

# **Decision-Related Feedback Influences the Structure of Correlated Variability in Visual Cortex**

A thesis by  
Adrian Gopnik Bondy

Thesis Committee:

Bruce G. Cumming, National Eye Institute, *Advisor*

Robert H. Wurtz, National Eye Institute, *Chair*

David L. Sheinberg, Brown University

Michael N. Shadlen, Columbia University, *Outside Reader*

Defended April 21<sup>st</sup>, 2016

Submitted in partial fulfillment of the requirements for the  
degree of *Doctor of Philosophy* in the Program in  
the Department of Neuroscience at  
Brown University

May, 2016

© Copyright 2016 by Adrian Gopnik Bondy.  
All rights reserved.

This dissertation by Adrian Gopnik Bondy is accepted in its present form  
by the Department of Neuroscience as satisfying the  
dissertation requirements for the degree of Doctor of Philosophy.

Date \_\_\_\_\_

\_\_\_\_\_  
Bruce G. Cumming, M.D., Ph.D., Advisor  
Laboratory of Sensorimotor Research  
National Eye Institute  
National Institutes of Health

Recommended to the Graduate Council

Date \_\_\_\_\_

\_\_\_\_\_  
Robert H. Wurtz, Ph.D., Reader  
Laboratory of Sensorimotor Research  
National Eye Institute  
National Institutes of Health

Date \_\_\_\_\_

\_\_\_\_\_  
David L. Sheinberg, Ph.D., Reader  
Department of Neuroscience  
Brown University

Date \_\_\_\_\_

\_\_\_\_\_  
Michael N. Shadlen, M.D., Ph.D., Reader  
Zuckerman Mind Brain Behavior Institute  
Columbia University

Approved by the Graduate Council

Date \_\_\_\_\_

\_\_\_\_\_  
Peter Weber, Dean of the Graduate School

# Curriculum Vitae

**Adrian Gopnik Bondy, BA & Sc., M.S.**

Born November 13, 1985

## Contact

Laboratory of Sensorimotor Research, National Eye Institute, NIH

Building 49, Rm. 2A50, MSC 4435

Bethesda, MD 20892

Email: [adrian.bondy@gmail.com](mailto:adrian.bondy@gmail.com)

## Education

**2010 – current** Brown University & National Institutes of Health Neuroscience Graduate Partnership Program, Providence, Rhode Island and Bethesda, Maryland

**2003-2007** BA & Sc., Cognitive Science, McGill University, Montreal, Canada

## Peer-Reviewed Publications

**Bondy, A.** & Cumming, B. (2016). Choice-Related Feedback Influences the Structure of Correlated Variability in Visual Cortex. *Submitted to Nature*.

**Bondy, A.** & Cumming, B. (2015). Synchronous Spikes Are More Effective (but not for long). *Neuron*, 87 (4), 676-678. DOI: 10.106/j.neuron.2015.08.011.

McFarland, J., **Bondy, A.**, Cumming, B., Saunders, R., & Butts, D. (2015). Saccadic modulation of stimulus processing in primary visual cortex. *Nature communications*, 6. DOI: 10.1038/ncomms9110.

Menzer, D., Rao, N., **Bondy, A.**, Truccolo, W., & Donoghue, J. (2014). Population Interactions Between Parietal and Primary Motor Cortices During Reach. *Journal of Neurophysiology*, 112 (11), 2959-2984.

McFarland, J., **Bondy, A.**, Cumming, B., & Butts, D. (2014). High-resolution eye tracking using V1 neuron activity. *Nature communications*, 5. DOI: 10.1038/ncomms5605.

## **Lectures**

Columbia University, Department of Neuroscience, March 2016

NYU Center for Neural Science, January 2016

University of Pennsylvania, Department of Psychology, September 2015

Rochester University, Brain and Cognitive Sciences Department, June 2015

Harvard Medical School Systems Journal Club, April 2015

NYU Center for Neural Science, Noise Workshop, June 2014

## **Conference Abstracts**

**Bondy, A.**, and Cumming, B. (2016). The Impact of Noise Correlations in Visual Cortex on Perceptual Performance Depends on their Origin. *Vision Sciences Society Meeting Abstract*.

Lange, R., **Bondy, A.**, Cumming, B. and Haefner, R. (2016). On the neural basis of probabilistic inference during perceptual decision making. Computational and Systems Neuroscience.

**Bondy, A.**, and Cumming, B. (2015). Monkeys behaving badly: probing macaques' internal task strategies with psychophysical reverse correlation. Society for Neuroscience Abstracts.

**Bondy, A.**, and Cumming, B. (2015). Choice-Related Activity in Macaque Primary Visual Cortex Reflects Feedback. NIH Graduate Student Research Symposium.

McFarland, J., **Bondy, A.**, Cumming, B. and Butts, D (2015). Saccadic modulation of stimulus processing in primary visual cortex. Computational and Systems Neuroscience.

McFarland, J., **Bondy, A.**, Cumming, B. and Butts, D (2014). V1 response variability driven by fixational eye movements. Society for Neuroscience Abstracts.

**Bondy, A.**, and Cumming, B. (2013). Top down signals influence the distribution of noise correlations amongst sensory neurons. Society for Neuroscience Abstracts.

McFarland, J., **Bondy, A.**, Cumming, B. and Butts, D (2013). Temporal windowing of stimulus processing in primary visual cortex by saccade-driven alpha oscillations. Computational and Systems Neuroscience.

McFarland, J., **Bondy, A.**, Cumming, B. and Butts, D (2013). Nonlinear modeling of foveal V1 neurons using model-based eye-tracking. Society for Neuroscience Abstracts.

## **Awards, Honors, and Societies**

Brown University International Affairs Travel Fund Award, 2011

McGill University Lorne Gales Scholar, 2003 - 2007

## Preface

If I were a benevolent deity, my first act of compassion to mankind would be to create a biological machine for visual perception that was easy for neuroscientists to reverse engineer. I would take my inspiration from a camera. First, I would include an image-forming lens. Then, a sensor array: a set of neurons whose rate of discharge represented the presence of different visual features. Finally, a third component that cameras lack: a set of algorithms for using the state of the sensors to infer the state of the world.

The benevolence, as I see it, lies in the clear delineation between the components, particularly the last two: a “representational” component and an “inferential” component. This lets the poor scientist trying to reverse engineer the system take the problem in turns. First, by observing the state of the sensors while varying the input, he could begin to understand the computations underlying the selectivity of the sensors. Then, by observing the state of the inference while monitoring the state of the sensors, he could begin to understand the inferential algorithms used to transform the sensor’s representation into perception and action.

Neuroscientists have generally proceeded as though such a bipartite scheme exists and that activity in sensory brain areas (like the well-studied visual cortices) constitutes a part of the sensory representation. Rarely is this laid out as explicitly as it was by Kenneth Johnson in 1980, in the first of series of influential theoretical papers:



“The processes that intervene between a relatively peripheral array of neural activity and a subject’s decision in a discrimination task are split into two sections: *a)* the ascending sensory processes that provide the final patterns of neural activity on which discrimination is based, and *b)* a process that yields decision of the type required by the experimental design used in the psychophysical study.” (Johnson, 1980)

Johnson’s two sections correspond closely to the “representational” and “inferential” components I just described. (He would go on to call them the sensory-neuronal (SN) and neuronal-decision (ND) components.) But even in cases where such a scheme is not as explicitly proposed, the same thinking is frequently apparent.

The primary contribution of the work presented in this thesis is neuronal evidence for a significant flaw in this approach. The flaw is precisely the lack of a division of labor, particular at the levels of sensory processing beyond the periphery where the idea is typically invoked. In Chapter 1, I lay the groundwork for this argument, by introducing current thinking about how sensory signals are “decoded” to inform perceptual judgments. In Chapter 2, I present evidence from multi-electrode array recordings in behaving macaque monkeys showing that, even at the earliest stages of visual processing, the “sensors” (in this case V1 neurons) appear to receive input from downstream areas that are closely related to an “inference”-like procedure. This breaks the camera analogy and suggests a relationship between brain areas in sensory processing that is considerably more complex than is often imagined.

In Chapter 3, I discuss a related, albeit more practically-oriented, topic: how to control the allocation of animal subjects’ internal resources (that is, the “inference”-like procedure) in service of addressing the question posed earlier. I discuss a novel finding of deviations in the strategies monkeys employ in a discrimination task from those implied by task instruction, and present methods for detecting and controlling the strategies they

actually use. In Chapter 4, I summarize the main results and present new avenues for future research.

## Acknowledgements

I am indebted to so many people who contributed to the work described in this dissertation. First, to a large team of individuals that maintains the facilities in Building 49, and an expert staff of veterinarians, among whom Denise Parker, Irina Bunea, and Beth Nagy merit specific acknowledgement. Then, to the Laboratory of Sensorimotor Research community which has provided me invaluable support over the years. I must also recognize the administrative efforts of Rosanna Cabral, Denise Camara, Phil Wang, Phil Ryan, Jason Shockey, Chris McBain, Katherine Roche, Jerome Sanes, Anne Hart and Sharon Milgram, all of whom help me succeed as a Brown graduate student at NIH.

Intellectually, the work reflects the input of a number of scientific collaborators: James McFarland, Xue-Xin Wei, James Herman, Ralf Haefner, Hendrikje Nienborg, Geoff Vargish, Doug Ruff, Incheol Kang, Ali Moeeny, Robbe Goris, Ilya Monosov, and Ethan Bromberg-Martin, to name a few. It also reflects the mentorship and kindness of a number of investigators, especially my committee members Bob Wurtz, Mike Shadlen, and David Sheinberg, but also Rich Krauzlis, whose door was always open, and Wilson Truccolo and John Donoghue at Brown. It reflects the excellent teaching of Gilad Barnea, John Stein, David Berson, and Matthew Harrison. Most significantly, it reflects the intellect and dedication of my primary advisor, Bruce Cumming, whose insights permeate the pages of this dissertation.

I should also mention the significant animal contributions to this dissertation, including the cooperation of the macaque subjects of the experiments (Lemieux, Rufus, Big Guy, and Junior Barnes) and the affection of my pets (J.J., Medea, and Heart).

Finally, this dissertation would not be possible without three decades of tireless devotion on the part of my family, especially my parents Morgan and Thomas, who encouraged my sense of curiosity from the very beginning. And of course, it represents the love of my dearest companion, Hanna Sherrill, whom I met on the first day of the journey that led to this dissertation, and without whom the joy and enlightenment it has brought me would be immeasurably diminished.

# Table of Contents

Curriculum Vitae .....	iv
Preface.....	viii
Acknowledgements .....	xi
Table of Contents .....	xiii
List of Figures.....	xvi
<b>Chapter 1: Introduction .....</b>	<b>1</b>
1.1 Variability in the Responses of Sensory Neurons .....	1
1.2 The Origin of Sensory Neuronal Variability .....	3
1.3 Correlated Variability and Pooling .....	7
1.3.1 Correlations with Choice and Between Neurons .....	8
1.3.2 “Information-limiting” Noise Correlations.....	14
1.4 Controlling the Psychophysical Strategies of Animal Subjects .....	17
1.5 Summary .....	18
<b>Chapter 2: Choice-Related Feedback Influences the Structure of Correlated     Variability in Visual Cortex.....</b>	<b>21</b>
2.1 Summary .....	21
2.2 Introduction .....	22
2.3 Approach.....	23
2.4 Results .....	29
2.4.1 $R_{sc}$ Structure Changes with Task Instruction .....	29
2.4.2 Task-Dependent Changes in $R_{sc}$ Structure Relate to Perceptual Choice .....	35

2.4.3 Task-Dependent Modulation of $R_{sc}$ Structure	
Introduces “Differential” Correlations .....	39
<b>2.5 Discussion</b> .....	41
<b>2.6 Methods</b> .....	47
<b>2.7 Appendix</b> .....	59
2.7.1 Possible Confounding Retinal Effects .....	59
2.7.2 The Timecourse of $R_{sc}$ Structure Within the Trial .....	66
2.7.3 Correlation vs. Covariance .....	68
2.7.4 Laminar Distribution of CP and $R_{sc}$ .....	69
2.7.5 The Square Lattice Model.....	71

### **Chapter 3: Identifying Animal Subjects’ Internal Task Strategies with**

<b>Psychophysical Reverse Correlation</b> .....	<b>74</b>
<b>3.1 Introduction</b> .....	74
<b>3.2 A Quantitative Motivation of the Problem</b> .....	76
<b>3.3 Approach: Measuring Psychophysical Kernels</b> .....	78
<b>3.4 Results</b> .....	81
3.4.1 Shape of the Kernels .....	81
3.4.2 Decision Strategies are Fully Described by 1 <sup>st</sup> -Order Kernels .....	82
3.4.3 Macaque Subjects Update Task Strategy Slowly .....	83
3.4.4 Kernel Misalignment Impairs Performance .....	87
3.4.5 Comparison with Human Subjects .....	90
3.4.6 Are Subjects Really at Threshold? .....	92
<b>3.5 Discussion</b> .....	94

3.6 Methods .....	97
<b>Chapter 4: Summary and Future Directions .....</b>	<b>101</b>
4.1 Chapter 2 Summary and Future Directions.....	101
4.2 Chapter 3 Summary and Future Directions.....	110
4.3 Summary of Contributions .....	115
<b>Bibliography .....</b>	<b>117</b>

## List of Figures

Figure 1.1 $R_{sc}$ depends on signal correlation and pairwise distance .....	5
Figure 1.2 Influence of the structure of noise correlations on Choice Probability .....	11
Figure 1.3 “Information-limiting” noise correlations .....	16
Figure 2.1 Orthogonal orientation discrimination task.....	25
Figure 2.2 Structure of feedback suggested by the presence of lattice-like correlation structure in V1.....	27
Figure 2.3 Subjects’ decision strategies change with task instruction.....	28
Figure 2.4 $R_{sc}$ structure in V1 depends on task instruction.....	30
Figure 2.5 Two-component multilinear regression model.....	32
Figure 2.6 The Task-dependent component of $r_{sc}$ can account for CP.....	34
Figure 2.7 $R_{sc}$ structure depends on choice distribution .....	36
Figure 2.8 Task-aligned $r_{sc}$ structure is low dimensional.....	38
Figure 2.9 The task-dependent component of $r_{sc}$ resembles stimulus-driven (“differential”) correlations.....	40
Figure 2.10 Models of perceptual decision making consistent with our results .....	43
Figure 2.11 Lattice-like $r_{sc}$ structure depends on active task engagement .....	59
Figure 2.12 $R_{sc}$ structure is not influenced by stimulus history .....	60
Figure 2.13 $R_{sc}$ structure is similar after short and long inter-trial intervals.....	62
Figure 2.14 $R_{sc}$ structure is not generated by fixational eye movements.....	63
Figure 2.15 $R_{sc}$ structure is not driven by stimulus variability .....	65
Figure 2.16 Temporal dynamics of structured noise correlations.....	67
Figure 2.17 The task-aligned structure of spike-count covariance .....	69



Figure 2.18 Laminar Distribution of CP and $r_{sc}$ .....	70
Figure 2.19 The square lattice model .....	72
Figure 3.1 Simulated effect of task strategy “misalignment” on performance.....	77
Figure 3.2 Measuring psychophysical kernels, Task 1 .....	79
Figure 3.3 Measuring psychophysical kernels, Task 2 .....	80
Figure 3.4 Using the measured kernel to predict choices .....	84
Figure 3.5 Timecourse of kernel rotation .....	86
Figure 3.6 Observed effect of task strategy “misalignment” on performance .....	88
Figure 3.7 Performance with interleaved orientation pairs .....	89
Figure 3.8 Human subjects rapidly update task strategy .....	91
Figure 3.9 Self-consistency underpredicts overall performance .....	93
Figure 4.1 The Haefner model .....	107

# Chapter 1: Introduction

## 1.1 Variability in the Responses of Sensory Neurons

While the inner workings of the brain remain largely mysterious, a great deal of progress has been made in understanding the neural circuits that encode sensory inputs. Much of this progress has been made by measuring spiking activity in single neurons in sensory areas in response to controlled patterns of sensory stimulation. This work dates back to the early days of neuroscience, and has shed light on the exquisitely selective responses of neurons to various features of the external world, from the orientation of contrast edges in primary visual cortex (V1; Hubel & Wiesel (1959)) to single-whisker stimulation in the rodent barrel cortex (Welker, 1976) to complex visual forms like faces in regions of the primate inferotemporal cortex (Desimone, Albright, Gross, & Bruce, 1984).

One consistent feature of the data is the variability of stimulus-driven spiking responses, particularly in the cortex (Dean, 1981; Henry, Bishop, Tupper, & Dreher, 1973; Tomko & Crapper, 1974). In other words, sensory cortical neurons generally do not respond the same way when confronted twice with the same sensory stimulus. This observation has potentially profound implications for the nature of the sensory representation available to downstream areas. Ideas about these implications have earned this variability the monicker “noise.” The early studies are the most explicit in laying out the motivation for this, which have become so widespread as to be almost taken for granted. For instance, Tolhurst, Movshon, & Dean (1983) state:

“It is generally thought that performance on psychophysical tasks is probabilistic ... because identical physical stimuli elicit neural responses that vary randomly in amplitude from presentation to presentation. Since this variable neural representation of the sensory event alone is accessible to the observer, he must make a *statistical decision* as to whether a particular level of neural activity is more likely to be the response to a stimulus than a random fluctuation in some ongoing background activity.”

[emphasis in the original]

In other words, Tolhurst et al. are saying that the sensory neuronal variability that they measure (in area V1 in their case) is equivalent to a stochastic representation of the environment that constrains the quality of the inference that a subject might go on to make. In other words, this *is* the noise that constrains the subject’s performance. This is a natural interpretation given the point of view by which the neural underpinnings of perception can be neatly divided into a component that objectively represents sensory input and a component which interprets/decodes that representation in the service of perception and action.

If this is correct, the practical implications are profound. Measurements of the statistical properties of sensory neurons would reveal the noise contaminating the brain’s sensory representation at any given moment in time. These could then be used to make testable predictions about causal effects on downstream computations, opening a pathway for investigating the neural mechanisms that transform the brain’s sensory representation into perception and action.

An alternate view is that variable activity in sensory neurons at least partially reflects changes in non-sensory signals (“top-down”, “feedback”) that are not necessarily treated as “noise” by downstream areas. This would suggest a fundamentally different way of thinking about the relationship between “sensory” and “non-sensory” brain areas, and would challenge the view that variability in the activity of recorded neurons gives

experimenters access to the stochasticity that limits perceptual performance. Investigating this possibility is the goal of the work described in this dissertation. A driving hypothesis is that the processes that generate perceptual decisions themselves also act as an input to sensory brain areas.

## **1.2 The Origin of Sensory Neuronal Variability**

Sensory neuronal variability presumably reflects changes over time in synaptic inputs to sensory neurons. But which synaptic inputs? Under the “variability=noise” view, the possibilities are, in principle, quite constrained. The only way to guarantee that sensory neuronal variability is “noise” is if it is an inherent feature of the feedforward pathways conveying sensory input. Variability in these pathways, and only these pathways, is fundamentally indistinguishable from noise. Variability in non-sensory inputs may contribute to uncertainty in the sensory representation as it is effectively read out downstream, or it may not. This depends on the origin of the non-sensory signals and the nature of the read out.

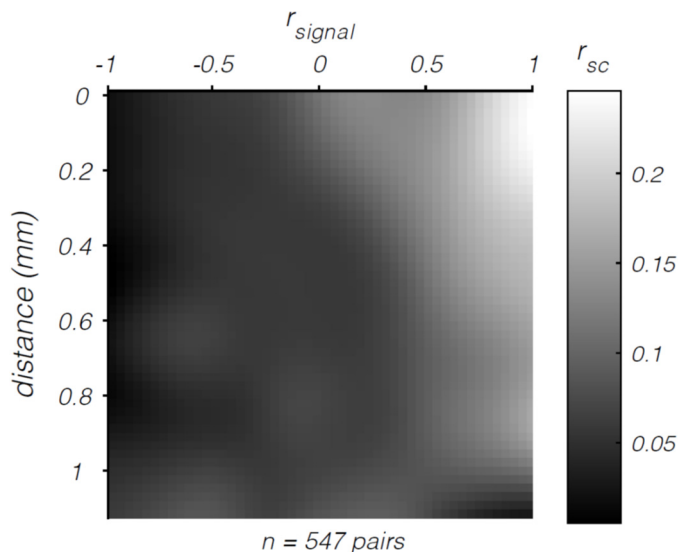
We know from several lines of research that there is a diversity of inputs to sensory cortical areas that do not derive from afferent sensory input. First, direct anatomical evidence suggests massively convergent patterns of connectivity throughout the neocortex, most of which cannot be characterized as purely feedforward (Callaway, 2004; Sillito, Cudeiro, & Jones, 2006). Even in primary visual cortex, it has been estimated that only about 5% of excitatory inputs come directly from the retina through the lateral geniculate nucleus (LGN) of the thalamus (Olshausen & Field, 2005). And this “feedforward” input relies on only a subset of the diverse neurotransmitter systems that

are known to play a role in cortical processing (Herrero et al., 2008; Hruby, Ott, Nieder, Pourriahi, & Nienborg, 2015; Paukert et al., 2014). This implies that we continue to lack a detailed understanding, even at the earliest stage of sensory cortical processing, of a great deal of the input.

Second, the temporal statistics of sensory neuronal variability also suggest a role for central signals. The precise timing of spikes in response to a fixed stimulus tends to have more structure than would be predicted based on the average rate of discharge (Dean, 1981; Henry et al., 1973; Tomko & Crapper, 1974). This implies that some of the variability in firing rate is due to changes in the inputs, rather than (Poisson) randomness associated with spike generation, even in the presence of fixed sensory input. A recent study (Goris, Movshon, & Simoncelli, 2014) has proposed a model of sensory neuronal responses which incorporates both of these elements. This “modulated Poisson” model provides a much better fit to neuronal variability in visual cortex and the LGN, and demonstrates that the non-Poisson component increases in magnitude through the visual hierarchy, consistent with an increase in the proportion of sensory neuronal variability that is not generated by stochasticity in spike generation but rather variability in the inputs. This suggests that the component of sensory neuronal variability that does not derive from sensory afferents increases through the visual hierarchy.

Another crucial feature of sensory neuronal variability, which places anatomical constraints on its origin, is the fact that it tends to be weakly correlated amongst pairs of neurons. The average Pearson correlation between pairs of spike counts across repeated stimulus presentations tends to lie in the range of 0.1-0.2 (Cohen & Kohn, 2011). This suggests that a substantial fraction of the variability reflects common inputs. The

magnitude of spike-count correlation ( $r_{sc}$ ) is higher for pairs that are in close physical proximity and amongst those which share similar stimulus selectivity (Fig. 1.1; Bair, Zohary, & Newsome, 2001; Smith & Kohn, 2008). This has been used as evidence in favor of a “feedforward” origin, since anatomically, such pairs are more likely to share common afferent input. However, while the anatomical details of “feedback” pathways are less well understood, they too display marked selectivity, as evidenced by the effects of spatial and feature-based attention on visual cortical neurons that are highly specific (Maunsell & Treue, 2006; Reynolds & Chelazzi, 2004). In addition,  $r_{sc}$  tends to remain positive even over large cortical distances, suggesting a component of sensory neuronal activity that is broadly coherent and inconsistent with the local structure of sensory inputs. Thus, a significant portion of  $r_{sc}$  is likely to be due to common input from other sources.



**Fig. 1.1.  $R_{sc}$  depends on signal correlation and pairwise distance.** Pseudocolor plot illustrating the positive dependence of noise correlations on signal correlation (in this case this is based on similarity in orientation tuning), and the negative dependence on physical distance. These measurements were made using 24-contact linear multi-electrode arrays implanted in macaque V1 (see Chapter 2 Methods). The image is smoothed with a Gaussian with s.d. of 10 pixels.

Several studies have also shown that non-sensory signals directly modulate the firing rate of sensory neurons, thus likely contributing to the “variability” observed in other studies where these factors were not controlled. For instance, reward, arousal, attention, anaesthesia and even locomotion appear to modulate the activity of neurons in visual cortex (Ecker et al., 2014; Goard & Dan, 2009; Niell & Stryker, 2010; Shuler & Bear, 2006). However, these studies have typically not analyzed how these factors relate to choices subject make in perceptual tasks, so whether these sources of variability are treated as noise by subjects is not known.

The most well studied source of non-sensory inputs to visual cortex is visual attention. Attention refers to the “filtering” of sensory evidence in order to improve sensitivity to a particular stimulus feature of behavioral relevance (Carrasco, 2011). Visual attention has been proposed to act at an early stage in visual processing by modulating the sensory representation through feedback in order to increase signal-to-noise in the encoding of the attended stimulus feature (Cohen & Maunsell, 2009; J. H. R. Maunsell & Treue, 2006; Mitchell, Sundberg, & Reynolds, 2009; Reynolds & Chelazzi, 2004). This is broadly consistent with its known effects on visual cortical neurons: increases in response gain (Treue & Martínez Trujillo, 1999) and changes in correlated variability (Cohen & Maunsell, 2009), both of which are thought to improve the reliability of pooled sensory signals. However, recent thinking has cast doubt on the idea that visual attention acts by reducing the “noise” in the sensory representation. For one, there is no evidence that these modulations underlie the increased performance associated with attention. In fact, a recent study showed that these modulations persist in primate visual cortex despite functional lesions of the superior colliculus that strongly disrupt

visual attention (Zénon & Krauzlis, 2012). Another recent study implies that the reduction in noise correlation is due to an attenuation of ongoing variability in feedback signals (Rabinowitz, Goris, Cohen, & Simoncelli, 2015). This may mean that neuronal effects due to attention relate to changes in self-generated signals, rather than modulations of the statistical structure of the sensory representation. Thus, the effects of attention on visual cortical neurons do not necessarily bear on the fundamental question posed here: does variability in sensory neurons really function as noise in the sensory representation?

### **1.3 Correlated Variability and Pooling**

In the absence of suitable data with which to address this question, the assumption that sensory neuronal variability does indeed act like “noise” in the sensory representation provides a powerful and tractable framework for thinking about the relationship between the activity of sensory neurons and the perceptual behavior of subjects. This framework has generated a number of important quantitative insights that have greatly influenced the approach taken in the experiment described in Chapter 2.

