistical distribution with $\mu_{prior} = 0°$ and $\sigma_{prior} = 1°$. Participants completed 1000 trials each. They were provided performance-based feedback in the first 600 trials to train them on the prior. Data from the remaining 400 hypothesis-testing trials with no

feedback provided were analyzed. There were 4 sensory uncertainty conditions: low noise (spatial Gaussian $\sigma_t = 0.1°$), medium noise ($\sigma_t = 0.5°$), high noise ($\sigma_t = 1°$), and an "infinite noise" condition in which the target did not reappear postsaccadically. Fig. 22 (left) illustrates the distributions used in the experiment, with the prior's statistical distribution centered at 0 and the likelihood (Gaussian width) distributions centered, for purpose of illustration, on a displacement of 1°.

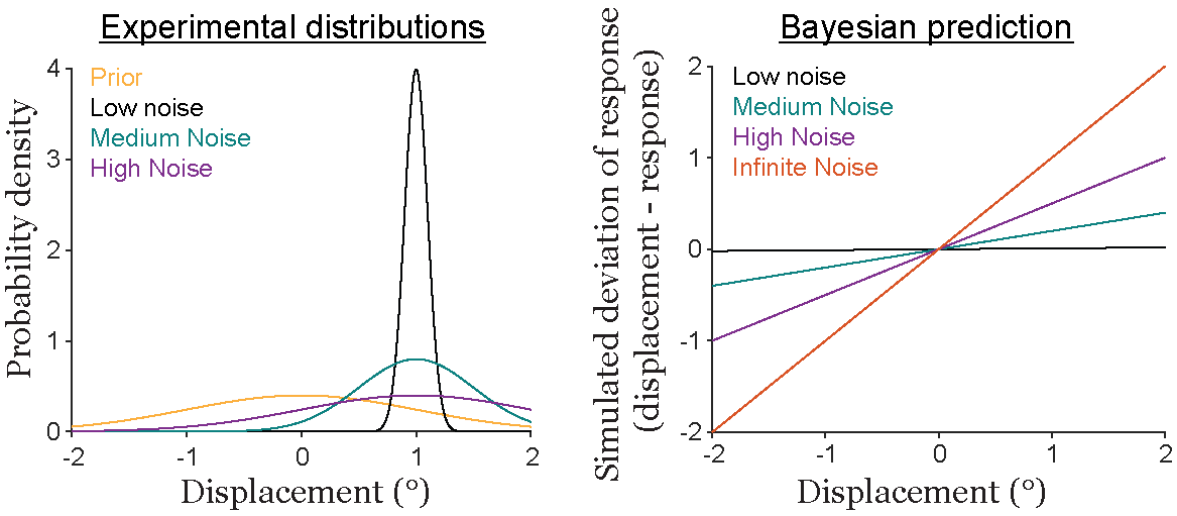

Figure 22. Distributions used in Experiment 3 (left) and Bayesian predictions for the results (right).

Fig. 22 (right) shows the predicted deviation in response from the presented displacement (displacement - response) for a Bayesian ideal observer using the true experimental distributions. If the sensory uncertainty is much smaller than that of the prior as in the lowest noise condition (black line), then the deviation of the posterior

(response) from the true displacement should be very close to 0 for all presented displacements. Conversely, for maximal sensory uncertainty as in the infinite noise condition (orange line), the response should always be the mean of the prior. Since μ_{prior} = 0°, the deviation for each displacement value equals the displacement itself. The slopes of the lines for the medium (teal) and high (purple) noise conditions are predicted to fall in between the low and infinite noise conditions. In summary, the slope of the deviation (response – displacement) line increases with increasing sensory uncertainty. We fit lines to empirical responses provided by 11 human participants for comparison with the performance of a Bayesian ideal observer. Note that if participants were not reliability-weighting their responses, i.e., if they were instead reporting the true displacement with a fixed internal sensory uncertainty, then the four lines would overlap and have the same slope. Instead, we found that the data from participants matched the predictions of the Bayesian ideal observer (Fig. 23, left), and the slopes of the deviation lines increased with increasing sensory uncertainty (Fig. 23, right).

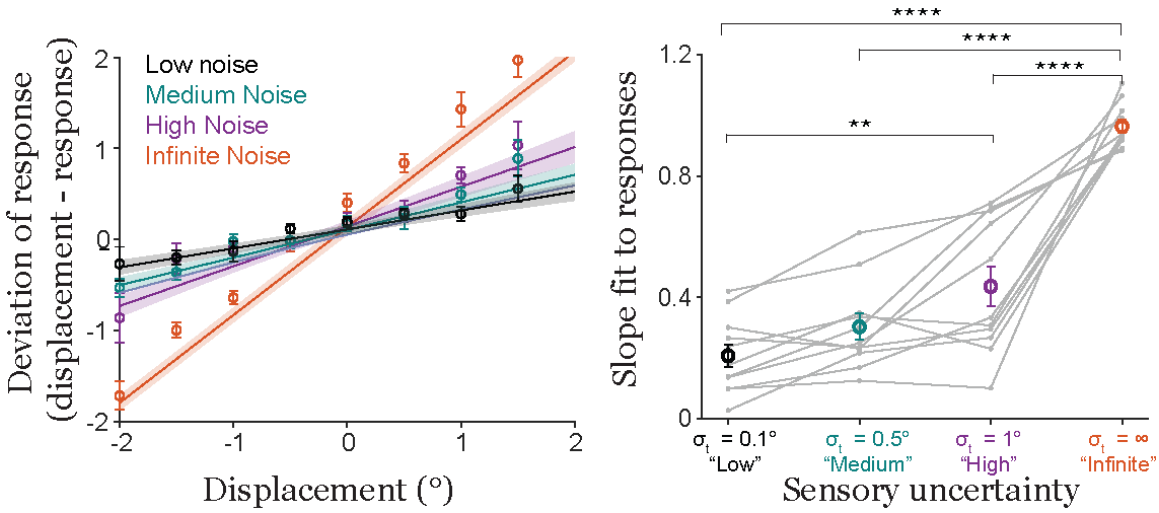

Figure 23. Results for the continuous Bayesian experiment. Lines fit to participant responses match Bayesian predictions well (left), quantified by their slopes (right).

A repeated-measures ANOVA on the slopes yielded a significant main effect of noise level ($F(3) = 78.01$, $p = 2.87 \times 10^{-11}$). A post-hoc comparison of conditions using a Tukey Honest Significant Difference (HSD) test showed that slopes in the low-noise ($\sigma_t = 0.1°$) condition (mean = 0.21, SE = 0.04) were significantly lower than in the high-noise ($\sigma_t = 1°$; mean = 0.44, SE = 0.07, $p = 0.0011$) and infinite noise conditions (mean = 0.96, SE = 0.02; $p = 6.98 \times 10^{-14}$). Also, slopes in the medium noise condition ($\sigma_t = 0.5°$, mean = 0.30, SE = 0.04) were significantly lower than in the infinite noise condition ($p = 2.05 \times 10^{-12}$), and slopes in the high-noise condition were significantly lower than in the infinite-noise condition ($p = 4.65 \times 10^{-10}$).

We fit participants' responses to a Bayesian ideal observer model by minimizing squared error to infer their used prior and sensory likelihood distributions (Fig. 24, left).

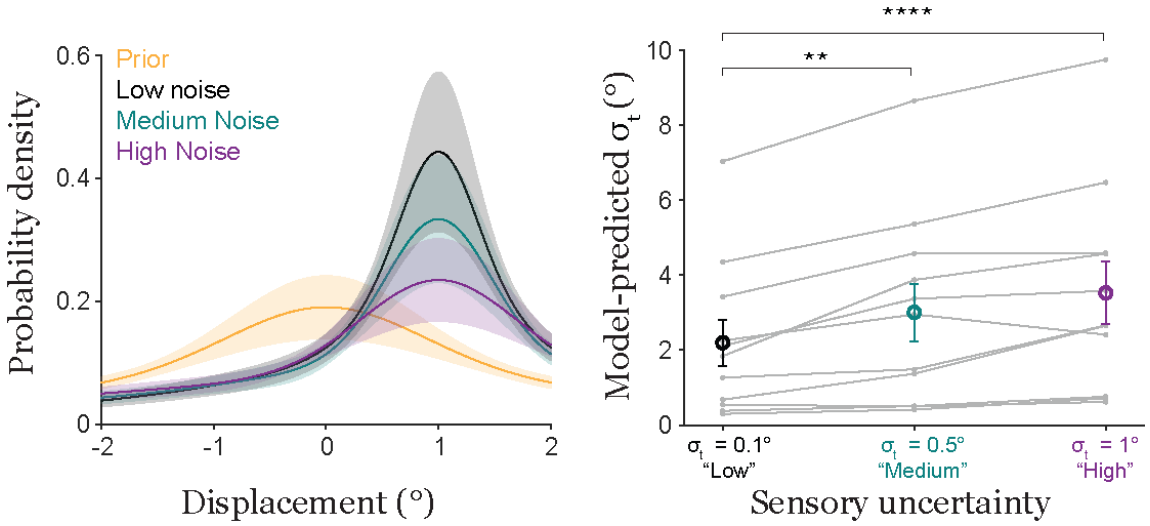

Figure 24. Parameter estimates for Bayesian model fit to the data in Experiment 3.

The prior mean and standard deviation, and standard deviations of the low-, medium- and high-noise parameters were fit simultaneously by pinning their uncertainties in the infinite noise condition to be "infinite". That is, the model assumed that the response in the infinite noise condition is the prior mean = 0. Model outputs for the prior mean (mean = 0.05, SE = 0.04) were not significantly different from 0, the true prior ($p = 0.22$ on a one-sample t-test). Fit parameters for likelihood standard deviations increased with increasing noise, suggesting that participants' internal uncertainty corresponded with the target width. Repeated-measures ANOVA on the fit standard deviations yielded a main effect of noise level ($F(2) = 16.72$, $p = 0.0003$) (Fig. 24, right). A post-hoc comparison of conditions using a Tukey HSD test showed that SD in the low

noise condition (mean = 2.12, SE = 0.62) was significantly lower than in the medium (mean = 3.01, SE = 0.77; p = 0.0062) and high noise (mean = 3.53, SE = 0.83; p = 3.71 X 10^{-5}) conditions. Finally, we assessed the correlation between model responses and participants' responses (Fig. 25) to confirm that the fit parameters explained the data on a participant-by-participant basis.

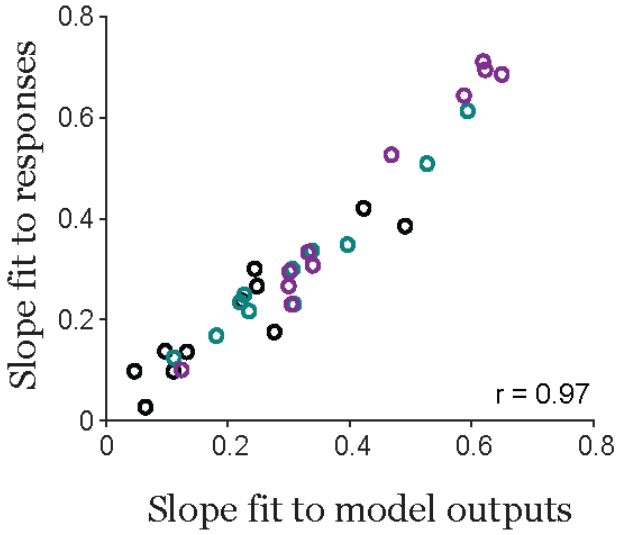

Figure 25. Correlation between model-predicted data and empirical data from human participants.

The correlation between the slopes of lines fit to *model-generated* responses with the slopes of lines fit to empirical responses (same as in 23, left) was strong and highly significant (r = 0.97, p = 1.48 x 10^{-20}), suggesting that the Bayesian ideal observer model explained the data well. Overall, the results in Experiment 3 showed that participants'

responses systematically moved closer to the prior with increasing sensory noise (Fig. 23) and that a Bayesian ideal-observer model explained the results well (Fig. 24-25).

2.6 Discussion

Overall, we found that continuous reports of the object displacement across saccades are Bayesian but categorical reports in the same task are, surprisingly, Anti-Bayesian. The latter, Anti-Bayesian result was independent of the underlying statistics of object movement, as shown with the monkey control data for that experiment. One potential interpretation is that the Anti-Bayesian behavior observed in Experiment 2 is a consequence of the specific conceptualizations of the prior and sensory uncertainty in our Bayesian ideal observer model. While this is possible, we chose the parameters of the model (e.g., the jump and non-jump likelihoods, and the sensory uncertainty distribution) to directly match those used in our experiment. Accounting for Anti-Bayesian behavior using the current model would involve either assuming a mischievous prior that moved in the opposite direction as the task statistic it was associated with, or sensory information that somehow became *more* certain with a greater target width (going against the initial confirmation of its effects in Experiment 1). Neither of these possibilities seem likely. We suggest that a more rational alternative is that the system uses a Bayesian decoder for continuous information but uses a

Discriminative decoder for categorical information in this task. This potential explanation is investigated in Chapter 3.

What gives rise to this puzzling dichotomy between Bayesian use in the continuous task but not the categorical task? One simplistic possibility is that continuous perception of stimulus quantities is Bayesian, but categorical perception is not. Indeed, stimulus estimates as reliability-weighted combinations of prior and sensory likelihoods (as schematized in Chapter 1, Figure 5 and Chapter 2: Figure 11) typically refer to stimuli along a single continuous dimension (e.g., Battaglia et al., 2003; Ernst and Banks, 2002; Kording and Wolpert, 2004; Jazayeri and Shadlen, 2010; Girschick et al., 2011; Fetsch et al., 2012; Darlington et al., 2017). This interpretation, however, does not account for the demonstration of Bayesian behavior in other categorical tasks, notably, the random dot kinematogram (RDK) task. In the RDK task, subjects are presented with a random dot pattern. A subset of the dots moves in the same direction while the remaining dots move randomly. The proportion of dots that is displaced in the same direction determines the coherence of the stimulus. In the categorical, two-alternative forced choice version of this task, subjects are required to report whether the dots moved towards the right or the left, for example. Behavior in this task is well-explained by models known as the Drift Diffusion Model or the Sequential Probability Ratio Test (Wald, 1945; Ratcliff, 1978; Roitman and Shadlen, 2002), both of which are formulations of an optimal, Bayesian solution to the task (Bitzer et al., 2014).

We considered an alternative, more biologically grounded possibility. Object location information in the visual system is coded via the organization of spatial receptive fields. Spatial receptive fields are continuous all the way from the retina to higher order visual areas (Colby et al., 1988; Engel et al., 1997; Golomb and Kanwisher, 2012a; Arcaro and Livingstone, 2017) and maintain their retinotopic properties even across eye movements (Golomb and Kanwisher, 2012b; Zimmerman et al., 2013). Moreover, our previous work found that neurons in the frontal eye field use continuous tuning to represent object displacements across saccades (Crapse and Sommer, 2012), the stimulus quantity we probed directly. Even on the motor side, presaccadic remapping of visual receptive fields in the direction of the saccade vector (reviewed in Hall and Colby, 2011; Neupane et al, 2020; Golomb and Mazer, 2021) is in continuous coordinates. Thus, it is possible that the visual system defaults to processing information across saccades in continuous coordinates. For this inherently prioritized computation, it uses a Bayes optimal strategy. Requiring the system to categorize this continuous information might lead it to adopt an alternative strategy that is revealed when probing it with image noise. This explanation, however, would mean two things.

First, the continuous = Bayesian, categorical = Anti-Bayesian dichotomy was mostly being driven by the organization of the *visual* system rather than motor-sensory interactions. Second, the system should be Bayesian even for *categorical* computations that are inherent to its functioning. Each saccade has an inherent categorical sensory

consequence in that visual processing is suppressed (in the form of Saccadic Suppression as discussed in Chapter 1) when a saccade is made, and not suppressed when there is no saccade. If the explanation that inherently prioritized computations are Bayesian is true, then Anti-Bayesian behavior in the categorical task must have been driven largely by *visual* noise, and that the trade-off between priors and motor-driven uncertainty will be Bayesian.

First, we performed a post-hoc analysis on the data in the current experiments to evaluate the prediction. Recall that the perception of displacement across saccades involves comparing the external sensory input with a motor-driven sensory prediction. In both the continuous and categorical experiments, the Gaussian "blob" stimulus was both the saccade target and the visual probe. Thus, in addition to increasing visual uncertainty, blurring the target might also have increased motor-driven uncertainty. To assess this, we compared the standard deviation of the scatter of saccade endpoints across noise levels. If blurring the target were increasing motor uncertainty in addition to visual uncertainty, then endpoint scatter should be greater at higher noise levels. We found no significant difference in endpoint SD across noise levels either parallel or perpendicular to the saccade in either the categorical (Fig. 26) or continuous experiment (Fig. 27).

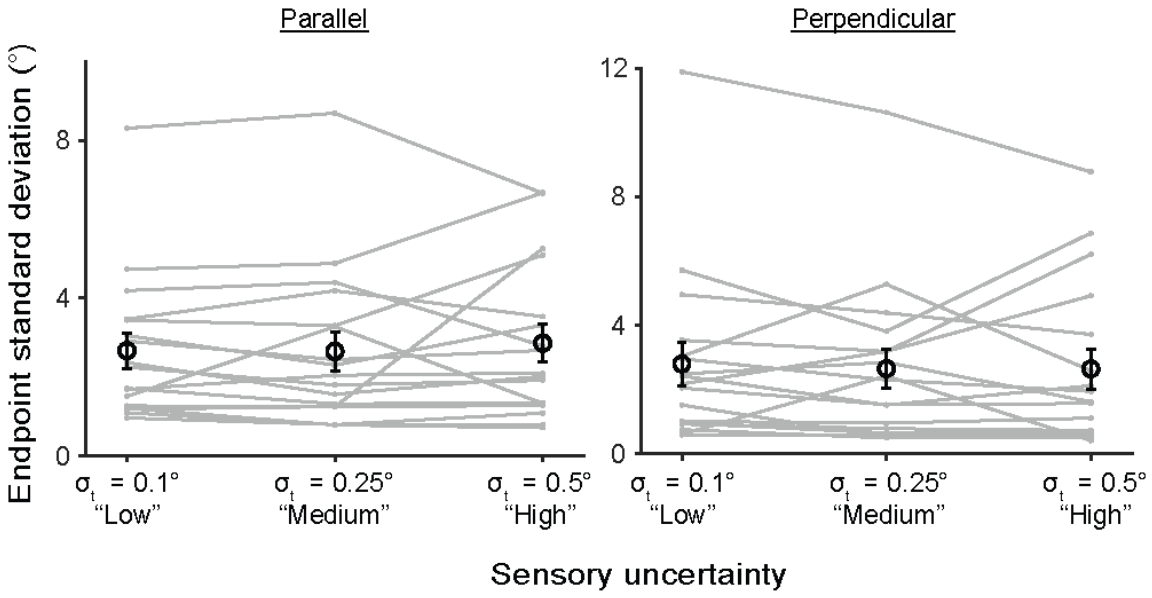

Figure 26. Saccadic endpoint scatter in the categorical experiment.

That is, repeated-measures ANOVAs did not reveal a main effect of noise levels in either experiment: $F(2) = 0.32$, $p = 0.7306$ in the categorical experiment and $F(3) = 1.55$, $p = 0.2230$ in the continuous experiment for endpoints parallel to the saccade; $F(2) = 0.14$, $p = 0.8693$ in the categorical experiment and $F(3) = 2.6$, $p = 0.1129$ in the continuous experiment for endpoints perpendicular to the saccade).

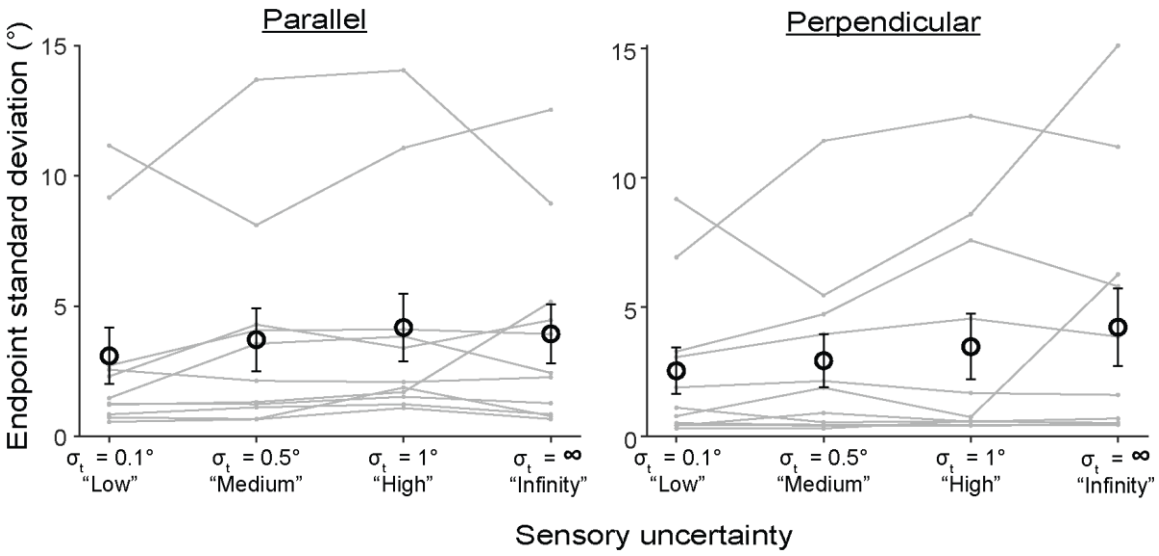

Figure 27. Saccadic endpoint scatter in the continuous experiment.

This suggests that blurring the target primarily increased external sensory uncertainty, thus corroborating our hypothesis that Anti-Bayesian behavior for categorical judgments was primarily driven by external uncertainty in the visual input. In Chapter 3, we test this hypothesis directly.

Overall, the current results are a first step towards establishing that under some conditions, active visual perception across saccades is Bayesian while in other conditions it is not. One main direction of future investigations, as suggested above, is to evaluate the relative contributions of visual and motor-driven uncertainty to the use of priors. Another major open question is how Anti-Bayesian categorical behavior might be implemented. An alternative framework to Generative, Bayesian models is that the system uses a Discriminative model, where it learns to draw a boundary between

categories of interest (Rumelhart et al., 1986; Hinton, 1992; Murphy, 2013; Ng and Jordan, 2002). Qualitatively, the current results suggest that the system learns the boundary between "jumped" and "did not jump" differently for each prior context (such that each colored psychometric curve is one such boundary), and that blurring the target blurs those boundaries towards each other. In Chapter 3, we identify the quantitative structure of such a Discriminative model.

3. Relative contributions of Bayesian and Discriminative models to Active Categorical Perception

3.1 Introduction

Distinguishing self-generated sensory inputs from externally generated sensory signals is a major aspect of active perception. In the saccadic system, distinguishing the saccade-induced retinal displacement of an object from object movement in the world involves comparing the predicted consequence of a saccade with the incoming sensory input. Rao et al. (2016) previously found that humans use priors about the probability of an object moving for this task. We found in Chapter 2, however, that those priors are not used in a Bayesian manner. That is, they are not used more with greater sensory uncertainty. Instead, they are used less. The Anti-Bayesian prior use suggested the use of a Discriminative model in the categorical task. The overall goal of this chapter is to understand the relative contributions of Bayesian and Discriminative models to the categorical perception of object displacement across saccades.

Based on an analysis of saccadic endpoint scatter in humans in Chapter 2, we hypothesized that Anti-Bayesian prior use might largely be driven by visual rather than motor noise. In Experiments 1 and 2 in this chapter, we directly tested this hypothesis in rhesus macaques by separating the saccade target from the visual probe. We used macaques instead of humans for these studies because their implanted scleral search coils provided superior spatial and temporal eye position precision. This is particularly

important when considering the impact of motor variability in a task. The animals made a saccade to a target square and reported whether a visual probe (Gaussian blob) in the periphery moved during the saccade. Additionally, we identified a biologically plausible neural network model that can implement a Discriminative category boundary in this task and account for the Anti-Bayesian behavior.

In Experiment 1, we selectively manipulated external sensory uncertainty by varying the width of the Gaussian blob with low, medium, or high noise (Fig. 28, left). The structure of Experiment 1 was nearly identical to the categorical experiment in humans: there were 3 noise levels (low, medium, and high) (Fig. 28, right).

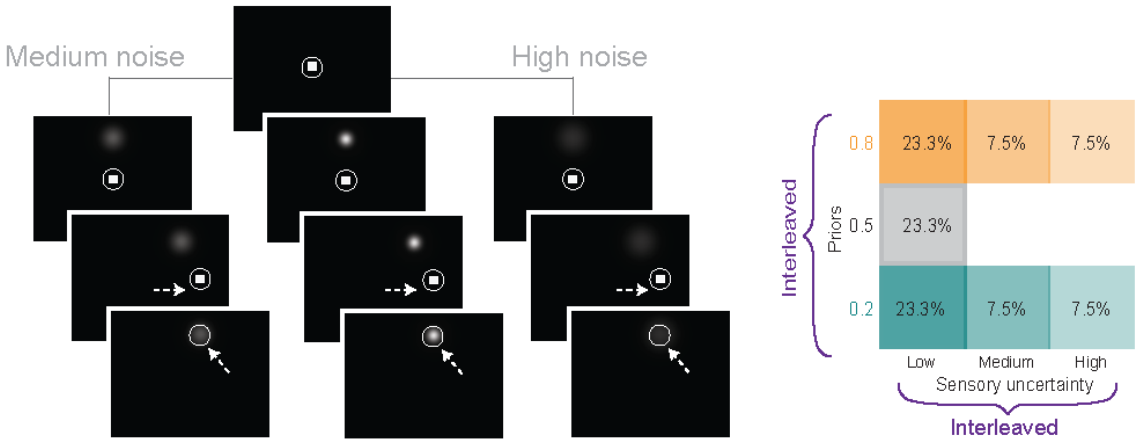

Figure 28. Experiment testing the Bayesian hypothesis selectively for external, visual uncertainty. Left: The low sensory noise condition used for prior training is shown in the middle, medium noise on the left, and high noise on the right. Right: Task structure.

Low-noise trials comprised 70% of all trials and were "prior training" trials, with 30% medium- and high-noise hypothesis-testing trials. All trial types were randomly interleaved. In contrast to the control task using monkeys that we presented in the section on Experiment 2, in which the statistics of the target displacements matched the priors, the data presented here were collected separately under conditions – just as in humans – in which the statistics of target jumps were neutral in the prior testing trials.

In Experiment 2, we introduced two levels of motor-driven uncertainty (Fig. 29, left). For this we used the two extremes: making a saccade or not. In the "high uncertainty" condition, animals made a saccade to a target and reported whether a probe moved or not. In the "low uncertainty" condition, they withheld the saccade. That is, they remained fixated on a central square throughout the trial. The temporal structure of the trials was otherwise identical. With- and no-saccade trials at 3 prior levels each were randomly interleaved (Fig. 29, right).

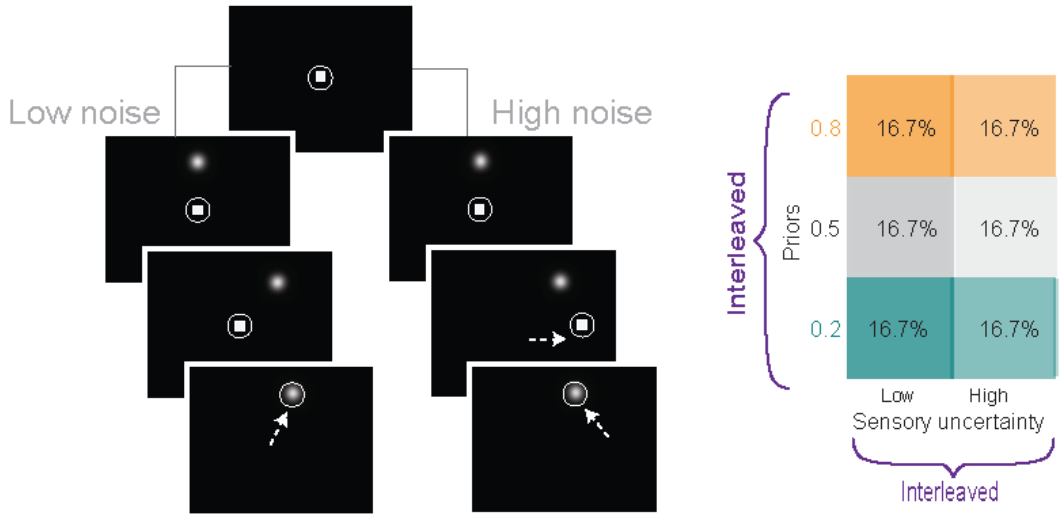

Figure 29. Experiment testing the Bayesian hypothesis for motor-driven uncertainty. Left: The low motor noise condition with no intervening saccade is shown on the left, and the high noise condition with a saccade on the right. Right: Task structure.

3.2 General Methods: Non-Human Primate Psychophysics

3.2.1 Materials and Paradigm

Two rhesus macaques (Monkey S and Monkey T, both males) were run in a modified Saccadic Suppression of Displacement paradigm, similar to the human participants. For each daily session, animals were brought into the lab in custom-made chairs (Crist Instruments, Hagerstown, PA) and their heads were stabilized using a head-post that attached to both the chair and a surgically implanted socket (Crist Instruments, Hagerstown, PA) on the skull. The socket was implanted in an aseptic

surgical procedure with the structural support of ceramic bone screws and dental acrylic. Eye position was measured using a surgically implanted scleral search coil (Robinson, 1963; Judge et al, 1980) in one eye, with its leads soldered to a connector plug embedded in the implant. All surgical and experimental procedures were performed in accordance with protocols approved by the Duke Institutional Animal Care and Use Committee.

In a typical experimental session, the animals performed the behavioral task in a dark experimental rig. They were positioned 60cm from an LCD monitor (1920 x 1080, 144Hz). To dissociate external sources of sensory noise from internal, motor-driven sources, the saccade target was dissociated from a visual probe (a Gaussian blob) which was displaced intrasaccadically on some trials. In the human experiments, the saccade target (also the visual probe) always appeared in one of two locations on the screen and only moved horizontally. For Experiments 4 and 5 in monkeys, it could appear in one of 4 locations, ±10° horizontally or ±10° vertically. The saccade target was always positioned along the orthogonal cardinal direction (e.g., if the probe appeared at ±10° horizontally, the saccade target would be at ±10° vertically), and the probe was displaced in a direction parallel to the saccade vector. For the control experiment (with results presented in Chapter 2, Figure 20), we simultaneously recorded from neurons while the animals performed the sessions (neural data presented in Chapter 4). Since we

placed the probe within the mapped receptive field of the neuron, the probe appeared in a different location during each session.

On each trial, a fixation square (1° x 1°) first appeared at the center of the screen. After fixation had been acquired and maintained for a variable duration of 300-500ms, the visual probe appeared at one of the 4 locations on the screen for a variable delay of 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 (1° x 1°) indicating to the animal that it was allowed to make a saccade. Saccade initiation (defined as the time the eye crossed a threshold set at 20% of the saccade length, i.e., 2°, in the direction of the saccade) triggered the displacement of the probe on some trials. The probe was displaced in a direction parallel to the saccade. Animals were further required to maintain postsaccadic fixation for 700ms 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 (occluded from the monkey's view) was flashed within the same frame as the measured stimulus.

Displacements were drawn from relatively broad and narrow Gaussian distributions in the "movement" and "non-movement" conditions, respectively. Both distributions had a mean of 0°. The "movement" distribution had a standard deviation of 2.5° and the "non-movement" distribution had a standard deviation of 0.2°. Positive displacements were either rightward or upward, and negative displacements were leftward or downward. Each prior, or the probability of target movement, was cued by the color of the fixation and target squares. For monkey S, a green square meant that the probe had a 0.2 probability of being displaced while a magenta square indicated a 0.8 probability of displacement. To avoid misinterpreting color-related visual responses as prior-related responses in the neural recording experiments of Chapter 4, we changed the color-prior associations for monkey T: a blue or orange square cued the 0.2 or 0.8 probability of displacement, respectively. Animals were trained on priors using performance-based feedback like human participants, except with liquid reward providing the feedback instead of smiling or frowning faces.

3.2.2 Data Preparation and Analysis Measures

Psychometric curves were fit to binary responses using a 4-parameter logistic regression model:

$$y = max + \frac{min - max}{1 + \left(\frac{x}{thresh}\right)^{slope}} \tag{3.1}$$

where x is the absolute value of the presented displacement, y is the value of the psychometric function, min is the minimum value of the function (i.e., y at $x = 0$), max is the maximum value, $thresh$ is the inflection point, and $slope$ is the slope of the psychometric function. The min, max, $thresh$, and $slope$ terms were fit to binary data by minimizing mean squared error. In the low-noise, prior training trials for macaques, displacements were drawn from continuous distributions. In these conditions, we used the value of min as the intercept. For the medium and high noise hypothesis testing trials in macaques, displacements were discretized, including a displacement = 0 condition. In these conditions, the intercept is simply the proportion of "moved" responses in the displacement = 0 condition. All analyses were repeated using Criterion from Signal Detection Theory as a measure of prior use and the results were the same. Since animals performed the exact same task with the same parameters every day, we pooled trials across days to maximize statistical power and include sessions in which the animal did not perform a large number of trials. For statistical estimates of confidence in the measures and differences across conditions, 95% confidence intervals were calculated using a bootstrap procedure.

3.3 Experiment 1: Trade-off between categorical priors and external, image noise

3.3.1 Methods

To measure performance as a function of external sensory uncertainty, the visual probe in Experiment 1 was a Gaussian "blob" with one of three possible standard deviations: 0.5° ("low noise"), 1.25° ("medium noise"), and 2° ("high noise") for Monkey S and 0.5° ("low noise"), 1.25° ("medium noise"), and 1.75° ("high noise"), for Monkey T (Fig. 28, left). The relative frequencies of all trial 7 types (2 priors X 3 noise levels + baseline), shown in Figure 28, right, were the same as in the categorical experiment for humans (Experiment 2 of Chapter 2). Baseline trials with white squares and 0.5 probability of displacement all had "no noise" visual probes. In the 0.2 and 0.8 prior conditions, 70% of trials had no noise and conformed to the displacement probability indicated by the prior. The remaining 30% of trials with low and high noise comprised a neutral "test" condition with a veridical jump probability of 0.5. All 7 trial types were pseudorandomly interleaved.

3.3.2 Results

Results from Experiment 1, the image noise manipulation, replicated the results from humans in Chapter 2, Experiment 2: the behavior was Anti-Bayesian. We first confirmed that the noise manipulation was selective to external, visual uncertainty in that there was no difference in the saccade endpoint scatter across noise levels (Fig. 30)

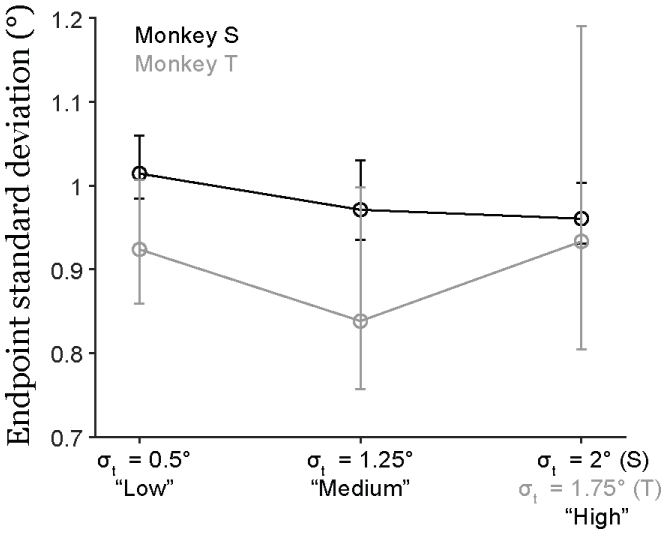

Figure 30. Saccade endpoint scatter in the direction of the saccade for rhesus macaques.

For monkey S (n = 10130 total trials from 9 sessions), the standard deviation of saccade endpoints [with 95% bootstrap confidence intervals] across noise levels was 1.01 [0.98, 1.06], 0.97 [0.94, 1.03], and 0.96 [0.93, 1.00], and monkey T (n = 9958 total trials from 18 sessions), it was 0.92 [0.86, 1.07], 0.84 [0.76, 1.00], and 0.93 [0.80, 1.19]. Note that 95% bootstrap confidence intervals report the confidence of the measure directly and can be interpreted akin to the standard error (SE). Thus, quantities and their confidence intervals can be directly compared to evaluate their overlap or separation without requiring a statistical hypothesis test.

Both animals learned the priors presented (P(J) = 0.2, 0.5, or 0.8) in the prior-training trials, leading to an upward shift in psychometric functions for the high (P(J) =

0.8) prior and a downward shift for the low (P(J) = 0.2) prior (Fig. 31, left for Monkey S and Fig. 31, right for Monkey T).

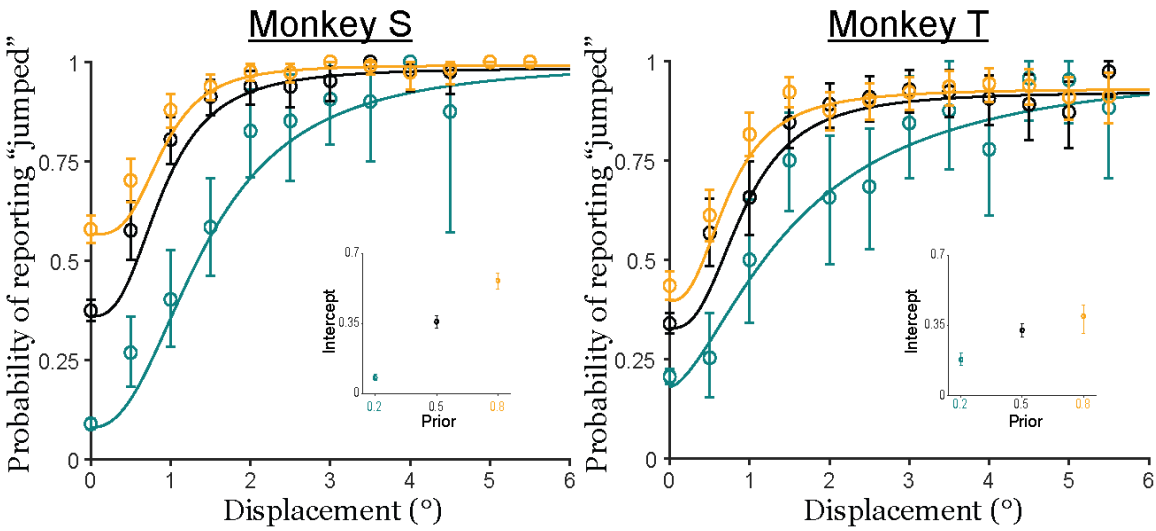

Figure 31. Prior learning in the visual noise experiment.

Note that prior learning in this experiment can be compared to the Bayesian prediction for prior use in Chapter 2, Figure 15. Intercepts for the curves [95% bootstrap confidence intervals] increased with increasing priors: 0.08 [0.07, 0.10], 0.36 [0.33, 0.39] and 0.57 [0.53, 0.60] for Monkey S (Fig. 31, left, inset) and 0.18 [0.15, 0.21], 0.33 [0.30, 0.36] and 0.40 [0.31, 0.45] for Monkey T (Fig. 31, right, inset).

Also consistent with human participants, prior use *decreased* with increasing noise (Fig. 32). Psychometric functions for the 0.2 and 0.8 prior conditions got closer to each other with increasing noise for both Monkey S (Fig. 32, top) and Monkey T (Fig. 32,

bottom), in contrast to the greater separation with noise predicted by a Bayesian model (Chapter 2, Fig. 18).

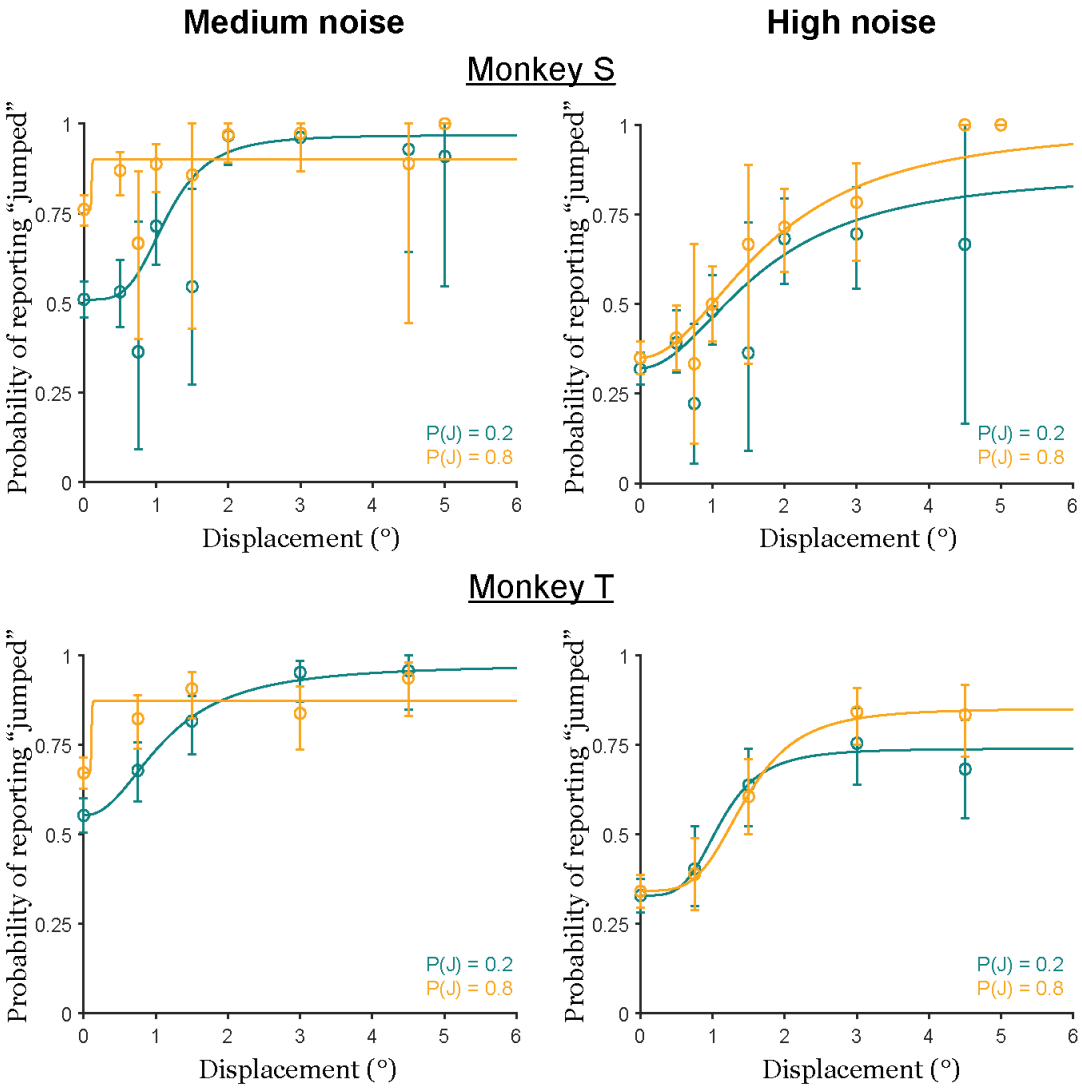

Figure 32. Prior use decreases with increasing external, sensory noise in an Anti-Bayesian manner.

This collapsing of the curves was quantified by the intercept differences between the high- and low-prior conditions (Fig. 33). Monkey S had an intercept difference [95% confidence interval] of 0.25 [0.19, 0.32] in the medium noise condition and 0.03 [-0.04, 0.09] in the high noise condition. For Monkey T, it was 0.12 [0.05, 0.18] for medium noise and 0.01 [-0.05, 0.08] for high noise.

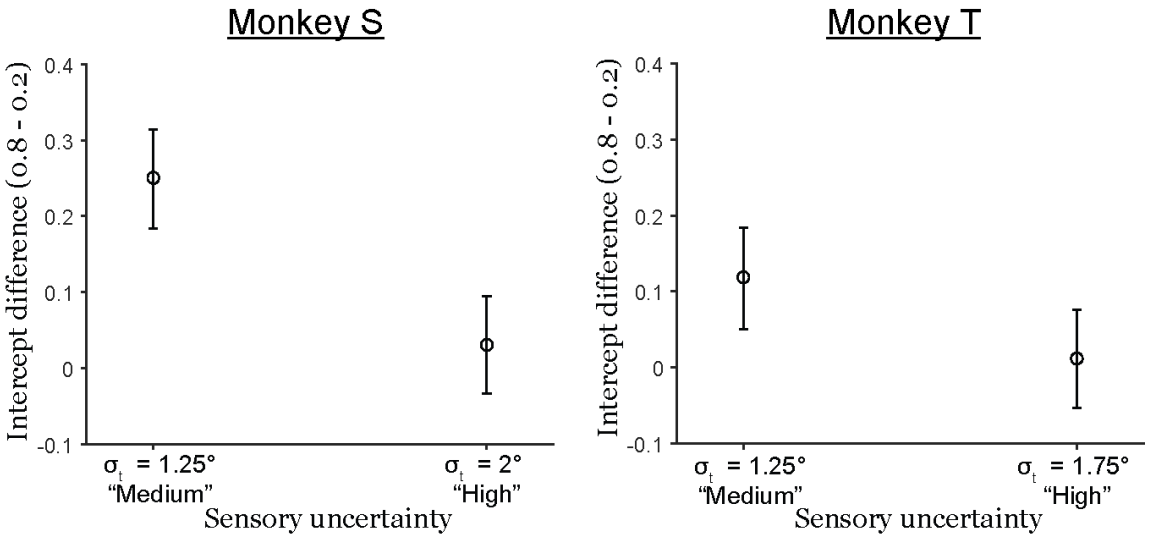

Figure 33. Prior use quantified by intercept differences decreases with increasing noise.

Overall, the results showed that both monkeys used their priors less with increasing external, sensory uncertainty. In addition, the control experiment presented in Chapter 1 also selectively varied the width of the blob but not the saccade target, thus showing that prior use with increasing external, visual uncertainty was Anti-Bayesian regardless of the veridical statistics of jump probabilities in the task structure.

3.4 Experiment 2: Trade-off between categorical priors and internal, motor noise

3.4.1 Methods

To measure performance as a function of internal, motor-driven sensory uncertainty, we added a condition to the experiment where monkeys did not make a saccade. They 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 temporal structure of the no-saccade trials was identical to the trials with a saccade. No-saccade trials were implemented by assigning the location of the "saccade target" to be the same as the fixation square. There were three prior conditions: 0.2, 0.5, and 0.8. Colors indicating the priors were the same as in Experiment 1. The visual probe had a standard deviation of 0.5°, the lowest noise condition, for all trials. All 6 trial types (3 priors x 2 noise levels) were randomly interleaved.

3.4.2 Results

The results from Experiment 2 in which motor noise was manipulated (saccade vs. no saccade) matched the predictions of a Bayesian, ideal observer model well. Saccadic suppression, where early visual processing and sensitivity to displacements are

suppressed around the time of saccades, is a main form of sensory uncertainty imposed by a saccade. Our main prediction, therefore, was that larger displacements would be perceived as "non-jumps" in the with-saccade condition than in the no-saccade condition. We simulated saccadic suppression in the Bayesian ideal observer model by increasing the standard deviation of the non-jump likelihood distribution in the with-saccade condition (σ_{NJ} = 1°) relative to the no-saccade condition (σ_{NJ} = 0.25°) while holding the standard deviation of the jump likelihood distribution constant (σ_J = 5°).

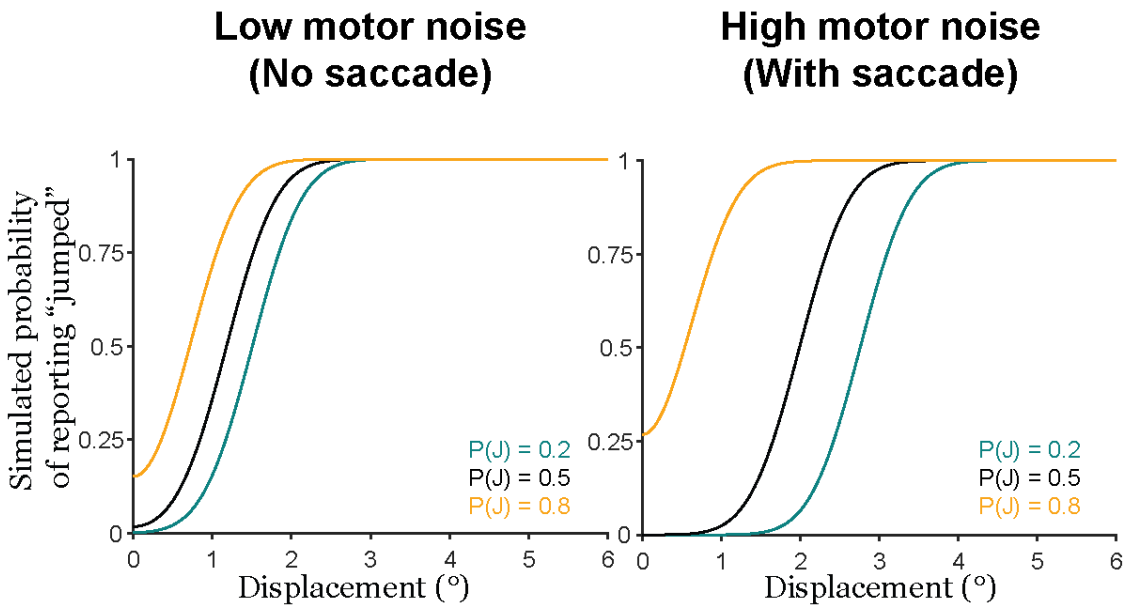

Figure 34. Predictions of the Bayesian ideal observer model for increasing motor noise.

The Bayesian model predicts that psychometric functions in the 0.2 and 0.8 prior conditions would separate further (Fig. 34) and that the difference in intercepts between

them would be greater (Fig. 35) with greater sensory noise in the with-saccade condition than in the no-saccade condition.