These insights all relate to a fundamental observation: spike-count correlations between sensory neurons are essential to any account of psychophysical performance based on pooling. This insight was first reached by studies that observed that the performance of an ideal observer on a psychophysical task, if its responses were based simply on the spike counts of single recorded neurons, frequently approached or even exceeded the performance of the subjects themselves (Britten, Shadlen, Newsome, & Movshon, 1992; Celebrini & Newsome, 1994; Newsome, Britten, Movshon, & Shadlen,



1989; Romo, Hernández, Zainos, Brody, & Salinas, 2002; Tolhurst et al., 1983). This implies either that subjects' choices are based on pooling very small numbers of neurons or that correlated variability vitiates any improvement associated with large pools. While the empirical finding that noise correlations were significantly positive (on average) demonstrated a way subjects' performance could be reconciled with large pools (Britten et al., 1992; Zohary, Shadlen, & Newsome, 1994), the fact that these studies did not directly measure noise correlations at the population level nor have access to the decoding strategies subjects actually use renders these results difficult to interpret.

The impact of correlated variability on performance in a common psychophysical paradigm—the two-alternative forced choice (2AFC) discrimination task—has been explored in theoretical studies in some detail. In a 2AFC discrimination task, subjects are required to classify a noisy stimulus as belonging to one of two categories. In the predominant theoretical model, decisions are made by linearly weighting the spike counts of a population of sensory neurons elicited by stimulus presentations (Haefner, Gerwinn, Macke, & Bethge, 2013; Moreno-Bote et al., 2014; Shadlen, Britten, Newsome, & Movshon, 1996; Zohary et al., 1994) and then applying a threshold to the resulting weighted sum to generate choices. I will refer to this as the “linear pooling model.” I wish to highlight two findings about the behavior of this model, which I will take in turn. Both of these findings point to the importance of a particular pattern of correlated variability for the perceptual judgments that arise.

### *1.3.1 Correlations with Choice and Between Neurons*

The first finding gives an account of how variability in sensory neurons affects subjects' reports on a trial-by-trial basis, traditionally thought to underlie the experimentally observed correlations between single neurons and choices (choice-related activity). Choice-related activity was first observed in experiments using unit recordings in motion-selective areas in the macaque superior temporal sulcus (Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Celebrini & Newsome, 1994). In those studies, subjects were trained to report the direction of coherent motion in a field of otherwise randomly moving dots. The two possible directions were always opposed, for instance leftward and rightward. The experimenters reasoned that if the variable neuronal responses they recorded provide the evidence supporting the animal's choices, then that variability should correlate with choice, even across repeated trials using the same stimulus. Indeed, they found precisely such a correlation, a finding which has been repeated many times in similar experiments (for review, see Hendrikje Nienborg, Cohen, & Cumming, 2012). They quantified this effect as the probability with which an ideal observer could predict the subject's choices using only the spike counts of a single neuron, a quantity known as "Choice Probability" (CP). Average CPs in many sensory areas across a variety of tasks tend to lie the range of 0.55-0.6, demonstrating a modest but above-chance correlation with choice in many neurons.

This finding has led to the widespread use of CP (and a related measure used in detection tasks) as a proxy for the causal influence of single neurons on perceptual decisions (Cook & Maunsell, 2002; de Lafuente & Romo, 2005; Gu, Angelaki, & DeAngelis, 2008; Liu & Newsome, 2005; Price & Born, 2010; Smolyanskaya, Haefner,

Lomber, & Born, 2015; Uka & DeAngelis, 2004). However, Shadlen et al. (1996) pointed out a quantitative problem in this interpretation: given the vast numbers of sensory neurons in the brain presumably contributing to perceptual decisions, the influence of any one neuron on choice should be immeasurably small. Instead, CPs seem to be almost everywhere one puts an electrode.

They identified a solution to the problem using a theoretical model. In their model, a subject's choices on a given trial were the outcome of a comparison between two signals: one signal represented the summed spike counts of all the leftward tuned neurons and the other the summed spike counts of the rightward tuned neurons (a simplified two-pool version of linear pooling). The only way to get widespread CPs, they found, is if the trial-to-trial variability within the two pools is correlated. Then the correlated variability cannot be averaged away by pooling and the resulting choice is a comparison between the correlated component of the signal in the two pools (Fig. 1.2). This introduces widespread correlations between choice and individual neurons whose activity also reflects the correlated signal.

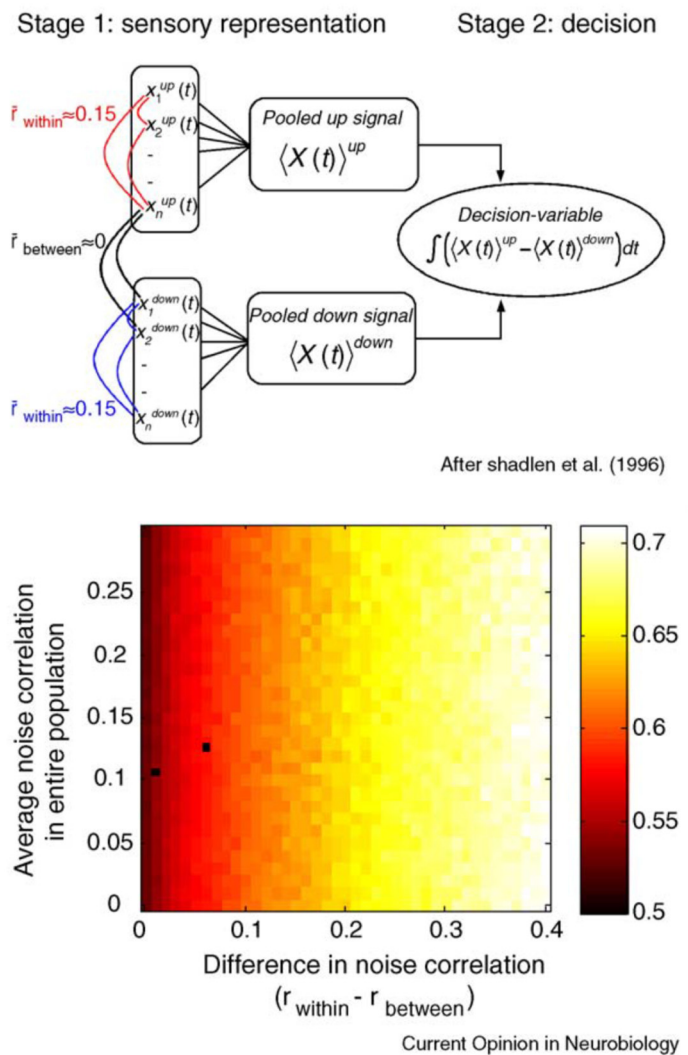
More recent work (Haefner et al., 2013; Hendrikje Nienborg & Cumming, 2010) has confirmed and extended these findings. In the general case of linear pooling, where each neuron is assigned an arbitrary weight (positive or negative depending on which choice it supports), a neuron's CP is given by:

$$CP_k = \frac{1}{2} + \frac{2}{\pi} \operatorname{sgn}(\xi_k) \arctan \sqrt{2\xi_k^{-2} - 1}^{-1} \quad \text{with} \quad \xi_k = \frac{(\mathbf{C}\boldsymbol{\beta})_k}{\sqrt{C_{kk}\boldsymbol{\beta}^T\mathbf{C}\boldsymbol{\beta}}} \quad (1.1)$$

$CP_k$  is the CP of neuron  $k$  with respect to choice 1,  $\boldsymbol{\beta}$  is the vector of read out weights and  $\mathbf{C}$  is the covariance matrix (Haefner et al., 2013). The critical term in the equation is the

numerator of  $\xi_k$ . This term is the element-wise product between two quantities: first, the spike-count covariance between neuron  $k$  and all others contributing to the decision, and second, the weights of all other neurons. Neuron  $k$ 's CP is monotonically related to their product. In the simplified two-pool model considered by Shadlen and colleagues, Haefner et al. (2013) showed using Eq. 1.1 that mean CP only depends on the average difference between within- and between-pool correlations, as first discovered by Nienborg & Cumming (2010) using simulations (Fig. 1.2).

**Figure 1.2. Influence of the structure of noise correlations on Choice Probability.** Top panel: Illustration of the two-pool version of the pooling model, developed by Shadlen et al. (1996) to explain the observation of CP in MT in a direction discrimination task. The decision variable is computed as the difference between the sums of neuronal activity in two pools, one containing the neurons supporting direction 1 and the other those supporting direction 2 (here “up” and “down”). Bottom panel: Simulated mean CP in the model population as a function of two quantities: 1) difference in average noise correlation between members of the same pool ( $r_{within}$ ) and between members of opposite pools ( $r_{between}$ ), and 2) the average overall noise correlation. Average CP depends only on the first quantity, demonstrating that noise correlation structure, not magnitude, determines CP. For this simulation, the noise correlations were assumed to be uniform within a pool. (Figure reproduced from Nienborg & Cumming (2010)).



These theoretical studies of choice-related activity in the linear pooling model have enriched the interpretation of CP in two important ways. First, this work demonstrates that CP reflects more than a single neuron's causal effect on choice. It mostly reflects the causal effect on choice of the other neurons with which a given neuron's activity is correlated. Indeed, a neuron can systematically show CP with the wrong sign if its variability is correlated with neurons supporting its anti-preferred choice, no matter what read out weight is actually assigned to it. Nonetheless, this correlated "noise", in the aggregate, still has a causal impact on choice, consistent with the traditional interpretation of CP. Furthermore, if one had sufficient information about the structure of the noise correlations, one could in principle identify the read out weights of single neurons.

Second, this line of work has united a long debate about the correct interpretation of choice-related activity (Crapse & Basso, 2015; Bruce G. Cumming & Nienborg, 2016; Hendrikje Nienborg et al., 2012) with the central question at stake in this dissertation: the origin of sensory neuronal (co)variability. This provides a key starting point for the work discussed in Chapter 2, whose main goal is to identify whether the pattern of noise correlations in a sensory neuronal population reflects noise in sensory afferents or changes in central signals. The latter could come about as an effect of choice on sensory neurons.

Several studies on choice-related activity in single neurons suggest this may be the case (for review, see: Cumming & Nienborg, 2016; Nienborg et al., 2012). One piece of evidence comes from the timecourse of CP within single trials, which typically shows an early increase followed by a plateau (Nienborg & Cumming, 2009; Shadlen et al.,

1996). If CP reflects the causal influence of sensory neurons on choice, this must mean that (monkey) subjects tend to evenly weight evidence presented throughout the trial, and perhaps even discount early evidence. However, psychophysical reverse correlation shows that the reverse is true: on average, subjects more strongly weight early evidence in fixed-duration trials (Kiani, Hanks, & Shadlen, 2008; Nienborg & Cumming, 2009). This suggests a component of CP may be a feedback effect of the evolving decision process itself.

Another line of evidence for this is the pronounced CPs observed in studies involving a perceptually bistable stimulus: the structure-from-motion cylinder (Dodd, Krug, Cumming, & Parker, 2001; Krug, Cumming, & Parker, 2004; Parker, Krug, & Cumming, 2002). This stimulus consists of two planes of dots drifting in opposite directions. The velocity of the dots follows a sinusoidal profile, consistent with the orthographic projection of a rotating cylinder. In the absence of disambiguating disparity cues, the direction of rotation is ambiguous and is perceived bistably. Dodd et al. (2001) trained monkeys to report the direction of rotational motion in the stimulus, while the amount of disambiguating disparity was varied to manipulate task difficulty. Because MT neurons are both motion- and direction-selective, they often display strong preference for one direction of rotation in the stimulus. On the ambiguous zero-disparity trials, they found an average CP of 0.67 in MT neurons, considerably higher than previously reported in the standard dot-motion task (0.55; Britten et al., 1996).

This increase is *inconsistent* with the prediction of a feedforward origin of CP, because the strong perceptual stability elicited by the stimulus implies a weak influence of sensory neuronal variability on choice. A more satisfactory explanation is that the

strong variations in the subject's perceptual state generate strong CPs via feedback. This view is bolstered by unpublished results from experiments in our lab. In these experiments, 500-ms pulses of disambiguating disparity were introduced at the beginning of zero-signal trials to bias subject's perception of rotation direction. This manipulation introduced large changes in the firing rate of MT neurons that persisted until the end of the trial, as though the disambiguating disparity cue had not disappeared (Ali Mooeny and Bruce Cumming, *personal communication*). This further supports the view that choice-related signals in sensory neurons are driven by the subject's perceptual state.

Finally, it has been noted that the observation of CP across tasks that require the same set of neuronal signals to be pooled in contrasting ways is difficult to reconcile if the underlying pattern of noise correlations is fixed (Krug et al., 2004; Nienborg et al., 2012). Rather, this points to noise correlations that change dynamically in a context-dependent manner. This would be particularly strong evidence in favor of the view that CP does not reflect the effect of sensory noise, but rather reflects reorganization of functional connectivity that depends on the task context. Importantly, however, none of these studies has directly measured the structure of noise correlations simultaneously in a large population. While one study did demonstrate a task-dependent element to noise correlation structure (Cohen & Newsome, 2008), this was done using paired recordings, making the results too limited to directly test this idea.

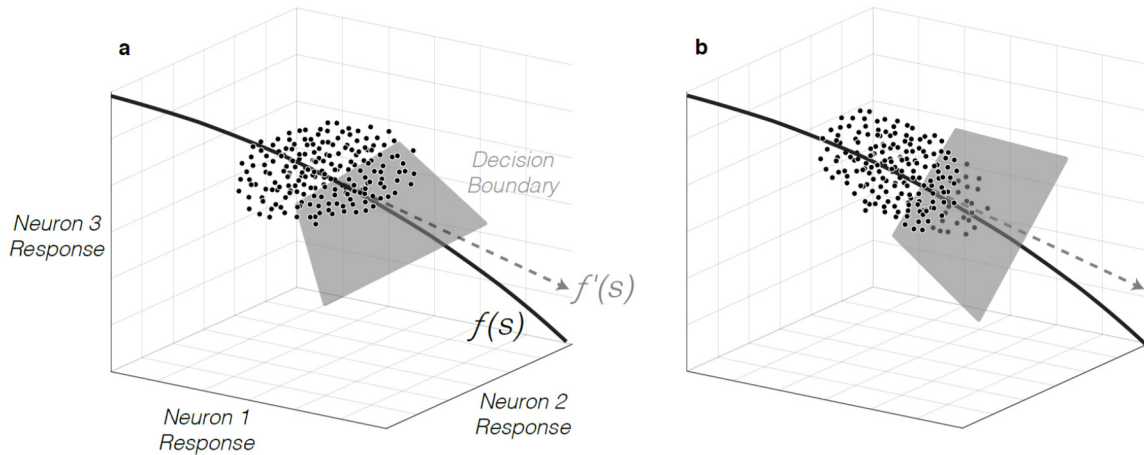
### 1.3.2 “Information-limiting” Noise Correlations

The second major finding of the pooling model is a quantitative account of the effect of sensory neuronal variability on psychophysical performance. As mentioned

earlier, variability that is correlated between pairs of sensory neurons has a much greater impact on a pooled signal than uncorrelated variability. This insight has been explored quantitatively in a number of studies, which have converged on similar results (Abbott & Dayan, 1999; Averbeck, Latham, & Pouget, 2006; Johnson, 1980; Moreno-Bote et al., 2014; Snippe & Koenderink, 1992). In the context of a 2AFC task, performance is degraded when correlations are proportional to the similarity in tuning between neuronal pairs. Another way of describing these correlations is that they are easily confused with the neuronal response to a change in the stimulus (Fig. 1.3). This pattern of  $r_{sc}$  has recently been termed “differential” or “information-limiting” correlations (Moreno-Bote et al., 2014). This study demonstrated that, in the limit of infinite neurons and assuming subjects can use optimal linear decoders, *only* differential correlations limit performance. However, given finite numbers of neurons and/or suboptimal decoding, a range of correlations structures that approximate “differential” correlations will impact performance.

It has been observed that these “differential” correlations are extremely similar to the pattern that gives rise to CP. For CP to exist, correlations must be higher within than between pools. For correlations to be “differential”, they must be proportional to similarity in task preference between pairs. For typical discrimination tasks, appropriate read out weights are closely related to a neuron’s tuning for the task. Thus there is a deep connection between the two findings about  $r_{sc}$  structure in the pooling model. In fact, a recent study showed that, assuming a subject uses the optimal linear read out weights for





**Figure 1.3. “Information-limiting” noise correlations.** The discriminability of a sensory input is sensitive to the structure of noise correlations in sensory neurons. **a.** The response of a set of three sensory neurons to stimuli in a discrimination task. The solid line— $f(s)$ —shows the average response of the neurons to the stimuli as they vary along the dimension being discriminated. The cloud of points represents the response of the neurons to repeat presentations of a near-threshold stimulus at a fixed value of  $s$ . The shape of this cloud is determined by the noise correlations. When this cloud is not oriented parallel to  $f(s)$ —that is, its projection onto  $f'(s)$  is small—an optimal decision boundary can be placed such that the decisions are still sensitive to changes in the stimulus but insensitive to the noise. This means that the noise correlations have a minimal effect on choice, attenuating CP. **b.** When the correlations lie on  $f'(s)$  there is no decision boundary that maintains sensitivity to stimulus changes without being subject to noise contamination. These “differential” correlations thus necessarily have an effect on choice (i.e. introduce CP). (after Moreno-Bote et al. (2014).)

a given discrimination task, CPs can only be observed if there are differential correlations (Pitkow, Liu, Angelaki, DeAngelis, & Pouget, 2015a). Alternatively, if subjects suboptimally decode their sensory neurons, CP can be observed with a broader range of noise correlation structures (Fig. 1.3).

The relationship between noise correlations and encoding of stimulus information described in these studies relies on the strong assumption that all sensory neuronal variability is noise. If it even partially reflects changes in centrally generated signals that are not interpreted as noise in the sensory representation, it would be unclear if any pattern of observed correlations actually limits performance. This would depend on

whether the decoding machinery can dissociate the effects of central signals on sensory neurons from those caused by the stimulus. Thus, the central question posed in this dissertation also has deep implications for interpreting the consequences of correlated variability on the information capacity of the brain's sensory representation.

#### **1.4 Controlling the Psychophysical Strategies of Animal Subjects**

One way of defining noise is variability in a measurement unexplained by an experimental manipulation. Thus, one could argue on purely semantic grounds that sensory neuronal variability is “noise”, as experimental manipulations that probed the responses of sensory neurons have typically focused exclusively on the sensory inputs. This is for good reason. Testing other potential sources of input requires some way of manipulating them.

Our driving hypothesis is that signals related to the decision process during a perceptual discrimination task contribute to structured noise correlations in sensory neurons. Testing this hypothesis requires identifying an effect of the subject's decision process. As we will discuss in more detail in Chapter 2, our approach involved changing the set of stimuli monkey subjects need to discriminate to receive reward while maintaining a fixed retinal input. This allowed us to dissociate  $r_{sc}$  structure that is fixed from that which depends on the task context.

A critical assumption of this approach is changes in task instruction elicit appropriate changes in the subject's internal task strategy. This assumption is shared with most prior psychophysical studies, that is typically only verified by showing that subjects are able to perform the task reasonably well after a change in instruction. However, in

principle, subjects can achieve high rates of reward using a task strategy that differs from the one suggested by task instruction. This possibility will not seem foreign to anyone who has experience training monkey subjects to perform psychophysical tasks.

To address this potential confounding factor, we turn to psychophysical reverse correlation (PRC). PRC is an approach developed for use in human psychophysics to objectively probe the strategies subjects use in a psychophysical task. In Chapter 2, we briefly discuss our use of PRC to ensure that subjects performed the task as instructed. In Chapter 3, we provide a more detailed analysis. We reveal a novel and striking difference between human and macaque performance on the task. While human subjects can flexibly adopt new task strategies when the discriminanda change, macaque subjects require days or even weeks of retraining. During the retraining period, they slowly and continuously update their internal task strategy to match new instruction. Importantly, this slow timecourse is masked by the ability of the subjects to achieve reasonably high rates of reward while performing the “wrong” task. This reveals an interesting feature of macaque behavior that also demonstrates a potentially serious behavioral confound likely to be present in past studies.

## **1.5 Summary**

Sensory neurons respond unpredictably to a fixed sensory input. Some of this response variability is correlated between neuronal pairs. A predominant view of this correlated variability is that it can confound any attempt to reconstruct the nature of the stimulus that elicited the response, because downstream brain areas must make a statistical decision based on pooling many sensory neurons. A lack of independence in

the individual neuronal signals that are pooled can degrade the reliability of pooling. However, a portion of the correlated variability may in fact reflect changes over time in modulatory/feedback inputs. How such common inputs impact the pooled neuronal signals that provide the sensory evidence for perceptual judgments is not well understood.

An important experimental constraint can be provided by quantifying the portion of the structured spike-count correlations thought to impact perceptual performance in fact derive from feedback. This is the goal of Chapter 2. In that chapter, we present results from population recordings in macaque V1 while subjects perform a coarse orientation discrimination task using filtered noise stimuli. Given current feedforward models, the presence of choice-related activity implies that the structure of sensory neuronal variability takes a particular form. In Chapter 2, we empirically demonstrate the existence of structured noise correlations compatible with these predictions in area V1. However, we show quantitatively that these are mostly due to common inputs which change dynamically with the subject's task (i.e. the set of orientations being discriminated), strongly implying they have a central origin. We also show quantitatively that CP is mostly due to the portion of spike-count correlation structure that changes dynamically with the task. This finding is compatible with the view that signals related to a subject's perceptual choice are fed back to neurons in V1 supporting that choice, and that CP is mostly due to this feedback. In addition, we show that the structured spike-count correlations that change dynamically with the subject's task degrade the perceptual performance of an ideal observer of the V1 population. This is because they introduce fluctuations in the V1 representation that mimic the effect of changing the stimulus along

the dimension being discriminated. We discuss what our results imply about the nature of the common feedback inputs that appear to underlie the V1 spike-count correlation structure, and discuss possible feedforward effect the spikes generated by the implied feedback signals may have on the perceptual decision.

An important aspect of the approach taken in Chapter 2, shared with a number of past studies, is the attempt to control a subject's allocation of internal resources by changing the task they perform. It is difficult to know that this approach has worked, particularly in animal subjects. Objective measures of task strategy, like psychophysical reverse correlation, can be used to address this problem. In Chapter 3, I discuss results from a PRC analysis of the behavior of the monkey subjects performing the orientation discrimination task used in the experiments discussed in Chapter 2. This demonstrates a fundamental and previously undocumented challenge in attempting to use task instruction to alter a subject's allocation of internal resources and suggests strategies for overcoming this challenge.

## Chapter 2:

# Choice-Related Feedback Influences the Structure of Correlated Variability in Visual Cortex

### 2.1 Summary<sup>1</sup>

The responses of neurons in sensory cortex typically covary weakly even with fixed sensory input (Cohen & Kohn, 2011). This correlated variability can limit how much information about the outside world can be extracted from sensory neurons (Abbott & Dayan, 1999; Averbeck & Lee, 2006; Cohen & Maunsell, 2009; Moreno-Bote et al., 2014; Zohary et al., 1994). It is also thought to mediate the influence of firing rate fluctuations on perceptual decisions, in the form of choice-related activity (Choice Probability; CP; Britten et al., 1996; Haefner et al., 2013; Shadlen et al., 1996). These hypotheses rely on the common assumption that  $r_{sc}$  reflects unreliability in the sensory representation (Tolhurst et al., 1983). However, the origin of  $r_{sc}$  is poorly understood. It may partially reflect variation in signals that do not derive from the periphery (i.e. “top-down”, “feedback”) and which are not treated as noise in the sensory representation. Here we show, using array recordings from populations of neurons, that the structure of spike-count correlations in primary visual cortex (V1) of behaving monkeys changes systematically with task instruction. This structure implies variability in feedback related to choice, clarifying longstanding uncertainty about the origin of CP (Cumming & Nienborg, 2016; Nienborg et al., 2012). This effect of feedback also implies an impairment of perceptual performance, if it is subsequently read out as sensory evidence

---

<sup>1</sup> This chapter is based on a manuscript of the same title by Adrian Bondy and Bruce Cumming.

(Averbeck et al., 2006; Moreno-Bote et al., 2014; Ruff & Cohen, 2014). Crucially, as we show these correlations reflect signals related to the subject's awareness of the task context, they need not necessarily be treated as part of the sensory representation and therefore may not limit performance. Taken together, our results fundamentally change our understanding of the origin and implications of stimulus-independent variability in sensory brain areas, and demonstrate a need for models of perceptual decision making that include feedback onto sensory neurons.

## **2.2 Introduction**

Decisions about sensory input have been described as the outcome of pooling the activity of large populations of noisy sensory neurons (Averbeck et al., 2006; Haefner et al., 2013; Moreno-Bote et al., 2014; Nienborg & Cumming, 2010; Shadlen et al., 1996; Zohary et al., 1994). In the context of perceptual discrimination, in which a subject must classify a noisy stimulus, the effect of correlated variability on the performance of pooling models has been explored in detail (Cohen & Maunsell, 2009; Ecker, Berens, Tolias, & Bethge, 2011; Haefner et al., 2013; Moreno-Bote et al., 2014; Nienborg & Cumming, 2010; Pitkow, Liu, Angelaki, DeAngelis, & Pouget, 2015b; Shadlen et al., 1996; Sompolinsky, Yoon, Kang, & Shamir, 2001; Zohary et al., 1994). This work has given rise to two important conclusions: 1) noise correlations can dramatically impair performance (Abbott & Dayan, 1999; Averbeck et al., 2006; Moreno-Bote et al., 2014; Zohary et al., 1994) and 2) they give rise to a weak correlation between variability in single neurons and perceptual reports (Choice Probability; CP), consistent with the notion that CP observed in real neurons reflects the causal influence of correlated sensory

neuronal variability on perception (Haefner et al., 2013; Nienborg & Cumming, 2010; Shadlen et al., 1996). Both results depend crucially on how noise correlations are distributed amongst neuronal pairs with different stimulus preferences ( $r_{sc}$  structure).

These studies assumed that spike-count correlations in sensory neurons reflect variability in shared inputs that can lead to a less reliable representation of the sensory input, compared to the case of independent variability. However, while  $r_{sc}$  implies that neuronal pairs share common input, some of this may derive from feedback and could even be under voluntary control. If so, its role in perceptual judgments will depend on the poorly understood nature of the pathways conveying this feedback, and in particular their relationship to ongoing decision processes. In principle, it would be possible for the spikes generated by feedback to be identifying as unrelated to the sensory representation. Thus, a feedback origin for noise correlations in sensory neurons may undermine both of the above conclusions. For this reason, identifying the sources of correlated noise in sensory neurons remains a major problem in systems neuroscience.