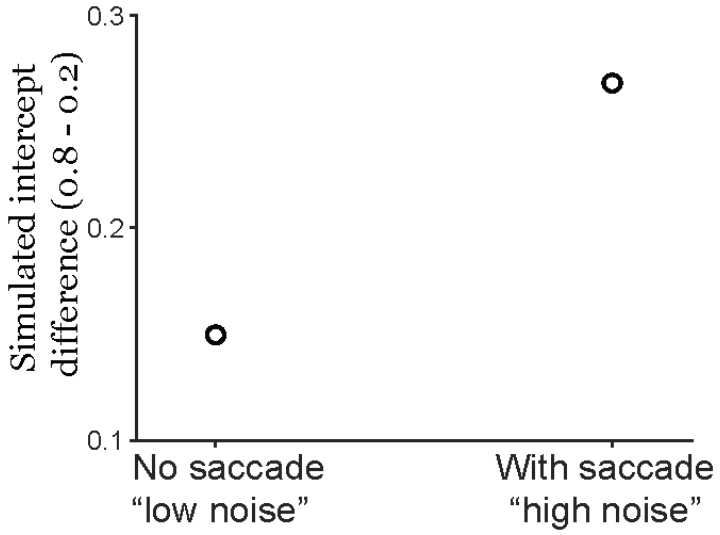

Figure 35. The Bayesian model predicts that prior use, quantified by intercept differences, increases with increasing motor noise.

Results from both monkeys matched model predictions well. Psychometric functions for the different priors showed greater separation when animals made a saccade (Monkey S: Fig. 36, top right, Monkey T: Fig. 36, bottom right; n = 6000 analyzed trials for both animals) than in the condition without a saccade (Monkey S: Fig. 36, top left, Monkey T: Fig. 36, bottom left).

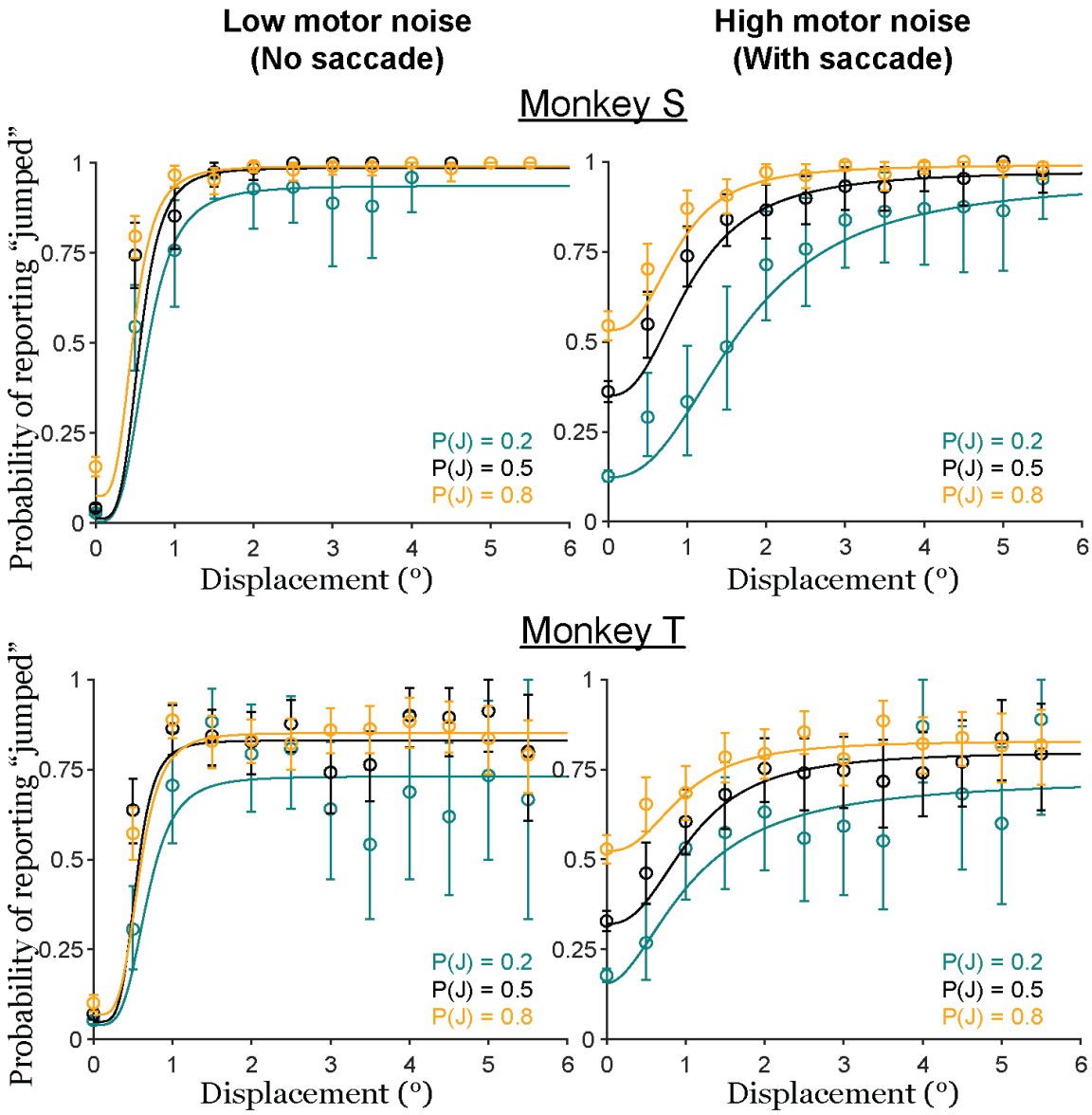

Figure 36. Empirical psychometric curves in the motor noise experiment.

The intercept difference between priors was 0.07 [0.04, 0.10] in the no-saccade condition and 0.41 [0.36, 0.46] when a saccade was made for Monkey S (Fig. 37, left), and 0.03 [-0.0003, 0.06] and 0.37 [0.30, 0.42] respectively for Monkey T (Fig. 37, right).

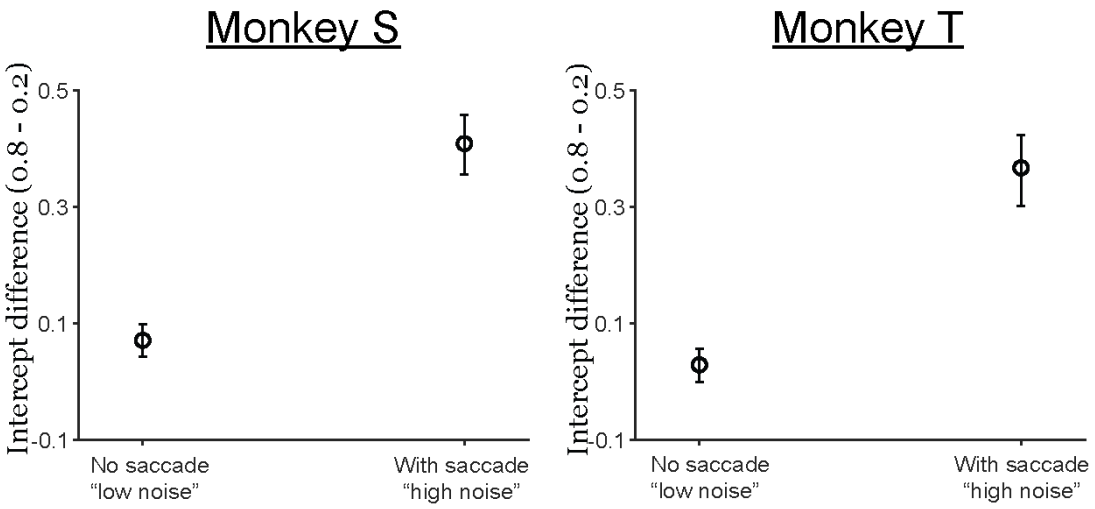

Figure 37. Empirical prior use, quantified by the intercept difference, increases with increasing motor noise.

3.5 Two-layer Perceptron: A candidate Discriminative model

3.5.1 Rationale

The results so far show that *continuous* estimates of object displacement across saccades are made in a Bayesian manner (Chapter 2), but a Bayesian model is not sufficient to account for *categorical* reports. Instead, categorical reports appear to be the result of two models as implied by the data in Experiments 1 and 2, above. Prior use for internal, motor-driven noise matched the predictions of a Bayesian ideal observer model, but prior use for external, image noise was Anti-Bayesian. These results still leave open the question of how Anti-Bayesian behavior might be generated.

We set out to identify an alternative model that explains those aspects of the results that violate Bayesian predictions. There are two such aspects. First, of course, was the decreasing use of priors with increasing sensory noise (recapped in Fig. 38 for human participants).

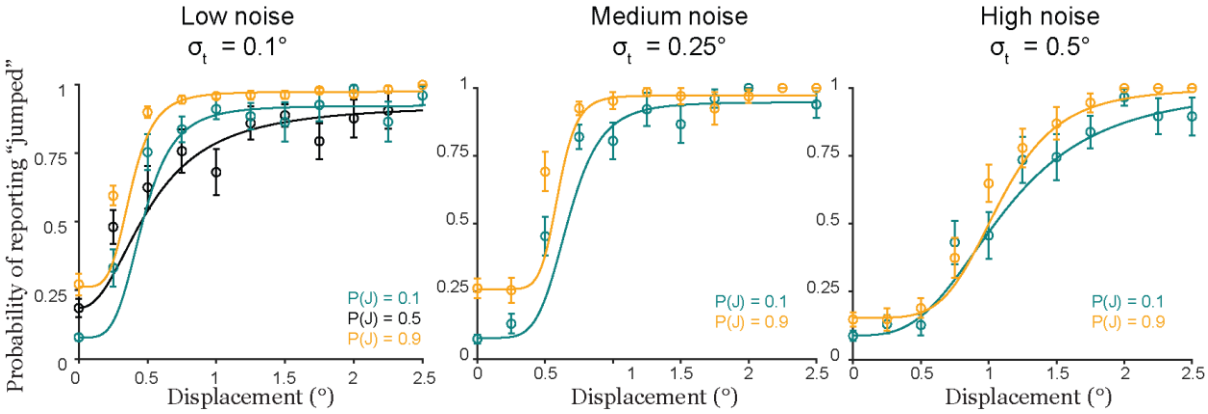

Figure 38. Prior use in humans decreases with increasing sensory noise (recapped from Chapter 2)

Second, for human participants, the proportion "jumped" responses for larger displacements in the low prior (P(J) = 0.1) condition were higher than in the baseline, P(J) = 0.5, condition for prior-training trials (the teal curve ascends above the black curve in Fig. 38, left). As noted in Chapter 2, this disproportionate increase in proportion "jumped" responses violates the Bayesian prediction that the three prior psychometric curves would be parallel to each other (Chapter 2, Fig. 15). This pattern was not true, however, for the monkey data in the visual and motor experiments (Fig. 31 and Fig. 36,

right, respectively). One difference was that the monkeys were extensively trained on priors before we ran the visual and motor noise experiments. Human participants were likely much earlier in the prior-training process, since they performed the task only for a couple hours in a single day. Early prior-training data from both monkeys showed the same pattern as for humans (surprisingly fast-rising teal, low-prior curves), supporting this hypothesis (Fig. 39).

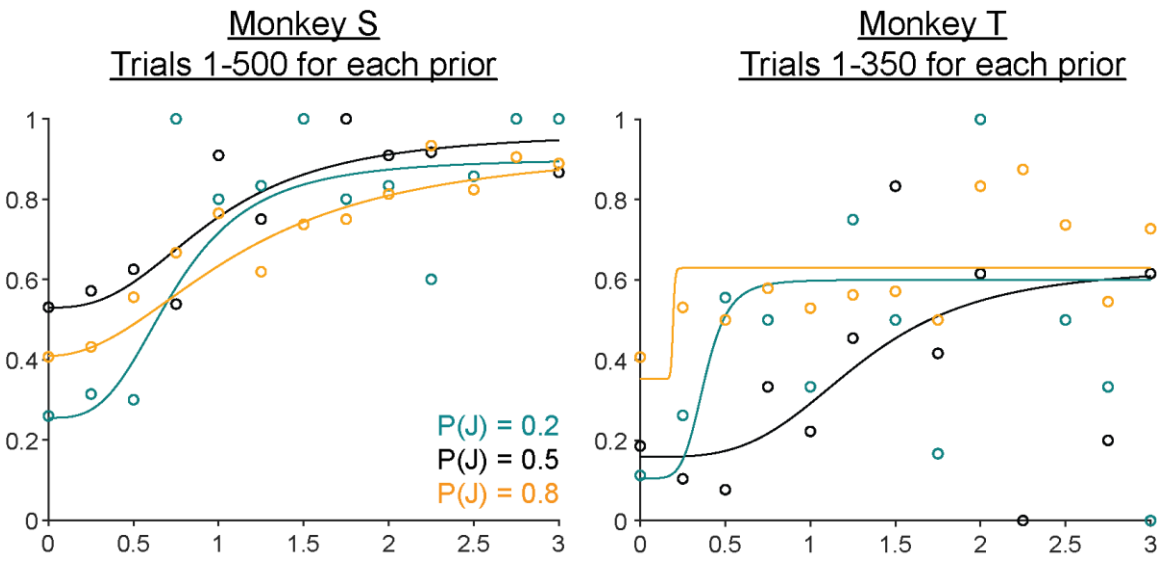

Figure 39. Psychometric curves for the two rhesus macaques early in prior training.

Overall, an alternative model would have to explain the decrease in prior use with increasing sensory uncertainty (collapse of the curves with increasing visual noise) and the disproportionately high probability of responding "jumped" in the low-prior

condition early in training but not late in training (sharply rising teal curves in the figures). The goal of this section is to identify a biologically plausible model that encapsulates these two ways in which the results deviate from Bayesian predictions.

3.5.2 Basic model structure: Perceptron

What would such a model look like? In Chapters 1 and 2, we noted that an alternative framework to Bayesian models is that of Discriminative models (Murphy, 2013; Ng and Jordan, 2002). Discriminative models directly learn boundaries between classes of stimuli. In other words, they learn to classify stimuli. For our categorical task, a Discriminative model would seek to classify continuous displacements into two categories: "jumped" and "did not jump" to meet its two-alternative forced choice demands. We first set up a simple two-layer neural network where the input layer consisted of units representing continuous displacements and the output layer has two units: "jump (J)" and "no jump (NJ)" (Fig. 40). For ease of computation, continuous input displacements were discretized into bins of 0.1°.

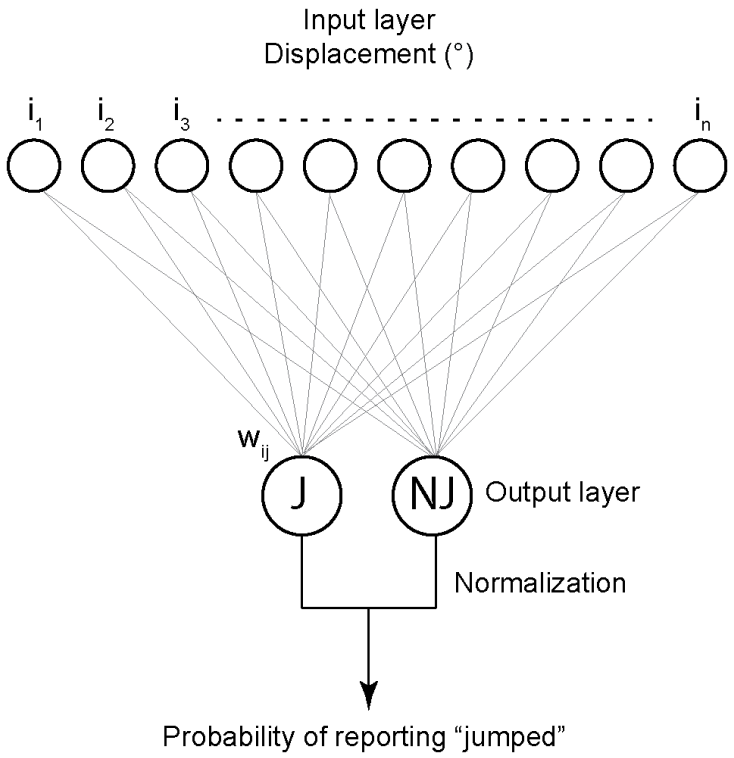

Figure 40. Schematic of a two-layer Discriminative neural network.

Each input unit was connected to both output units. On each trial, the input unit representing the displacement was activated. The activation of each output unit was the *weighted* sum of inputs, i.e.:

$$a_j = \sum_{i=1}^{n} a_i \, w_{ij} \tag{3.2}$$

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 is the normalized activation of the "jump" and "no jump" output units such that the output for each unit is bounded between 0 and 1:

$$o_{jump} = \frac{a_{jump}}{a_{jump} + a_{no\ jump}} \tag{3.3}$$

and

$$o_{no\ jump} = \frac{a_{no\ jump}}{a_{jump} + a_{no\ jump}} \tag{3.4}$$

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.

The "knowledge" of the two categories is thus stored in the weights between the input and output units, and the learned, binary category boundary takes the form of a psychometric function reporting the probability that the target "jumped" given an input displacement. In other words, the shape of the psychometric function is determined by the activation of the inputs and the weights between inputs and outputs (simulations in Section 3.5.3).

How does the model *learn* such a boundary? That is, by what rule are the weights between inputs and outputs updated? One possibility is that they are updated by a simple Hebbian-like associative learning rule (Hebb, 1949). That is, the weights between two units are updated in a manner proportional to the activity of the given input and output units. We chose a slight variation of this rule based on work by Gluck

and Bower (1988). They showed that learning category boundaries using an *error*-based learning rule, rather than a purely associative learning rule, leads to a disproportionate overweighting of infrequent events early in training.

The learning rule is given by:

$$\Delta w_{ij} = \beta a_i (d - o_j) \tag{3.5}$$

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. The term $(d - o_j)$ is therefore the error between the current output and the desired state. In summary, the change in weights or learning on each trial is proportional to the activity of the input and the error of the model on that trial. This is equivalent to the Perceptron learning rule (Rosenblatt, 1957; Marvin and Papert, 1969). Since errors are large early in training, weight changes are relatively large too (Fig. 41; red dashed arrows depict the error). Over time, as the errors get smaller, the weights asymptote towards the desired state.

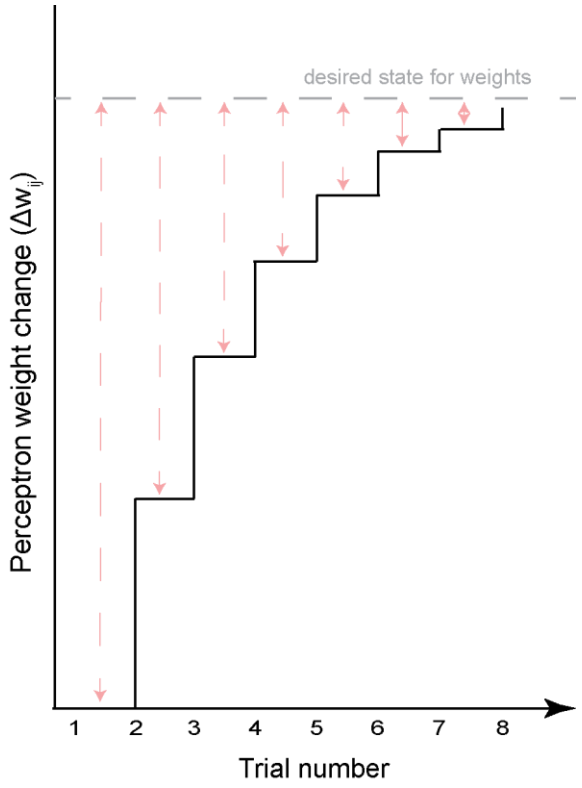

Figure 41. Schematic of weight changes between input and output units.

For infrequent events early in training, the weight changes between the event and the output are relatively large (left side of Fig. 41). Taking a snapshot of performance at this stage would thus result in an apparent overweighting of their contribution to the output. This rule, however, predicts that once weights asymptote towards the desired state late in training (right side of Fig. 41), events should contribute to the model's performance in a manner proportional to their relative frequencies. This is what we see in our categorical tasks. Large displacements in the low prior condition are infrequent events and they are overweighted early in training (Fig. 38, left and Fig.

39). Late in training for the monkeys, however, this overweighting disappears, and the psychometric curves become parallel to one another (Fig. Fig. 31 and Fig. 36, right). Simulated results recapitulating these effects are shown in Section 3.5.3.

3.5.3 Simulation of task parameters and results

First, we simulated sensory noise as a Gaussian distribution of input unit activation. The distribution was centered on the true displacement for the trial and the total activation of input units was always 1. The width of activation was determined by the sensory noise level on the given trial. Fig. 42 illustrates the effects of sensory uncertainty at two levels, 0.1° (left) and at 0.7° (right), on the output psychometric functions. As expected, adding sensory uncertainty flattened the psychometric functions.

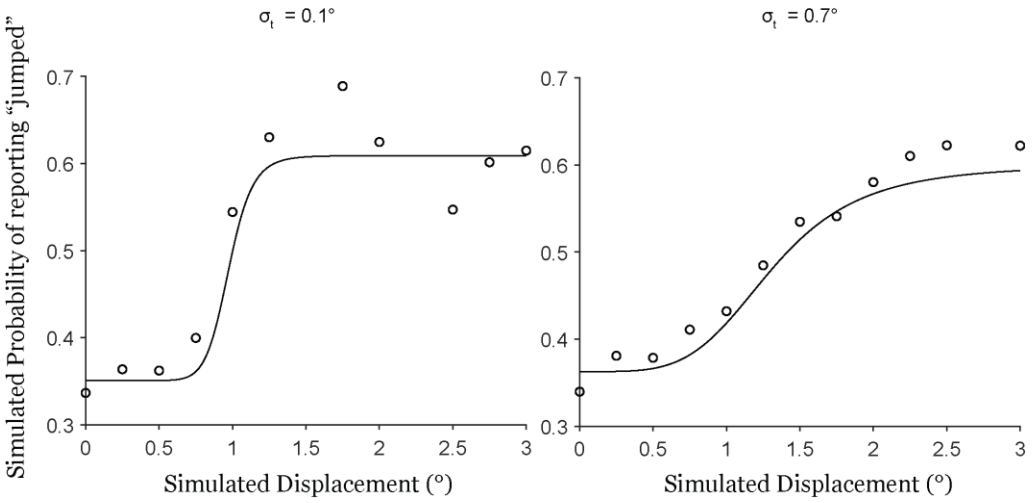

Figure 42. Effect of simulated sensory noise the Discriminative (Perceptron) model output.

Next, we considered how priors would be represented. Recall that the output psychometric function is the learned category boundary, and that its shape is determined by the activation of inputs and the weights between inputs and outputs. One clue for how priors might be represented came from the fact that early in training, for the same, low sensory noise level, i.e., the same input activation, psychometric curves in the low and high prior conditions had different shapes (slopes). To account for this, we assumed that the connections between inputs and outputs (and their corresponding weights) are *prior-dependent.* In other words, the learning of category boundaries happens in a context-dependent manner, and context-dependence is a reasonable biological assumption (reviewed in Maren et al., 2013). This means that the input-output connections (and therefore their corresponding category boundaries or psychometric functions) are learned separately for each prior level. Qualitatively, this means that blurring the input with sensory noise would blur the distinction between priors leading psychometric functions to collapse towards one another in an Anti-Bayesian manner.

We simulated the prior-dependence of the input-output relationship by simply setting up separate sets of inputs for each prior (Fig. 43).

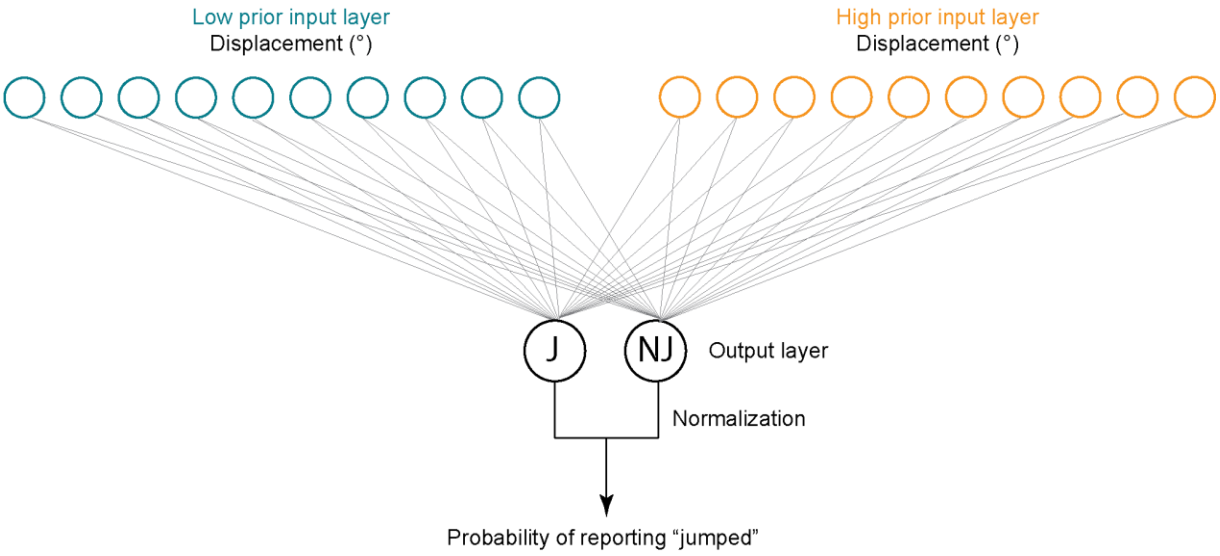

Figure 43. Prior-dependent input-output mapping. Illustrated for the low and high prior conditions.

This structure then allowed us to first simulate what prior learning would look like early in training. Specifically, does it capture the overweighting of large jump sizes in the low prior condition (sharply rising teal curves in the figures)? We simulated early training by following the same experimental structure as for the human experiments: a baseline block at P(J) = 0.5 followed by two 600-trial blocks at P(J) = 0.8 and P(J) = 0.2 respectively. Of those trials, 70% were prior training trials at the lowest noise level, σ_{target} = 0.1°. The remaining 30% were medium (σ_{target} = 1°) and high-noise (σ_{target} = 2°) testing trials with a neutral movement statistic of 0.5 but simulated using the same inputs as the prior condition. Displacements were drawn from overlapping Gaussian distributions as in the experiments. Jumps were drawn from a distribution with μ_{jump}= 0°, σ_{jump} = 2.5°, and non-jumps were drawn from a distribution with $\mu_{non\text{-}jump}$ = 0°, $\sigma_{non\text{-}jump}$ = 0.5°. On

trials where the target jumped, the desired state was set to 1 for the "jump" output unit and 0 for the "no jump" output unit. On trials where the target did not jump, it was set to 0 for the "jump" unit and 1 for the "no jump" unit. The learning rate was set at 0.5.

Fig. 44 shows the model output on prior-training trials (comparable to the data in Fig. 38, left).

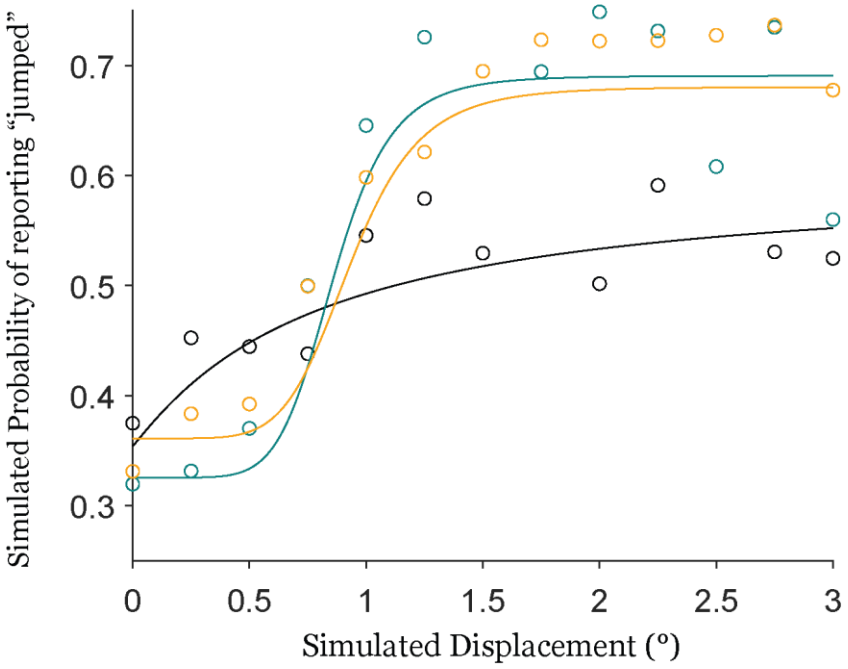

Figure 44. Simulated early prior training data from the Perceptron model.

As expected, the simulated output recapitulates the disproportionately high response rate for large displacements in the low prior condition. It also, however, *downweights* small displacements in the high prior conditions, such that the intercept of the orange curve in Fig. 44 is lower than the intercept in Fig. 38, left. To account for this,

we considered that, as noted at the start of Section 3.5.1, reports in the categorical task are the product of both a Discriminative and a Bayesian model. The Bayesian use of priors to compensate for high motor uncertainty imposed by the saccade drives high prior intercepts up (Fig. 34, right), and could thus compensate for the downweighting.

Therefore, we next simulated a weighted *combination* of outputs from the Perceptron and Bayesian models in the prior training condition. Motor uncertainty in the Bayesian model was simulated by setting the width of the non-jump distribution, $\sigma_{\text{non-jump}} = 1°$, i.e., double the width of the simulated experimental distribution to mimic saccadic suppression. Combining the outputs of the Perceptron model and the Bayesian model at relative weights of 0.9 and 0.1 respectively generated the data in Fig. 45, where the pattern of the curves better matches data from human participants.

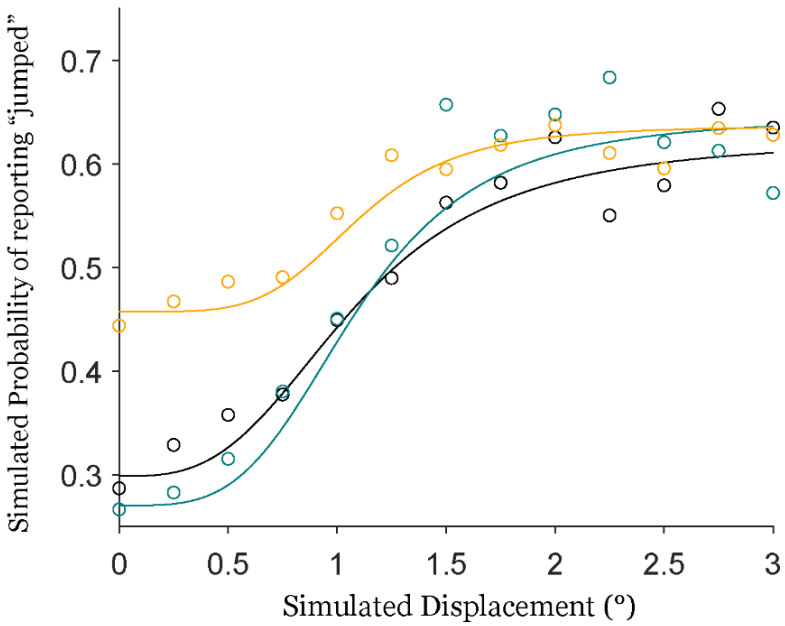

Figure 45. Simulated prior learning in the combined Perceptron and Bayesian model.

We then looked at responses in the medium and high noise hypothesis-testing trials. Fig. 46 depicts psychometric functions in the medium- and high-noise conditions of the simulated experiments.

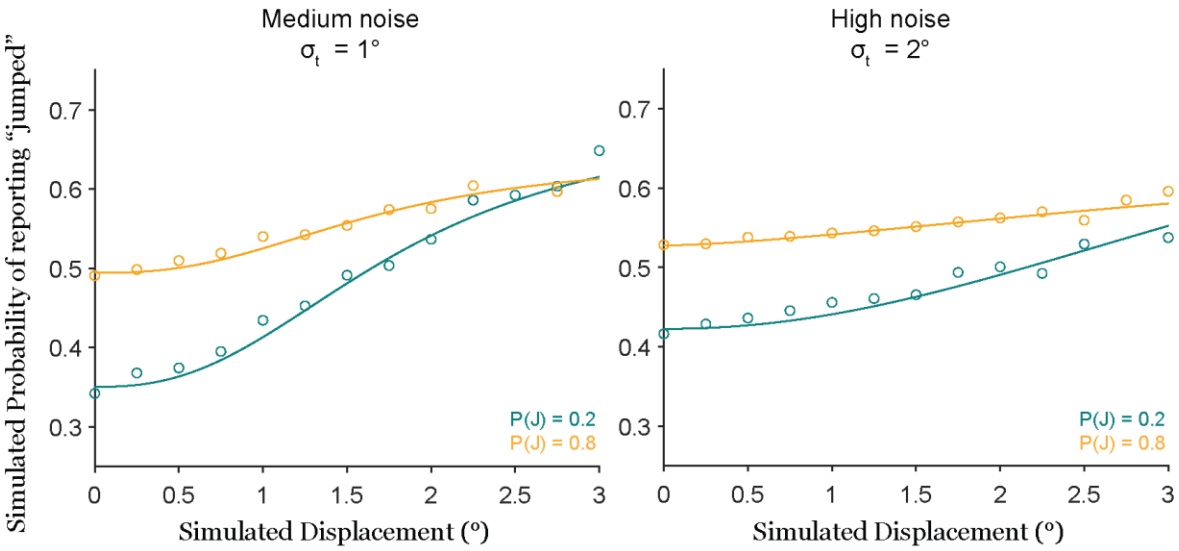

Figure 46. Medium and high noise psychometric curves for the simulated experiments.

As with all the empirical data in Chapters 2 and 3, we compared intercept differences between the P(J) = 0.8 and P(J) = 0.2 conditions across the three noise levels (Fig. 47), and they decrease with increasing noise.

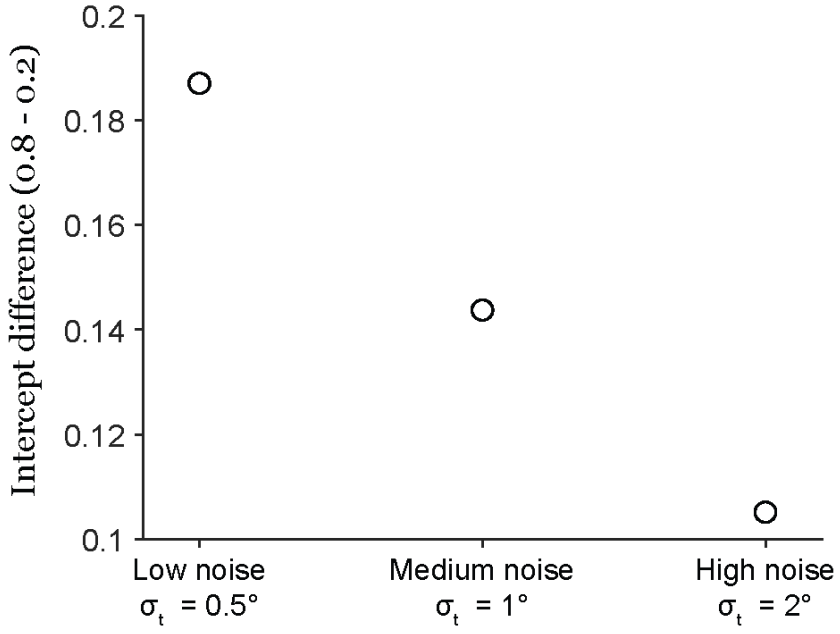

Figure 47. Intercept difference between the high (0.8) and low (0.2) prior conditions

Finally, recall that a prediction of the error-based learning rule is that once weights asymptote towards a relatively stable desired state for all input units, then the apparent overweighting of infrequent inputs should disappear. That is, inputs should contribute to the output in a manner proportional to their relative frequencies. For our task, this would mean that psychometric curves for the three prior conditions become parallel to one another. Late-training data from monkeys matched this prediction (Fig. 31 and Fig. 36). To confirm that our models could reproduce it, we let them run for 5000 trials. Fig. 48 shows simulated data from trials 3000-5000 (as a proxy for late training) for the combined (Perceptron + Bayesian) model (left) and Perceptron model alone (right).

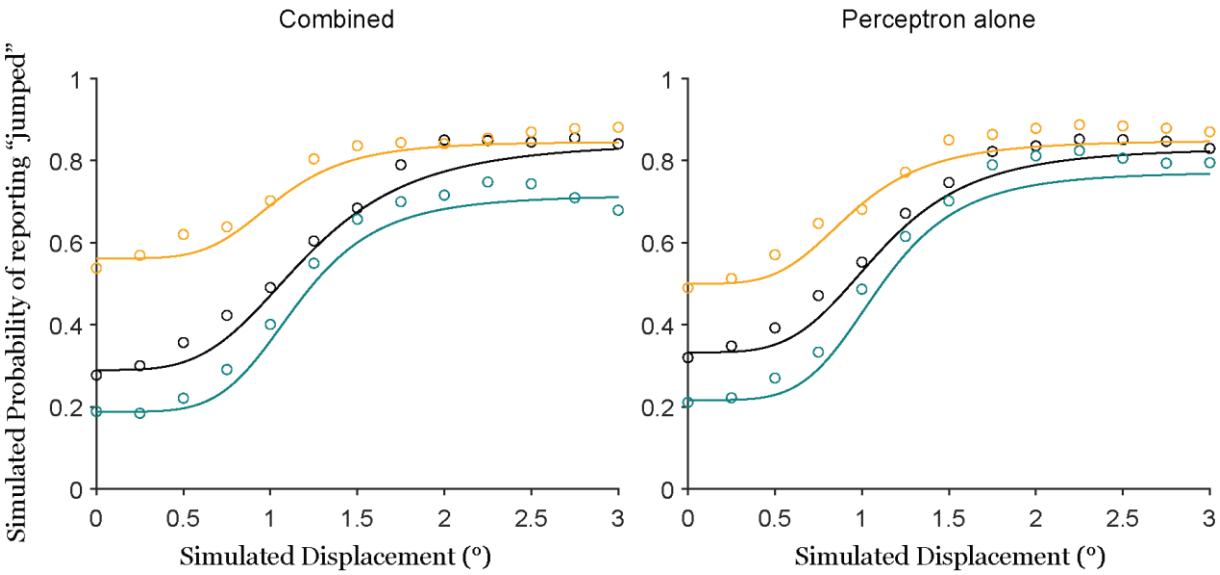

Figure 48. Simulated late-training results.

In summary, a neural network model (Perceptron) when combined with the normative Bayesian model, recapitulates several otherwise unexplained features of the data including the Anti-Bayesian prior use with increasing noise, the apparent overweighting of large displacements in the low prior condition early in training, and the eventual proportional weighting of displacements across priors.

3.5.4 Limitations and Future Directions

The goal of this section was to investigate whether a biologically plausible alternative to the normative, Bayesian model could reproduce the patterns of results observed in our categorical tasks. We identified a broad model structure that accounts

for unexplained aspects of the data and chose simulation parameters to roughly match those used in experiments wherever possible. This provided proof-of-principle that the novel, combined model, which includes a small Bayesian component but is not solely Bayesian, can recapitulate the empirical data. However, the parameters were still chosen arbitrarily, including the widths of the "jump" and "non-jump" distributions from which displacements were drawn, the scaling factor for increasing the width of the "non-jump" distribution to simulate motor noise (saccadic suppression), the three levels of sensory noise, the learning rate, and the combination weights of the Perceptron and Bayesian outputs. All these internal parameters likely vary on a subject-by-subject basis. We also made minimal additional assumptions about how these parameters interact with one another under different conditions.

Thus, while the chosen parameters and set of assumptions were sufficient to recapitulate *patterns* in the data, many simulated results are "compressed" and do not match the empirical data numerically. For example, psychometric curves move closer with increasing visual noise as predicted (Figs. 47 and 48), but the magnitude of the "collapse" was smaller than seen in the empirical data. Further, simulated psychometric functions early in training are limited to a smaller range of output probabilities. A major future direction for this work is to fit the models to empirical results to identify best-fit parameters on a subject-by-subject basis, and to determine the interaction assumptions

they necessitate. Necessary assumptions for the model will make predictions that can be tested in future behavioral or neural experiments.

Another future direction for the combined model is to account for the fact that continuous reports of object displacement are Bayesian. Currently, the activation of the input layer which represents continuous displacements is centered on the true displacement with the width determined by the sensory uncertainty. This activation is equivalent to the sensory likelihood distribution. One way to account for Bayesian inference at this level would be for the input layer to represent the *posterior* instead. Alternatively, the distribution of the input layer may be modified to better represent the prior distribution of displacements (Fischer and Peña, 2011). Overall, the goal of this modeling work is to develop a unifying model that both 1) explains the mixed Discriminative and Bayesian aspects of active visual perception across saccades, and 2) makes further, testable predictions for the saccadic system.

3.6 Discussion

Overall, we found that human participants were Bayesian for *continuous* reports of object displacement across saccades, but this was not the case for categorical reports (Chapter 2). The latter, Anti-Bayesian result was independent of the underlying statistics of object movement, as shown with the monkey control data for that experiment. The task included both image noise and motor noise but based on a post-hoc analysis of saccadic endpoint scatter, we reasoned that the Anti-Bayesian effect was primarily due

to image noise. Further investigation on monkeys in this Chapter supported the hypothesis. In other words, categorical prior use for active perception can be summarized as being positively modulated by internal, motor-induced noise and negatively modulated by external, image noise. (Fig. 49).

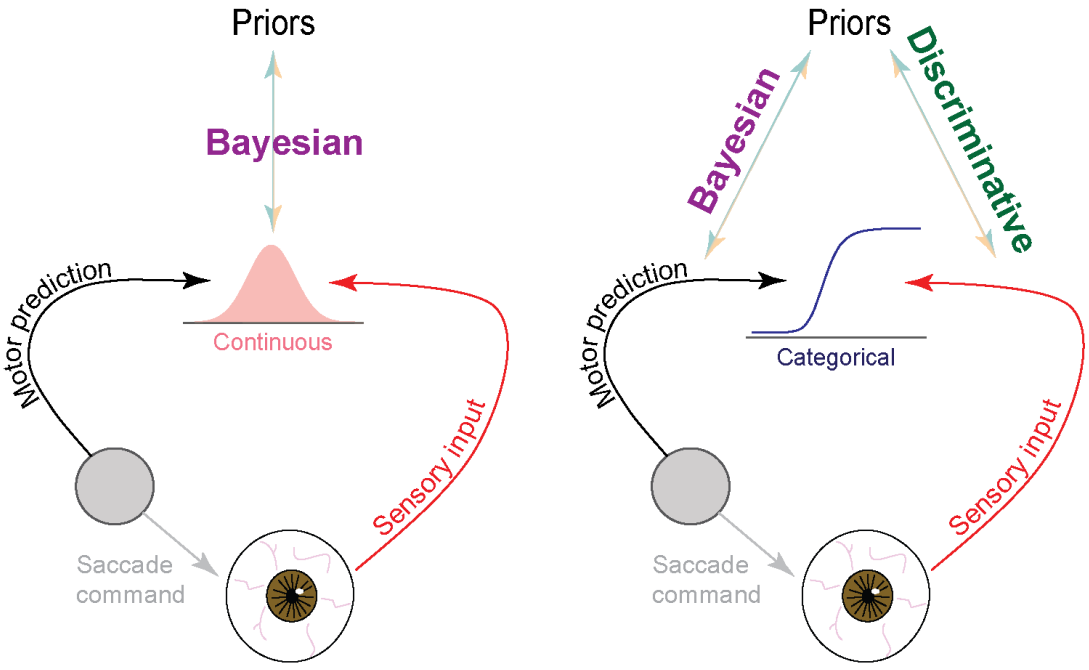

Figure 49. Summary of results in Chapters 2 and 3.

The negative modulation of priors by external, visual noise can be explained by a two-layer, Perceptron model, which implements a Discriminative category boundary between "jumped" and "did not jump" in the task instead.

Limitations of the experimental design and modeling choices made in this study should be considered while interpreting the results. First, as raised in Chapter 2, one

concern is that our empirical results in Chapters 2 and 3 are a consequence of the specific conceptualizations of the prior, the sensory uncertainty, and motor noise in the normative, Bayesian model for the categorical task. While this is possible, to our best estimate, there was no unifying alternative model that explained both the increasing prior use with motor noise and the decreasing prior use with image noise parsimoniously. Second, we limited the simulation of motor-induced noise to just one phenomenon, i.e., saccadic suppression. We did this by increasing the width of the non-jump likelihood. That is, there was a greater likelihood that larger displacements were perceived as non-jumps. We focused on saccadic suppression since it is a gross, categorical form of uncertainty that is present when a saccade is made, and not otherwise. It was thus sufficient for testing our hypothesis in the current study. However, there are other ways in which saccades influence vision both at the level of behavior and neurons. Such effects include the compression of space towards the saccade target under illumination (Honda, 1993; Awater and Lappe, 2006; Hamker et al., 2011; Pola, 2011) and the shifting or smearing of visual receptive fields around the time of saccades (reviewed in Neupane et al., 2020; Golomb and Mazer, 2021). Indeed, the magnitude of saccadic suppression may vary with the amplitudes of saccades (Stevenson et al., 1986; Li and Matin, 1990) and the direction in which the probe moves relative to the saccade vector (Niemeier et al., 2003). The current results do not preclude the inclusion of additional, fine-grained motor-induced phenomena into the normative

model, and future studies may be designed to specifically test their predictions. Another consideration for Experiment 2 is that in the no-saccade, "low motor noise" condition of the task, animals stay fixated on a central square for the duration of a trial. We did not prevent the animal from making fixational eye movements such as microsaccades (reviewed in Martinez-Conde et al., 2004), and saccadic suppression has been observed around the time of microsaccades too (Bair and O'Keefe, 1998; Hafed and Krauzlis, 2010; reviewed in Martinez-Conde et al., 2013; Hafed et al., 2015). Although we did not control for microsaccades in our study, the displacement of the stimulus was not timed to their onset as it was in the "high motor noise" condition with a saccade. Thus, at least on average, we expect that the effects of (micro)saccadic suppression were lower in the no-saccade condition. Finally, several experimental parameters such as the categorical priors and sensory uncertainty levels were discretized to allow for sufficient data in each condition. Indeed, even the categorical experiments in monkeys that tested prior use against image and motor noise were independent. That is, we varied each type of noise while holding the other constant. Future studies may explore additional combinations of the various parameters to provide a more comprehensive picture of how Bayesian and non-Bayesian computations combine for active perception.

One interesting outcome of the Discriminative modeling in our task, is that even when modeling learning purely with a Perceptron model, psychometric curves for the three priors eventually become parallel late in training (Fig. 48, right). Recall that

parallel curves are predicted by the normative, Bayesian model. Thus, after sufficient training, even a Discriminative model approximates the "optimal" solution in this task. However, it fails the critical test of the Bayesian hypothesis, i.e., the learned priors are not used more to compensate for uncertainty. This outcome highlights: 1) the importance of probing the use of Bayesian vs. Discriminative models with uncertainty, and 2) that there are several ways for a system to perform a task well enough to meet its goals.

Finally, the results in this chapter pose the question: how are Bayesian and Discriminative behavior might be implemented by the oculomotor system? There have been theoretical and empirical investigations into how Bayesian inference is implemented neurally, both for continuous (Ma et al., 2006; Fischer and Peña, 2011, Fetsch et al., 2012; Darlington et al., 2018; reviewed in Sohn and Narain, 2021) and categorical tasks (Tajima et al., 2016). The current psychophysical findings provide motivation for adapting concepts from this literature to the saccadic system. In Chapter 5, we investigate how one brain region which is important for visual continuity across saccades, the Frontal Eye Field (FEF), supports Discriminative vs. Bayesian processing in the categorical task.

4. Neuronal Correlates of Bayesian and Discriminative models in the Frontal Eye Field

4.1 Introduction

When primates view a scene, at least two sources of sensory uncertainty affect their perception of visual objects: 1) uncertainty in the visual input (image noise), and 2) uncertainty caused by suppression of visual input during a saccade (motor noise). In the psychophysical experiments and modeling of Chapters 2 and 3, we tested whether the primate visual system compensates for image noise and motor noise by using expectations about the environment, or "priors," in a Bayesian manner. We found that the trade-off between priors and *motor* noise is Bayesian, but the trade-off between priors and *image* noise is Anti-Bayesian. Decreasing prior use with image noise was instead explained by a Discriminative model which learned to categorize continuous displacement values. Here we ask: how are these distinct models implemented neurally by the primate visuosaccadic system?

In the current study, we focused on a critical node in the system, the Frontal Eye Field (FEF) and asked whether the activity of individual neurons in FEF are predictive of Bayesian or Discriminative prior use for the perception of object displacement across saccades. We chose FEF as a candidate region because it either receives or generates information about all the ingredients necessary to distinguish self-generated retinal displacement from true object movement in the world. First, it receives a copy of the

saccade command or corollary discharge, which contributes to the presaccadic remapping of retinal receptive fields of some FEF neurons (Sommer and Wurtz, 2002; 2004a; 2004b; 2006; 2008). As discussed in Chapter 1, since remapping allows neurons to sample the same region of space (the future field) before and after a saccade, it can serve as the basis for making a prediction about the sensory consequence of the saccade (reviewed in Hall and Colby, 2011; Wurtz, 2008; Wurtz, 2018). Next, FEF receives visual information (Umeno and Goldberg, 1997) that it can compare to the prediction to determine whether there is a discrepancy. A discrepancy would indicate that an external event such as object movement took place. Finally, a previous study (Crapse and Sommer, 2012) showed that neurons in FEF report this discrepancy between the prediction and the visual input. That is, they report whether an object moved across a saccade. The FEF is thus a natural candidate for supporting either Bayesian or Discriminative prior use in this task, or both.

Two rhesus macaques performed the motor- and image-noise experiments detailed in Chapter 3, with all trial types interleaved during each session. Recall that behavioral prior use is Bayesian for motor noise but Discriminative for image noise. This allowed us ask whether the activity of individual neurons selectively correlates with one or the other computation. We first investigated whether FEF neurons report the prior, a key variable for Bayesian or Discriminative computations in our task that has not previously been reported in FEF. We found that many neurons in FEF selectively

reported one prior over the other. We next asked if the prior selectivity of these neurons is modulated by image noise or motor noise. If so, does the noise modulation predict the Bayesian prior use we found for motor noise, the Discriminative prior use we found for image noise, both, or neither? We found that the FEF neurons predicted the animals' Discriminative prior use in the image noise trials but not their Bayesian prior use in the motor noise trials. These findings 1) corroborate predictions from the behavioral results in Chapter 3 suggesting that Bayesian and Discriminative computations are dissociable within the saccadic system, and 2) situate FEF along a pathway that selectively supports a Discriminative model.