### **2.3 Approach**

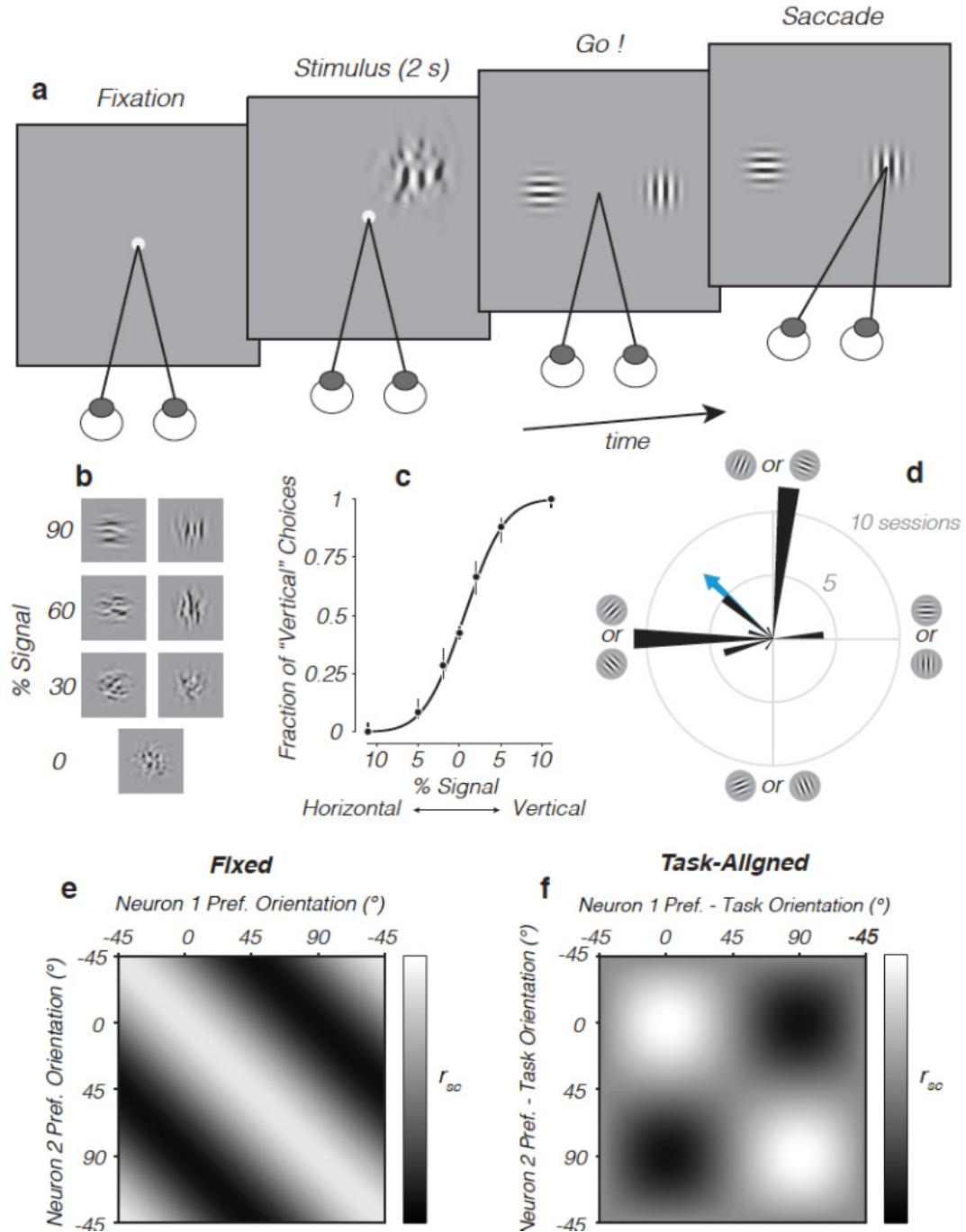
We sought to directly test the hypothesis that the structure of noise correlations in visual cortex reflects signals of central origin that vary during discrimination of a visual input. We measured  $r_{sc}$  amongst populations of neurons in the primary visual cortex (V1) of two macaque (*Macaca mulatta*) subjects, in contexts with changing task instruction and fixed retinal input. Subjects performed a coarse orientation task (Nienborg & Cumming, 2014; Fig. 2.1a), discriminating two orthogonal orientations (for instance vertical versus horizontal), which were fixed in a given recording session, but varied



between sessions (Fig. 2.1d). The stimulus (Fig. 2.1b) consisted of dynamic white noise, filtered to vary the range of orientations present. On “zero signal” trials, the stimuli contained a uniform distribution of orientations that was independent of the discriminanda orientations. Measuring  $r_{sc}$  across the zero-signal isolates any effect of task instruction on  $r_{sc}$ . In practice, we found that combining trials of all signal levels did not qualitatively alter our results (Fig. 2.7) and increased signal-to-noise, so this is what we report. The design of this task was in part conceived as an orientation analogue to the random-dot motion task (Newsome & Paré, 1988) used in many prior studies. Similar to that task, power in the stimulus is broadly distributed over the preferences of the neuronal population at low signal, while at high signals it is concentrated on a subpopulation.

We used psychophysical reverse correlation (PRC; Ahumada J., 1996; Nienborg & Cumming, 2014; Nienborg & Cumming, 2007) to ensure that animals internal task strategies correctly reflected the discriminanda orientations. We found that subjects required multiple days of training to update their task strategies after a change in the discriminanda, so recordings were only performed after enough retraining for their task strategy to be appropriate (Fig. 2.3). This represents a significant improvement over past studies (Cohen & Newsome, 2008) that have not used these behavioral measures to confirm that animals are doing the task as assigned. In Chapter 3, we present more detailed results of the PRC analysis.

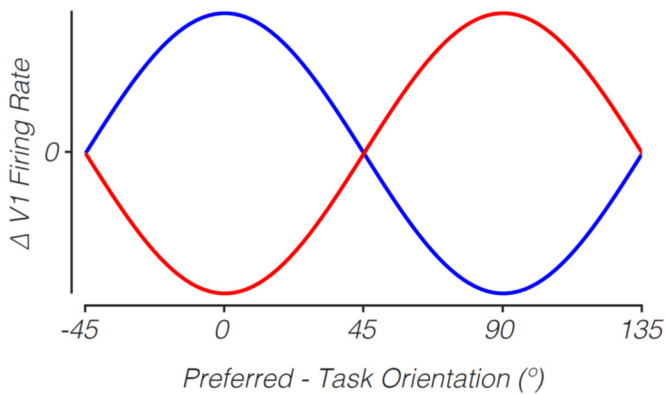
We replicated the previous finding (Nienborg & Cumming, 2014) of a significant mean CP in V1 for this task (population average CP of 0.54; Fig. 2.6e). Assuming linear pooling, this requires a distribution of  $r_{sc}$  such that neurons contributing to the same choice be more highly correlated, on average, than pairs contributing to opposite choices



**Figure 2.1. Orthogonal orientation discrimination task.** **a.** The task, with the cardinal orientations as discriminanda. **b.** Single frames of the stimulus, at various signal strengths for both discriminanda orientations. **c.** Example psychometric function for monkey ‘*lem*’. Solid line is a probit fit, and error bars are 95% confidence intervals, assuming choices are binomially distributed. **d.** The distribution of task orientations used across all 41 sessions in both subjects. Blue arrow is the mean resultant vector. **e.** Hypothetical “fixed”  $r_{sc}$  matrix for orientation that could generate CP for any pair of discriminanda. **f.** Hypothetical “task-aligned” matrix that would generate CP for the cardinal ( $0^\circ$  and  $90^\circ$ ) task configuration.

(Haefner et al., 2013; Nienborg et al., 2012; Nienborg & Cumming, 2010; Shadlen et al., 1996). To achieve this with a fixed correlation structure, for all task orientations, requires a relationship between  $r_{sc}$  and pairwise orientation preference like that shown in Fig 2.1e. This pattern can be described using a correlation matrix indexed by pairwise orientation preference, which takes the form of a diagonal ridge. This can also be viewed as an extrapolation of the previously measured relationships between tuning similarity and  $r_{sc}$  (Bair et al., 2001; Cohen & Kohn, 2011; Cohen & Maunsell, 2009; Kohn & Smith, 2005; Smith & Kohn, 2008; Zohary et al., 1994). An alternative possibility is that just those neurons involved in a given task are affected: correlations are high for pairs contributing to the same choice, low for pairs contributing to opposite choices, and unchanged for the rest. This produces a correlation matrix that resembles a square lattice (Fig. 2.1f). Such task-dependent changes would implicate feedback as the source of the structured correlations. This lattice-like pattern would suggest a common feedback input that increased the firing rate of neurons supporting one choice while decreasing the firing rate of neurons supporting the other choice on some trials, and having the reverse effect on other trials. This would be similar to a prediction based on fluctuations in feature-based attention between the two discriminanda orientations (Fig. 2.2; Ecker, Denfield, Bethge, & Tolias, 2016).

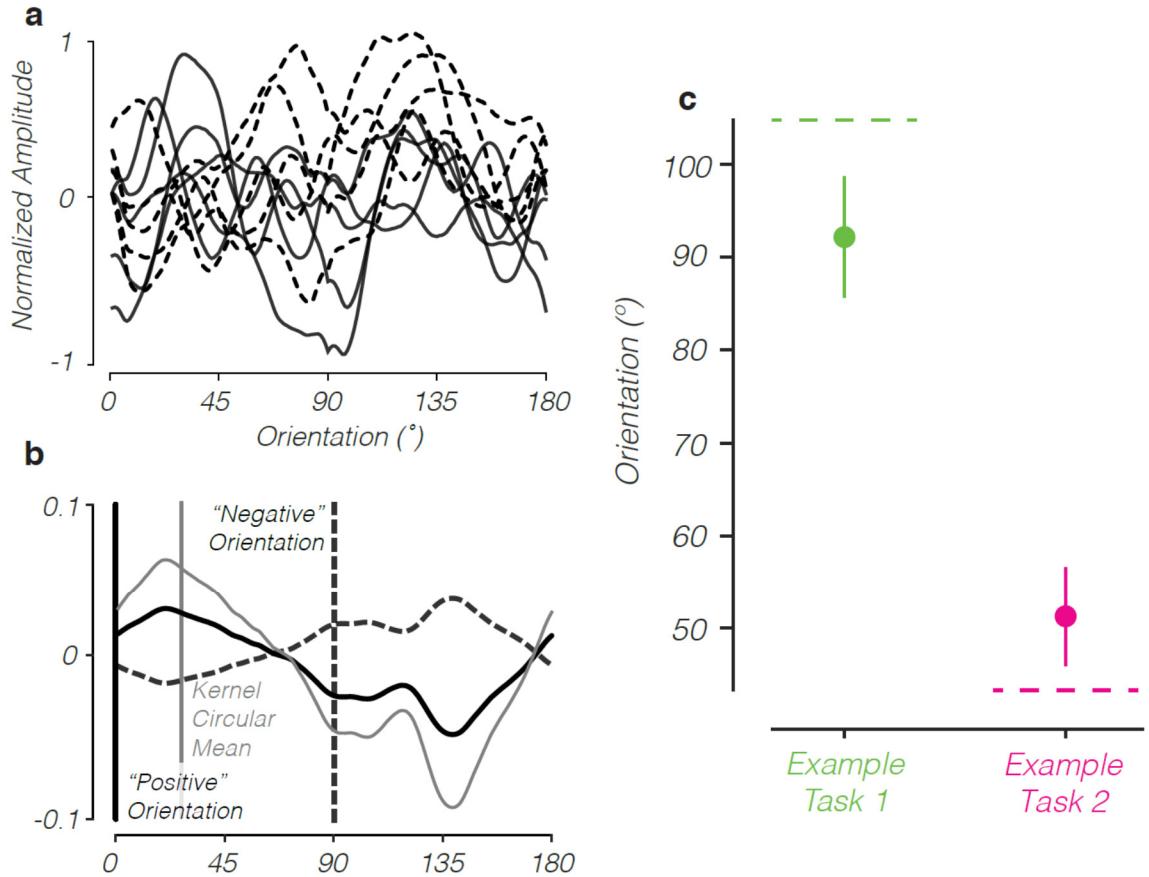
Prior studies have been unable to fully distinguish these possibilities because they did not measure the full correlation matrix (in the task-relevant stimulus dimension). Without directly manipulating the feedback by controlling the task the subject performs, it would be impossible to distinguish these. Indeed, a number of past studies have measured noise correlation structure simply as the relationship between  $r_{sc}$  and similarity



**Figure 2.2. Structure of feedback suggested by the presence of lattice-like correlation structure in V1.** The presence of task-aligned lattice structure in the V1 correlation matrix for orientation (Fig. 2.1f) would suggest a common input to the V1 population whose effect was to increase the firing of neurons supporting choice 1 (e.g. 0°) and decrease the firing rate of neurons supporting choice 2 (e.g.

90°) on some trials (blue curve) and on other trials have the reverse effect (red curve). In fact, the matrix in Fig. 2.1f is generated simply as the outer product of one of these sine waves. Variability in a common input of this sort correlates the spike counts of neurons preferring the same choice and decorrelates the spike counts of neurons preferring opposite choices. This is similar to a prediction based on fluctuations in feature-attention between the two discriminanda (Ecker et al., 2016).

in stimulus preference, independent of the task being performed (Bair et al., 2001; Cohen & Maunsell, 2009; Cohen & Newsome, 2008; Kohn & Smith, 2005; Smith & Kohn, 2008; Zohary et al., 1994). This is equivalent to estimating the diagonal marginal of these matrices. Since both matrices have similar diagonal marginals (correlation is higher near the diagonal, on average) such measurements make it impossible to distinguish these possibilities. While Cohen & Newsome (2008) did measure noise correlations while directly manipulating the task context, they used paired recordings, making it difficult to extrapolate the shape of the full correlation matrix from their data.



**Figure 2.3. Subjects' decision strategies change with task instruction.** **a.** To perform psychophysical reverse correlation, we first summarized each stimulus as the radial sum of its 2D Fourier amplitude spectrum, averaged across frames, to remove information about spatial frequency and phase. Examples shown here for several zero-signal trials, with dashed and solid lines corresponding to different choices. **b.** The psychophysical kernel (gray) is calculated as the difference between the two choice-conditioned average radial sums ("positive" minus "negative" choices). Axis labels and y-axis units are the same as in (a). Note that the peaks and troughs of the kernel are offset from the discriminanda orientations, indicating the subject is using a slightly misaligned task strategy (example session, monkey *lem*). **c.** Average kernel circular means ( $\pm 1$  bootstrap s.e., obtained by resampling sessions) for the two groups of sessions shown in Fig. 2.4. The average (nominally "positive") task orientations for the two groups are shown by the dashed lines. The two average kernel circular means are significantly different ( $p < 10^{-3}$ , bootstrap test) and closely aligned to their respective task orientations, demonstrating that subjects did use distinct decision strategies for the two groups of sessions.

## 2.4 Results

### 2.4.1 $R_{sc}$ Structure Changes with Task Instruction

Single sessions yielded between 2 and 18 orientation-tuned units, too few to measure the full “noise correlation matrix for orientation” for each session. Instead, we first separated all 41 recording sessions into two subsets employing similar task orientations, and measured the smoothed, average correlation matrix associated with each subset. We found that the two matrices differed dramatically, but were each consistent with a fixed, lattice-like pattern that was simply offset diagonally by a distance reflecting the task orientations (Fig 2.4a,b). To summarize this, we aligned the data from each session relative to the discriminanda, and then summed these aligned matrices, producing a task-aligned population correlation matrix for orientation. This matrix shows a clear lattice structure (Fig. 2.4e), demonstrating that  $r_{sc}$  structure changes with the subject’s task. This lattice structure was not present during separate blocks of trials during which the subject simply fixated passively but the same set of stimuli was shown (Fig. 2.11). This demonstrates that the dynamic changes depended on feedback deployed during active task engagement, and could not be explained, for instance, by an effect of task experience on local V1 circuitry. We also ruled out a number of other stimulus-driven effects as the source of the task-dependent noise correlations (see §2.7.1).

The task-aligned average matrix showed symmetries characteristic of a particular type of square lattice (p4m; Schattschneider, 1978; also see Fig. 2.4g and more detailed discussion in §2.7.5). The best fitting lattice of this type explained 86% of the variance in the observed data, significantly higher than could be explained for shuffled data (Fig. 2.4h). The shuffling procedure randomly translated each data point along the diagonal, ruling out the possibility that the structure is simply a noisy manifestation of a diagonal ridge.

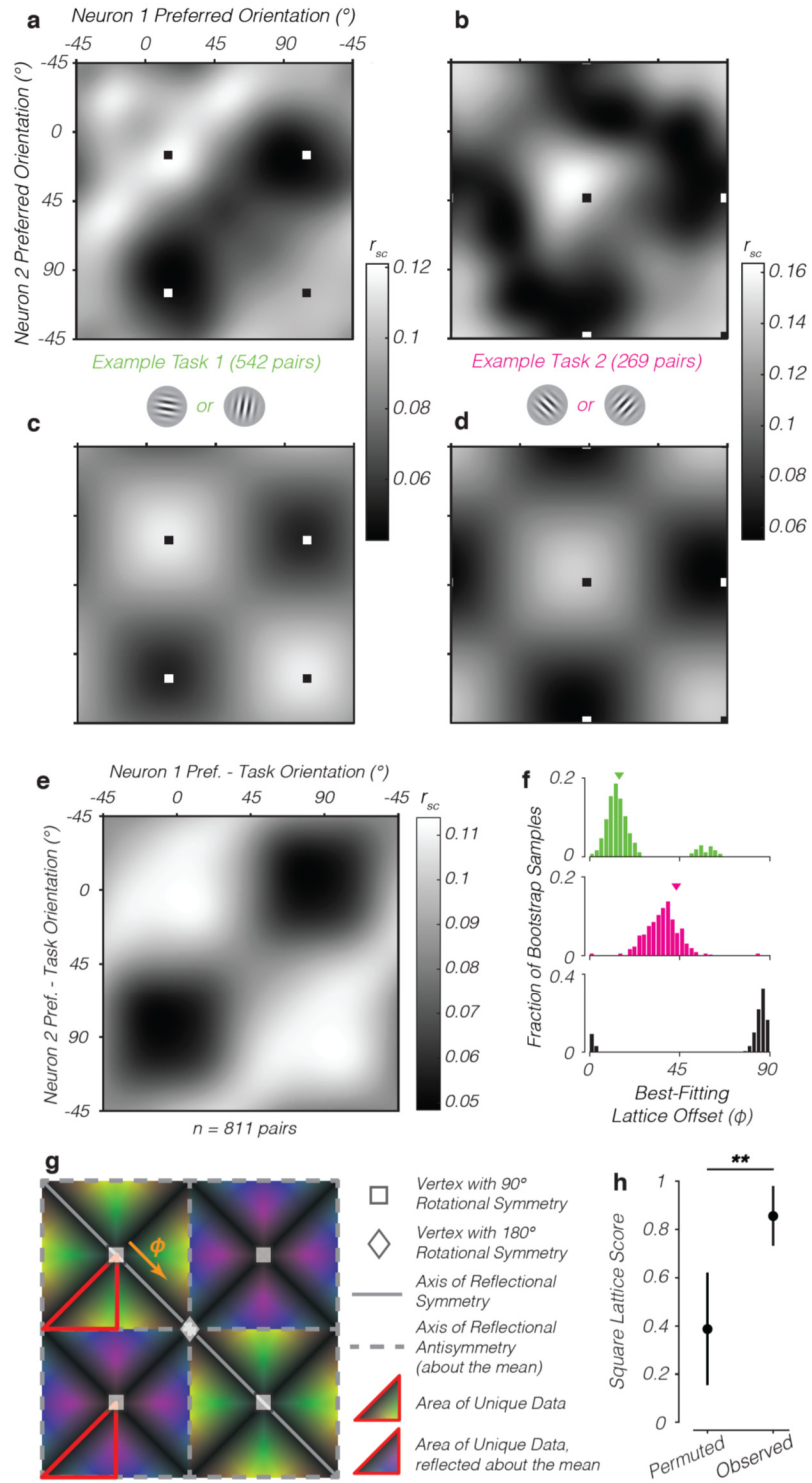


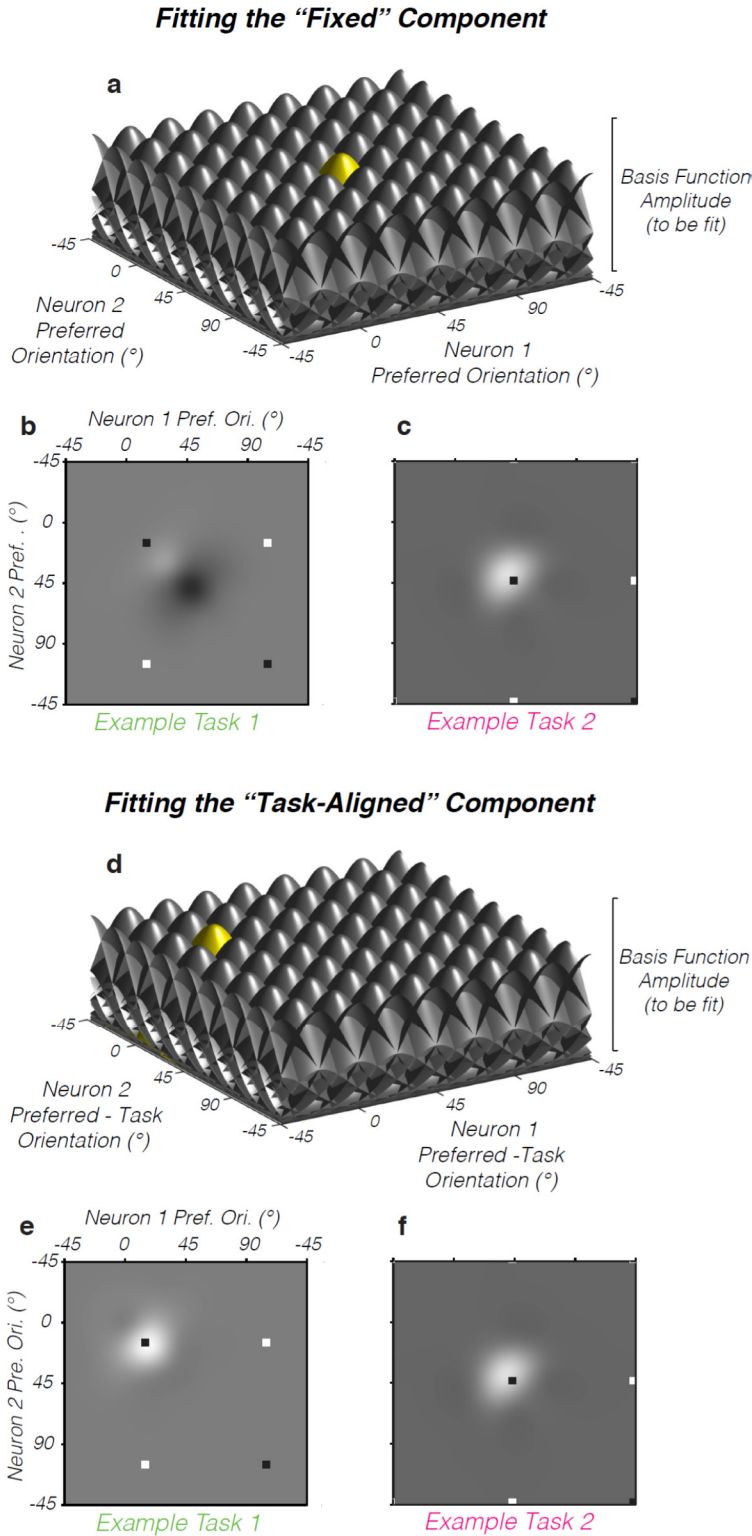
Figure 2.4 (caption on next page)

**Figure 2.4.  $R_{sc}$  structure in V1 depends on task instruction a-b**, Average  $r_{sc}$  matrices for orientation observed for two subsets of sessions with mean discriminanda given by the Gabor icons. These determine the within-pool (black squares) and between-pool (white squares) centers in the matrices. Peaks and troughs in observed  $r_{sc}$  are aligned to these locations. **c-d**. Square lattices fit to the data in (a) & (b) and obeying the symmetries in (g). A fixed diagonal offset ( $\phi$ ) was added to each lattice to optimize the fit. This value closely matched the discriminanda. **e**. Average  $r_{sc}$  matrix, combined across all sessions, with preferred orientations expressed relative to the discriminanda orientations for each session. (Because there were two discriminanda, there were two possible alignments for each session. We used the average.) **f**. Histograms of the best-fitting lattice offsets, obtained by resampling from the observed correlations. Color indicates which sessions were used, with the combined, task-aligned data in black. The best-fitting lattice offsets were not significantly different from the discriminanda (triangles). **g**. Lattice symmetries used to model the observed matrices. The lattice has circular boundary conditions and a  $180^\circ$  period. Color and saturation gradients indicate regions containing unique data. (For more detailed discussion, see §2.7.6). **h**. The combined data in (e) displayed a significantly higher lattice score than shuffled data (with no offset;  $p < 0.01$ ). Error bars indicate  $\pm 1$  bootstrap s.e.

Allowing the lattice vertices to be offset from the discriminanda orientations did not significantly improve the fit, indicating a good alignment between the instructed task and the dynamic noise correlation structure of V1 (Fig. 2.4f). The lattice structure implies variability in a feedback signal that increases the spike rate of neurons contributing to one choice while it decreases the spike rate of neurons contributing to the other choice, similar to theoretical predictions based on fluctuations in feature attention (Ecker et al., 2016) or a Bayesian prior (Haefner, Berkes, & Fiser, 2014).