4.2 Experimental Methods

We recorded from 90 FEF neurons in 2 rhesus macaques (54 in Monkey S and 36 in Monkey T) while they performed the SSD task that included randomly interleaved trials of the image- and motor-noise conditions. The purpose was to identify the neuronal correlates of Discriminative (image noise experiment) and Bayesian (motor noise experiment) computations. In each experimental session, we first mapped the visual, motor, or visuomotor spatial receptive field (RF) of an identified neuron (4.2.2). The animal then performed the SSD task for the rest of the session. The probe was placed such that its final location was in the center of the identified RF on every trial (details in 4.2.3).

4.2.1 Single-unit Electrophysiology

After behavioral training was completed, a recording chamber (Crist Instruments, Hagerstown, MD) was stereotaxically implanted over the FEF in an aseptic surgical procedure. After at least two weeks for recovery from surgery, animals were placed under water control in their home cage and experiments commenced. In each experimental session, 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 animals received drops of water as reward for performing the tasks. All surgical, experimental, and water control procedures were performed in accordance with protocols approved by the Duke Institutional Animal Care and Use Committee.

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).

4.2.2 Spatial Receptive Field Mapping

After an FEF 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 it was rewarded with a drop of water. 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°. If the neuron

had a direction preference, we then ran the amplitude task, where the target was restricted to one direction but presented at one of 4 amplitudes: 2°, 5°, 10°, and 20°. The overall extent of the neuron's responsiveness and the putative center of its receptive field were inferred based its response properties in the direction and amplitude tasks.

4.2.3 Saccadic Suppression of Displacement (SSD) task for recording

The structure of the task performed by animals during each recording session was the same as that in Chapter 3. On each trial, a fixation square (1° x 1°) first appeared on the screen. After fixation had been acquired and maintained for 300-500 ms, the visual probe appeared for 500-700ms, during which the monkey maintained fixation. The fixation square was then replaced by the saccade target (1° X 1°). Saccade initiation (defined as the time the eye crossed a threshold set at 20% of the saccade length) triggered the displacement of the probe. Tests that compared photocell measurements (described in Chapter 3) to eye position traces confirmed that the probe displacements occurred during the saccade, typically when it was ~2/3 completed. The probe was displaced in a direction parallel to the saccade. Animals were further required to maintain postsaccadic fixation for 700ms 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. Priors (0.8, 0.5, and 0.2 probability of probe movement) were cued by the same colors used in the behavioral experiments for each animal. Animals performed both experiments detailed in Chapter 3 – one manipulating image noise, and one manipulating motor noise – simultaneously every session. Three levels of image noise were introduced by increasing the width of the Gaussian visual probe with the same values used for each animal in Chapter 3. Two levels of motor noise were determined by the requirement to make an intervening saccade or not. Note that the lowest image noise condition in which an intervening saccade was required was common to both experiments; we thus use it as a baseline condition for both the image and motor noise analyses. Unlike in the behavioral sessions of Chapter 3, the prior statistic remained valid (i.e., it reflected the underlying probe displacement probability) for all noise conditions. There were thus 10 prior x noise conditions: 3 priors with and without an intervening saccade respectively for a total of 6 conditions, and 2 priors (0.8 and 0.2) for two additional image noise levels. All conditions were presented at an equal frequency and randomly interleaved throughout each session.

The main difference between the recording session tasks and the behavioral tasks of Chapter 3 was the placement of the probe and saccade target (Fig. 50). While for behavioral sessions, both were placed in the same location every day, in neural recording sessions they were placed to match the identified spatial receptive field of the

neuron. On trials where the animal did not make a saccade (Fig. 50, left), the final probe location was simply at the center of the receptive field (magenta dashed circle). On trials in which the animal made a saccade (Fig. 50, right), the final probe location was at the center of the *postsaccadic* receptive field of the neuron. Pre-saccadically, that was the "future field" of the neuron (blue dashed circle).

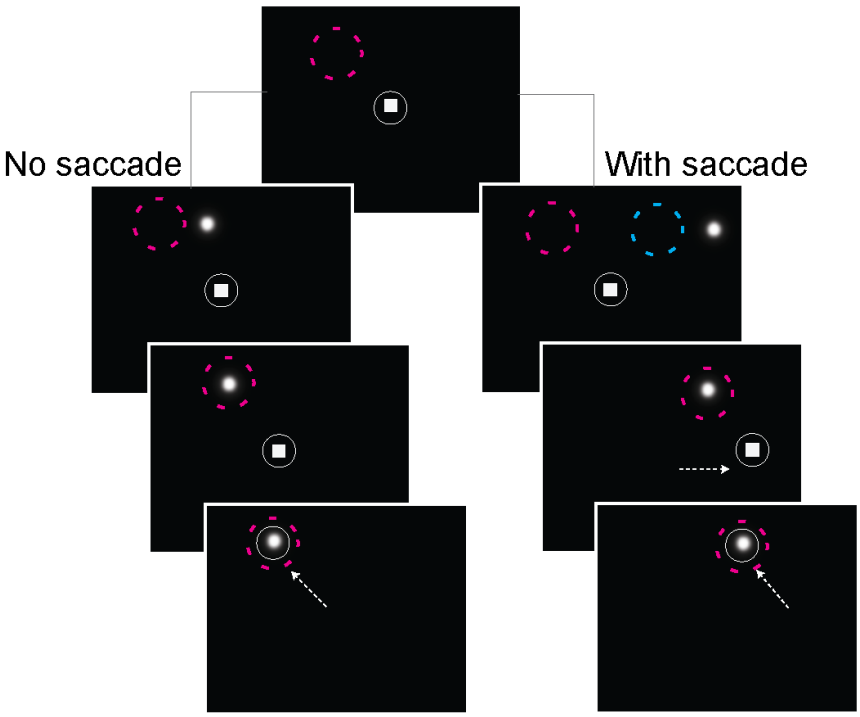

Figure 50. Schematic of the SSD task during recording sessions.

The probe was initially placed at a location determined by the jump size. In trials in which the probe moved during a saccade, 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

amplitude of the saccade was jittered by 1°, so that an absolute target location did not correspond with a jump or non-jump response. Note, however, that while this was possible while still maintaining that the target was displaced to the center of the postsaccadic RF, it was not possible to do on trials with no intervening saccade (Fig. 51, left). Thus, the post-jump location of the probe was always the same on trials with no saccade. Displacements were drawn from the same Gaussian distributions as used in the behavioral experiments in Chapter 3, but those distributions were maintained *across sessions*. For each session, four to six discretized displacements were chosen to allow for a comparison of neural responses across prior and noise conditions with variability minimized across all other parameters.

4.3 Analysis methods

4.3.1 Epochs

There were 5 epochs during each SSD trial in which neuronal activity was predicted to potentially support the performance of the task:

1) The fixation epoch, which was aligned to the acquisition of initial fixation on the square. During this epoch, the animal only had information about the prior, i.e., the probability that the probe may move.

2) The probe onset epoch when the probe first appeared in its pre-saccadic, pre-jump position. For trials in which the animal made an intervening saccade, the probe

was located near the center of the future field of the neuron. For trials in which the animal did not make a saccade, it was near the center of the receptive field.

3) The intervening saccade epoch, which was aligned to the onset of the intervening saccade on trials where there was one. For the no-saccade (low motor noise trials), there was no intervening saccade.

4) The reafferent epoch, which was aligned to the time that the probe moved to the center of the (post-saccadic) receptive field after the jump. On trials in which the animal made an intervening saccade, this epoch began at the end of that saccade. At that point the center of the receptive field had moved to the probe location. On trials where the animal did not make an intervening saccade (in the low motor noise condition), the epoch was aligned to when the probe moved to the center of the receptive field.

5) The response epoch, which was aligned to the final response. On trials in which the animal responded "jumped" with a response saccade, this epoch was aligned to the onset of the saccade. On trials in which the animal remained fixated (responded "did not jump"), the epoch was aligned with the appearance of the cue indicating that a response saccade was permitted. All response saccades were made into the receptive field of the neuron.

Our main quantitative analyses, as reported here, focused only on the epochs for which neural activity was directly comparable across task conditions both in the image noise experiment and the motor noise experiment. This criterion disqualified the probe

epoch (since the probe was near the receptive field for no-saccade condition but near the future field for all others) and the intervening saccade epoch (since the intervening saccade was not made in the no-saccade condition). Thus, the focus of Chapter 4 will be on neural activity during the fixation, reafferent, and response epochs (Fig. 52).

The fixation epoch spanned 0-150ms after fixation acquisition. The reafferent epoch was defined as the range 0-150ms after the end of the saccade or, if there was no intervening saccade, the displacement of the probe. The response epoch was defined as -150 to 50ms relative to the onset of the response saccade or, if there was no response saccade, the cue to make a response saccade.

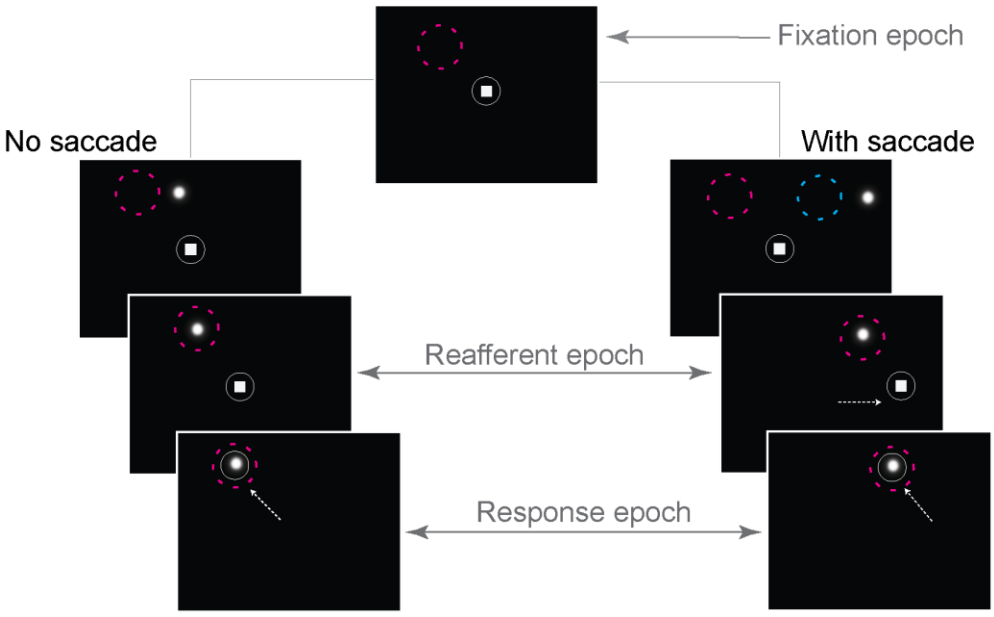

Figure 51. Task epochs used for further analysis.

4.3.2 Analysis measures

We first quantified the neuron's firing rate in each epoch. The firing rate for each trial was quantified based on the total spike count for that trial in each epoch:

$$Firing\ rate\ (spikes/s) = Spike\ count\ X\ \frac{1000}{Duration\ of\ epoch\ (ms)} \qquad (4.1)$$

For each neuron, we next assessed if there was a statistically significant difference between its firing rate on high prior trials (P(J) = 0.8) and on low prior trials (P(J) = 0.2) in the lowest sensory noise condition where the animal made an intervening saccade. This condition served as a baseline for evaluating prior effects because it was the common condition between the image- and motor-noise experiments. We tested for statistical significance with a Wilcoxon rank sum test at $p<.05$ criterion level. For all statistical tests, corrections for multiple comparisons were included as appropriate.

For trials in which there was a significant difference, we then used the Wilcoxon effect size between the high and low prior for further analyses and comparison with behavior (details in Results). 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 (Mann and Whitney, 1947).

4.4 Results

4.4.1 Behavior

We first confirmed that each animal's behavioral performance during the recording sessions replicated the findings in Chapters 2 and 3. We performed 40 recording sessions in Monkey S and 24 recording sessions in Monkey T. Behavioral data were analyzed from sessions in which the animal performed at least 250 total trials (40 sessions for Monkey S, 18 sessions for Monkey T). Figure 52 shows the results from the image noise experiment. These results were presented as a control in Chapter 2, to show there that the Anti-Bayesian effect occurred for image noise regardless of whether the underlying statistics were neutral (0.5 probability of probe jump) as in the human experiments or matched to the prior as done here during the recording experiments.

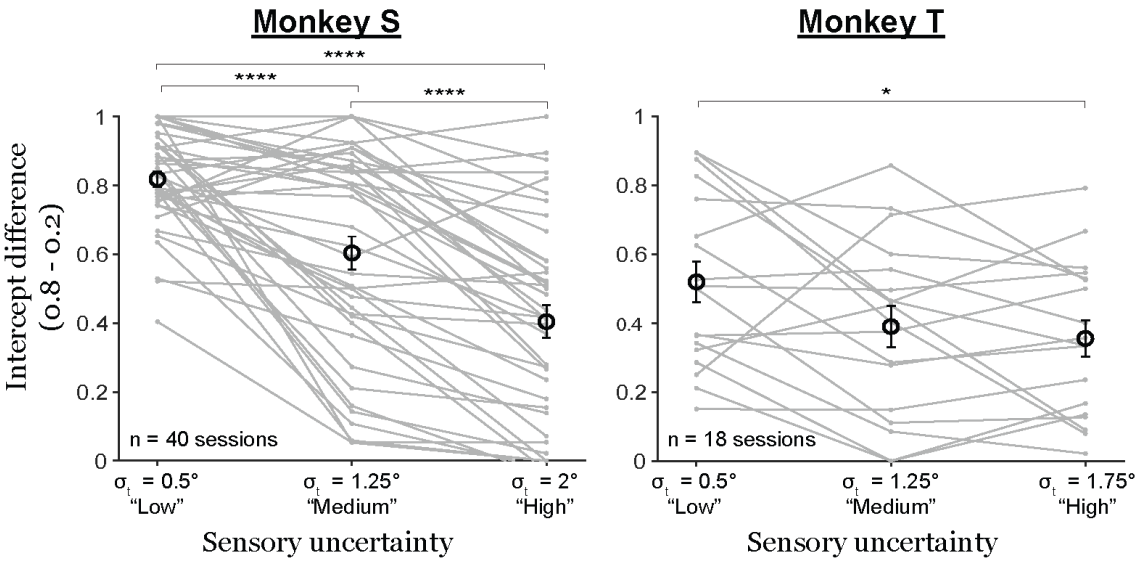

Figure 52. Prior use in the image noise experiment was Anti-Bayesian,

To briefly recapitulate the results, we found that prior use measured by intercept differences, or the difference in proportion "jumped" responses at displacement = 0, between the P(J) = 0.8 and P(J) = 0.2 conditions *decreased* with increasing sensory noise for both animals. A repeated-measures ANOVA with intercept difference as the main within-subjects factor yielded significant effects (Monkey S: $F(2) = 51.75$, $p = 4.97 \times 10^{-15}$; Monkey T: $F(2) = 4.56$, $p = 0.0176$). For monkey S (n = 40 sessions), post-hoc comparisons using a Tukey HSD test showed that intercept differences in the low noise condition (σ_t = 0.5°; mean = 0.82, SE = 0.02) were significantly higher than in the medium noise (σ_t = 1.25°; mean = 0.60, SE = 0.05; $p = 3.49 \times 10^{-6}$) and high noise (σ_t = 2°; mean = 0.40, SE = 0.05; p = 0) conditions. Intercept differences in the medium noise condition were also higher than in the high noise condition ($p = 1.51 \times 10^{-5}$). For Monkey T (n = 18 sessions), there

was a significant difference between the low (σ_t = 0.5°; mean = 0.52, SE = 0.06) and high noise (σ_t = 1.75°; mean = 0.36, SE = 0.05; p = 0.0190) conditions. Intercept differences in the medium noise condition (σ_t = 1.25°; mean = 0.39, SE = 0.06) fell in between the low and high noise conditions but were not statistically different from either.

Similarly, the results in the motor noise trials during neural recordings were consistent with the motor noise data from separate sessions reported in Chapter 3. For both Monkey S and Monkey T, intercept differences between the P(J) = 0.8 and P(J) = 0.2 conditions increased with increasing motor noise (Fig. 53).

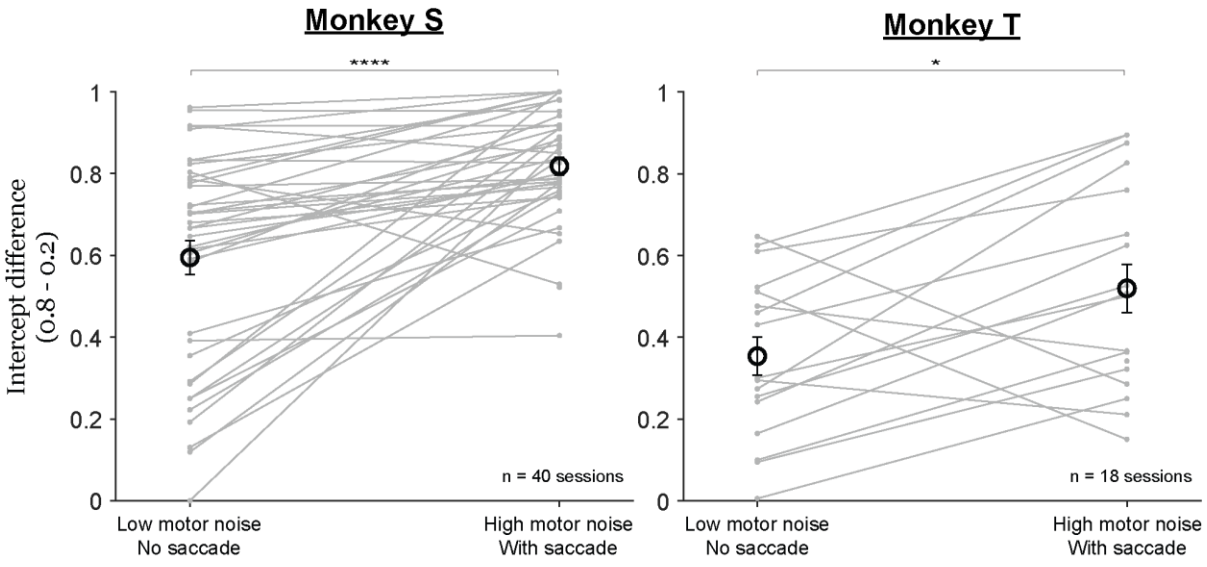

Figure 53. Prior use in the motor noise experiment was Bayesian.

For both Monkey S and Monkey T, intercept differences in the low-noise, no-saccade condition (mean: 0.59, SE = 0.04 for Monkey S; mean: 0.35, SE = 0.05 for Monkey

T) were lower than in the high-noise, with saccade condition (mean: 0.82, SE = 0.02 for Monkey S; mean: 0.52, SE = 0.06 for Monkey T). For both, the differences were statistically significant ($p = 2.12 \times 10^{-6}$ for Monkey S and $p = 0.0125$ for Monkey T on a paired t-test). In sum, the behavioral results from the recording sessions were consistent with all the psychophysical results in the behavior-only experiments in Chapters 2 and 3.

4.4.2 Prior selectivity

The first question for investigating the neuronal correlates of Bayesian vs. Discriminative processing in FEF was whether neurons in FEF encode the prior at all. To address this, we tested whether the neurons showed a significant difference between the high and low prior conditions in the SSD task during the condition we considered to represent baseline performance: the low sensory noise, with-saccade trials. For those trials, the firing rate of each neuron was compared between all high prior trials and all low prior trials, in the fixation, reafferent, and response epochs (Wilcoxon rank sum tests, $p < 0.0167$ criterion as Bonferroni correction for the three tests for each neuron). We found that 51 of the 90 neurons (42/54 from Monkey S, 9/36 in Monkey T) showed a significant difference in firing rate between priors in one or more of the epochs. In the previous section 4.4.1, we performed the *behavioral* analyses only on trials for which the displacement = 0° (i.e., to quantify intercept differences as in previous chapters).

However, we did not observe a difference in neuronal prior selectivity across displacements (not shown). Therefore, for the neural analyses and to compare neural activity to behavior, we pooled data across displacements to increase statistical power. Additionally, since displacements were drawn from broad and narrow Gaussian distributions for the high and low prior respectively, pooling across displacements allowed for a roughly equal number of trials analyzed in each prior condition.

We then assessed the epochs during which the 51 neurons were prior-selective, and the direction of their selectivity. Fig. 54 summarizes the results.

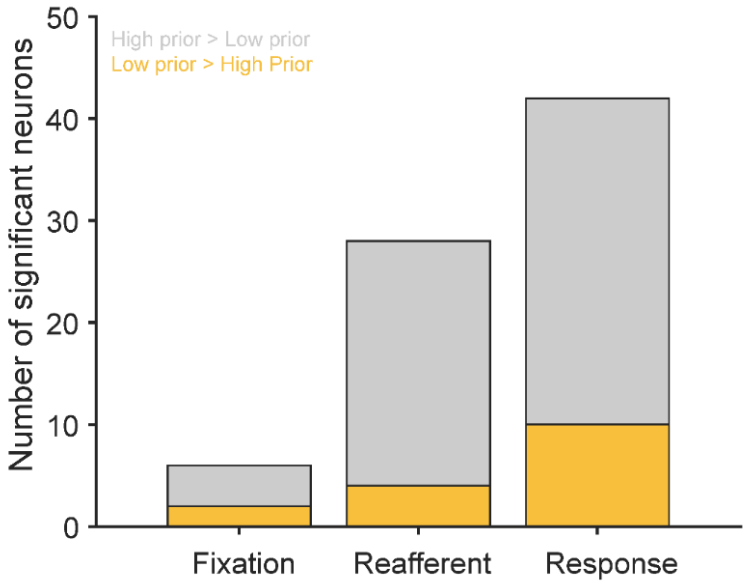

Figure 54. Distribution of prior selectivity by epoch.

We found that 6 neurons were selective during the fixation epoch, with 4 showing a positive effect size (i.e., firing rate in the 0.8 prior condition greater than in the 0.2 prior condition; Fig. 54, left). Because prior-selective activity was so rare during the fixation epoch, we do not analyze it further. In contrast, 28 cells were prior-selective during the reafferent epoch, 24 of which had a positive effect size (Fig. 54, middle), and 42 cells were selective during the response epoch, 32 of which had a positive effect size (Fig. 54, right). Since they were the majority conditions, we focused on cells which showed a positive effect size in the reafferent and response epochs for further analyses. Specifically, we asked if these cells predict behavior in the image noise experiment (Discriminative model), motor noise experiment (Bayesian model), or both.

4.4.3 Correlates of Discriminative vs. Bayesian behavior

4.4.3.1 Predictions

We first considered the possible ways in which prior selectivity may be modulated by image and motor noise respectively. The predictions are schematized in Figures 55-58. As noted above, we designated the condition that was in common between the image noise and motor noise experiments, the low sensory noise, with-saccade condition, as the "baseline" to which we compared the other trial types (purple box in Figs. 55-58). This baseline condition had the highest behavioral prior use across the two experiments, as shown by the leftmost symbols in Fig. 52 and the right symbols

in Fig. 53). Since we also used the baseline condition to screen neurons for high prior selectivity, this condition always had some prior selectivity. That is, firing rates in the high prior condition (orange lines in Figs. 55-58) were higher than in the low prior condition (teal lines). To the right, in the red box for Figures 55-58, the two image noise conditions are depicted. Behavioral prior use *decreases* with increasing image noise, consistent with a Discriminative model as addressed in Chapters 2 and 3. If prior selectivity in the neurons also decreased relative to the baseline with increasing image noise, that would indicate that the activity of FEF neurons follows the patterns of a Discriminative model. To the left, in the black box, is the low motor noise condition. Behavioral prior use also decreases in this condition relative to the baseline, high motor noise condition, and we showed in Chapters 2 and 3 that this was consistent with a Bayesian model. Thus, if the FEF neurons also showed decreasing prior selectivity in the low motor noise condition relative to the baseline, it would indicate that the neurons follow the patterns of a Bayesian model. In general, the neurons could exhibit activity consistent with neither model, the Bayesian model only, the Discriminative model only, or both models. Specifically, we predicted that the prior selectivity of FEF neurons might be modulated by:

1) Neither motor nor image noise, suggesting that FEF neurons report the prior but do not directly support either Bayesian or Discriminative models (Fig. 55):

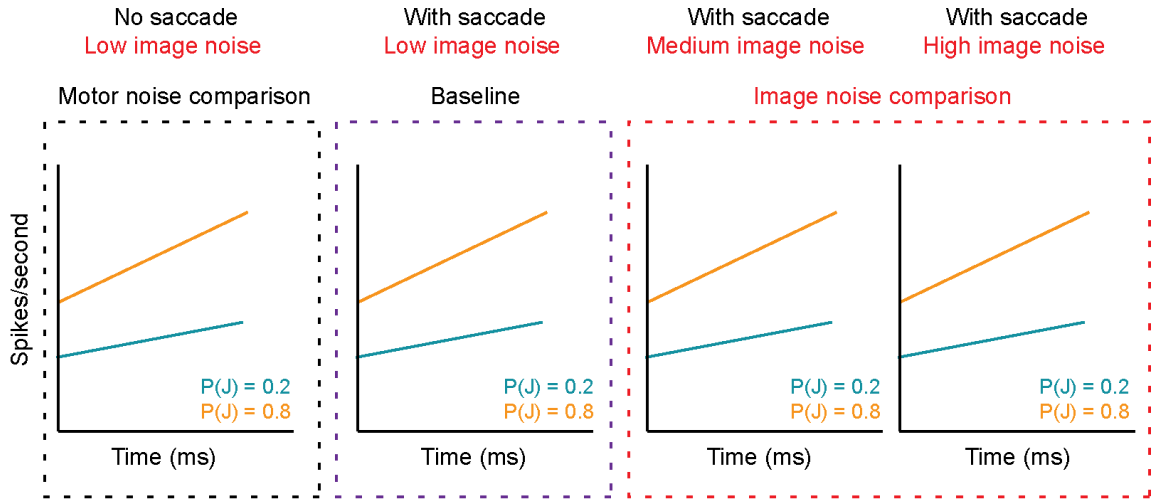

Figure 55. Predicted modulation of prior selectivity if FEF activity reflects neither Bayesian nor Discriminative behavior.