We wondered whether aligning the data relative to the task obscured correlation structure that did not vary across sessions. To estimate the value of any fixed component in the correlation structure, we used multilinear regression to describe the  $r_{sc}$  values recorded across all sessions as the sum of two components: a “fixed” component and a “task-aligned” component. The fixed component was a correlation matrix for orientation that remained





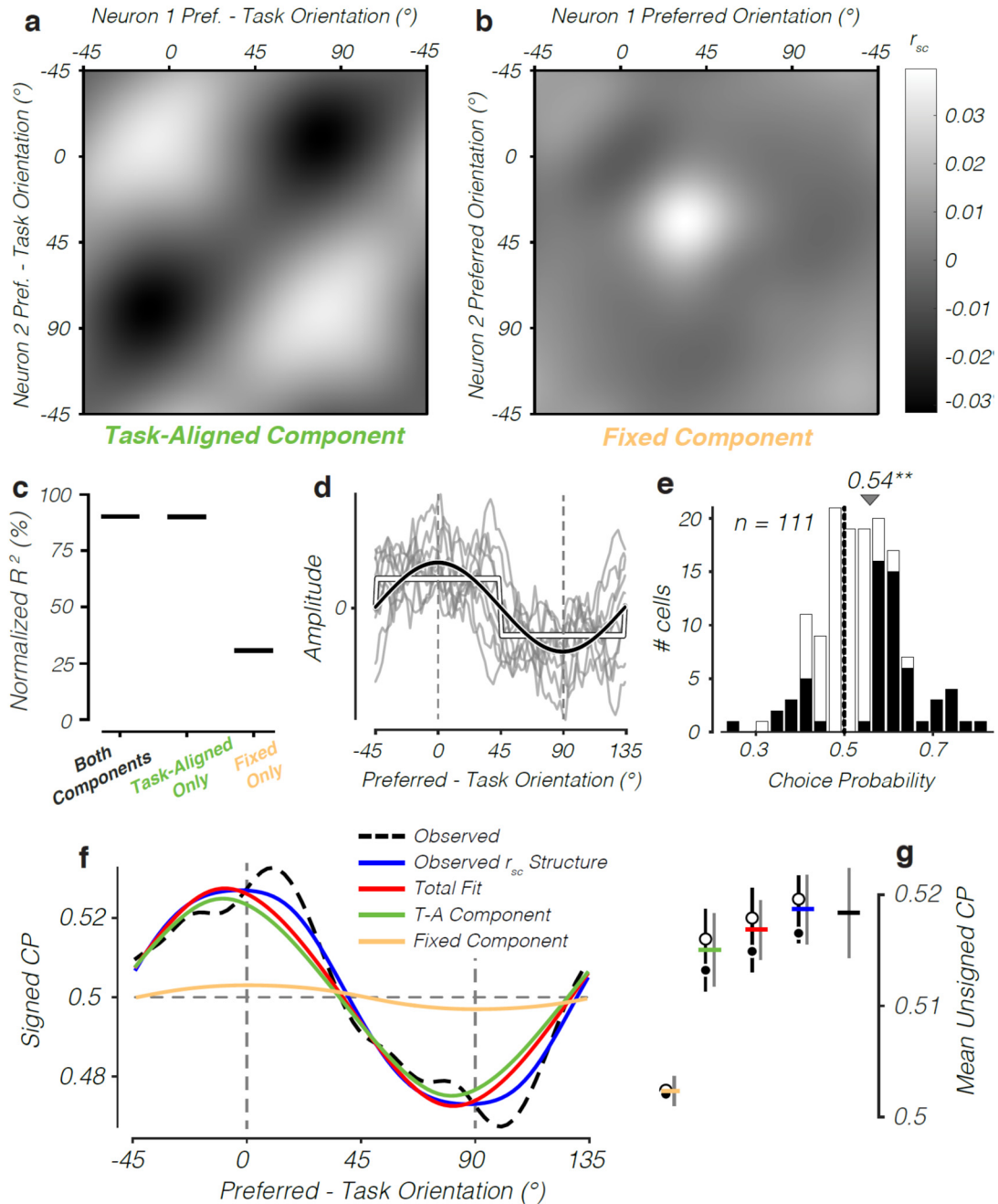
**Figure 2.5. Two-component multilinear regression model.** We used a multilinear regression model with two components (“fixed” and “task-aligned”) to explain the observed noise correlations. The components were parameterized as the sum of  $8 \times 8$  von Mises basis functions. The centers and dispersion of the basis functions were fixed. (Because the data are periodic, edge effects are visible in this illustration.) The model parameters to be fit were the amplitudes of the basis functions. **a.** The fixed component was a function of pairwise preferred orientation alone. Example basis function centered on  $(45^\circ, 45^\circ)$  in yellow. **b,c.** Observed correlation matrices reproduced from Fig. 2.4a,b for two subsets of sessions. The impact of the correlation values on the amplitude assigned to this basis function depends on their distance from the basis function center, illustrated by applying a contrast envelope to the data, centered on  $(45^\circ, 45^\circ)$ . Because the fixed component is invariant across sessions,

the task on a given session plays no role. We found that the “fixed” structure tends to be inconsistent across sessions, leading to little variance explainable by this component. **d.** The “task-aligned” component was parameterized in the same way as the “fixed”

component, except that it is a function of pairwise preferred orientation *relative to* the task. Example basis function centered on  $(0^\circ, 0^\circ)$  in yellow. In the task-aligned component,  $0^\circ$  and  $90^\circ$  always refer to the discriminanda orientations. **e,f.** Observed correlation matrices reproduced from Fig. 2.4a,b. The impact of the correlation values on the amplitude assigned to this basis function depends on their distance from the basis function center, illustrated by applying a contrast envelope centered on  $(0^\circ, 0^\circ)$  in the task-aligned space. Because the correlations tended to change dynamically with the task, this component explained the greater part of the variance in the data.

invariant across sessions. This was a function of pairwise orientation preference alone, and reflected aspects of the correlation structure that did not change with the subject's task. The task-aligned component was also fixed, but its position changed dynamically with the subject's task. This component was a function of preferred orientation relative to the task, and reflected correlation structure that changed dynamically. To confer smoothness on the components and to reduce the number of parameters in the model, each of these was fit as the sum of an  $8 \times 8$  array of von Mises (closely similar to wrapped Gaussian) basis functions (See Fig. 2.5 and Methods). The only free parameters were the amplitudes of the basis functions.

If the correlation structure changed completely with the task, the “task-aligned” component would explain all of the variance in the data. If the correlation structure was invariant across sessions, the “fixed” component would explain the variance. In fact, we found that the task-aligned component explained most of the variance (79%) and captured the lattice-like structure in the observed data (Fig. 2.6a). The fixed component had a markedly smaller amplitude, with a less organized structure (Fig. 2.6b). Removing the fixed component had little effect, while removing the task-aligned component dramatically impaired fit quality (Fig. 2.6c). This implies that the majority of structured noise correlations in V1 change dynamically with task instruction and are therefore top-down in origin.



**Figure 2.6. The Task-dependent component of  $r_{sc}$  can account for CP.** **a-b.** Components of observed  $r_{sc}$  structure, jointly estimated using multilinear regression. Average  $r_{sc}$  values are close to zero because the model also included a constant term, reflecting the mean of the entire population. **c.** Goodness-of-fit for the joint model and two single-component models. **d.** Examples from a large distribution of random read out weight profiles consistent with task performance, used to predict CP. Uniform and sinusoidal weight profiles are illustrated in white and black, respectively. **e.** Histogram of observed CPs. The mean was significantly above chance (bootstrap test, cell resampling,  $p < 0.01$ ). CPs that were independently significant ( $p < 0.05$ ; bootstrap test, trial resampling) are shown in black. **f.** Observed (dashed) and mean

predicted (solid) CP profiles, signed with respect to the choice associated with positive weights. Observed profile smoothed with a von Mises kernel that approximated a wrapped Gaussian with  $10^\circ$  s.d. The two components are the same as in (a) and (b). **g.** Mean unsigned CPs associated with the profiles in (f),  $\pm 1$  s.e. obtained from the distribution of read out weights (black bars) and by resampling from the data (gray bars). The predictions associated with the uniform and sinusoidal weight profiles in (c) are shown with white and black circles, respectively. Note that the mean observed CP is lower here than in (e) because all neurons are included, regardless of their selectivity in a particular task.

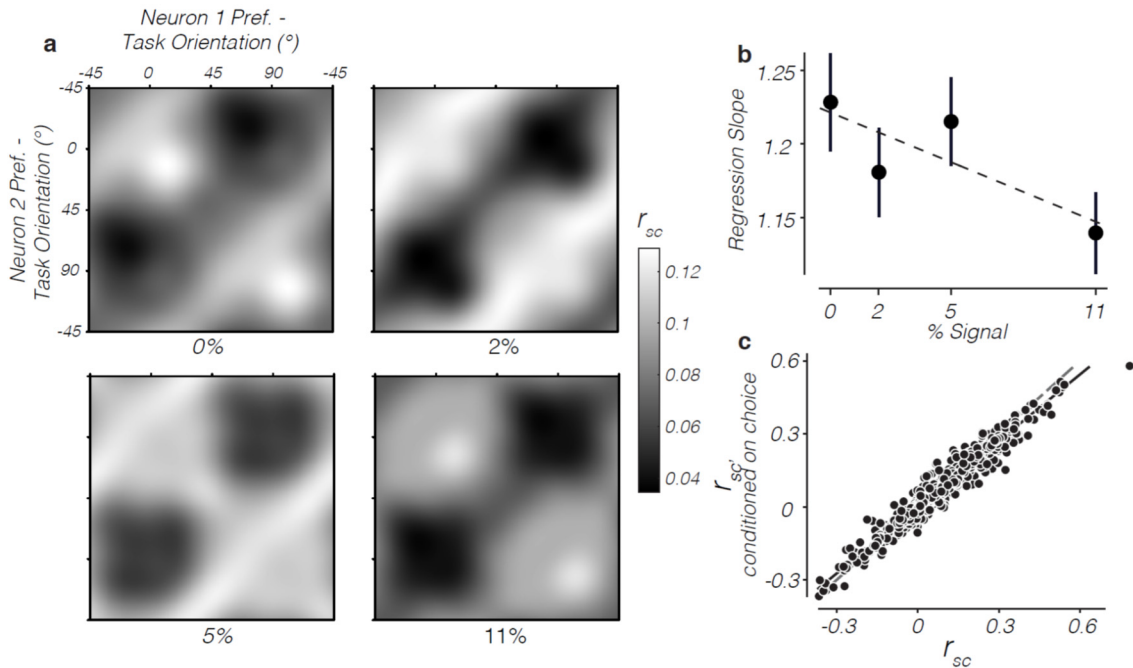
#### *2.4.2 Task-Dependent Changes in $R_{sc}$ Structure Relate to Perceptual Choice*

Several observations suggest that the signal producing  $r_{sc}$  structure is also related to the subject's choices. First, when we measured  $r_{sc}$  separately by choice, the amplitude of the structure was attenuated (Fig. 2.7c). Second, the amplitude was lower in high signal trials, where the presence of a strong signal inevitably reduces variability in choice (Fig. 2.7a,b).

We also found that the average, task-aligned matrix (Fig. 2.4e) was low dimensional, in the sense that it could be well approximated (86% variance explained) by the Gram matrix consisting only of its first eigenvector, significantly better than could be replicated with shuffled data (Fig. 2.8a). (Before calculating the eigenspectrum, we removed the matrix mean and normalized by its trace. This removed any flat eigenvectors reflecting mean correlation and ensured that the eigenvalues summed to 1. To correctly determine the estimation error of the eigenvectors given their arbitrary sign, we inverted the sign of resampled eigenvectors whose dot product with the observed eigenvector was negative.)

The low dimensionality implies that V1 population activity, along the dimension spanned by preferred orientation relative to the discriminanda, was driven largely by a single principal component. This principal component closely resembled a sinusoid

whose phase was aligned to the task (Fig. 2.8b), and can be interpreted as reflecting the effect on V1 firing rates of a task-dependent top-down input: on trials when the firing rates of neurons tuned for task orientation 1 were above average, the firing rates of neurons tuned for task orientation 2 were below average, and vice versa. This is similar to the pattern of common input used to generate the prediction of a task-aligned correlation matrix in Fig. 2.1f. The low dimensionality of the observed correlation matrix (in the



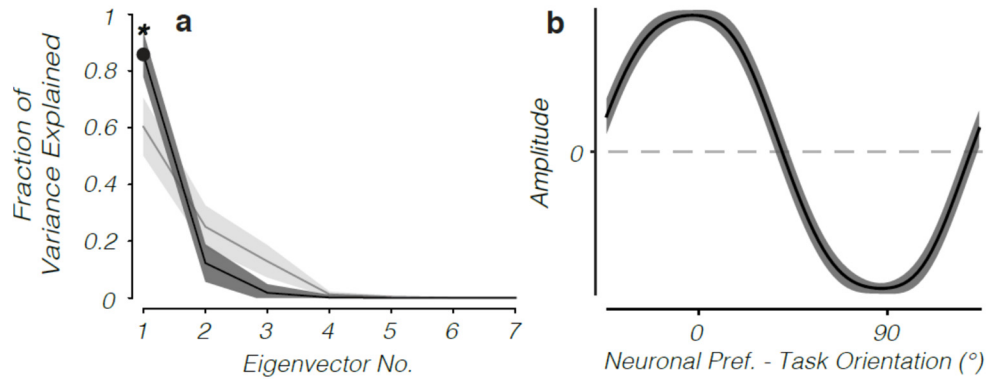
**Figure 2.7.  $R_{sc}$  structure depends on choice distribution.** **a.** The task-aligned noise correlation matrix for orientation, averaged across sessions (as in Fig. 2.4e), computed separately by signal level. The matrix associated with the 0% signal trials reflects task-dependent changes in noise correlation despite identical retinal input. A qualitatively similar task-aligned structure was apparent at non-zero signal levels. **b.** We quantified the slope of a type 2 regression of  $r_{sc}$  at each signal level against  $r_{sc}$  measured across all signal levels. We observed a significant inverse relationship ( $p < 0.01$ , bootstrap test) between this slope and signal level (error bars are  $\pm 1$  bootstrap s.e.). **c.** Scatter plot of  $r_{sc}$  across all trials against  $r_{sc}$  measured separately by choice and averaged. Identity line and type 2 regression line are in dashed gray and solid black, respectively. The slope is 0.97, significantly less than 1 ( $p < 0.05$ , bootstrap test), indicating a modest attenuation of the noise correlation structure, when measured across trials with no variability in choice.

space defined by preferred orientation relative to the task) implies that V1 population activity can be well described as a scalar multiple of this eigenvector. Given that the amplitude of correlation structure is related to variability in the subject's choices, we speculate that this scalar variable may be correlated with the decision variable itself.

We note that this finding is only partially related to more general attempts to understand the dimensionality of neuronal populations (for instance, Rigotti et al., 2013). Our analysis is specific to variability along the dimension defined by orientation preference. This allows us to identify the task-relevant source of variability, but ignores other aspects of V1 responses.

Next, we used the assumption of linear pooling to quantify the respective influence on choice of the task-aligned and fixed components of  $r_{sc}$  structure identified using multilinear regression. Because the readout weights are not directly observable, we randomly generated a large distribution of weight profiles that could support performance of the task (Fig. 2.6d) and used them to generate a distribution of predicted CPs. The uncertainty in CP prediction introduced by variability in the hypothetical read out weights was no greater than uncertainty in our measurements of the correlations (Fig. 2.6g), so our conclusions do not depend on knowing these readout weights. The relationship between predicted CP and orientation was similar to our data (Fig. 2.6f) - the first quantitative demonstration that the linear pooling model is consistent with experimentally observed measurements of  $r_{sc}$  and CP. However, the task-aligned component alone was sufficient to account for most of observed CP (82%), while the fixed component could explain only a small fraction. This is, in a sense, not surprising, given that we hypothesized the existence of a task-aligned lattice structure (Fig. 2.1f) based on the

observation of CP in the first place. And as we also showed, the common input suggested by that correlation structure increases the firing rate of neurons supporting one choice while decreasing the firing rate of neurons supporting the other choice, and vice versa. This is precisely the structure of trial-to-trial variability in population activity that would generate CP under the linear pooling model.



**Figure 2.8. Task-aligned  $r_{sc}$  structure is low dimensional.** **a.** Normalized eigenspectra associated with the average, task-aligned matrix in Fig. 2.4e and shuffled data, in black and gray respectively. Shuffling procedure is the same as in Fig. 2.4h. Error bars are  $\pm 1$  s.e. Significantly more of the variance in the observed matrix could be explained by its first eigenvector ( $p < 0.05$ , bootstrap test). **b.** The first eigenvector of the matrix in Fig. 2.4e,  $\pm 1$  bootstrap s.e.

Another possibility is that the decision may ignore the effect of feedback on V1 entirely. Given that this feedback has an effect on V1 firing rates that follows a sinusoidal pattern whose amplitude may be related to a simple scalar value like a decision variable, we speculate that it would not be difficult for downstream decoding processes to use knowledge of this scalar value to discount the portion of V1 firing rates due to feedback. If this obtained, CP would not reflect an effect of neuronal activity on choice, but would reflect the effect of choice-related feedback on V1 neurons, which is subsequently ignored in decoding. Alternatively, the feedback we observe may both convey

information about the subject’s choice as well as contribute to changing it—a combination of both interpretations. This could be a useful mechanism of ensuring perceptual stability, and is similar to a theoretical implementation of probabilistic inference (Haefner et al., 2014; further discussion in §4.1).

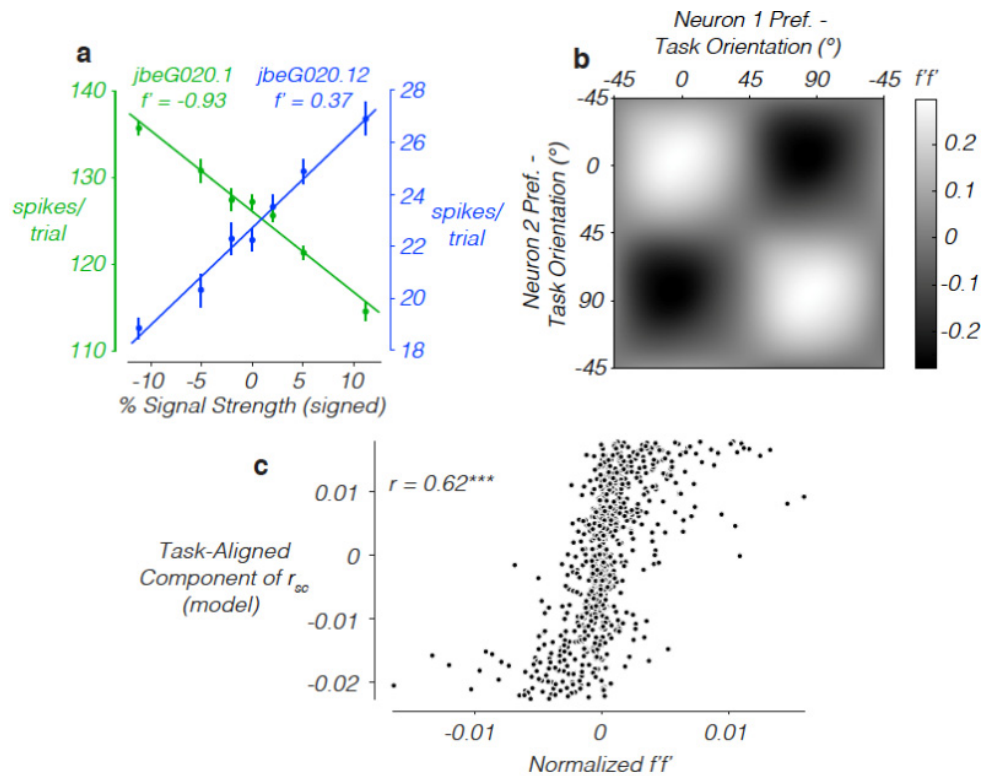
#### *2.4.3 Task-Dependent Modulation of $R_{sc}$ Structure Introduces “Differential” Correlations*

Lastly, we found that the task-dependent modulation of  $r_{sc}$  would impair the performance on the task of an ideal observer of the V1 population. This is because the lattice-like  $r_{sc}$  structure matches so-called “differential” correlations (Moreno-Bote et al., 2014; compare Fig. 2.9b to Fig. 2.4e), which place an upper limit on the information capacity of a sensory neuronal population (assuming linear pooling). This is because the  $r_{sc}$  structure describes variability in the V1 population between a state in which neurons preferring one choice are firing more than average and neurons preferring the other choice are firing less than average: precisely the structure of variability that mimics changes in the stimulus along the dimension being discriminated.

While there remains no proven method of identifying “differential” correlations given the limitations of real data, a range of correlation structures that approximate them can affect the amount of stimulus information decodable from finite populations (Moreno-Bote et al., 2014). Consequently, the task-dependent modulation of  $r_{sc}$  structure we observe would have a detrimental effect on perceptual performance. Quantitative studies would need to be undertaken to assess the magnitude of this effect. Nonetheless, the apparent contamination of the sensory representation by feedback is a novel finding, and contrasts with the reported beneficial effects of spatial attention on information



capacity (Cohen & Maunsell, 2009; Mitchell et al., 2009). However, our finding that  $r_{sc}$  reflects variability in centrally-generated inputs highlights an important problem in making inferences about the effect of  $r_{sc}$ : downstream areas may be able to discount the influence of feedback on sensory neurons. Consequently, measuring the information capacity of sensory neurons may require knowledge of the state of internal signals generated elsewhere in the brain.



**Figure 2.9. The task-dependent component of  $r_{sc}$  resembles stimulus-driven (“differential”) correlations.** **a.** An example pair of neuronal responses (mean  $\pm$  1 s.e.) to the task stimuli at various signal strengths. The product ( $f'f'$ ) of the regression line slopes (solid lines) is used as a measure of similarity in task-tuning. **b.** The matrix of average  $f'f'$  values, as a function of pairwise orientation preference relative to the task. This structure is extremely similar to the structure of  $r_{sc}$  we observed during task performance. **c.** Scatter plot of the task-aligned (putatively top-down) component of  $r_{sc}$  (Fig. 2.6a) against normalized  $f'f'$  values (see Methods), plotted for each recorded neuronal pair.

## 2.5 Discussion

In this chapter, I presented evidence showing that the distribution of stimulus-independent covariability in V1 neurons changes dynamically with task instruction during performance of a 2AFC orientation discrimination task, despite fixed retinal input. We observed a lattice-like noise correlation structure (in the space defined by pairwise orientation preference), such that  $r_{sc}$  was high for pairs preferring the same discriminandum orientation and low for pairs preferring opposite discriminanda orientations. The fact that this structure changed dynamically with the task implies that it primarily reflects input from central sources, not noise in sensory afferents. Indeed, we could not detect a component of noise correlation structure that remained fixed across task contexts. We found that the task-dependent source of noise correlations predicted CP that was consistent with our data, under the assumption that the spikes generated by feedback are read out as though they were sensory evidence. This is the first experimental support for the predicted relationship of CP and  $r_{sc}$  in the traditional feedforward framework. (While prior studies did address the relationship between these, the fact that the full correlation matrix was not measured does not provide a full test for the model).

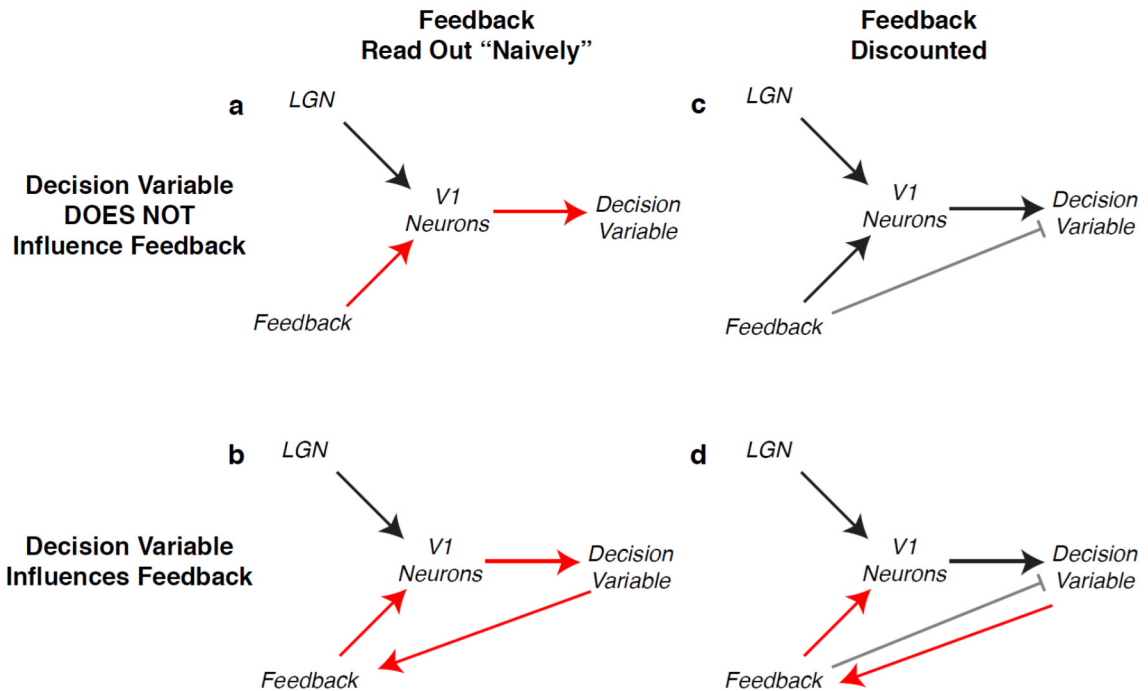
We also observed that the structured noise correlations introduced through feedback are closely similar to “differential” correlations: those which resemble stimulus fluctuations along the task axis. This implies that, if this variability is naively read out as sensory evidence, it would effectively “contaminate” the sensory representation, and reduce the psychophysical performance of the subject. This would be a significant departure from current hypotheses about the role of feedback to sensory areas, which

typically take the view that they function to improve the quality of the relevant sensory representation (Maunsell & Treue, 2006; Reynolds & Chelazzi, 2004).

Several models for the decision making process would be consistent with our data. These can be categorized based on two factors: 1) whether downstream computations treat the spikes that arise through feedback as sensory noise (i.e. are “naively” read out); and 2) whether the feedback signal that introduces the structured noise correlations is itself correlated with the state of the decision process (that is, whether the amplitude of the sinusoidal modulation of firing rates relates to the decision variable).

The first possibility (Fig. 2.10a) is that the feedback signals *are* naively read out and are not correlated with the state of the decision variable. In this scheme, CP arises entirely as a causal effect of the feedback on the sensory neurons, which is subsequently read out to inform choice. In a sense, this scheme is compatible with the original proposal for CP: that it reflects the causal effect of sensory neuronal variability on choice. While this was originally thought to be subserved by noise inherent in the sensory representation, to be consistent with our data this variability now derives from feedback. The origin of feedback, under this interpretation, while clearly constrained by the task being performed, would be unknown to the observer and not related to choice directly.