2) Only motor noise but not image noise, suggesting that FEF neurons selectively support a Bayesian but not Discriminative model (Fig. 56):

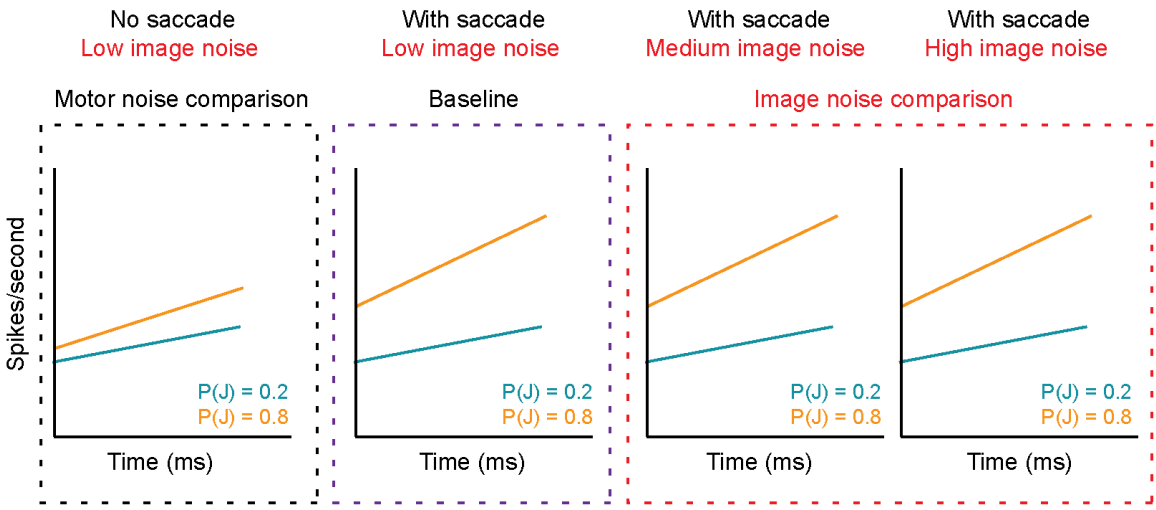

Figure 56. Predicted modulation of prior selectivity if FEF activity reflects Bayesian but not Discriminative behavior.

3) Only image noise but not motor noise, suggesting that FEF neurons selectively support a Discriminative but not Bayesian model (Fig. 57):

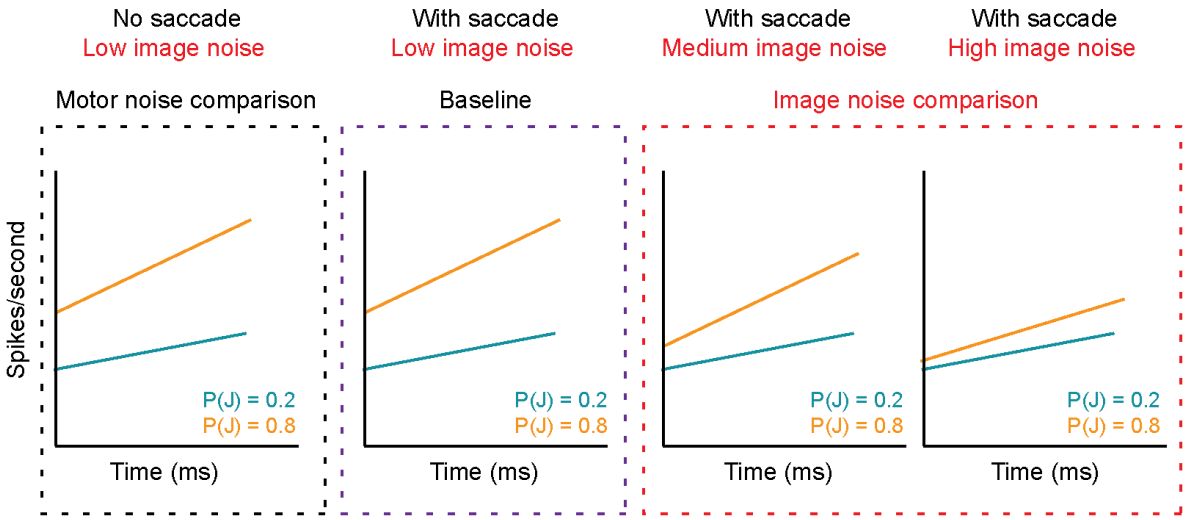

Figure 57. Predicted modulation of prior selectivity if FEF activity reflects Discriminative but not behavior.

4) Both motor noise and image noise, suggesting that FEF neurons either support both models or inherit their outputs and signal the final behavioral output (Fig. 58):

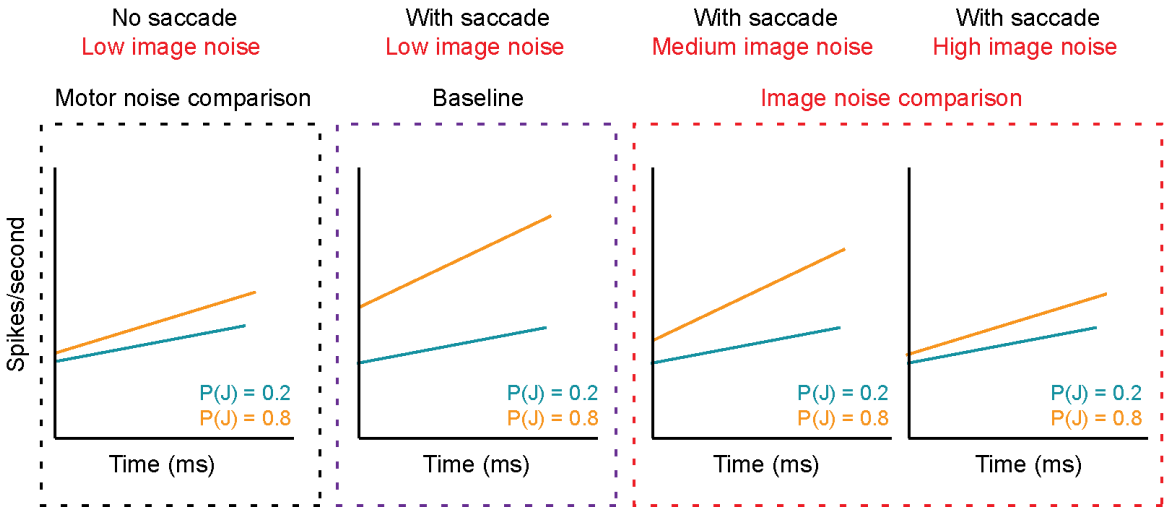

Figure 58. Predicted modulation of prior selectivity if FEF activity reflects both Bayesian and Discriminative behavior.

We tested these predictions for FEF neural activity in the reafferent epoch and the response epoch (fixation epoch excluded due to lack of prior selectivity), as described next.

4.4.3.2 Reafferent epoch

Fig. 59 shows spike density functions (convolved with a Gaussian kernel with σ = 20ms) in the reafferent epoch for the population of FEF neurons that had significant, positive effect size prior selectivity for this epoch. The order of the panels matches the rows in Fig. 55-58. From left to right, the panels depict firing rates for the no saccade (low motor noise, low image noise), with saccade (high motor noise) but low image

noise, with saccade (high motor noise) and medium image noise, and with saccade (high motor noise) and high image noise conditions.

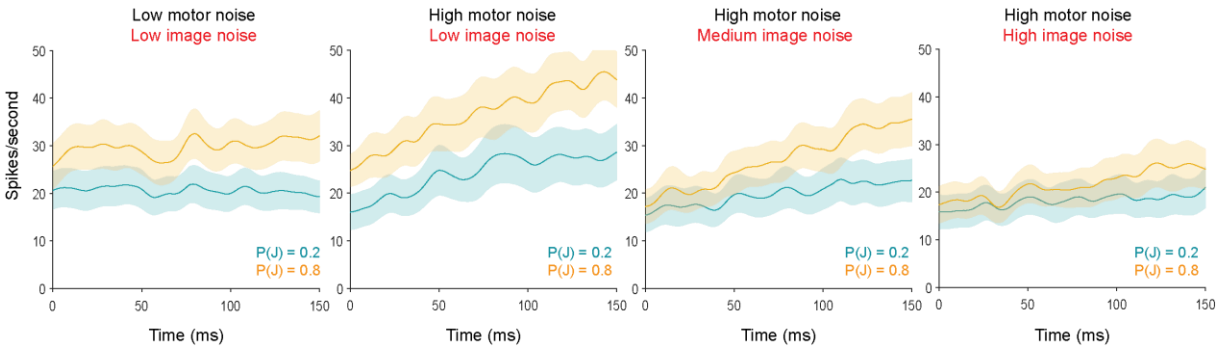

Figure 59. Spike density functions in the reafferent epoch (n = 24 cells).

Qualitatively, we see that prior selectivity in the spike density functions (the difference between orange and teal lines) collapses with greater image noise (looking rightward from the 2nd panel, baseline condition) but not with lower motor noise (looking leftward). That is, the results are consistent with the prediction that FEF neurons selectively support a Discriminative but not Bayesian model (cf. Fig. 57).

To quantify the results, we asked if the pattern observed in the spike density functions predicts behavioral prior use on a session-by-session basis. We assessed the correlation between the modulation of neuronal prior selectivity and behavioral prior use with the two types of noise, motor and image, for each session. The exact same trials, pooled across displacements, were used to compute the neuronal and behavioral measures. For both measures, neuronal prior selectivity and behavioral prior use, we

quantified the noise modulation by normalizing the measure in the noise condition with its value in the baseline condition. To quantify the magnitude of neuronal prior selectivity, we used the Wilcoxon effect size between firing rates in P(J) = 0.8 and P(J) = 0.2 conditions. Thus, the normalized effect size in each condition was:

$$Normalized\ effect\ size = \frac{Effect\ size}{Effect\ size\ in\ baseline\ condition} \tag{4.2}$$

For behavioral prior use, we used the difference in proportion "jumped" responses between the P(J) = 0.8 and P(J) = 0.2 conditions and normalized it in the same way as for the firing rate effect size. Additionally, we excluded sessions where the normalized firing rate in the given noise condition was greater than 3 Median Absolute Deviations (MAD) away from the median before assessing correlations (Wilcox, 2011; Leys et al., 2013). This excluded 3 cells in the low motor noise condition, none in the medium image noise condition, and 1 cell in the high image noise condition. We also performed the correlation analyses without excluding outliers and the results were the same. As suggested by the spike density functions, if FEF activity selectively supports behavior that is explained by a Discriminative, but not Bayesian, model, we would expect to see a positive correlation in the medium and high image noise conditions (Fig. 60) but not in the low motor noise condition (Fig. 61).

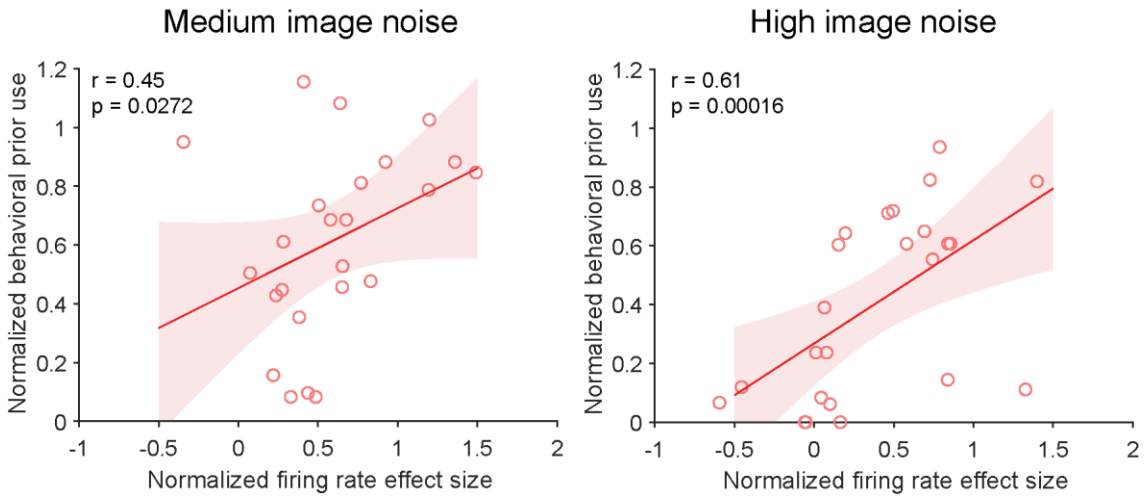

Figure 60. Comparison of normalized firing rate effect size across priors in the reafferent epoch with normalized image noise behavior on a session-by-session basis.

That is what we found. There was a statistically significant positive correlation between normalized firing rate effect sizes in the medium (Spearman r = 0.45, p = 0.0272) and high image noise (r = 0.61, p = 0.0016) conditions, but not in the low motor noise condition (r = 0.06, p = 0.79).

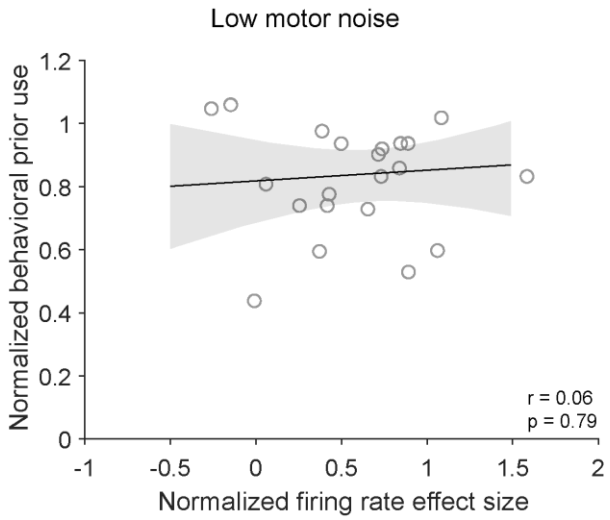

Figure 61. Comparison of normalized firing rate effect size across priors in the reafferent epoch with normalized motor noise behavior on a session-by-session basis.

We then considered that for all conditions in the image noise experiment, the probe was initially placed near the center of the *future field*, i.e., the cell's receptive field after the saccade, and displaced to the center of the post-saccadic receptive field during the saccade. Thus, the first time that the probe occupied the center of the receptive field of the neuron was after the saccade, i.e., in the reafferent period. For the low motor noise condition without a saccade, however, the probe initially appeared near the center of the cell's receptive field. The displacements used in the experiment were typically small (with the most frequent displacements falling in the 0-1° range) and FEF receptive fields are typically much larger than that. Thus, although the probe appeared at the true center of the receptive field *after* it was displaced, i.e., during the defined reafferent period, its first appearance in the receptive field was during the *probe* epoch. This meant that priors

might interact with this initial visual response in no-saccade trials. Therefore, we performed the same correlation analysis using firing rate effect sizes in the probe epoch normalized to baseline firing rate effect sizes in the reafferent epoch (Fig. 62). No cells were excluded as outliers in this epoch. We found no significant correlation between normalized firing rate effect sizes in the probe epoch and behavioral prior use either in the low motor noise condition ($r = -0.09$, $p = 0.68$).

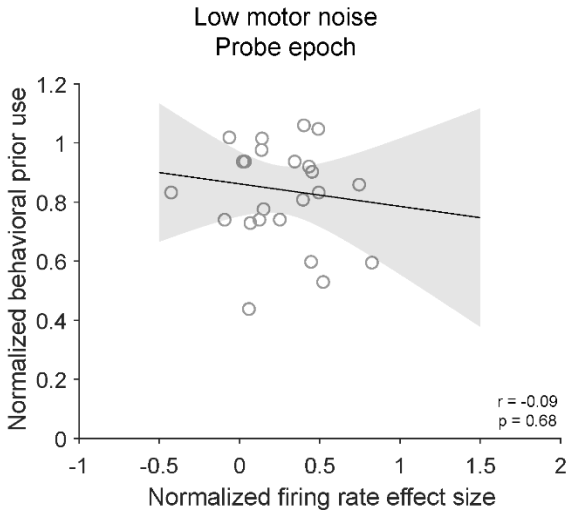

Figure 62. Comparison of normalized firing rate effect size across priors in the probe epoch with normalized motor noise behavior on a session-by-session basis.

Finally, we performed the same analyses for non-normalized data in the same noise conditions to confirm that the selective modulation of firing rates with behavioral prior use was not an artifact of the normalization process (Fig. 63). One cell each was

excluded as an outlier in the medium and high image noise conditions. No cells were excluded in the low motor noise conditions.

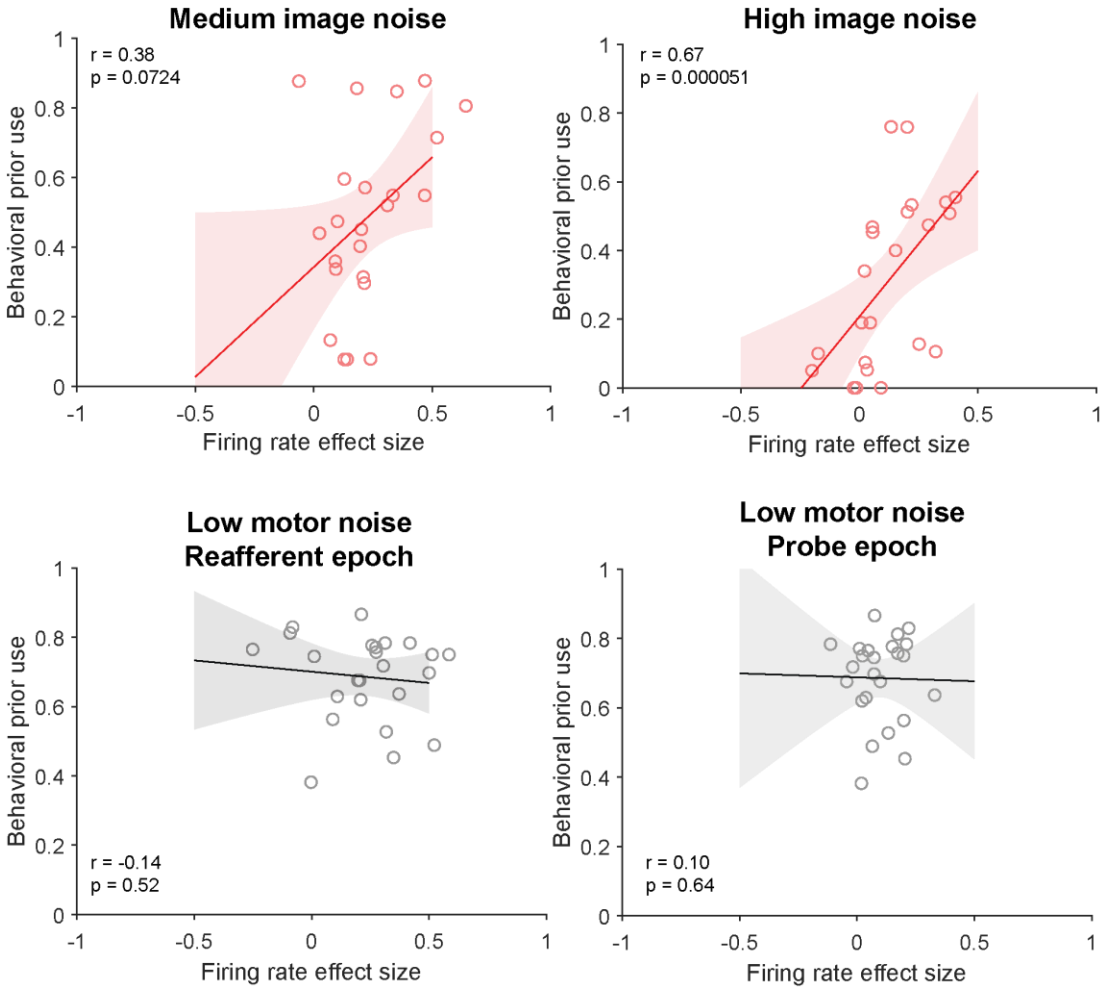

Figure 63. Comparison of non-normalized firing rate effect sizes across priors in the reafferent epoch with behavioral prior use.

We found the same general results. Firing rates showed a positive correlation with behavioral prior use that that did not quite reach significance in the medium image

noise condition (r = 0.38, p = 0.0724) but was highly significant in the high image noise condition (r = 0.67, p = 0.00051). In contrast, there were no hints of correlation at all between neural activity and behavior in the low motor noise condition in the reafferent epoch (r = -0.14, p = 0.52) or in the probe epoch (r = 0.10, p = 0.64).

Overall, the pattern of FEF neural activity during the reafferent period, which is when the cells exhibit visual responses to the probe, was clear. On a session-by-session basis, the reafferent activity predicted Discriminative prior use in the image noise experiment but not Bayesian prior use in the motor noise experiment.

4.4.3.3 Response epoch

We performed the same analyses as for the reafferent epoch on cells that had a significant difference and positive effect size between firing rates across the high (0.8) and low (0.2) priors in the *response* epoch of the task (n = 32 neurons). This epoch is well after the initial visual response and around the time of the response saccade to the probe in the receptive field. Consistent with results in the reafferent epoch, we qualitatively found that prior selectivity in the spike density functions decreased with increasing image noise more than with lower motor noise (Fig. 64).

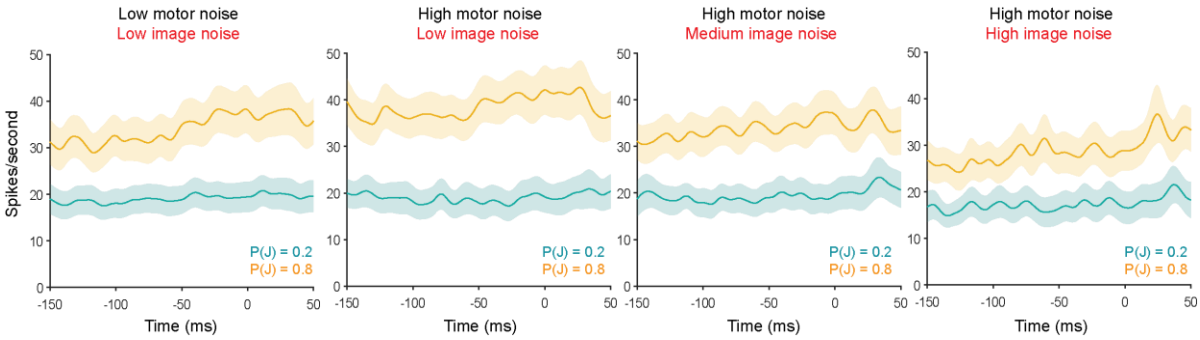

Figure 64. Spike density functions in the response epoch (n = 32 cells).

Additionally, there was a significant positive correlation between normalized firing rate effect sizes and normalized prior use in the medium (n = 30 cells after excluding outliers; r = 0.68, p = 0.00013) and high (n = 32 cells; r = 0.65, p = 0.000047) image noise conditions (Fig. 65).

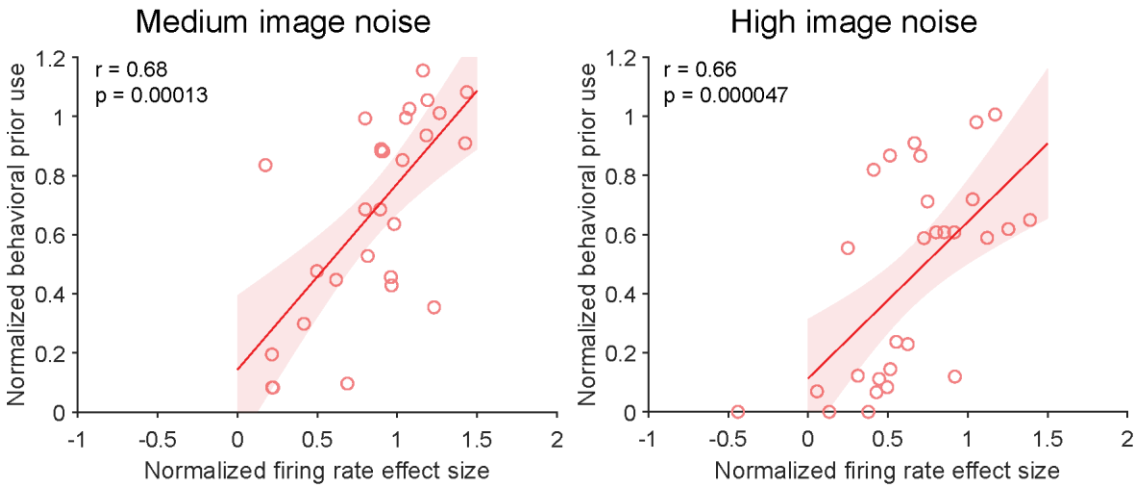

Figure 65. Comparison of normalized firing rate effect size across priors in the response epoch with normalized image noise behavior on a session-by-session basis.

There was no significant correlation, however, in the low motor noise condition (n = 28 cells; r = -0.04, p = 0.83) (Fig. 66).