The second possibility (Fig. 2.10b) is that both are true, such that the state of the decision process is correlated with the sensory neurons via feedback, and the resulting sensory neuronal variability influences the decision process. This creates a self-reinforcing loop. As a consequence, CP emerges both as a causal effect of sensory neurons on choice and an effect of choice on sensory neurons.



**Figure 2.10. Models of perceptual decision making consistent with our results.** Schematic illustrations of the potential role of feedback in perceptual decision making. In all subpanels, the red lines indicate the pathways that generate CP. (The afferent input from LGN never contributes to CP because it cannot generate noise correlations with the appropriate structure given our results.) **a.** CP emerges through the causal influence of feedback on sensory neurons, which is read out as sensory evidence. Because feedback is not influenced by the decision variable, CP does not reflect the effect of choice on sensory neurons. **b.** CP emerges through the correlation between the decision variable and the feedback on the sensory neurons *and* the effect of sensory neurons on the decision variable, generating a self-reinforcing loop. (The correlation between the decision variable and feedback is illustrated as though the decision variable directly influences feedback, one plausible way this could arise. Another possibility is that the decision variable and the feedback share a common input). **c.** Feedback is discounted and is also not correlated with choice. As a result, no CP is observed in the sensory neurons. This would be inconsistent with the experimental data. **d.** CP is generated solely through the effect of choice-related feedback on the sensory neurons. The effect is subsequently discounted by directly comparing the feedback and the sensory neuronal activity as part of the readout. (This comparison is implemented schematically by adding a direct inhibitory influence of the feedback on the decision). As a result, there is no detrimental effect of the introduction of “differential” correlations.

In these first two interpretations, the naïve decoding of feedback implies that performance is degraded by the introduction of “differential” correlations. Under the third scheme (bottom right, Fig. 2.10d), the state of the decision process influences feedback, but this component of the sensory neuronal response is ignored by the decoder. This can be implemented simply by comparing sensory neuronal activity with the feedback signal directly (here illustrated by introducing a direct inhibitory effect of feedback on the decision). Under this scheme, the structured noise correlations introduced by feedback have no influence on choice. This has two consequences. First, observed CP is exclusively an effect of feedback, and second, the “differential” structure of the correlations has no impact on performance. This scheme would have the curious consequence of introducing feedback to sensory neurons only to ignore it downstream, and so for this reason it seems unlikely.

In the final scheme (top right, Fig. 2.10c), feedback is ignored and is also not itself correlated with choice. This implies that the noise correlations which could influence choice are only those derived from the sensory input. In our experiment this corresponds to the fixed component of the correlation matrix. We showed in Fig. 2.6 that this can only generate very small CP. Because the feedback is not influenced by choice, there is also no feedback source of CP. Consequently, this scheme predicts no CP, and is therefore inconsistent with our data.

We argue that the first (purely feedforward) scheme (Fig. 2.10a) is problematic for several reasons. First, under this view, CP entirely reflects the influence of sensory neuronal variability on choice. Therefore, the magnitude of CP must correlate with the influence of sensory evidence on the animal’s choices (i.e. the amplitude of the

psychophysical kernel). However, we found, similar to past studies, that CP stayed constant or even increased over the course of the trial (see §2.7.3; Nienborg & Cumming, 2009; Shadlen et al., 1996) while subjects' weighting of sensory evidence tends to decrease in fixed-duration trials (Kiani et al., 2008; Hendrikje Nienborg & Cumming, 2009). This mismatch suggests that CP results from more than the causal influence of sensory neuronal variability on choice.

Second, this interpretation requires that the state of feedback signals during perceptual decision making not be influenced by choice itself. Evidence from other studies involving perceptual discrimination suggests that this is not the case. One strong line of evidence for this is the pronounced CPs observed in studies involving perceptually bistable stimuli (Dodd et al., 2001; see discussion in §1.3.1). In these tasks, zero-signal trials do not look ambiguous. Rather the percept fluctuates strongly between two alternative interpretations with equal probability. The perceptual stability elicited by these stimuli implies a relatively weak effect of sensory neuronal variability on choice, suggesting that this cannot be the origin of the large CPs. A more parsimonious explanation is that CP relates to variability in feedback related to the subject's strong alternations in perceptual state. We also found evidence that the feedback inputs that manifest as structured spike-count correlations in V1 appear to scale with variability in subject's choices.

Thus, we conclude that CP at least partially derives from the fact that the common feedback input generates the structured correlations we observe in V1 is correlated with the state of the subject's decision variable. However, we cannot say what downstream effect this feedback has. Either 1) it is ignored by the decoder, in which case

CP is entirely due to the effect of choice on sensory neurons and there is no detrimental effect of the “differential” structure (as in Fig. 2.10d) , *or* 2) spikes generated by feedback is read out naively as sensory evidence, in which case CP is the result of a self-reinforcing loop of feedback and feedforward signals, and the structure of the resulting noise correlations would appear to limit the information capacity of the population (as in Fig. 2.10b). In either case, the relationship between sensory neuronal variability and the “noise” that limits perceptual performance is a complicated affair indeed. It is not well experimentally constrained, particularly in light of our results indicating  $r_{sc}$  relates to poorly understood feedback signals. This poses serious problems for experimental attempts to measure the information capacity of sensory neurons without knowledge of the state of internally-generated feedback signals. This also undermines attempts to use CP as a reflection of the causal influence of sensory neurons on choice.

Finally, we consider how the results may generalize to other tasks and brain areas. One reason for choosing to perform our experiment in V1 is because top-down effects have been reported to be weaker in early visual cortex (Maunsell & Cook, 2002; McAdams & Maunsell, 1999; Mehta, Ulbert, & Schroeder, 2000; Reynolds & Chelazzi, 2004). Therefore, it seems reasonable to conclude from our results that CP is also generated by feedback later in the visual hierarchy, where feedback effects are typically stronger. Indeed, a more limited version of our results, using paired recordings, has previously been made in MT using a motion discrimination task (Cohen & Newsome, 2008). Their measurements of noise correlations are consistent with a dynamic structure similar to our data, but with greater amplitude. Along the same lines, preliminary data from population recordings in area MT during a cylinder direction discrimination task

indicate task-dependent structured noise correlations similar to those reported here, but a great deal larger in amplitude (Incheol Kang and Bruce Cumming, *personal communication*), consistent with a feedback origin of the much larger CPs reported previously in MT using the same task (Dodd et al., 2001; Krug et al., 2004; Parker et al., 2002). Nonetheless, further experiments would need to be performed to directly test the replicability of these results in other brain areas and using other experimental paradigms.

## **2.6 Methods**

### *Electrophysiology*

We recorded extracellular spiking activity from populations of V1 neurons in two awake, head-fixed rhesus monkeys (*Macaca mulatta*). Both were implanted with a head post and scleral search coils under general anaesthesia (Judge, Richmond, & Chu, 1980). In Monkey ‘*lem*’, a recording chamber was implanted over a craniotomy above the right V1 operculum, as described previously (Cumming & Parker, 1999), by which we introduced linear microelectrode arrays (U- and V-probes, Plexon; 24-contacts, 50 or 60  $\mu\text{m}$  spacing) at an angle approximately perpendicular to the cortical surface with a custom micro-drive. We positioned the linear arrays so that we roughly spanned the cortical sheet, as confirmed with current-source density analysis, and removed them after each recording session. In monkey ‘*jbe*’, a planar “Utah” array (Blackrock Microsystems; 96 electrodes 1mm in length inserted to target supragranular layers, 400  $\mu\text{m}$  spacing) was chronically implanted, also over the right V1 operculum. All procedures were performed in accordance with the US Public Health Service Policy on the humane care and use of



laboratory animals and all protocols were approved by the National Eye Institute Animal Care and Use Committee.

Broadband signals were digitized at 30 or 40 kHz and stored to disk. Spike sorting was performed offline using custom software in MATLAB. First, spikes were detected using a voltage threshold applied to high-pass filtered signals. Next, triggered waveforms were projected into spaces defined either by principal components or similarity to a template. Clusters boundaries were finally estimated with a Gaussian mixture model, and then rigorously verified and adjusted by hand when needed. In the linear array recordings, spike sorting yield and quality was substantially improved by treating sets of three or four neighboring contacts as “n-trodes”. As this was not possible with the Utah array, we excluded pairs of neurons recorded on the same electrode to avoid contamination by misclassification. Neurons from separate recording sessions were treated as independent. To reduce the possibility that a single neuron from the Utah array contributed to two datasets, we included only sessions that were separated by at least 48 hours (with a median separation of 5 days). We excluded from analysis those neurons whose mean evoked firing rate did not exceed 7 spikes/second.

### *Visual Stimuli*

All stimuli were presented binocularly (at zero disparity) on two gamma-corrected cathode ray tube (CRT) monitors viewed through a mirror haploscope, at 85 or 100Hz. The monitors subtended  $24.1^\circ \times 19.3^\circ$  (1280x1024 pixels). The stimuli presented during performance of the discrimination task consisted of bandpass filtered dynamic white noise, as described previously (Nienborg & Cumming, 2014). Stimuli were filtered in the Fourier domain with a polar-separable Gaussian. The peak spatial frequency was

optimized for the recorded neuronal population (1 and 4 cpd medians for '*lem*' and '*jbe*', respectively) while the peak orientation could take one of two orthogonal values the animal had to discriminate in a given session. The angular s.d. of the filter modulated the orientation bandwidth (i.e. task difficulty) and was varied trial to trial. A 2D Gaussian contrast envelope was applied to the stimulus so that its spatial extent was as small as possible while still covering the minimum response fields of the neuronal populations being recorded. The median envelope s.d. was  $0.6^\circ$  for both animals. The median stimulus eccentricity was  $5.4^\circ$  for '*lem*' and  $0.5^\circ$  for '*jbe*'. In Fig. 2.1, we quantify orientation bandwidth as % signal strength. This was calculated as  $100 * R$ , where  $R$  is the length of the resultant vector associated with the angular component of the stimulus filter. A two-pass presentation procedure was used. Each instance of a stimulus (generated with a given noise seed) was shown twice. This allowed us to control for any effect of fluctuations in the stimulus on firing rate (see §2.7.1 and Fig. 2.15).

We estimated the preferred orientation of the neurons in separate blocks of trials, using 400-ms presentations of the following types of stimuli, presented at a range of orientations: 1) an orientation narrowband version of the stimulus described above ( $10^\circ$  angular s.d.); 2) sinusoidal gratings; and 3) dynamic 1D noise patterns (random lines). The preferred orientation of a neuron was calculated as the circular mean of its orientation tuning curve. From among the set of tuning curves elicited by the different stimulus types described above, we chose as the final estimate of preferred orientation the one with the smallest standard error, obtained by resampling trials. We excluded from further analysis all neurons where this exceeded  $5^\circ$ . On a subset of sessions, we also used these orientation-tuning blocks to present examples of the zero-signal orientation-filtered

noise stimuli. These were presented at the same location and size as during task performance, allowing us to calculate noise correlation structure in the absence of task engagement but identical retinal input.

### *Orientation Discrimination Task*

The animals performed a coarse orientation discrimination task using orientation-filtered noise stimuli, described previously (Nienborg & Cumming, 2014). To initiate a trial, the subject had to acquire a central fixation square. After a delay of 50ms, the stimulus appeared for a fixed duration of 2 seconds. The trial was aborted if the subject broke fixation at any point during the stimulus presentation, defined as either 1) making a microsaccade covering a distance greater than a predefined threshold (typically  $0.5^\circ$ ) or 2) a deviation in mean eye position from the center of the fixation point of more than a predefined threshold, typically  $0.7^\circ$ . At the end of the stimulus presentation, two choice targets appeared. These were Gabor patches of  $2\text{-}3^\circ$  in spatial extent, oriented at each of the two discriminandum orientations. The locations of the choice targets depended on the task. For discriminanda pairs near horizontal and vertical ( $-22.5^\circ - +22.5^\circ$  and  $67.5^\circ - 112.5^\circ$ ), the choice targets were positioned along the vertical meridian, at an eccentricity of about  $3^\circ$ , with the more vertically-oriented target appearing always in the upper hemifield. For orientation pairs near the obliques ( $22.5^\circ - 67.5^\circ$  and  $112.5^\circ - 157.5^\circ$ ), the choice targets were positioned along the horizontal meridian, at the same range of eccentricities, with the smaller of the two orientations always appearing in the left hemifield. (We use the convention that horizontal is  $0^\circ$  and that orientation increases with clockwise rotation.) To penalize random guessing, the amount of liquid reward delivered after correct choices was doubled with each consecutive correct choice,

up to a maximum of four times the initial amount. Since our hypothesis related to an effect of task engagement on neuronal activity, we applied a behavioral criterion to our data, excluding sessions where the subject's psychophysical threshold (defined as the signal level eliciting 75% correct performance) exceeded 14% signal.

### *Psychophysical Reverse Correlation*

We performed PRC to objectively measure the weights subjects applied to different stimulus orientations to make their choices. To do this, we first summarized the stimulus on each trial as the radial sum of its 2D Fourier amplitude spectrum, averaged across frames, to remove information about spatial frequency and phase. Psychophysical kernels were calculated as the difference between the two choice-conditioned radial sums. This was performed separately for each signal level and then the resulting kernels were averaged.

### *Noise Correlation Measurements*

Noise correlations were calculated as the Pearson correlation between spike counts, counted over the 2 second duration of the stimulus, with a 50ms delay to account for the typical V1 response latency. These spike counts were first z-scored separately for each experimental block (typically a set of 100-200 trials lasting about 10 minutes) and each stimulus condition. This normalization within blocks removes correlations related to long-term firing rate nonstationarities. Normalizing each stimulus condition allowed us to combine trials at different signal levels without introducing correlations related to similarity in stimulus response. We used a balanced z-scoring method proposed recently to prevent bias related to differences in choice distributions across signal levels (Kang &

Maunsell, 2012). We excluded pairs that were not simultaneously isolated for at least 25 trials (total across signal levels).

A main goal of the study was to measure how noise correlation varies with pairwise orientation. We describe this relationship as a smoothed function estimated from measures of  $r_{sc}$  combined across multiple recording sessions, which we then sampled discretely with  $1^\circ$  resolution. The smoothed estimates were obtained using kernel smoothing. We used a bivariate von Mises kernel, with zero correlation and equal dispersion in both dimensions. A point in an estimated correlation matrix for orientation  $\mathbf{C}$  was given as:

$$\mathbf{C}(x, y) = \tanh \frac{\sum_{i=1}^n z_i K(x, y, \theta_i, \phi_i)}{\sum_{i=1}^n K(x, y, \theta_i, \phi_i)}, \text{ where } K(x, y, \theta_i, \phi_i) = e^{\kappa (\cos(\theta_i - x) + \cos(\phi_i - y))} \quad (2.1)$$

$z_i$  is the  $i^{\text{th}}$  (Fisher z-transformed) noise correlation measurement,  $\theta_i$  and  $\phi_i$  are the preferred orientations of the  $i^{\text{th}}$  pair, and  $\kappa$  is the von Mises dispersion parameter. We set  $\kappa = 1.3\pi$ , yielding a smoothing kernel closely approximating a bivariate wrapped Gaussian with  $15^\circ$  s.d. We also estimated the noise correlation matrix in a task-aligned coordinate frame (e.g. Fig. 2.4e), for which the preferred orientations of the  $i^{\text{th}}$  pair relative to the task orientation were used for  $\theta_i$  and  $\phi_i$ . Since there were always two orthogonal task orientations, we averaged across both choices of coordinate frame, such that  $\mathbf{C}(x, y) = \mathbf{C}(x + 90^\circ, y + 90^\circ)$ . All angular quantities were doubled, as orientation has a period of  $180^\circ$ .

### *Square Lattice Approximations*

We observed that the task-dependent noise correlation structure resembled a square lattice. In other words, it was characterized by a set of identity-preserving transformations (i.e. a symmetry group), schematized in Fig. 2.4g. We wanted to identify the lattice  $\hat{\mathbf{X}}$  obeying these symmetries that best described a given smoothed matrix  $\mathbf{X}$ . The problem of identifying  $\hat{\mathbf{X}}$  is easily decomposed into a series of independent problems for each element of  $\hat{\mathbf{X}}$ :  $\hat{x}_k = x_i + \varepsilon_i$ ,  $\forall x_i$  that are those elements of  $X$  mapped onto  $\hat{x}_k$  by the symmetry group. Assuming normal error, we can easily find the maximum likelihood estimate of  $\hat{x}_k$  by taking the sample means of the  $x_i$ 's. To include the axis of anti-symmetry about the mean, we first subtracted the mean of the matrix  $\mathbf{X}$  and then inverted the sign of the data in the inverted regions. We then inverted the sign of these regions again after arriving at  $\hat{\mathbf{X}}$ .

### *Regression Model*

We used a multilinear regression model to identify “fixed” and “task-aligned” components of the structured correlations we observed. Our approach was to describe the set of observations (811 individual pairwise noise correlation measurements, Fisher z-transformed to produce normal error) in terms of a set of two underlying correlation structures: one defining  $r_{sc}$  as a function of pairwise preferred orientation alone (“fixed”) and the other defining  $r_{sc}$  as a function of pairwise preferred orientation relative to the task (“task-aligned”). Because the fixed matrix need not obey the lattice symmetries outlined above, these were not imposed in the regression model. We simply fit a description to the full correlation matrix. In order to provide a continuous and smooth

description of the data, each component was parameterized as the sum of an array of  $n \times n$  evenly spaced basis functions. Each observation,  $y_i$ , was expressed as:

$$y_i = x_i^{fixed} \cdot \beta^{fixed} + x_i^{task} \cdot \beta^{task} + \beta_0 + \varepsilon_i \quad (2.2)$$

$x_i^{fixed}$  and  $x_i^{task}$  are length- $n^2$  vectors of loadings onto the basis functions, which were given by evaluating the basis functions at the location corresponding to the pairwise orientation preference of the  $i^{th}$  pair.  $\beta^{fixed}$  and  $\beta^{task}$  are the length- $n^2$  vectors of amplitudes of the basis functions (coefficients to be fit),  $\beta_0$  is a model constant, and  $\cdot$  is the element-wise product. For the basis functions, we used bivariate von Mises functions, with zero correlation and equal dispersion in both dimensions. Thus the  $k^{th}$  loading ( $x_i^{fixed}(k)$  or  $x_i^{task}(k)$ ) was given by:

$$x_i(k) = \frac{e^{\kappa(\cos(\theta_i - \mu_k^1) + \cos(\phi_i - \mu_k^2))}}{Z} \quad (2.3)$$

where  $\theta_i$  and  $\phi_i$  are the preferred orientations of the  $i^{th}$  pair,  $\mu_k$  is a pair of orientations defining the location of the  $k^{th}$  basis function,  $Z$  is a normalization constant such that the sum of all loadings for observation  $i$  ( $x_i^{fixed} + x_i^{task}$ ) is 1, and  $\kappa$  is the von Mises dispersion parameter. (For the loadings onto the task-aligned component,  $\theta_i$  and  $\phi_i$  were expressed relative to the task. Because there were two discriminanda, we were forced to make an arbitrary choice of alignment, or to constrain the task-aligned component to have identical values at  $(0^\circ, 0^\circ)$  and  $(90^\circ, 90^\circ)$ . The latter approach would have introduced an asymmetry in the number of degrees of freedom between the two components. Instead, we chose to make an arbitrary choice of alignment, and expressed orientation relative to the discriminanda reported by either an upward or leftward saccade.) Again, angular quantities were doubled and  $\kappa$  was set to  $1.3\pi$ . We found that arrays of  $8 \times 8$  were

sufficient to describe the structure of the two components. Because the observations were pairwise correlations, it was sufficient only to fit the upper triangular portion of the array of basis functions. Thus, the two-component model contained 73 parameters (36 for each component, plus the model constant).

We fit the model by finding the parameters ( $\beta^{fixed}$ ,  $\beta^{task}$  &  $\beta_0$ ) that minimized the L1 error (to encourage sparseness) plus two additional terms that encouraged smoothness and symmetric positive semi-definiteness (as the two components were meant to represent correlation matrices). Thus the solution was obtained as:

$$\hat{\beta}^{fixed}, \hat{\beta}^{task}, \hat{\beta}_0 = \underset{\beta^{fixed}, \beta^{task}, \beta_0}{\operatorname{argmin}} \sum_i |\varepsilon_i| + \alpha_1 \Gamma(\beta^{fixed} + \beta^{task}) + \alpha_2 D_{SPD}(\beta^{fixed} + \beta^{task}) \quad (2.4)$$

$\Gamma$  is the discrete Laplace operator in two dimensions, corresponding to circular

convolution with the following kernel:  $\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$  and  $D_{SPD}(X)$  is the 2-norm between

$X$  and the nearest symmetric positive semidefinite matrix  $\hat{X}$ , which is given by

$(B + H)/2$  where  $H$  is the symmetric polar factor of  $B = \frac{(A+A')}{2}$  (Higham, 1988). The

$\alpha$ 's controlled the strength of regularization and were chosen to produce the best fit (as measured with  $R^2$  under 50-fold cross-validation). The solution was obtained by gradient descent using the MATLAB function *fminunc*.

While this model did not explain more than a small percentage (3.2%) of the variance of the observed noise correlations, this is not surprising as the raw correlation data do not vary smoothly with preferred orientation (reflecting both noise, and the fact that  $r_{sc}$  is known to depend on parameters other than orientation (Cohen & Kohn, 2011; Kohn & Smith, 2005; Smith & Kohn, 2008)). We therefore compared model performance



with the variance explained simply by a smoothed version of the raw data (sum of values in fixed and task aligned matrices was 3.6%).

### *Choice Probability*

Choice Probability was calculated in the standard way, using zero-signal trials (K. H. Britten et al., 1996). To estimate the average CP of the population, and to generate the histogram in Fig. 2.6e, we included only those neurons ( $n=111$ ) which could discriminate the task orientations ( $d' > 1.2$  at the highest presented signal level).

In the linear pooling model, the relationship between the covariance matrix for a population of neurons, the readout weight of each neuron, and the CP of each neuron is given in Eq. 1.1, reproduced from Haefner et al. (2013). We used this predicted relationship to quantify the CPs that would be associated with the correlation structure we observed and the fixed and task-dependent components we identified. CPs, correlations, and read out weights were described as functions of preferred orientation, relative to the task. We assumed a homogenous population of infinite size spanning this space, that we sampled discretely with  $1^\circ$  resolution. Since the read out weights were unknown, we generated a random distribution of plausible read out weights that could support task performance. To do this, we started with a vector of standard normal deviates and applied the  $90^\circ$  symmetry inherent in the task, such that  $\beta_\theta = -\beta_{\theta+90}$ , where  $\beta_\theta$  is the weight assigned to neurons with task-relative preferred orientation  $\theta$ . Then, we smoothed the read out weight profiles with a wrapped Gaussian kernel with  $15^\circ$  s.d. and excluded profiles which did not have a circular mean within  $22.5^\circ$  of choice 1 ( $0^\circ$ ). To guarantee real-valued CPs on  $[0,1]$ , we performed the calculations using a symmetric positive

definite approximation (Higham, 1988) of the observed correlation matrices, which introduced negligible error. For the CP predictions and for Fig. 2.6a, we expressed the task-aligned fit component as the average of both possible alignments (i.e. so that data at  $(0^\circ, 0^\circ)$  and  $(90^\circ, 90^\circ)$  are identical), although this was not directly imposed by the fitting procedure.

Estimating mean covariance for a population of neurons is necessarily more error-prone than estimating mean correlation, as the former is sensitive to sampling error in measurements of average firing rate, so for this reason we preferred to perform the calculations using correlations (see §2.7.4). We can use correlations interchangeably with covariances in Eq. 1.1, under the simplifying assumption that the variance is uniform as a function of preferred orientation: If  $\Sigma$  is the correlation matrix for a population with uniform variance  $\alpha$ , then it follows that:

$$\xi_k = \frac{\alpha(\Sigma\beta)_k}{\sqrt{\alpha\Sigma_{kk}\beta^T(\alpha\Sigma)\beta}} = \frac{(\Sigma\beta)_k}{\sqrt{\Sigma_{kk}\beta^T\Sigma\beta}} \quad (2.5)$$

where  $\Sigma_{kk} \equiv 1$  for all  $k$ . We felt that spike-count variance that depended systematically on preferred orientation was unlikely to be a feature of the V1 representations, and thus that the advantages of using correlations outweighed the cost of making this assumption.

Pooling noise has the effect of uniformly scaling down CPs, such that  $\xi_k$  in Eq. 1.1 is substituted with:  $\frac{(\mathbf{C}\beta)_k}{\sqrt{C_{kk}(\beta^T\mathbf{C}\beta + \sigma_{pool}^2)}}$ , where  $\sigma_{pool}^2$  is the variance of the pooling noise<sup>6</sup>. We found that non-zero pooling noise was needed to avoid overestimating the magnitude of CP from the observed correlation structure. We used a fixed amount of pooling noise in our predictions such that the average squared difference between the CP profile predicted from the observed correlation matrix and the observed CP profile was

minimized. Empirically, we found that pooling noise variance of 0.6 was optimal. Since our spike counts were normalized to have unit variance, this implies pooling noise whose variance is 60% of the average spike-count variance of single neurons. However, the overestimation of CPs may also be related to the assumption of a homogeneous population (Haefner et al., 2013), so the need to invoke pooling noise may be partially artifactual.

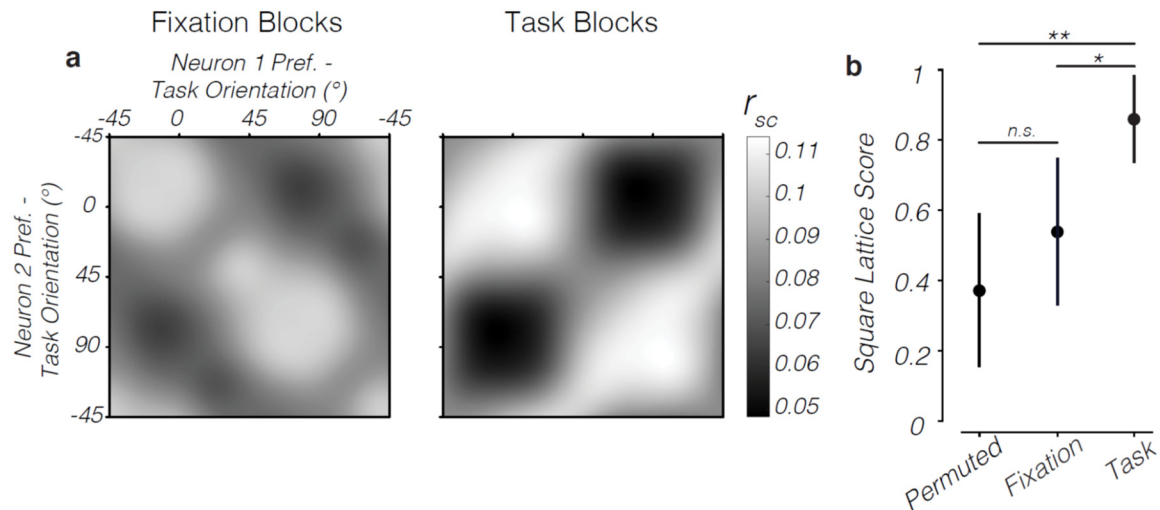
### *Calculating Task-Tuning Similarity*

We compared the structure of task-aligned noise correlations with the structure of similarity in preference for the two discriminanda orientations. The latter was quantified using the product of the slopes of the tuning curves ( $f'f'$ ), where the tuning curves refer to changes in firing rate as a function of signal strength along the task axis. Recent theoretical work has shown important limitations on stimulus information conveyed by sensory neurons in which the covariance structure matches the structure of task-tuning similarity, measured in this way (Moreno-Bote et al., 2014). Since we made use of correlations, rather than covariances, in the present study, we normalized the task-tuning similarity metric by the product of the standard deviations of the stimulus-independent variability of each pair, averaged across stimulus conditions.

## 2.7 Appendix

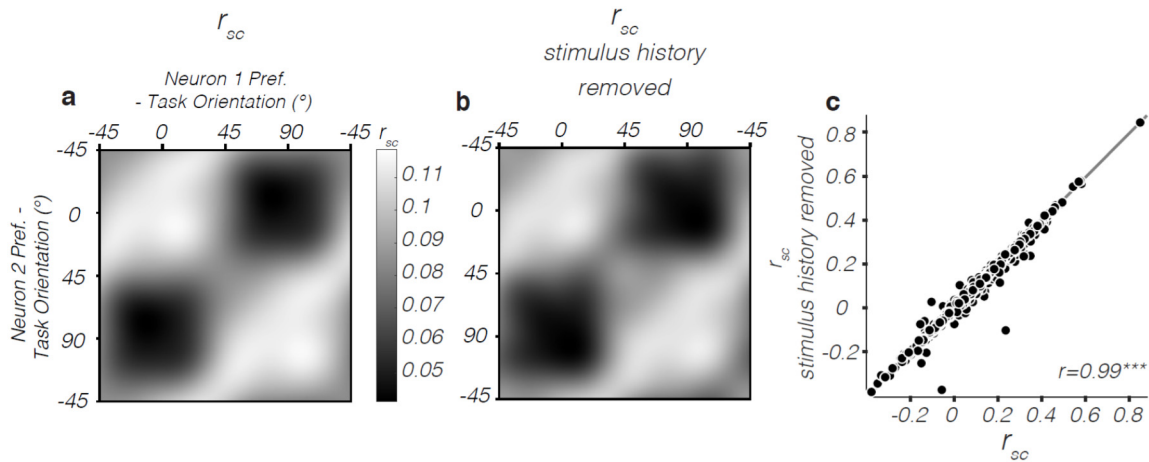
### 2.7.1 Possible Confounding Retinal Effects

We considered whether the task-dependence of the noise correlation structure we observed could be explained without invoking central mechanisms. First, we wondered whether the task-dependent changes in noise correlation structure were due simply to task experience and did not require active task engagement, for instance as a result of rewiring of local circuitry in visual cortex after training on a fixed set of orientations for several consecutive days. To address this, we interleaved presentations of the “zero-signal” filtered noise stimuli in the blocks of trials used to probe the orientation tuning of the neurons, during which the animal simply fixated passively. When we calculated  $r_{sc}$  using



**Figure 2.11. Lattice-like  $r_{sc}$  structure depends on active task engagement.** **a.** Task-aligned lattice-like noise correlation structure was present during performance of the orientation discrimination task, but was absent in separate blocks of trials used to measure neuronal orientation tuning, during which the animal fixated passively for reward. Noise correlations were calculated across interleaved presentations of the zero-signal orientation-filtered noise stimulus in the fixation blocks. During these blocks, the noise correlation matrix more closely resembled a diagonal ridge, demonstrating that the task-aligned lattice structure depends on active task engagement. Only 556 pairs were used, as not all recording sessions included fixation blocks with the zero-signal filtered noise stimuli interleaved. **b.** During the fixation blocks, the square lattice score associated with the correlation matrix was not significantly above chance at the  $p < 0.05$  level, unlike during task performance (error bars are  $\pm 1$  bootstrap s.e.).

these trials, we found that the task-dependent pattern of noise correlations was abolished. Instead, we observed a matrix more closely resembling a diagonal ridge (Fig. 2.11). This confirms that the source of the task-aligned correlation structure indeed depends on active engagement in the task.



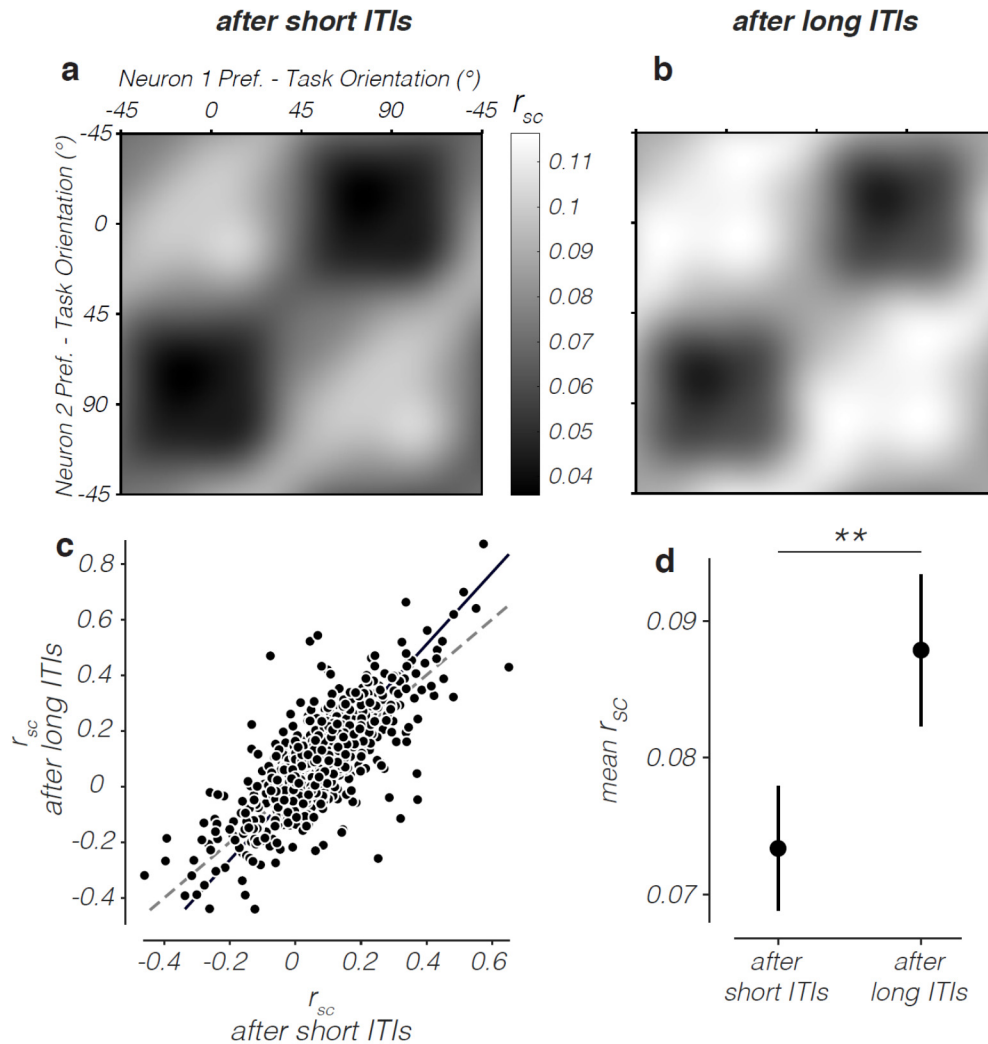
**Figure 2.12.  $R_{sc}$  structure is not influenced by stimulus history.** **a.** Average, task-aligned noise correlation matrix for orientation (identical to Fig. 2.4e except without using z-scored counts). **b.** Average, task-aligned noise correlation matrix for orientation after normalizing spike counts to remove the influence of the stimulus on the preceding trial. Color scaling and axis labels as in (a). **c.** Scatter plot of the noise correlations before and after the normalization procedure, showing almost no effect. This demonstrates that stimulus history does not significantly influence observed noise correlations.

Next, we considered the possibility that stimulus history influenced the firing rate of the neurons in ways that depended systematically on their tuning and thus created structured noise correlations that changed with the task. To rule this out, we used a different z-scoring procedure from the one described in the Methods. We broke trials down into groups for which the identity of the stimulus presented on that trial *and* on the preceding trial were the same. After z-scoring spike counts within these groups separately, counts no longer contained information about either stimulus. If coordinated firing rate fluctuations driven by stimulus history led to task-dependent noise correlation

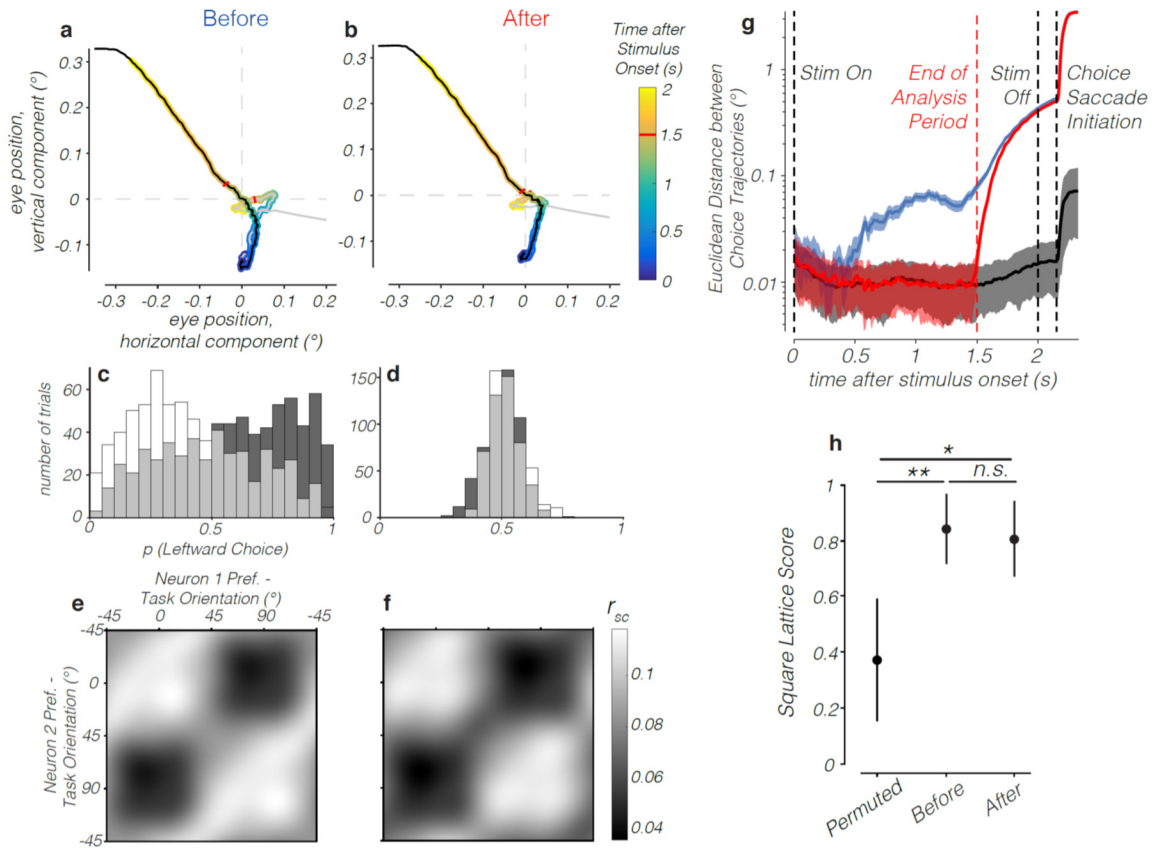
structure, then this normalization should attenuate it. However, we found that this procedure had virtually no effect on the correlation structure (Fig. 2.12).

We also reasoned that, if a mechanism related to stimulus history accounted for our results, we should see a marked attenuation of the correlation structure on trials following a long intertrial interval (i.e. a long delay before the subject acquired fixation). Following this reasoning, we performed a median split of the trials on each session based on the preceding ITI duration. The short ITI subset of trials was preceded by an average ITI of 1.1 seconds, while the long ITI subset was preceded by an average ITI of 4.8 seconds. We found that the lattice-like correlation structure was present in both subsets of trials. In fact, the amplitude of the noise correlation structure and mean  $r_{sc}$  was slightly increased for the long-ITI subset (Fig. 2.13), inconsistent with  $r_{sc}$  structure being driven primarily by stimulus history.

Both animals tended to make anticipatory microsaccades near the end of the trial that could predict the direction of their upcoming choice saccade (Fig. 2.14a), as reported in a prior study using the same task (Nienborg & Cumming, 2014). We were concerned that this introduced variability in the retinal image that influenced our measures of noise correlation. However, the direction of the choice saccade was not directly related to the orientations being reported, so it is unclear how this could generate structured noise correlations that depend on preferred orientation. Nonetheless, to rule out a role for eye movements, we measured noise correlations across a subset of trials on each session for which fixational eye movements were not predictive of choice. To identify these trials, a linear classifier was trained to predict the subject's choices using the time series of mean



**Figure 2.13.  $R_{sc}$  structure is similar after short and long inter-trial intervals. a,b.** The average, task-aligned correlation matrix for orientation for two subsets of trials, obtained from a median split based on the duration of the preceding inter-trial interval (ITI) for each session. The lattice-like noise correlation structure is present in both subsets. Axis labels and color scale in (b) are the same as in (a). **c.** A scatter plot of the noise correlations for the half of the trials preceded by long ITIs against the other half. The dashed gray line is the identity line, and the solid black line is the (type II) regression line. The slope is significantly greater than 1 ( $p < 0.05$ , bootstrap test), demonstrating an increase in the amplitude of the noise correlation structure after long ITIs. **d.** The mean noise correlation was also greater after long ITIs ( $p < 0.01$ , bootstrap test). Error bars represent  $\pm$  bootstrap s.e. This increase is in the opposite direction of any change that could be explained by correlations introduced by the identity of the preceding stimulus.

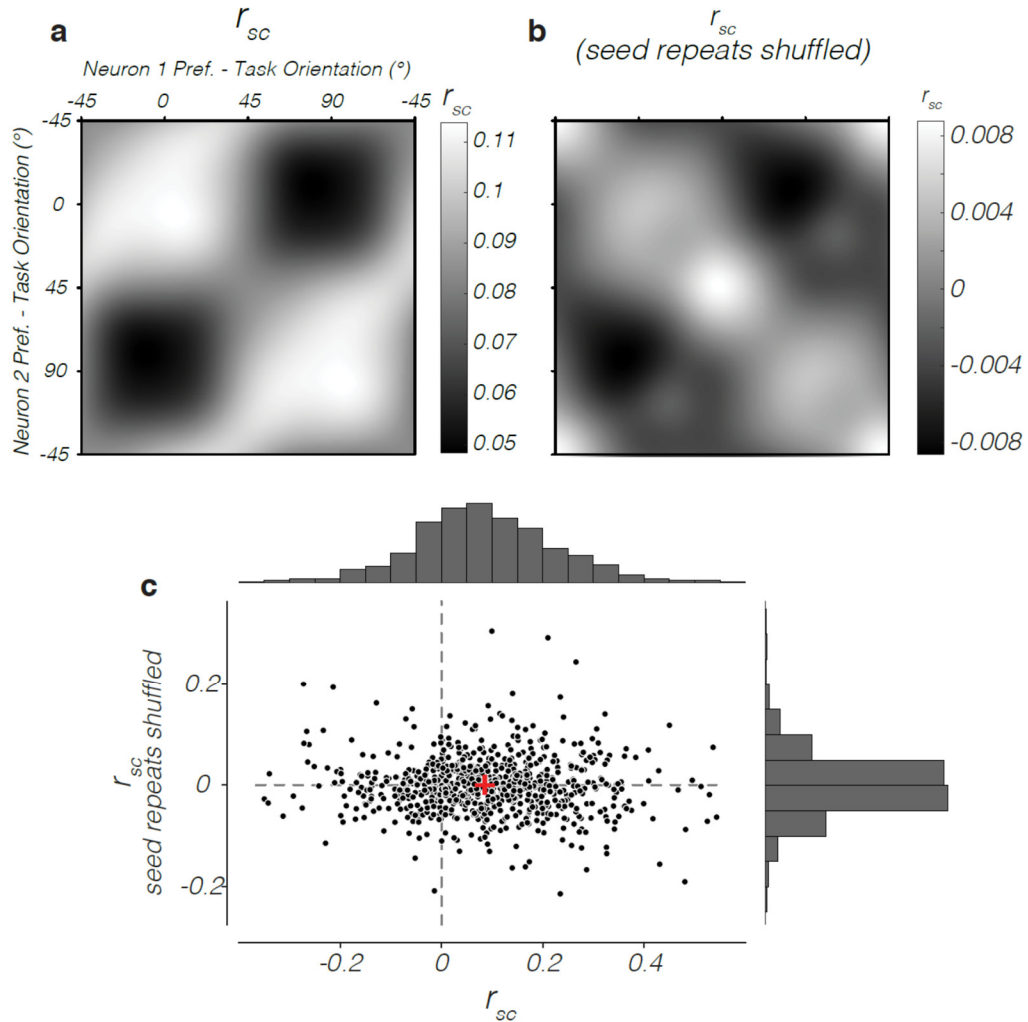


**Figure 2.14.  $R_{sc}$  structure is not generated by fixational eye movements.** **a.** The trial-averaged trajectories of mean binocular eye position, from one example session from monkey ‘*lem*’. Black and gray traces correspond to trials preceding leftward and rightward choice saccades, respectively. Red lines indicate eye position 1.5 seconds after stimulus onset, around which the trajectories noticeably diverge. **b.** Eye position trajectories for a subset of trials from which choice could no longer be predicted, identified using a linear classifier after having removed the last quarter of the trial. **c,d.** Histograms of posterior probabilities, giving the chance that a trial was associated with a leftward choice saccade according to the classifier, before (c) and after (d) the trial-removal procedure. The white and black distributions represent trials followed by rightward and leftward choices, respectively. **e,f.** The task-aligned V1 noise correlation matrix for orientation (averaged across all sessions) was extremely similar before (e) and after (f) we applied the trial-removal procedure to each session, indicating that choice-related fixational eye movements were not responsible for the structure we observed. Axis labels and color scale are same for (e) and (f). **g.** The Euclidean distance between the eye position trajectories associated with the two choices, as a function of time after stimulus onset. The distances before and after the trial-removal procedure are shown in blue and red, respectively, with error bars indicating  $\pm 1$  s.e. obtained from resampling trials. The distance expected by chance, obtained by shuffling the choice labels, is shown in black with  $\pm 1$  s.e. shown. **h.** Square lattice scores for the average, task-aligned correlation matrices before and after the trial removal procedure were not significantly different and were both significantly above chance level ( $p < 0.01$  and  $p < 0.05$ , bootstrap test, respectively).



binocular eye-position recorded on each trial. Then, trials were iteratively removed, starting with those furthest from the separating hyperplane, until classification performance was no better than chance (Fig. 2.14a-d). This analysis was restricted to the first 1.5 seconds since considering later time points required discarding too many trials. We still observed the same pattern of structured noise correlations in the remaining trials (Fig. 2.14e-h), suggesting that choice-related eye movements did not play a significant role in generating our results.

Lastly, we considered the influence of fluctuations in the white noise underlying our stimuli on co-fluctuations in neuronal firing rates. We used a two-pass experimental design in which each exact stimulus sequence was presented on two separate trials in each block. This allowed us to specifically identify the portion of firing rate covariability due to the common influence of the white noise variability. To do this, we calculated  $r_{sc}$  for a pair of neurons after we permuted the indices of the paired repeat trials for one neuron, while preserving the indices for the other neuron. This preserved the exact spatiotemporal sequence of pixels on the monitor while destroying the temporal alignment of other common inputs that could drive noise correlations. We then computed the average, task-aligned matrix for orientation using these “stimulus-induced” noise correlations. We found that the lattice structure was abolished, with pairwise correlations becoming tightly clustered near zero (Fig. 2.15). Thus, we can safely rule out rate covariability introduced by the stimulus as a prominent source of the noise correlation structure we report here.

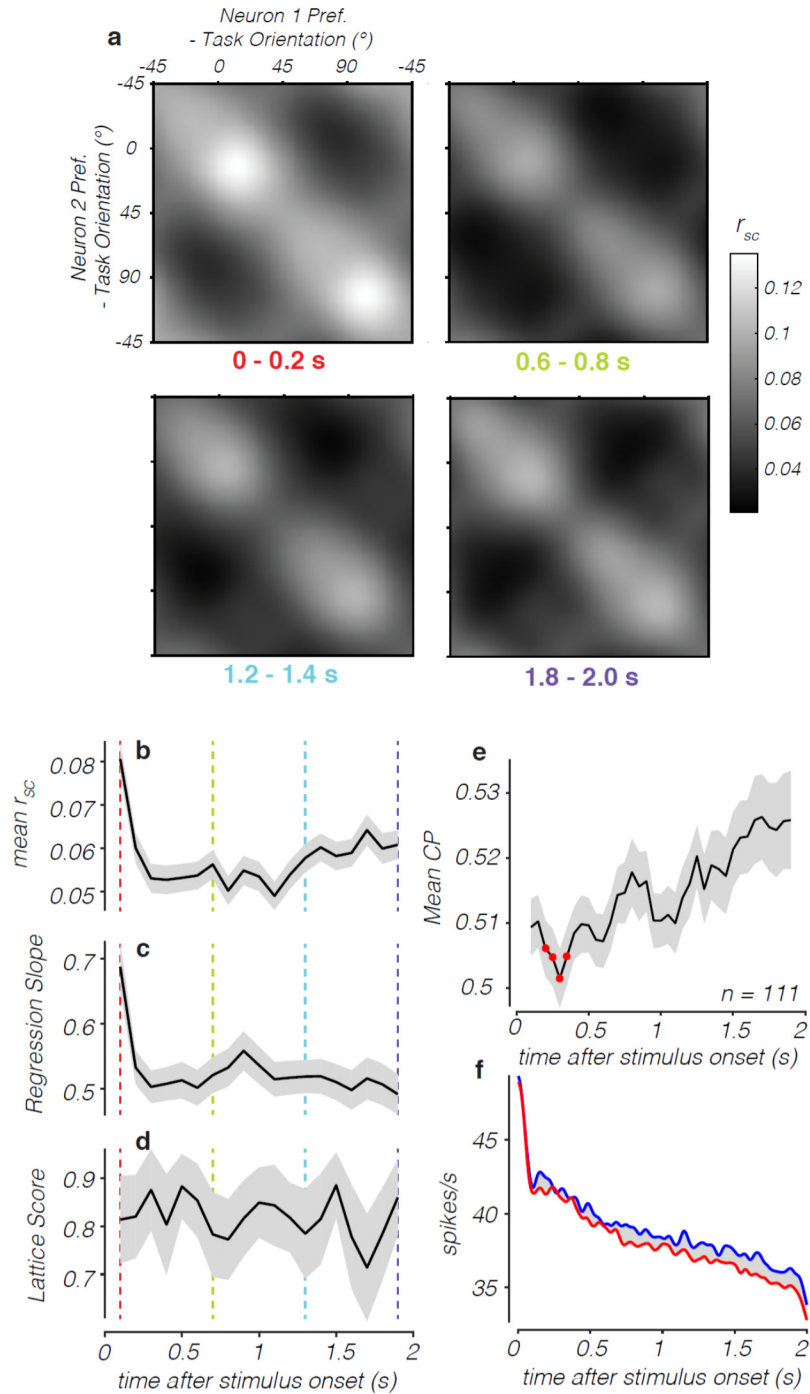


**Figure 2.15.  $R_{sc}$  structure is not driven by stimulus variability.** **a.** The average, task-aligned V1 noise correlation matrix for orientation (identical to Fig. 2.4e). **b.** The noise correlation matrix attributable to covariability introduced by trial-to-trial fluctuations in the stimulus, identified using a two-pass trial-shuffling procedure (see Methods). The task-aligned lattice-like structure is abolished, and the amplitude of noise correlations is drastically reduced (note expanded color scale). The remaining structure resembles a weak, diagonal ridge. (Data from only 801 pairs, because a few pairs lacked sufficient two-pass trials for the analysis.) **c.** A scatter plot of the noise correlations before versus after the trial-shuffling procedure. There was no significant correlation between the two and the noise correlations after shuffling were centered very close to zero (mean of  $-4 \times 10^{-4}$ ). Red cross indicates the population average.

### 2.7.2 The Timecourse of $R_{sc}$ Structure Within the Trial

We were interested to know how the structured noise correlations we observed evolved over the course of single trials. To examine this, we calculated the average, task-aligned correlation matrix using spike counts from 200-ms windows centered on various time points during the 2 second stimulus presentation, with a 50ms offset to account for the typical V1 response latency. We found that the lattice-like pattern was present at the very beginning of the trial and its amplitude stayed roughly constant throughout the trial (Fig. 2.16). This pattern suggests that the feedback signals driving these correlations do not simply arise after the animal forms a decision about each individual stimulus. (We also noticed abrupt changes in  $r_{sc}$  shortly after stimulus onset, similar to results reported in Churchland et al., 2010).

Given the close theoretical link between noise correlation structure and choice-related activity, we expected that correlation structure might follow a similar pattern to CP. However, we found that CP increased gradually over the course of the trial (Fig. 2.16e), unlike the amplitude of noise correlation structure. This might reflect changes in readout weight during the trial, or that choices are not solely determined by a linear readout of the population. Other proposed models of perceptual decision making that include feedback may be able to account for this apparent divergence. We also cannot rule out that the increase in CP over the last quarter of the trial may be related to the choice-related differences in fixational eye position which became particularly pronounced at the end of the trial (Fig. 2.14).