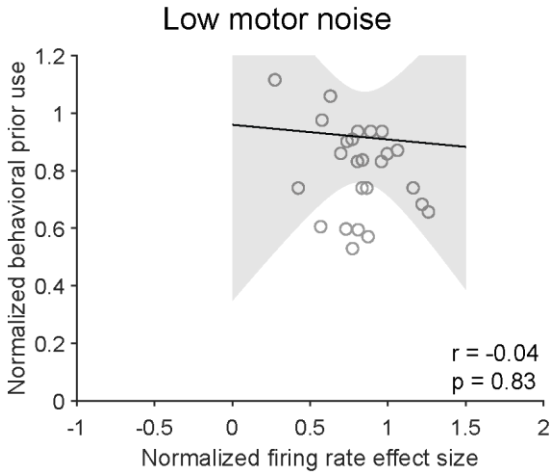

Figure 66. Comparison of normalized firing rate effect size across priors in the response epoch with normalized image noise behavior on a session-by-session basis.

Finally, consistent with results in the reafferent epoch, there was also a significant positive correlation between raw, non-normalized firing rates and prior use in the medium (r = 0.45, p = 0.01) and high (r = 0.48, p = 0.0064) image noise, but not for low motor noise (r = 0.13, p = 0.48) (Fig. 67). No cells were excluded as outliers in any condition.

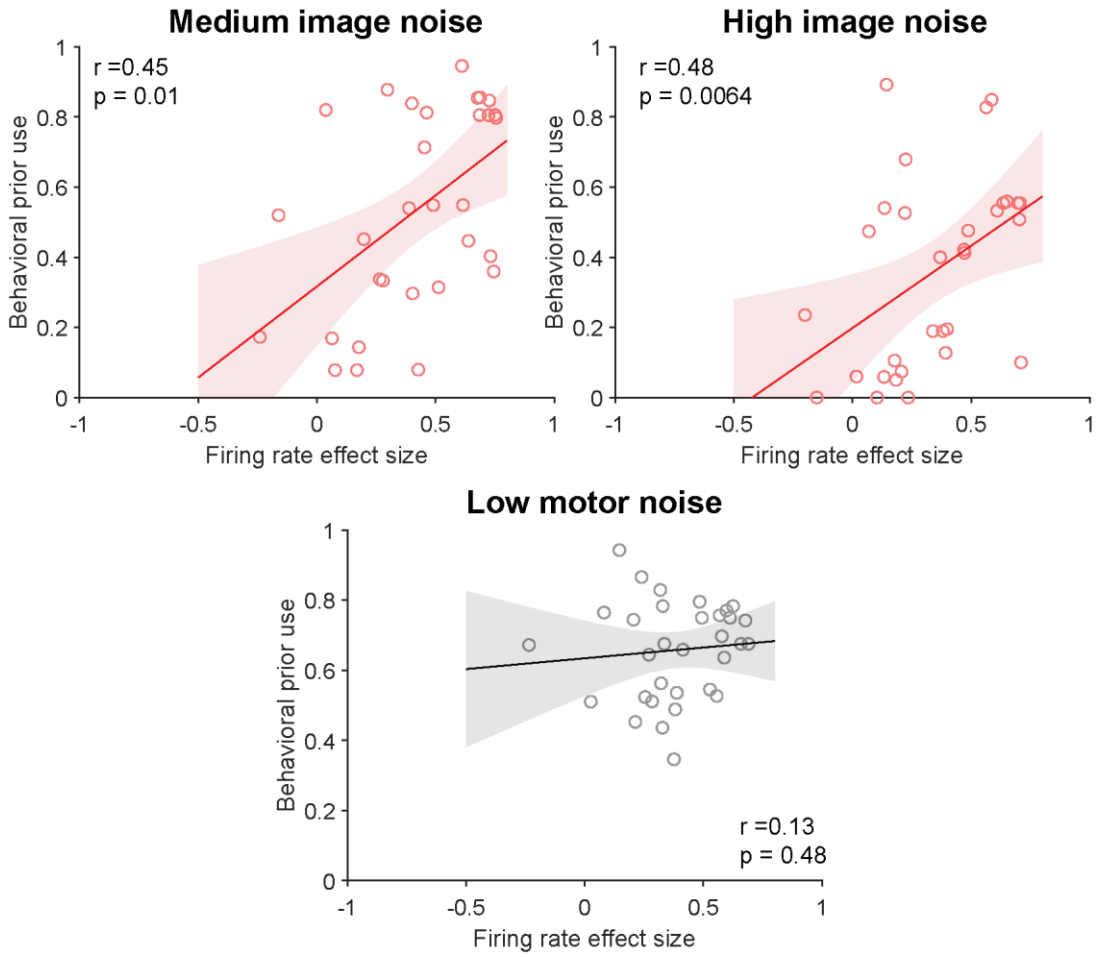

Figure 67. Comparison of non-normalized firing rate effect sizes across priors in the response epoch with behavioral prior use.

Overall, consistent with results in the reafferent epoch, firing rates in the response epoch of the SSD task also predicted Discriminative prior use in the image noise experiment but not Bayesian prior use in the motor noise experiment.

4.5 Discussion

The results presented in this chapter consisted of behavioral data collected during the FEF recording sessions, patterns of neural activity during the various epochs and conditions of the SSD task, and session-by-session comparisons of neural prior-selectivity with behavioral prior use. First, the behavioral data replicated our psychophysical and modeling results in Chapters 2 and 3. Specifically, prior use in the task increased with motor-induced noise as predicted by a Bayesian model and decreased with image noise as predicted by the Discriminative model described in Chapter 3. Second, neuronal activity in FEF signaled prior expectations. When compared to Bayesian and Discriminative behavior in the motor- and image-noise tasks respectively, we saw a clear distinction: FEF activity predicted Discriminative behavior on a session-by-session basis but not Bayesian behavior. These results situate FEF along a neural pathway which selectively supports a Discriminative model, implying that the Bayesian computation is implemented elsewhere in the system.

The scope of the analyses must be considered before interpreting the results. The analyses performed in this chapter were designed to address the main question: does FEF neuronal activity support a Bayesian or Discriminative model? Therefore, we have not discussed in detail the other factors that might drive the responses of these neurons. Understanding the total response profile of these neurons, however, would provide more insight into the computations they support. For example, how does the prior

selectivity of the neurons interact with their inherent visuomotor selectivity (identified in the receptive field mapping tasks)? Does prior activity, and its modulation by uncertainty, build on the cell's more fundamental functions or are they independent of one another? A richer characterization of this interaction constitutes one future direction for this work. Another major future direction is identifying a comprehensive panel of task variables to which the neurons respond. For example, to determine behavioral prior use (i.e., the difference in proportion "jumped" responses across priors), we included trials in which the animals responded "jumped" with a saccade into the cell's receptive field as well as trials on which they did not make a response saccade (reporting "did not jump"). In the high prior condition, they responded "jumped" more and, accordingly, made saccades more often. Thus, instead of *prior use*, the difference in firing rates across priors in the response period might have simply correlated with whether the animals made a saccade or not on that trial (Bruce and Goldberg, 1985; Bruce et al., 1985). Similarly, previous work showed that FEF neurons are responsive to the magnitude of object displacement across saccades in the reafferent epoch of this task (Crapse and Sommer, 2012). Since we pooled across displacements for our analyses, more trials with larger displacements are included in the high prior condition. Therefore, it is possible that the difference in firing rates across priors are driven in part by the difference in displacement magnitudes. Both possibilities are partially likely, in that we expect response epoch activity will also predict the saccade and its kinematics, and that

reafferent epoch activity will also reflect the magnitude of the displacement. However, the potential confounds are held in common across both the image- and motor-noise experiments. That is, if cells simply reflected the displacement or the response, their activity would be similarly (un)correlated with both Discriminative and Bayesian behavior in the image- and motor noise conditions, respectively. Therefore, the interpretation that their activity selectively supports a Discriminative but not Bayesian model is not diminished by other factors. Nevertheless, the next step is to quantitatively characterize the relative contribution of various task parameters and their interactions to the firing rates of neurons using a linear mixed-effects regression model.

In addition to the direct evidence for Discriminative over Bayesian processing in FEF provided by the neuron-behavior correlations, a second line of indirect evidence comes from comparing the pattern of activity observed with the assumptions made for the Bayesian and Perceptron models in Chapters 2 and 3. Specifically, consider the assumptions made about how the prior is represented by each model. In the Perceptron model, knowledge of the two categories – "jumped" and "did not jump" – is stored in the weights between inputs and outputs. Recall that to account for early-training data, we assumed in Chapter 3 that these weights are modified separately for each prior and thus assumed a distinct input layer for each prior (Chapter 3, Fig 44). In other words, each input layer represented compound prior + displacement units, leading the prior to be represented in an inherently intertwined manner with the sensory input about the

displacement. A change in sensory inputs would lead to a change in the manifestation of prior use with prior curves collapsing towards one another. In the Bayesian model, by contrast, the prior determines the location of the criterion (vertical lines in Chapter 2, Fig. 14). Although the final output combines the Criterion with the sensory input (black distributions in Chapter 2, Fig. 14), the location of the criterion itself is *independent* of the nature of the sensory input. Thus, if FEF activity signaled a Bayesian prior, it would have to be independent of the sensory input. This would predict neural prior selectivity in the fixation epoch when the animal has information about the prior but not yet the sensory input. In our study, a few FEF neurons showed such selectivity in the fixation epoch, but it was rare. The majority of FEF neurons signaled the prior in the reafferent and response epochs when the animal had already viewed the sensory input about displacement (in the reafferent epoch) and mapped it onto a categorical response (in the response epoch). This pattern is consistent with the assumptions made for the Perceptron model. Alternatively, even if a Bayesian prior is represented "just in time" later in a trial, to be sensory input-independent, it would have to remain unmodulated by noise (as depicted in Fig. 55). Of course, this is not the case. Taken together, the input-*dependence* of FEF prior selectivity additionally corroborates the finding that an FEF-inclusive pathway selectively supports Discriminative processing within the saccadic system.

These findings pose the question: where in the saccadic system is *Bayesian* processing in the categorical task supported? One possibility is the Superior Colliculus (SC), another critical sensorimotor node in the saccadic system. As with FEF, the SC generates or receives both saccade-related and visual information (Goldberg and Wurtz, 1972; Mays and Sparks, 1980; Wurtz and Optican, 1994). In addition, recall that we assumed in our modeling that the mechanism of motor-induced noise in our task was visual degradation caused by saccadic suppression. Neurons in the superficial layers of SC show suppressed visual activity around the time of saccades, consistent with the time course of saccadic suppression. Finally, recent work has shown that neurons in SC also signal a sensory input-independent prior (Crapse and Basso, 2018; Jun et al., 2021). Thus, the SC has all the pieces in place necessary to support a Bayesian model.

In conclusion, the results demonstrated a division of labor between Bayesian and Discriminative models for perception across saccades at the computational and neural levels. A detailed synthesis of this dissociation follows in Chapter 5.

5. General Discussion

5.1 Summary

The studies in this dissertation started with the hypothesis that Bayesian inference is a useful framework for understanding active perception. Specifically, we asked if priors about the environment are used in a Bayesian manner to compensate for sensory uncertainty for the perception of visual continuity across saccades. In Chapter 2, we found that human participants were Bayesian when providing *continuous* reports of object displacement across saccades. For categorical judgments, however, they were Anti-Bayesian. Based on an analysis of saccadic endpoint scatter in these experiments, we reasoned that the sensory manipulation used in our human experiments mainly imposed external sensory uncertainty. Thus, we hypothesized that Anti-Bayesian prior use for categorical judgments might be driven by external sensory uncertainty rather than motor-driven uncertainty. In Chapter 3, we directly tested this hypothesis in two rhesus macaques. The results supported the hypothesis. In other words, categorical prior use for active perception can be summarized as being positively modulated by motor-induced noise and negatively modulated by external, image noise. A Discriminative model that used a Perceptron learning rule accounted for the Anti-Bayesian prior use as well as otherwise unexplained aspects of the results. We then recorded from single neurons in FEF while animals performed the categorical image and motor noise

experiments with all trial types interleaved. We found that the FEF neurons predicted the animals' Discriminative prior use in the image noise trials but not their Bayesian prior use in the motor noise trials. The results have several implications for the study of perception.

5.2 What determines the use of Bayesian models for perception?

First, these studies add theoretical nuance to the debates about the importance and implications of Bayesian processing for perception. Bayesian models have been successful in explaining behavior in a wide range of sensorimotor and cognitive tasks (e.g., Jacobs, 1999; Ernst and Banks, 2002; Battaglia et al., 2003; Fetsch et al, 2012; Weiss et al, 2002; Kording and Wolpert, 2004; Darlington et al., 2017; Jazayeri and Shadlen, 2010; Girschick et al, 2011). Given their widespread success, Bayesian inference has been postulated to be a near-ubiquitous and fundamental principle of brain function. This notion is encapsulated by the Bayesian Brain Hypothesis (Knill and Richards, 1996; Knill and Pouget, 2004; Friston, 2012; Pouget et al., 2013). Strong interpretations of the Bayesian Brain hypothesis assume that Bayesian models make predictions for *how* the brain solves the task, i.e., that the brain literally represents probability distributions and computes posteriors (e.g., Beck et al., 2008; Ma et al., 2006; Deneve, 2008a, 2008b). A weaker interpretation is that the Bayes optimal solution to a task only reflects the evolutionary goals or the *teleology* of the system at hand (Griffiths et al., 2012). Thus,

Bayesian models make no mechanistic predictions (Jones and Love, 2011). On the other hand, limitations of the Bayesian framework for the study of sensorimotor and perceptual systems have been highlighted (Laquitaine and Gardner, 2018; Rahnev and Denison, 2018), and stronger critics of the Bayesian framework have argued that it is often so flexible in how parameters are modeled that it precludes falsifiability, limiting its relevance (Bowers and Davis, 2011; Danks and Eberhardt, 2011; Glymour, 2011; Rogers and Seidenberg, 2011). Despite strong views on both sides, there have been recent efforts to acknowledge the contribution of both Bayesian and non-Bayesian models, including various types of Discriminative models, to perception (Gardner, 2019; Sohn and Jazayeri, 2021; DiCarlo et al., 2021). Yet, the field lacks a synthesized, theoretical account of when Bayesian models are used and when they are not.

The current results provide one possible account, as first noted in Chapter 2. To briefly recap, we found that *continuous* estimates of object displacement across saccades are Bayesian. For categorical estimates, however, the trade-off between priors and motor-driven noise was Bayesian but not the trade-off between priors and sensory-driven noise. This pattern of results posed the question: what determines the use of Bayesian models for active perception? One possible explanation that is consistent with the results is that *computations that follow the inherent neural organization of a sensory system can implement the Bayes optimal solution for stimulus estimation*. Continuous estimates of object location (and their displacement, a linear derivative of location) may be the

default mode of the visual system. For example, spatial receptive fields are continuous all the way from the retina to higher order visual areas (Colby et al., 1988; Engel et al., 1997; Golomb and Kanwisher, 2012a; Arcaro and Livingstone, 2017) and maintain their retinotopic properties even across eye movements (Golomb and Kanwisher, 2012b; Zimmerman et al., 2013). Moreover, our lab found that neurons in the frontal eye field use continuous tuning to represent object displacements across saccades (Crapse and Sommer, 2012), the stimulus quantity we probed directly. Requiring the system to categorize this continuous information might lead it to adopt an alternative, Discriminative strategy that is revealed when probing it with image noise. Saccadic suppression, however, is 1) inherent to the system in that it is a necessary component of visual processing across saccades, and 2) an essentially binary way in which saccades influence visual processing, in that there is saccadic suppression when a saccade is made and not otherwise. Thus, if inherent, default sensory computations are privileged for Bayesian inference, then priors should compensate for saccadic suppression in a Bayesian manner, even when the report is categorical. The results in Chapter 3 supported this hypothesis. We found that priors were used to compensate for intrinsically generated, saccade-associated noise but not external noise added to the visual input.

Taken together, these results warrant further investigation into the link between the inherent neural organization of sensorimotor systems and Bayesian behavior.

Establishing a link between them would allow our understanding of each to constrain and advance our understanding of the other.

5.3 A division of labor for active perception: Revisited through the lens of Bayesian and Discriminative models

A second line of evidence linking Bayesian processing with the default organization and functioning of a neural system comes from revisiting our neural results from Chapter 4 given our understanding of how saccades influence visual processing. Recall that we noted in Chapter 1 that the visual system compensates for the effects of saccades on vision in two dissociable ways. First, it corrects for nuisance factors such as blur caused by the movement via saccadic suppression (Zuber and Stark, 1966; Bridgeman et al., 1975). Second, it stitches together visual snapshots from before and after each saccade to create a coherent, continuous percept of the world (reviewed in Wurtz, 2008; 2018). For both functions, the system uses a copy of the movement command called corollary discharge. There are distinct corollary discharge pathways that contribute to the two functions, respectively (Fig. 68).

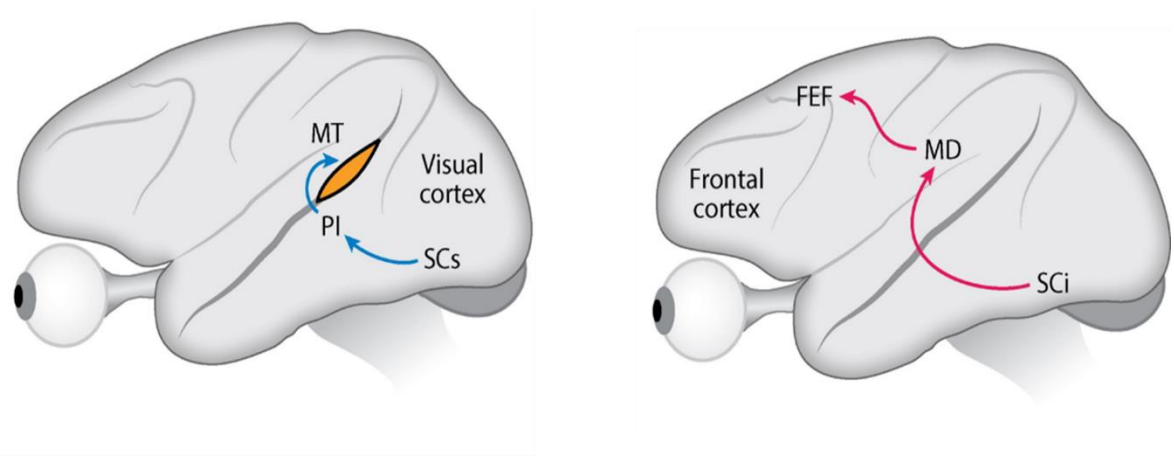

Figure 68. Distinct corollary discharge pathways in the saccadic system. Reproduced from Wurtz, R. H. (2018). Corollary Discharge Contributions to Perceptual Continuity Across Saccades. *Annual Review of Vision Science*, *4*, 215–237, Fig. 2b (right) and 2d (left)

First, a pathway from visually responsive neurons in superficial SC (SCs) to the thalamic Pulvinar nucleus (PI) to the Middle Temporal visual area (MT), which processes motion information, has been shown to carry information about saccadic suppression (Fig. 68, left) (Berman and Wurtz, 2010). While this pathway does not transmit a corollary discharge signal itself, it carries suppressed visual activity around the time of saccades (Goldberg and Wurtz, 1972; Robinson and Wurtz, 1976) and inactivation of the pathway causes a decrease in the suppression of MT activity (Berman et al., 2017). The suppression of activity in visually responsive neurons in SCs is postulated to arise within the SC itself, using corollary discharge signals that originate in the saccade-generating intermediate layers of the SC (SCi) (Wurtz, 2018). Since we find that saccadic suppression (motor-driven noise) is compensated for with categorical

priors in a Bayesian manner (Chapter 3), one candidate neural system for the integration of priors with motor-driven noise is the SCs-PI-MT pathway.

A first step towards this integration would be identifying the neuronal correlates of priors along this pathway. Additionally, recall that we noted in Chapter 4, that the implementation of a Bayesian model would require a sensory input-independent prior to be represented. Indeed, as also noted in Chapter 4, recent work has shown that neurons in the SCi signal such an input-independent prior (Crapse and Basso, 2018), and such priors play a causal role in decision-making (Jun et al., 2021). Thus, the SC has both components of the Bayesian computation, categorical priors and motor-driven uncertainty in the form of saccadic suppression.

An important caveat is that the neural signatures of saccadic suppression have been observed in a variety of regions in the visual system including MT, the Medial Superior Temporal area (MST), the Ventral Intraparietal region (VIP) and the Lateral Intraparietal region (LIP) (Bremmer et al., 2009). Additionally, visual processing starting in the retina itself is shown to be another source of saccadic suppression (Idrees et al., 2020; Castet et al., 2001). Thus, in addition to the SC-PI-MT pathway for saccadic suppression, there are other candidate systems which may play a role in the Bayesian behavior observed in our task.

The second function of corollary discharge across saccades is to maintain visual *continuity* across saccades. For this, a pathway from SCi to Mediodorsal thalamus (MD)

to FEF, discovered by Sommer and Wurtz (2002, 2004a, 2004b, 2006, 2008) has been shown to play a key role (Fig. 68, right). Corollary discharge, via its contribution to the remapping of visual receptive fields in FEF (Sommer and Wurtz 2006; 2008), can be used to predict the consequence of a saccade. This prediction can be compared to incoming visual inputs (Umeno and Goldberg, 1997) to judge whether there was a discrepancy.

As noted in Chapters 2 and 3, however, the FEF may be preferentially suited to supporting the perception of *continuous* displacements across saccades. Visual receptive fields in FEF maintain their retinal coordinates, combining information from across the visual system to represent parts of the scene relative to the fovea, and are thus continuous in space. The corollary discharge signal conveys continuous information about the saccade vector, which is then used to remap receptive fields for transsaccadic continuity. Remapped receptive fields are still retinotopic and are thus continuous in space. Critically, Crapse and Sommer (2012) found that FEF neurons are tuned continuously for the magnitude of displacements across saccades. Thus, if it is true that Bayesian processing requires and leverages the inherent organization of a neural system, then FEF neurons are most likely to support a *continuous* Bayesian model. This hypothesis is consistent with the finding that FEF activity supports the implementation of a continuous Bayesian model for a different type of oculomotor function, smooth pursuit of moving visual objects (Darlington et al., 2018).

If the FEF's default mode is continuous, then like with the behavior, we'd expect

to see that it categorizes that continuous information in a Discriminative manner to meet the demands of a two alternative forced choice task. Therefore, when probed under a categorical regime in this task, neural activity would reflect a Discriminative model. That is what we found in Chapter 4.

To summarize, the results in Chapter 3 showed that Bayesian and Discriminative processing for the categorical judgment of object displacement across saccades are dissociable from one another. This dissociation maps on to a dissociation between two functions of corollary discharge for perception across saccades, saccadic suppression, and visual continuity. Priors are used to compensate for saccadic suppression in a Bayesian manner. Further, Bayesian-like input-independent priors are observed in a node along the neural pathway for saccadic suppression. Visual continuity is thought to be primarily supported by a second corollary discharge pathway from SC-MD-FEF. This pathway is well-suited to supporting continuous computations suggesting that the default mode for maintaining visual continuity (and perceiving object displacement) across saccades may be continuous processing. When probed in a categorical task, this would manifest as the Discriminative categorization of continuous displacement values, as observed in our neural experiments in Chapter 4. These results, therefore, provide a second line of evidence linking Bayesian processing to the inherent organization of the neural system underlying active perception across saccades.

5.4 Future Directions

In addition to the predictions for Bayesian behavior that arise from the proposed division of labor for Bayesian vs. Discriminative models within the oculomotor system (Section 5.3), the results in this dissertation generate additional questions for how Discriminative models are implemented, and how the two combine for behavior. The questions include:

1) What is the underlying organization and physiological circuitry of the larger Discriminative model system that includes FEF? For one thing, we do not expect that FEF neurons instantiate a Perceptron-like model in isolation. Moreover, our conceptions of the connections, weights, and learning of discrimination between prior-modulated patterns of evidence, while sufficient to simulate behavioral data, were relatively simplistic. An important future direction would be to repeat the experiments of Chapter 4 in other brain areas to identify components of the Perceptron model for this task. For example, the FEF, in concert with other cortical and subcortical regions with which it is densely interconnected such as LIP and the SC (Stanton et al., 1995, Sommer and Wurtz, 2001; Lynch and Tian, 2006), might integrate signals from across the visual system and prefrontal cortex to signal the context of the priors. Thus, comparing the modulation of prior signals with sensory uncertainty across these regions would be a first step. Similarly, the basal ganglia (Jin et al., 2009; Murray and Escola, 2020) and the cerebellum (Marr, 1969; Albus, 1971) have neural architectures that are well-suited to implementing

perceptron-like learning rules, and the FEF connects to both (Künzle and Akert, 1977; Leichnetz, 2001; Huerta et al., 1986). Understanding the integrated physiological basis of the Discriminative model across the brain would be a step in the direction of advancing from descriptions of fundamental sensory responses to a network-based, algorithmic approach for understanding perception.

2) Are the neural signals in FEF that reflect Discriminative model processing causal to the behavior? Our studies revealed clear correlates of behavior in our tasks, in that neural representations of priors collapsed with sensory uncertainty just as behavioral use of priors did. Further, the neural collapse of priors predicted behavioral prior use on a session-by-session basis. Demonstrating this session-by-session link between Discriminative behavior and the activity of FEF neurons is an important first step, but it does not establish that FEF neurons are *causing* the behavior. For that we will need follow-up experiments that perturb FEF activity to observe the effects on behavior. One way to do this is to transiently inactivate FEF pharmacologically. Alternatively, we could optogenetically suppress activity specifically during the three epochs we tested: the fixation, reafferent, and response periods. From our neural data showing that priors and visual inputs interact during the reafferent and response periods of the SSD task, we predict that transient inhibition would disrupt behavioral prior use during those periods but not in the fixation period. That would test for necessity of the signals for the behavior. Optogenetic activation or electrical microstimulation would provide a

complementary test of sufficiency of the neural activity for supporting the behavior.

Finally, the current work sets the stage for investigating how the brain uses Bayesian vs. Discriminative models in other perceptual systems. Given our understanding of how the two forms of models compare at the behavioral, computational, and neurophysiological levels for one sensorimotor system, visual processing across saccades, we are better positioned to formulate and predict the outcomes of studies in other perceptual domains.

5.5 Conclusions

Overall, the work in this dissertation shows that the selection of Bayesian vs. Discriminative models for active visual perception across saccades depends on both task requirements (continuous vs. categorical) and the source of uncertainty (motor vs. sensory) for active perception. Further, the mechanisms underlying both models are dissociable at the computational and neural levels. Taken together, these results set the stage for understanding how Discriminative and Bayesian models interact for active perception.

Finally, the current results may generalize beyond the oculomotor system. Accounting for self-movement is a problem that spans almost all sensory modalities, and the integration of movement and sensory signals have been observed widely in the brain (e.g., Niell and Stryker, 2010; Keller et al., 2012; Schneider and Mooney, 2018;

Crapse and Sommer, 2008). Understanding the relative contributions of Bayesian and Discriminative computations may guide future studies on how expectations, self-movement, and external sensory information combine for active perception in healthy and disordered states.

References

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

Albus, J. S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, *10*(1), 25–61. https://doi.org/10.1016/0025-5564(71)90051-4

Arcaro, M. J., & Livingstone, M. S. (2017). A hierarchical, retinotopic proto-organization of the primate visual system at birth. *ELife*, 6, e26196.

Awater, H., & Lappe, M. (2006). Mislocalization of Perceived Saccade Target Position Induced by Perisaccadic Visual Stimulation. *Journal of Neuroscience*, *26*(1), 12–20. https://doi.org/10.1523/JNEUROSCI.2407-05.2006

Bair, W., & O'keefe, L. P. (1998). The influence of fixational eye movements on the response of neurons in area MT of the macaque. *Visual Neuroscience*, *15*(4), 779–786. https://doi.org/10.1017/S0952523898154160

Bajcsy, R. (1988). Active perception. *Proceedings of the IEEE*, *76*(8), 966–1005. https://doi.org/10.1109/5.5968

Bansal, S., Jayet Bray, L. C., Peterson, M. S., & Joiner, W. M. (2015). The effect of saccade metrics on the corollary discharge contribution to perceived eye location. *Journal of Neurophysiology*, *113*(9), 3312–3322. https://doi.org/10.1152/jn.00771.2014

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. *Sensory communication*, *1*(01).

Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *20*(7), 1391–1397. https://doi.org/10.1364/josaa.20.001391

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

Berman, R. A., & Wurtz, R. H. (2010). Functional identification of a pulvinar path from superior colliculus to cortical area MT. *The Journal of Neuroscience: The Official*

Journal of the Society for Neuroscience, *30*(18), 6342–6354. https://doi.org/10.1523/JNEUROSCI.6176-09.2010

Berman, R. A., Cavanaugh, J., McAlonan, K., & Wurtz, R. H. (2017). A circuit for saccadic suppression in the primate brain. *Journal of Neurophysiology*, *117*(4), 1720–1735. https://doi.org/10.1152/jn.00679.2016

Bitzer, S., Park, H., Blankenburg, F., Kiebel, S. J. (2014) Perceptual decision making: Drift-diffusion model is equivalent to a Bayesian model. *Frontiers in Human Neuroscience*, 8: 102. DOI: https://doi.org/10.3389/fnhum.2014.00102

Bowers, J. S., Davis, C. J. (2011) More varieties of Bayesian theories, but no enlightenment. *Behavioral and Brain Sciences*, 34(4): 193–194. DOI: https://doi.org/10.1017/S0140525X11000227

Bremmer, F., Kubischik, M., Hoffmann, K.-P., & Krekelberg, B. (2009). Neural Dynamics of Saccadic Suppression. *Journal of Neuroscience*, *29*(40), 12374–12383. https://doi.org/10.1523/JNEUROSCI.2908-09.2009

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

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*, *13*(1), 87–100. https://doi.org/10.1017/s095252380000715x

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

Bruce, C. J., Goldberg, M. E., Bushnell, M. C., & Stanton, G. B. (1985). Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements. *Journal of Neurophysiology*, *54*(3), 714–734. https://doi.org/10.1152/jn.1985.54.3.714

Castet, E., Jeanjean, S., & Masson, G. S. (2001). 'Saccadic suppression'- no need for an active extra-retinal mechanism. *Trends in Neurosciences*, *24*(6), 316–318. https://doi.org/10.1016/s0166-2236(00)01828-2

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

Colby, C. L., Gattass, R., Olson, C. R., & Gross, C. G. (1988). Topographical organization of cortical afferents to extrastriate visual area PO in the macaque: A dual tracer study. *Journal of Comparative Neurology*, *269*(3), 392–413. https://doi.org/10.1002/cne.902690307

Collins, T., Rolfs, M., Deubel, H., & Cavanagh, P. (2009). Post-saccadic location judgments reveal remapping of saccade targets to non-foveal locations. *Journal of Vision*, *9*(5), 29. https://doi.org/10.1167/9.5.29

Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience*, *9*(8), 587–600. https://doi.org/10.1038/nrn2457

Crapse, T. B., & Sommer, M. A. (2008). The frontal eye field as a prediction map. *Progress in Brain Research*, *171*, 383–390. https://doi.org/10.1016/S0079-6123(08)00656-0

Crapse, T. B., & Sommer, M. A. (2012). Frontal eye field neurons assess visual stability across saccades. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(8), 2835–2845. https://doi.org/10.1523/JNEUROSCI.1320-11.2012

Crapse, T. B., Lau, H., & Basso, M. A. (2018). A role for the superior colliculus in decision criteria. *Neuron*, *97*(1), 181–194.

Crevecoeur, F., & Kording, K. P. (2017). Saccadic suppression as a perceptual consequence of efficient sensorimotor estimation. *ELife*, *6*, e25073. https://doi.org/10.7554/eLife.25073

Danks, D., Eberhardt, F. (2011) Keeping Bayesian models rational: The need for an account of algorithmic rationality. *Behavioral and Brain Sciences*, 34(4): 197–197. DOI: https://doi.org/10.1017/S0140525X11000240

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

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

Deneve, S. (2008a) Bayesian spiking neurons I: Inference. *Neural Computation*, 20(1): 91–117. DOI: https://doi.org/10.1162/neco.2008.20.1.91

Deneve, S. (2008b) Bayesian spiking neurons II: Learning. *Neural Computation*, 20(1): 118–145. DOI: https://doi.org/10.1162/neco.2008.20.1.118

DiCarlo, J. J., Haefner, R., Isik, L., Konkle, T., Kriegeskorte, N., Peters, B., Rust, N., Stachenfeld, K., Tenenbaum, J. B., Tsao, D., & Yildirim, I. (2021). *How does the brain combine generative models and direct discriminative computations in high-level vision?* https://openreview.net/forum?id=zlTiwFtLlR4

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science (New York, N.Y.)*, *255*(5040), 90–92. https://doi.org/10.1126/science.1553535

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, 7(2), 181–192.

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

Feinberg, I., & Guazzelli, M. (1999). Schizophrenia—A disorder of the corollary discharge systems that integrate the motor systems of thought with the sensory systems of consciousness. *The British Journal of Psychiatry: The Journal of Mental Science*, *174*, 196–204. https://doi.org/10.1192/bjp.174.3.196

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, *1*(1), 1–47. https://doi.org/10.1093/cercor/1.1.1-a

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

Fischer, B. J., & Peña, J. L. (2011). Owl's behavior and neural representation predicted by Bayesian inference. *Nature Neuroscience*, *14*(8), 1061–1066. https://doi.org/10.1038/nn.2872

Ford, J. M., & Mathalon, D. H. (2005). Corollary discharge dysfunction in schizophrenia: Can it explain auditory hallucinations? *International Journal of Psychophysiology:*

Official Journal of the International Organization of Psychophysiology, *58*(2–3), 179–189. https://doi.org/10.1016/j.ijpsycho.2005.01.014

Friston, K. (2012). The history of the future of the Bayesian brain. *NeuroImage*, *62*(2), 1230–1233. https://doi.org/10.1016/j.neuroimage.2011.10.004

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

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Houghton Mifflin.

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

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 227–247. https://doi.org/10.1037/0096-3445.117.3.227

Glymour, C. (2011) Osiander's psychology. *Behavioral and Brain Sciences*, 34(4): 199–200. DOI: https://doi.org/10.1017/S0140525X11000276

Goldberg, M. E., & Wurtz, R. H. (1972). Activity of superior colliculus in behaving monkey. I. Visual receptive fields of single neurons. *Journal of Neurophysiology*, *35*(4), 542–559. https://doi.org/10.1152/jn.1972.35.4.542

Golomb, J. D., & Kanwisher, N. (2012a). Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex*, 22(12), 2794–2810.

Golomb, J. D., & Kanwisher, N. (2012b). Retinotopic memory is more precise than spatiotopic memory. *Proceedings of the National Academy of Sciences*, *109*(5), 1796–1801.

Golomb, J. D., & Mazer, J. A. (2021). Visual Remapping. *Annual Review of Vision Science*, *7*, 257–277. https://doi.org/10.1146/annurev-vision-032321-100012

Griffiths, T. L., Chater, N., Norris, D., & Pouget, A. (2012). How the Bayesians got their beliefs (and what those beliefs actually are): Comment on Bowers and Davis (2012). *Psychological Bulletin*, *138*(3), 415–422. https://doi.org/10.1037/a0026884

Hafed, Z. M., & Krauzlis, R. J. (2010). Microsaccadic Suppression of Visual Bursts in the Primate Superior Colliculus. *Journal of Neuroscience*, *30*(28), 9542–9547. https://doi.org/10.1523/JNEUROSCI.1137-10.2010

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

Hamker, F. H., Zirnsak, M., Ziesche, A., & Lappe, M. (2011). Computational models of spatial updating in peri-saccadic perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. https://doi.org/10.1098/rstb.2010.0229

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory* (pp. xix, 335). Wiley.

Hinton, G. E. (1992). How neural networks learn from experience. *Scientific American*, *267*(3), 144–151. https://doi.org/10.1038/scientificamerican0992-144

Honda, H. (1993). Saccade-contingent displacement of the apparent position of visual stimuli flashed on a dimly illuminated structured background. *Vision Research*, *33*(5–6), 709–716. https://doi.org/10.1016/0042-6989(93)90190-8

Hopp, J. J., & Fuchs, A. F. (2004). The characteristics and neuronal substrate of saccadic eye movement plasticity. *Progress in Neurobiology*, *72*(1), 27–53. https://doi.org/10.1016/j.pneurobio.2003.12.002

Huerta, M. F., Krubitzer, L. A., & Kaas, J. H. (1986). Frontal eye field as defined by intracortical microstimulation in squirrel monkeys, owl monkeys, and macaque monkeys: I. Subcortical connections. *Journal of Comparative Neurology*, *253*(4), 415–439. https://doi.org/10.1002/cne.902530402

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

Inaba, N., & Kawano, K. (2014). Neurons in cortical area MST remap the memory trace of visual motion across saccadic eye movements. *Proceedings of the National Academy of Sciences*, *111*(21), 7825–7830. https://doi.org/10.1073/pnas.1401370111

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

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), 1020–1026. https://doi.org/10.1038/nn.2590

Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009). Neural representation of time in cortico-basal ganglia circuits. *Proceedings of the National Academy of Sciences of the*

United States of America, *106*(45), 19156–19161. https://doi.org/10.1073/pnas.0909881106

Johnson, K. O. (2000). Neural Coding. *Neuron*, *26*(3), 563–566. https://doi.org/10.1016/S0896-6273(00)81193-9

Jones, M., Love, B. C. (2011) Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences*, 34(4): 169–188. DOI: https://doi.org/10.1017/S0140525X10003134

Jordan, M. I., & Rumelhart, D. E. (1992). Forward Models: Supervised Learning with a Distal Teacher. *Cognitive Science*, *16*(3), 307–354. https://doi.org/10.1207/s15516709cog1603_1

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

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.

Keller, G. B., Bonhoeffer, T., & Hübener, M. (2012). Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. *Neuron*, *74*(5), 809–815. https://doi.org/10.1016/j.neuron.2012.03.040

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

Knill, D. C., & Richards, W. (Eds.). (1996). *Perception as Bayesian Inference*. Cambridge University Press. https://doi.org/10.1017/CBO9780511984037

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*(6971), 244–247. https://doi.org/10.1038/nature02169

Künzle, H., & Akert, K. (1977). Efferent connections of cortical, area 8 (frontal eye field) in Macaca fascicularis. A reinvestigation using the autoradiographic technique. *Journal of Comparative Neurology*, *173*(1), 147–163. https://doi.org/10.1002/cne.901730108

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

Leichnetz, G. R. (2001). Connections of the medial posterior parietal cortex (area 7m) in the monkey. *The Anatomical Record*, *263*(2), 215–236. https://doi.org/10.1002/ar.1082

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Li, W. X., & Matin, L. (1990). The influence of saccade length on the saccadic suppression of displacement detection. *Perception & Psychophysics*, *48*(5), 453–458. https://doi.org/10.3758/bf03211589

Lynch, J. C., & Tian, J.-R. (2006). Cortico-cortical networks and cortico-subcortical loops for the higher control of eye movements. *Progress in Brain Research*, *151*, 461–501. https://doi.org/10.1016/S0079-6123(05)51015-X

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

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, *18*, 50–60. https://doi.org/10.1214/aoms/1177730491

Maren, S., Phan, K. L., & Liberzon, I. (2013). The contextual brain: Implications for fear conditioning, extinction and psychopathology. *Nature Reviews. Neuroscience*, *14*(6), 417–428. https://doi.org/10.1038/nrn3492

Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, *202*(2), 437-470.1.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, *5*(3), 229–240. https://doi.org/10.1038/nrn1348

Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2013). The impact of microsaccades on vision: Towards a unified theory of saccadic function. *Nature Reviews Neuroscience*, *14*(2), 83–96. https://doi.org/10.1038/nrn3405

Matsuda, K., Nagami, T., Sugase, Y., Takemura, A., & Kawano, K. (2017). A Widely Applicable Real-Time Mono/Binocular Eye Tracking System Using a High Frame-Rate Digital Camera. In M. Kurosu (Ed.), *Human-Computer Interaction. User Interface Design, Development and Multimodality* (pp. 593–608). Springer International Publishing. https://doi.org/10.1007/978-3-319-58071-5_45

Mays, L. E., & Sparks, D. L. (1980). Dissociation of visual and saccade-related responses in superior colliculus neurons. *Journal of Neurophysiology*, *43*(1), 207–232. https://doi.org/10.1152/jn.1980.43.1.207

Merriam, E. P., Genovese, C. R., & Colby, C. L. (2003). Spatial Updating in Human Parietal Cortex. *Neuron*, *39*(2), 361–373. https://doi.org/10.1016/S0896-6273(03)00393-3

Merriam, E. P., Genovese, C. R., & Colby, C. L. (2007). Remapping in Human Visual Cortex. *Journal of Neurophysiology*, *97*(2), 1738–1755. https://doi.org/10.1152/jn.00189.2006

Minsky, M., & Papert, S. A. (1969). *Perceptrons: An Introduction to Computational Geometry*. MIT Press.

Murphy, K. P. (2013). Machine learning: a probabilistic perspective. Cambridge, Mass. [u.a.]: MIT Press.

Murray, J. M., & Escola, G. S. (2020). Remembrance of things practiced with fast and slow learning in cortical and subcortical pathways. *Nature Communications*, *11*(1), 6441. https://doi.org/10.1038/s41467-020-19788-5

Nakamura, K., & Colby, C. L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences*, *99*(6), 4026–4031. https://doi.org/10.1073/pnas.052379899

Neupane, S., Guitton, D., & Pack, C. C. (2016a). Two distinct types of remapping in primate cortical area V4. *Nature Communications*, *7*(1), 10402. https://doi.org/10.1038/ncomms10402

Neupane, S., Guitton, D., & Pack, C. C. (2016b). Dissociation of forward and convergent remapping in primate visual cortex. *Current Biology*, *26*(12), R491–R492. https://doi.org/10.1016/j.cub.2016.04.050

Neupane, S., Guitton, D., & Pack, C. C. (2020). Perisaccadic remapping: What? How? Why? *Reviews in the Neurosciences*, *31*(5), 505–520. https://doi.org/10.1515/revneuro-2019-0097

Ng, A. Jordan, M.I. (2002) On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. Advances in Neural Information Processing Systems 14, 841.

Niell, C. M., & Stryker, M. P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron*, *65*(4), 472–479. https://doi.org/10.1016/j.neuron.2010.01.033

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

Pola, J. (2011). An explanation of perisaccadic compression of visual space. *Vision Research*, *51*(4), 424–434. https://doi.org/10.1016/j.visres.2010.12.010

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: Knowns and unknowns. *Nature Neuroscience*, *16*(9), 1170–1178. https://doi.org/10.1038/nn.3495

Rahnev, D., & Denison, R. N. (2018). Suboptimality in perceptual decision making. *The Behavioral and Brain Sciences*, *41*, e223. https://doi.org/10.1017/S0140525X18000936

Rao, H. M., Abzug, Z. M., & Sommer, M. A. (2016). Visual continuity across saccades is influenced by expectations. *Journal of Vision*, *16*(5), 7. https://doi.org/10.1167/16.5.7

Rao, H. M., Mayo, J. P., & Sommer, M. A. (2016). Circuits for presaccadic visual remapping. *Journal of Neurophysiology*, *116*(6), 2624–2636. https://doi.org/10.1152/jn.00182.2016

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Ratcliff, R. (1978) A theory of memory retrieval. *Psychological Review, 85*(2): 59–108.

Robinson, D. A. (1963). A Method of Measuring Eye Movemnent Using a Scieral Search Coil in a Magnetic Field. *IEEE Transactions on Bio-Medical Electronics*, *10*(4), 137–145. https://doi.org/10.1109/TBMEL.1963.4322822

Robinson, D. L., & Wurtz, R. H. (1976). Use of an extraretinal signal by monkey superior colliculus neurons to distinguish real from self-induced stimulus movement. *Journal of Neurophysiology*, *39*(4), 852–870. https://doi.org/10.1152/jn.1976.39.4.852

Rogers, T. T., Seidenberg, M. S. (2011) Distinguishing literal from metaphorical applications of Bayesian approaches. *Behavioral and Brain Sciences*, 34(4): 211–212. DOI: https://doi.org/10.1017/S0140525X11000434

Roitman, J. D., Shadlen, M. N. (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, 22(21): 9475–9489. DOI: https://doi.org/10.1523/JNEUROSCI.22-21-09475.2002

Rosenblatt, Frank (1957). "The Perceptron—a perceiving and recognizing automaton". Report 85-460-1. Cornell Aeronautical Laboratory.

Rumelhart, D. E., McClelland, J. L., & Group, P. R. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations* (Vol. 1). A Bradford Book.

Schall, J. D., Morel, A., King, D. J., & Bullier, J. (1995). Topography of visual cortex connections with frontal eye field in macaque: Convergence and segregation of processing streams. *Journal of Neuroscience*, *15*(6), 4464–4487. https://doi.org/10.1523/JNEUROSCI.15-06-04464.1995

Schneider, D. M., & Mooney, R. (2018). How Movement Modulates Hearing. *Annual Review of Neuroscience*, *41*, 553–572. https://doi.org/10.1146/annurev-neuro-072116-031215

Shin, S., & Sommer, M. A. (2012). Division of labor in frontal eye field neurons during presaccadic remapping of visual receptive fields. *Journal of Neurophysiology*, *108*(8), 2144–2159. https://doi.org/10.1152/jn.00204.2012

Sohn, H., & Jazayeri, M. (2021). Validating model-based Bayesian integration using prior–cost metamers. *Proceedings of the National Academy of Sciences*, *118*(25). https://doi.org/10.1073/pnas.2021531118

Sohn, H., & Narain, D. (2021). Neural implementations of Bayesian inference. *Current Opinion in Neurobiology*, *70*, 121–129. https://doi.org/10.1016/j.conb.2021.09.008

Sommer, M. A., & Wurtz, R. H. (2001). Frontal eye field sends delay activity related to movement, memory, and vision to the superior colliculus. *Journal of Neurophysiology*, *85*(4), 1673–1685. https://doi.org/10.1152/jn.2001.85.4.1673

Sommer, M. A., & Wurtz, R. H. (2002). A Pathway in Primate Brain for Internal Monitoring of Movements. *Science*, *296*(5572), 1480–1482. https://doi.org/10.1126/science.1069590

Sommer, M. A., & Wurtz, R. H. (2004a). What the brain stem tells the frontal cortex. I. Oculomotor signals sent from superior colliculus to frontal eye field via mediodorsal thalamus. *Journal of Neurophysiology*, *91*(3), 1381–1402. https://doi.org/10.1152/jn.00738.2003

Sommer, M. A., & Wurtz, R. H. (2004b). What the brain stem tells the frontal cortex. II. Role of the SC-MD-FEF pathway in corollary discharge. *Journal of Neurophysiology*, *91*(3), 1403–1423. https://doi.org/10.1152/jn.00740.2003

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

Sommer, M. A., & Wurtz, R. H. (2008). Brain Circuits for the Internal Monitoring of Movements. *Annual Review of Neuroscience*, *31*(1), 317–338. https://doi.org/10.1146/annurev.neuro.31.060407.125627

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, *43*(6), 482–489. https://doi.org/10.1037/h0055479

Stanton, G. B., Bruce, C. J., & Goldberg, M. E. (1995). Topography of projections to posterior cortical areas from the macaque frontal eye fields. *The Journal of Comparative Neurology*, *353*(2), 291–305. https://doi.org/10.1002/cne.903530210

Stevenson, S. B., Volkmann, F. C., Kelly, J. P., & Riggs, L. A. (1986). Dependence of visual suppression on the amplitudes of saccades and blinks. *Vision Research*, *26*(11), 1815–1824. https://doi.org/10.1016/0042-6989(86)90133-1

Tajima, C. I., Tajima, S., Koida, K., Komatsu, H., Aihara, K., & Suzuki, H. (2016). Population Code Dynamics in Categorical Perception. *Scientific Reports*, *6*(1), 22536. https://doi.org/10.1038/srep22536

Thakkar, K. N., & Rolfs, M. (2019). Disrupted Corollary Discharge in Schizophrenia: Evidence From the Oculomotor System. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(9), 773–781. https://doi.org/10.1016/j.bpsc.2019.03.009

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

van Opstal, A. J., & van Gisbergen, J. A. (1989). Scatter in the metrics of saccades and properties of the collicular motor map. *Vision Research*, *29*(9), 1183–1196. https://doi.org/10.1016/0042-6989(89)90064-3

von Helmholtz, H. (1924). *Helmholtz's treatise on physiological optics, Vol. 1, Trans. From the 3rd German ed* (pp. xxi, 482). Optical Society of America. https://doi.org/10.1037/13536-000

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften*, *37*(20), 464–476. https://doi.org/10.1007/BF00622503

Wald, A. (1945). Sequential Tests of Statistical Hypotheses. *The Annals of Mathematical Statistics*, *16*(2), 117–186. https://doi.org/10.1214/aoms/1177731118

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

Wilcox, R. R. (2011). *Introduction to Robust Estimation and Hypothesis Testing*. Academic Press.

Wolpert, D. M., & Miall, R. C. (1996). Forward Models for Physiological Motor Control. *Neural Networks: The Official Journal of the International Neural Network Society*, *9*(8), 1265–1279. https://doi.org/10.1016/s0893-6080(96)00035-4

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, *48*(20), 2070–2089. https://doi.org/10.1016/j.visres.2008.03.021

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

Wurtz, R. H., & Optican, L. M. (1994). Superior colliculus cell types and models of saccade generation. *Current Opinion in Neurobiology*, *4*(6), 857–861. https://doi.org/10.1016/0959-4388(94)90134-1

Zimmermann, E., Morrone, M. C., Fink, G. R., & Burr, D. (2013). Spatiotopic neural representations develop slowly across saccades. *Current Biology: CB*, *23*(5), R193-194.

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

Biography

Ms. Divya Subramanian attended Cornell University as an undergraduate. She graduated in 2013 with a B.A. in Psychology. She then worked as a Research Assistant in the laboratory of Dr. Javier Medina at the University of Pennsylvania for two years. She joined the Cognitive Neuroscience Admitting Program (CNAP) at Duke University in 2015. In 2016, she affiliated with the Department of Neurobiology and and joined the laboratory of Dr. Marc Sommer.

During her time at Duke, she was elected co-chair of the Gordon Research Seminar on Eye Movements, 2021 (moved to 2023). She has also received conference travel awards from the Duke Graduate School and the Gordon Research Conference and was awarded for presenting the best poster at the Neurobiology Departmental retreat. At the time of submission of this dissertation, she has published one peer-reviewed article called “Corollary discharge for action and cognition”. The work presented in this dissertation will comprise three additional articles which are in preparation.