**Figure 2.16. Temporal dynamics of structured noise correlations.** **a.** The average, task-aligned V1 noise correlation matrix for orientation obtained using spike counts from various 200-ms windows during the stimulus presentation. The task-aligned lattice structure is present at all time points (4 examples shown). **b-d.** Plots showing the temporal dynamics of four statistical measures of the observed noise correlations (mean  $\pm$  1 bootstrap s.e.). The colored lines indicate the example time points shown in (a). The population mean noise correlation (b) showed a sharp drop shortly after stimulus onset and

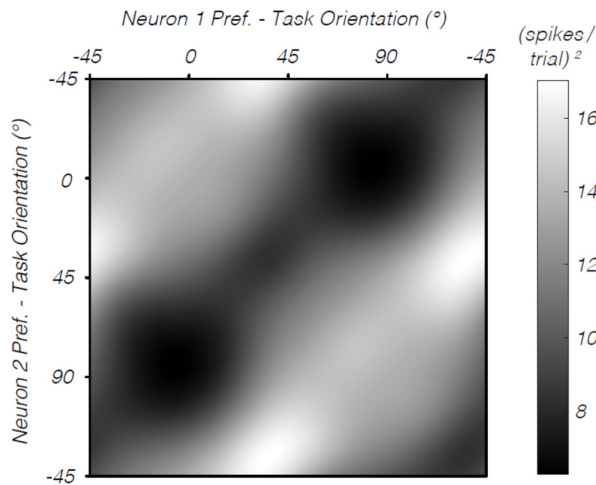
then a gradual recovery over the course of the trial. The slope of the regression line of the noise correlations obtained in each 200-ms window against the noise correlations obtained from trial-length spike counts is in (c). Apart from an increased slope at the first time point, likely due to the onset of the visual stimulus, this showed no significant modulation over the course of the trial. The lattice score (d) was not significantly modulated over the course of the trial. e. Population mean CP, measured in 200-ms bins, showed a gradual increase over the duration of the trial (slope significantly positive,  $p < 10^{-3}$ , bootstrap test). Some of this may be due to choice-related fixational eye movements, which were particularly pronounced at the end of the trial (Fig. 2.14). Red dots indicate time points for which CP is not significantly above chance at the  $p < 0.05$  level (bootstrap test, false-discovery rate of 0.05 ensured across comparisons using the Benjamini-Yekutieli procedure (Benjamini & Yekutieli, 2001)). Error bars indicate  $\pm 1$  bootstrap s.e. The same set of cells is used as in Fig. 2.6e. f. Population mean PSTH, shown separately by preferred (blue) and anti-preferred choices (red). The same set of cells is used as in Fig. 2.6e. *N.B. All times include a 50-ms offset to account for the average response latency of V1 neurons.*

### 2.7.3 Correlation vs. Covariance

The present study depended critically on comparing estimates of the spike-count covariability of subpopulations of neurons with different orientation tuning. We quantified this covariability using the mean spike-count correlation. A key advantage of using spike-count correlations is insensitivity to estimates of the spike-count variance. By contrast, measurements that do not normalize for spike-count variance will overweight more variable pairs. In addition, using spike-count correlation allowed us to combine counts across experimental blocks and stimulus conditions without introducing artifacts related to changes in mean firing rate or firing rate variance, by applying z-scoring. This substantially increased the signal-to-noise ratio of our measurements, both by removing the confounding effect of long-term firing non-stationarities and increasing the number of trials that the measurements were based on.

As a confirmation that this approach yielded results that generalize, we measured the average, task-aligned covariance matrix for orientation, using the same approach as

we used to generate the correlation matrix in Fig. 2.4e. To estimate covariance between a given pair of neurons, we used an average of the covariance measured separately by stimulus condition, weighted by the number of trials. We found that it displayed a qualitatively similar task-aligned lattice structure (Fig. 2.17).

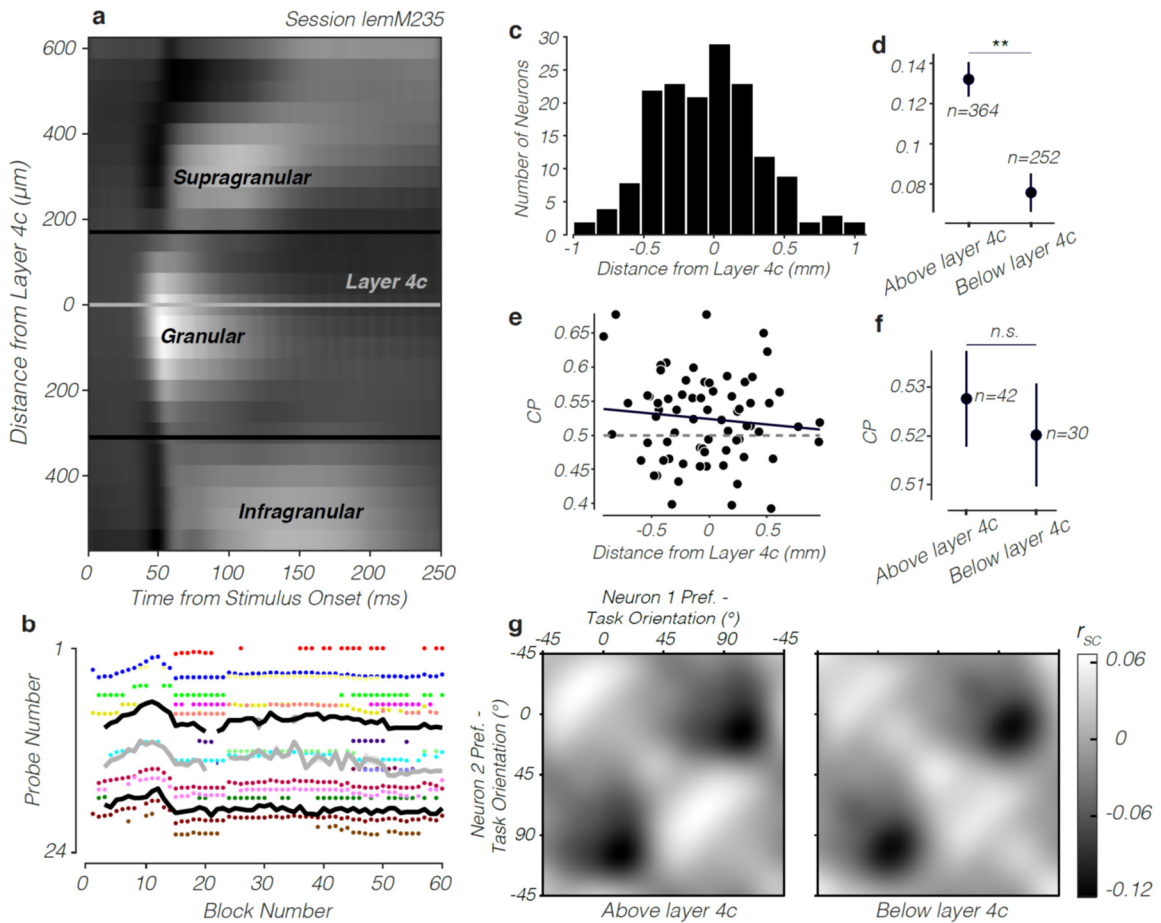


**Figure 2.17. The task-aligned structure of spike-count covariance.** The average, task-aligned V1 covariance matrix for orientation. This is closely similar to the average noise correlation matrix shown in Fig. 2.4e, demonstrating that the procedure we used in the main analysis to normalize spike-counts individually for each neuron before averaging led to generalizable results.

#### 2.7.4 Laminar Distribution of CP and $R_{sc}$

In monkey ‘*lem*’, recordings were performed using a 24-contact linear array, chronically inserted perpendicular to the dura. This allowed us to record from neuronal populations within individual cortical columns. We identified the laminar positions of the recorded units using current-source density (CSD) analysis (Mitzdorf & Singer, 1979), calculated as the second spatial derivative of the LFP voltages recorded across the 24 channels. We identified the supragranular-granular and granular-infragranular borders using stereotypical features of the CSD profile (Maier, Aura, & Leopold, 2011). We identified layer 4c by comparing the visual latencies across the electrodes (Fig. 2.18a). We presume that the minimum latency corresponds to layer 4c, although we do not have independent confirmation of this. We wondered whether CPs and  $r_{sc}$  differed across

laminae. Because V1 laminae receive distinct types of inputs, this could provide useful information about the source of the underlying signals. While there was a slight trend towards higher CPs for more superficial neurons, this was not significant (Fig. 2.18e). We did not see any noticeable difference in the noise correlation structure across laminae (Fig. 2.18g).



**Figure 2.18. Laminar distribution of CP and  $r_{sc}$ .** **a.** Pseudocolor plot of current-source density (CSD) profile following stimulus onset, example session. These were measured using a 24-contact linear array, inserted acutely in V1, roughly perpendicular to the dura. Light and dark regions correspond to current sinks and sources, respectively. Stereotypical features in the CSD profile mark the laminar boundaries of V1. Layer 4c is identified by the latency of the visual onset. **b.** Plot showing the temporal evolution of the laminar landmarks over the time elapsed during the example recording session. These were measured independently during each experimental block. Red lines correspond to the laminar boundaries, and the dashed blue line to layer 4c. These features remain in nearly fixed positions over the recording session relative to the linear array, with slow drift that is

consistent across the independently estimated laminar landmarks. Colored dots indicate the laminar positions of the recorded units. Their positions stay fixed relative to the estimated landmarks. **c.** Distribution of depths of recorded units across sessions, relative to layer 4c. (Negative values indicate units more superficial than layer 4c). **d.** Mean  $r_{sc}$  was significantly higher for pairs in which at least one member was above layer 4c, than pairs in which at least one member was below layer 4c. (We tried breaking the correlations into groups in which both members of each pair were confined to the same lamina, but this resulted in too few pairs to make quantitative conclusions.) **e.** CPs recorded during performance of the orientation discrimination task, plotted against depth relative to layer 4c. We observed a non-significant trend towards greater CPs at more superficial depths. (Negative values indicate units more superficial than layer 4c). **f.** CPs for neurons more and less superficial than layer 4c were not significantly different. Error bars are +/- 1 bootstrap s.e. **g.** Observed noise correlation matrices for orientation, aligned to the task, for the two groups of noise correlations in (d). We subtracted the mean correlation so that we could better visualize any differences in correlation structure. The two matrices appear qualitatively identical.

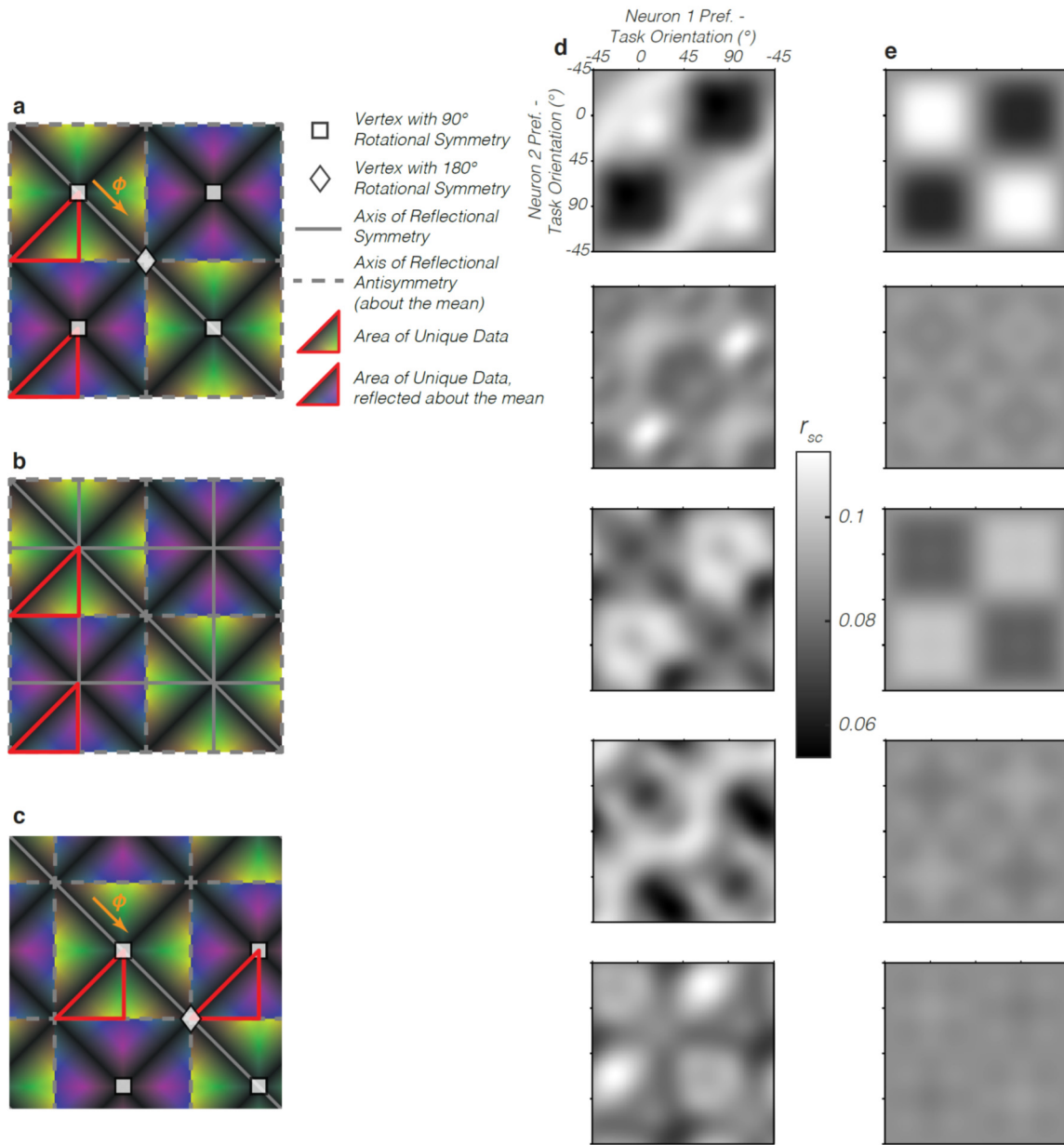
### 2.7.5 The Square Lattice Model

We observed that the task-aligned pattern of noise correlations obeyed a certain set of symmetries (Fig. 2.4g and reproduced in Fig. 2.19, below). We explicitly modeled these using a square lattice of type p4m (Schattschneider, 1978)<sup>2</sup>. Some of these symmetries were inherent in the data. For instance, the axis of diagonal symmetry is given by the fact that the data are Pearson correlations— $r(i,j)=r(j,i)$ , by definition. In addition, when the data is shown in a task-aligned coordinate frame (as in Fig. 2.4e), we averaged across both possible alignments (because there were always two orthogonal discriminanda) resulting in a matrix where the value at a given position  $(\phi, \phi)$  is necessarily identical to the value at  $(\phi+90^\circ, \phi+90^\circ)$ . This makes some of the symmetries

---

<sup>2</sup> The p4m lattice does not include an axis of diagonal symmetry, nor does it include any axes of anti-symmetry about the mean. The addition of these symmetries was used specifically to model the structure in our data.





**Figure 2.19. The square lattice model.** **a.** Illustration of the lattice symmetries used to model the correlation structure (reproduced from Fig. 2.4g). **b.** The same symmetries can be illustrated without explicitly using any vertices of rotational symmetry. **c.** Illustration of the symmetries after shifting the lattice along the diagonal by an angle  $\phi$ . To fit the raw correlation matrices (i.e. not task-aligned) in Fig. 2.4, we used the angle  $\phi$  that optimized the lattice score. **d.** Observed task-aligned correlation matrix (reproduced from Fig. 2.4e; top row) and surrogate data (other rows), obtained by randomly shuffling the correlation values but preserving the orientation preferences. **e.** Best-fitting square lattices (with no offset) for the correlation matrices in (d). Axis labels and color scale as in (d).

modeled by the lattice trivial given the way the data was processed. However, the critical feature of the data—peaks and troughs in spike-count correlation at the within- and between-pool regions—was in no way constrained by the data processing.

As an illustration, we generated several surrogate correlation matrices (Fig. 2.19d), produced by randomly shuffling the correlation values of the neurons but preserving the orientation preferences. (This is different than the diagonal shuffling used to produce the data in Fig. 2.4h). These surrogates have been processed identically to the observed data in Fig. 2.4e (and reproduced in the top row of Fig. 2.19), but simply include shuffled correlation values. We then fit task-aligned lattices to these (Fig. 2.19e). Relatively little of the variance in these data is captured by the best-fitting task-aligned lattices, demonstrating that the high lattice score for the real data is not an artifact of the way the data has been processed, but reflects real symmetric structure in the data. Second, the best-fitting lattices contain a diversity of forms, and do not necessarily resemble the best-fitting lattice for the real data. In other words, the lattice symmetries, while they capture the *symmetry* relationships in the observed correlation structure, do not constrain the fits to take on the particular values we observed.

# **Chapter 3:**

## **Identifying Animal Subjects' Internal Task Strategies with Psychophysical Reverse Correlation**

### **3.1 Introduction<sup>3</sup>**

In the preceding chapter, I discussed the finding that patterns of correlated variability in V1 change with a subject's allocation of internal resources. Our approach to testing this hypothesis involved changing the set of orientations they needed to discriminate to obtain reward. A crucial assumption underlying this approach is that the animals changed their task strategy appropriately when task instruction changed. The validity of this assumption relates to a broader issue faced by studies involving animal subjects trained to make psychophysical judgments: simply put, are subjects performing any psychophysical task as instructed? In most studies, assumptions about the task strategies employed by animal subjects are verified by measuring overall performance (in terms of the reward rate, shape of the psychometric function, percent correct, etc.) However, this is only a valid test if deviations from the assumed task strategy necessarily impose measurable performance costs. In many cases, they do not.

In this chapter, we discuss our efforts to determine our subjects' decision strategies objectively. To do this, we use psychophysical reverse correlation (PRC). While widespread in the human psychophysics literature, PRC has been less often used to analyze animal behavior. First, we show using a simple model that subjects can obtain relatively high reward rates using a task strategy that most highly weights orientations

---

<sup>3</sup> A subset of the data used for the analyses presented in this chapter was collected by Hendrikje Nienborg and Bruce Cumming.

other than the discriminanda orientations, thus generating psychophysical thresholds that appear consistent with subjects performing the task as instructed. We go on to show, using PRC, that our subjects frequently use such “misaligned” decision strategies. This is due to the striking finding that it requires days or even weeks of training to fully update their task strategy after a change in the discriminanda. On the timescale of single training sessions, subjects’ behavior can be well characterized using a single, fixed task strategy, even if the set of discriminanda is not fixed. By contrast, we show that human subjects performing the same task are able to rapidly update their task strategy on the order of tens of trials. We go on to show that knowledge of the subject’s task strategy can predict psychophysical performance, providing insight into otherwise mysterious day-to-day variability in behavior. Lastly, we use a novel approach to assess whether subjects are truly at their perceptual thresholds, and show that our subjects typically are not, despite the use of standard procedures to ensure threshold performance.

Taken together, our results demonstrate the utility of psychophysical reverse correlation as an approach to objectively measure the strategies employed by animal subjects in psychophysical tasks. We present evidence for a fundamental, and previously unappreciated, difference in the way humans and non-human primates perform discrimination tasks. This result highlights a serious potential confound in studies in which researchers seek to measure an effect on neuronal activity of changes in task instruction (for instance, Cohen & Newsome, 2008). Given our findings, it seems likely that subjects often employ a fixed strategy that lets them receive an acceptable rate of reward across multiple task contexts. We provide tentative suggestions for changes to typical training procedures and task designs that may mitigate this deficit. Whether or not

these will be successful, PRC can be used to obtain objective measures of the strategies animal subjects actually employ in future studies<sup>4</sup>.

### 3.2 A Quantitative Motivation of the Problem

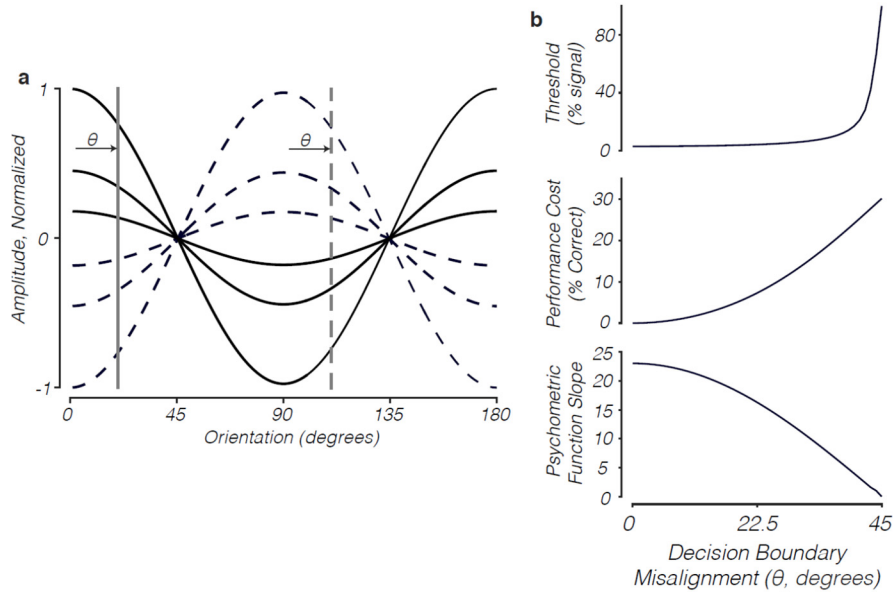
First, we wanted to understand, using a quantitative model, the impact on performance if our subjects were to use the “wrong” orientations to motivate their choices. Recall that orientation amplitude in the stimuli used in our study is distributed according to a wrapped Gaussian. As a result of this design, a range of orientations near the center of the filter will provide useful evidence about the correct choice. In other words, there would seem to be a relatively low cost associated with a “misaligned” task strategy. We quantified the performance of a model subject whose decisions are made by comparing the amplitude of two orthogonal orientations, and ignoring all others. Performance is based solely on the discriminability of stimulus energy at those two orientations. This is determined by their angular difference (call this  $\theta$ ) from the discriminanda orientations, the orientation bandwidth of the stimulus, as well as noise in the stimulus and the decision process. The latter is the only free parameter, and is set to match the thresholds of the monkey subjects (see Methods).

We found that the impact on performance of changing  $\theta$  (a misaligned strategy) is markedly sublinear, and as a result is quite weak for small amounts of misalignment (Fig. 3.1). Surprisingly, a task strategy misaligned by  $10^\circ$  will decrease percent correct by only 2%. As  $\theta$  increases, the difference in the mean amplitude at the two decision

---

<sup>4</sup> We applied PRC analysis in the previous chapter to confirm that the subjects’ internal task strategies closely matched task instruction. This required breaks in recording after a change in the discriminanda during which subjects were retrained.

orientations becomes small relative to noise, and performance starts to degrade significantly.



**Figure 3.1. Simulated effect of task strategy “misalignment” on performance. a.** The distribution of orientation in the stimulus at a variety of signal levels. Stimuli with 0° filter center are solid lines, and those with 90° filter center are dotted lines. The grey vertical lines indicate those orientations used by a hypothetical observer to inform choice. Offsets of these decision orientations from the true discriminanda orientations (decision boundary misalignment;  $\theta$ ) degrades performance. **b.** Quantification of performance as a function of  $\theta$ . For small amounts of misalignment, the effect is modest, but performance degrades steeply as  $\theta$  approaches 45°. Threshold refers to the % signal at which the subject is 75% correct. The psychometric function slope is inversely related to the threshold by a scale factor (see Methods for calculation of psychometric function slope).

Thus, it is reasonable to conclude that significant deviations from the optimal decision boundary impair performance by amounts smaller than would be detectable given other sources of day-to-day behavioral variability, such as arousal, motivation, and the like. It follows that overall performance is not a reliable test for whether a subject is performing the task as instructed. This conclusion is not specific to our task, but applies to any discrimination task in which changes in signal level affect the amplitude of a broad range of stimulus features. In fact, broad stimulus filters, rather than a broad distribution

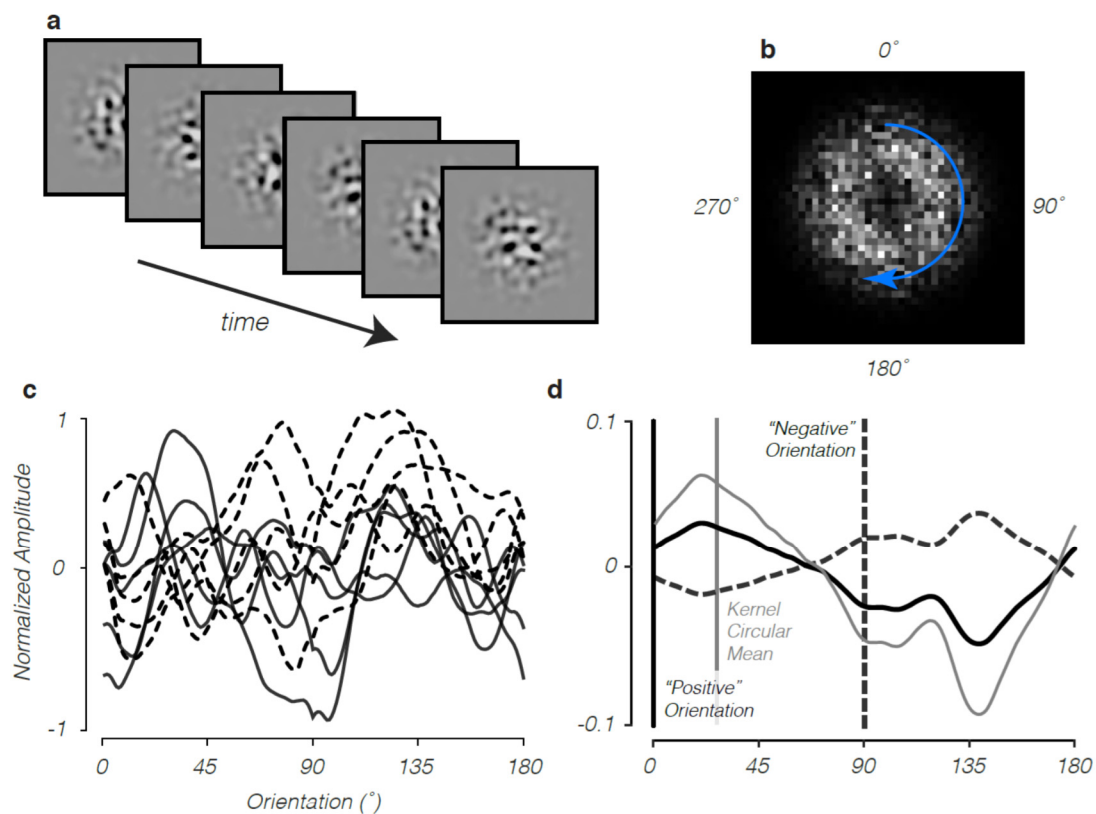
of stimulus amplitude, would yield identical results (see Methods), raising the possibility of serious behavioral confounds in an even wider variety of studies involving animal behavior.

### **3.3 Approach: Measuring Psychophysical Kernels**

We sought to measure the strategies subjects actually used to perform the task, particularly to verify that they update their strategies appropriately after a change in the discriminanda orientations. To do this, we borrowed an approach from human psychophysics that measures the small influence of random fluctuations in the stimulus on subject's choices, analogous to the way receptive fields are measured by spike-triggered averaging. This procedure is referred to as “psychophysical reverse correlation” (PRC) and the result is a “classification image” or “psychophysical kernel” (or simply “kernel”) which reflects the weights applied by the subject to various stimulus features (Eckstein & Ahumada, 2002). Assuming the animal's choices are made simply as a weighted average of the values of the stimulus features on a given trial, the (first-order) kernel is the maximum likelihood estimate of the linear weights. It is calculated as the difference between the average value of the stimulus features associated with the two choices.

We combined training data from two subjects (*'lem'*, who also contributed neuronal data to our study, and another monkey, *'ruf'*, who did not) across two variants of the orientation discrimination task, for a total of 461 training sessions (359 from *'lem'* and 102 from *'ruf'*). The first task (Task 1) used an orientation-filtered noise stimulus and is described in detail in the previous chapter. The second version of the task (Task 2)

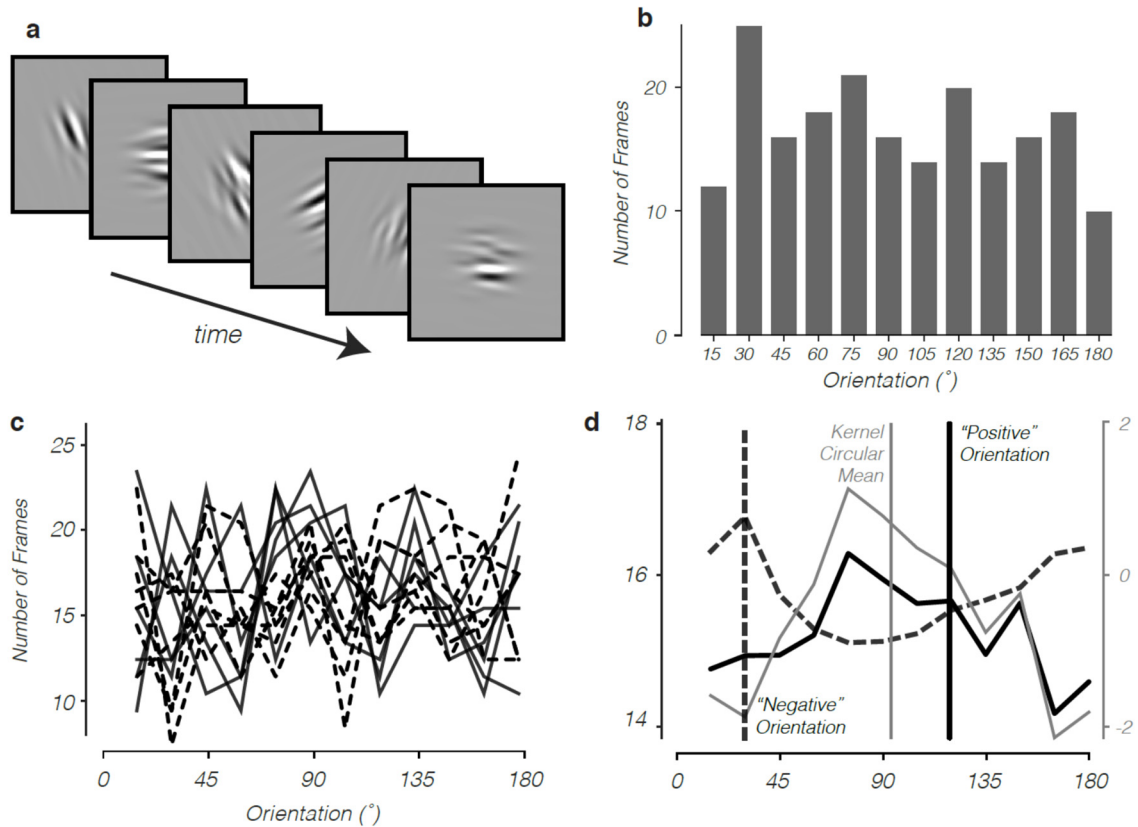
involved a stimulus constructed slightly differently, but the task was otherwise identical. This version of the task used a “subspace” stimulus (Hendrikje Nienborg & Cumming, 2007; Ringach, G., & Shapley, 1997). Individual frames were narrowband for orientation (10° s.d.), with center orientations sampled from a discrete set of orientations. On zero-signal trials, all orientations in the set were equally likely to appear. Otherwise, the “signal” orientation occurred with an elevated probability that defined the nominal signal



**Figure 3.2. Measuring psychophysical kernels, Task 1.** **a.** Example stimulus frames at zero-signal. Each stimulus frame consists of independent white noise, filtered uniformly for orientation. **b.** Example amplitude spectrum of a zero-signal stimulus, averaged across frames. **c.** For the PRC analysis, we summarized each stimulus using the radial sum of its amplitude spectrum, to describe stimulus energy as a function of orientation. Shown here are these functions for a sample of zero signal stimuli, illustrating the random fluctuations in orientation energy introduced by the use of filtered noise. **d.** The psychophysical kernel is calculated as the difference (gray) of the two choice-conditioned radial sums (“positive” choice average in solid black and “negative” choice average in dashed black). Choices were signed according to an arbitrary rule. Note that the peak in the kernel is not located at the nominally “positive” orientation, which would correspond to an optimal kernel.



level. One way of stating the difference between the two tasks is that, in Task 1, orientation power varied across space, while in Task 2, it varied across time. All results generalized across subjects and the two task variants.



**Figure 3.3. Measuring psychophysical kernels, Task 2.** **a.** Example stimulus frames for a zero-signal trial. Each stimulus frame consists of independent white noise, filtered narrowly for orientation (10-degree orientation filter s.d.) with a center orientation drawn randomly from a discrete uniform distribution. **b.** Histogram illustrating the distribution of orientations presented across frames on an example zero-signal trial. **c.** For the PRC analysis, we summarized each stimulus using the histogram of orientations presented. Several examples are shown here. Solid lines indicate the histograms associated with “positive” choices, and dashed lines are those associated with “negative” choices. **d.** The psychophysical kernel is calculated as the difference (gray) of the two choice-conditioned average orientation distributions (“positive” choice average in solid black and “negative” choice average in dashed black). Choices were signed according to the same arbitrary rule as was used in Task 1. Note that the peak in the kernel is again not located at the nominally “positive” orientation, which would correspond to an optimal kernel.

The motivation for using the subspace version of the task is that kernels required fewer trials to estimate. This is because the subspace task generated considerably larger fluctuations in orientation amplitude across trials. This generated a larger influence of stimulus variability on choice to measure with PRC. Kernels were calculated slightly differently for the two task variants. For Task 1, we summarized each stimulus as the radial sum of its 2D Fourier amplitude spectrum (to marginalize out spatial frequency and phase), averaged across frames. We then computed the kernel as the difference between the two choice-conditioned radial sums (Figure 3.2). For Task 2, we summarized each stimulus as the distribution of filter center orientations presented across frames. The kernel was then calculated as the difference between the two choice-conditioned distributions (Figure 3.3). For both variants of the task, kernels were calculated separately for each nominal signal level and then averaged, so that their shape reflected the influence of random orientation fluctuations on choice, not the influence of changes in nominal signal.

## **3.4 Results**

### *3.4.1 Shape of the Kernels*

We found that psychophysical kernels, expressed in this way as functions of orientation, typically followed a roughly sinusoidal shape (Figs. 3.2.d and 3.3.d). The smoothly varying shape of the kernels demonstrates that the animals integrated evidence broadly across a range of orientations. (Because the random stimulus fluctuations were independent across orientations, this is unrelated to the construction of the orientation-

filtered stimuli<sup>5</sup>). In addition, the bimodality of the kernel shows that the subjects' strategy involves comparing the relative strength of orientations at the positive and negative lobes of the kernel, rather than simply detecting the presence of a single orientation band. We confirmed using simulations involving a model observer that a strategy consisting of reporting the presence of a single orientation band, rather than a real binary discrimination, does not generate a bimodal kernel. Instead, the subtraction procedure used to generate kernels means that the resulting simulated kernel contains a single peak and is uniformly negative elsewhere (data not shown).

### *3.4.2 Decision Strategies are Fully Described by 1st-Order Kernels*

Next, we sought to confirm that the measured psychophysical kernels did indeed capture the influence of the stimulus on the subject's choices. This would not obtain under either of two conditions: 1) if we had too few trials to reliably estimate the kernel, or 2) if the animal's decisions were based on non-linear combinations of orientations (in which case, higher-order kernels would be needed to describe behavior). To rule these out, we used the measured 1<sup>st</sup> order kernel to predict choices, and compared the accuracy to an objective measure of the maximum variance in the animal's choices that is explainable by any model (Neri & Levi, 2006).

If the first-order kernel is sufficient to capture the task strategy, then choices can be predicted by the vector projections of the stimuli onto the kernel, thresholded to

---

<sup>5</sup> For Task 2, the random fluctuations were generated by sampling frames from a discrete set of orientations. Because we used filtered noise, this introduced fluctuations in orientation power that were not independent across orientations. However, kernels were not calculated by measuring the orientation power in the stimuli. Rather, they were calculated as the empirical distribution of frames at each nominal orientation. In this subspace, the random fluctuations were independent.

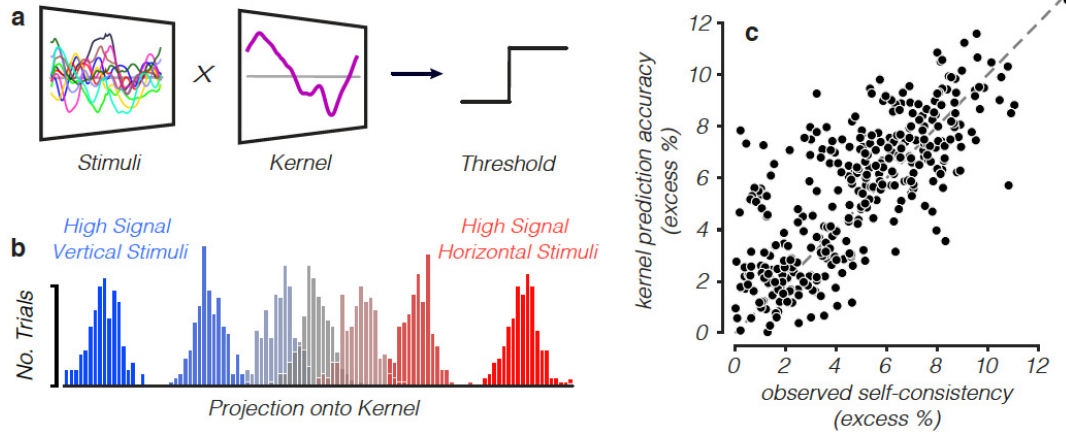
produce binary choices (Fig. 3.4a-b). We set this threshold to match the observed distribution of the animal's choices, and then compared the predicted choices to the animal's actual choices. The prediction accuracy was expressed as the amount in excess of the prediction accuracy based simply on the distribution of choices at each signal level. (Given an observed 90% rate of choosing choice 1 for a given signal level, prediction accuracy of a random kernel trivially achieves an accuracy of  $0.9^2$ , on average).

The maximum variance explainable by any model of the subject's choices can be determined by measuring the subject's consistency in responses to identical stimuli. This gives an estimate of the amount of stimulus-independent noise affecting the decision process, which is unexplainable variance. To do this, we used a two-pass task design, such that each exact stimulus (i.e. same noise seed) was presented twice per block, and expressed self-consistency as the fraction of repeat trials eliciting the same choice. This was then normalized relative to the trivial prediction given the choice distributions, as with the kernel prediction accuracy. We found that the accuracy of the kernel in predicting choices matched, on average, the observed self-consistency, demonstrating that the estimated first-order kernel fully captured the influence of stimulus fluctuations on choices (Fig. 3.4c).

### *3.4.3 Macaque Subjects Update Task Strategy Slowly*

We frequently observed significant deviations between the circular mean of the kernel and the (nominally positive) discriminandum orientation. In other words, the kernel was frequently not optimally aligned to the task. We found that these deviations were highly systematic and easily predicted from the subjects' training history, revealing

a striking novel finding that macaque subjects' task strategies are extremely slow to adapt to changing task instruction.

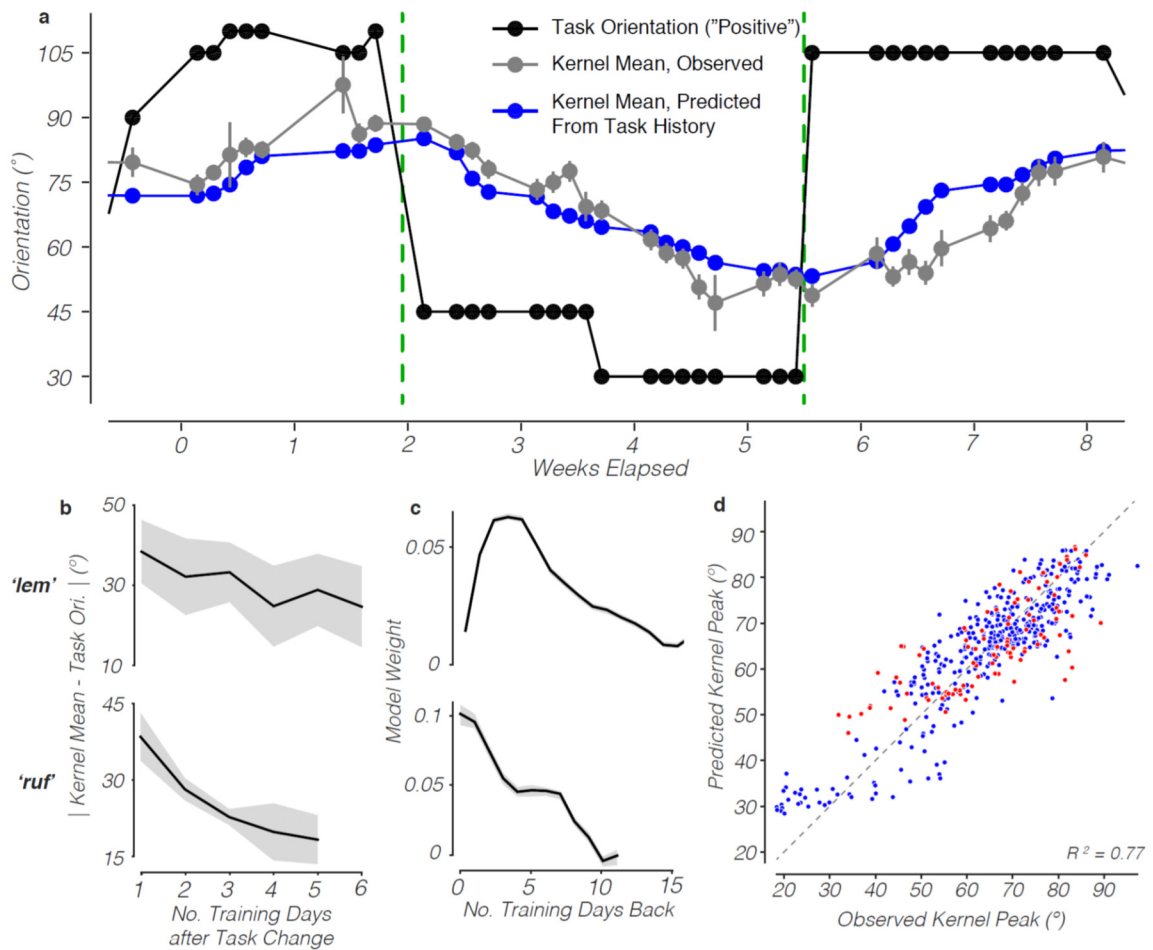


**Figure 3.4. Using the measured kernel to predict choices.** **a.** Assuming the animal's choices are based on a weighted average of the energy across stimulus orientations, choices can be modeled by projecting each stimulus onto the estimated kernel and applying a threshold to generate binary choices. Example kernel is the same as in Fig. 3.2d, observed when the subject was discriminating the cardinal orientations. **b.** Histogram of the projections of stimuli onto the example kernel in (a). Color indicates nominal signal level. Choices are determined by thresholding these projections. **c.** Scatter plot comparing the subject's self-consistency in response to paired repeats of identical stimuli and the accuracy of the first-order kernel in predicting choices. Each dot represents a single training session. Only sessions in which the two-pass procedure was used are shown ( $n=347$ ). Values are shown relative to those that would be predicted simply from the knowing the distribution of choices at each signal level.

This result is easily illustrated by the kernels observed during an example two-month training period from monkey 'lem', when he was performing the subspace variant of the task (Figure 3.5a). During this training period (at which point the subject was already expert), he first performed 9 day-long training sessions in which he had to discriminate orientations close to the cardinals, followed by a rapid switch to oblique discrimination which lasted 17 sessions, followed by another abrupt transition back to the cardinals.

When we measured kernels separately for each session, we found that they were remarkably slow to adapt to the abrupt task changes. In each of the three stable training periods, the kernel misalignment started at its maximum, reflecting the effect of the previous task, and then slowly decreased with exposure to the new task. In other words, the kernel slowly rotated in the direction of the new set of task orientations. This was a consistent feature of the behavior of both subjects (Fig 3.5b). Having observed this feature of the subjects' behavior, we used training sessions to rotate the subject's kernel, and then used tasks aligned with their kernels on sessions when we recorded.

To quantitatively describe the influence on a subject's kernel of prior task experience, we devised a multilinear regression model, fit separately to each of the two subjects, describing the circular mean (henceforth "peak") of the kernel on a given training session as a weighted average of the (nominally positive) discriminanda orientations presented during the subject's recent training history. The weights fit by these models describe the influence of training sessions as a function of number of days back. We found that such a model led to remarkably accurate predictions of day-to-day variability in kernel peak (80% and 62% variance explained; Fig. 3.5a,d). In both subjects, the fitted weights assigned to a session  $n$  training days back tended to decrease with  $n$ , demonstrating that training sessions that were most recent tended to have the most influence on the location of the kernel peak (Fig. 3.5c). However, the weights were significantly non-zero for at least ten training sessions back (equivalent to about 50,000 trials), demonstrating that the subjects' decision strategies are based on discriminanda that have not been presented for weeks, a striking quantitative demonstration of the hysteresis characterizing their task strategy. We even found that, in monkey '*lem*', the



**Figure 3.5. Timecourse of kernel rotation.** **a.** Dynamics of kernel circular mean (“peak”) orientation over a training period of nine weeks for monkey ‘*lem*’ performing Task 1. Each data point represents a single day-long training session. The nominally positive discriminandum orientation for each session is shown in black. The observed kernel peak  $\pm 1$  bootstrap s.e. is in gray. The kernel peak predicted by a multilinear regression model based on prior training history is in blue. Dashed green lines indicate abrupt task switches. **b.** The average kernel peak misalignment as a function of training sessions elapsed since an abrupt task switch. **c.** The weights assigned to prior training history by the multilinear regression model, as a function of number of days back. Error bars indicate  $\pm 1$  s.e. obtained by cross-validation. **d.** Scatter plot of observed versus predicted kernel peak, demonstrating the model’s success in accounting for the variability in kernel peak using prior training history. Blue and red dots indicate sessions from monkeys ‘*lem*’ and ‘*ruf*’, respectively (n=461 total).

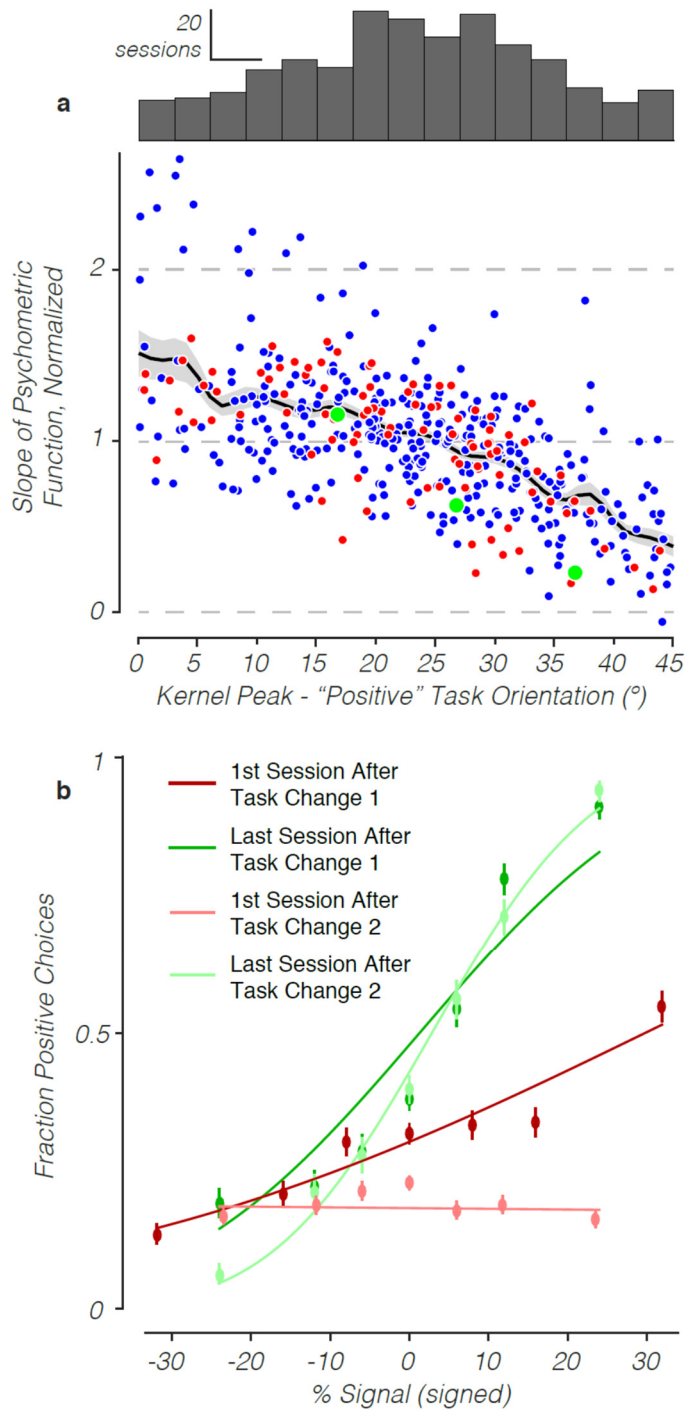
weight for  $n=0$  (i.e. the current session) was near zero, suggesting that, for this animal, the task instruction on a given day does not begin to exert much influence on the task strategy until the follow training session!

#### *3.4.4 Kernel Misalignment Impairs Performance*

Using only the angle of kernel misalignment, we could predict a significant portion of the day-to-day variability in the animal's performance (39%, cross-validated 2<sup>nd</sup>-order polynomial regression). The systematic relationship between the angle of kernel misalignment and performance was consistent with the prediction from the simple model presented earlier, showing relatively modest impairment with small angles of misalignment but profound impairments otherwise (Fig. 3.6a). After the abrupt task switches, chosen to specifically challenge the previous day's kernel, performance was frequently reduced to chance levels (Fig. 3.6b) for the entire training session, suggesting that the slow timecourse of kernel changes incurs severe reductions in reward rate on some sessions.

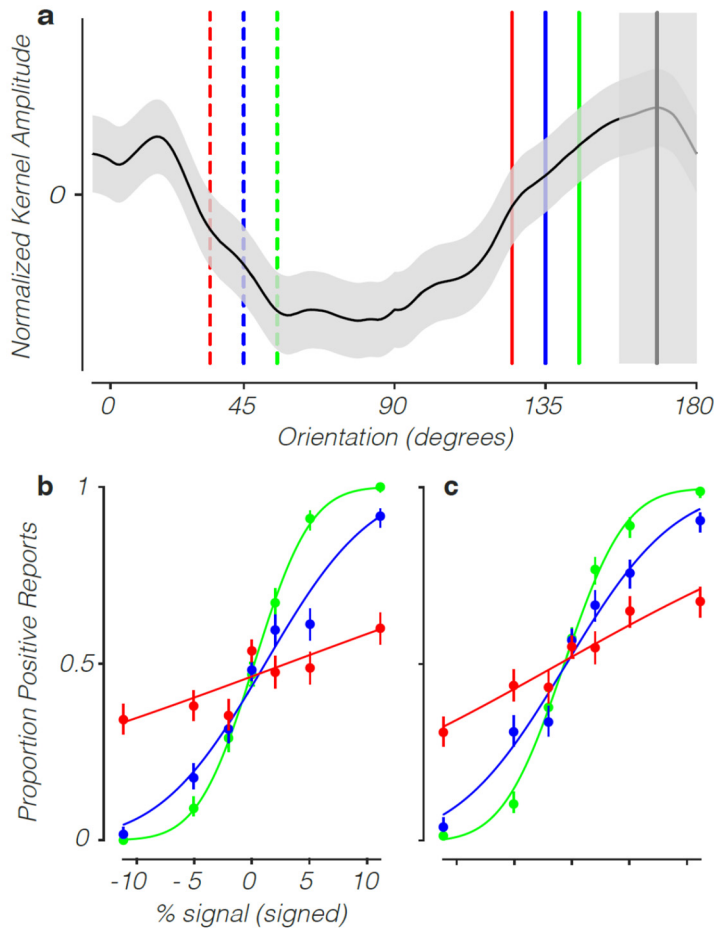
We wondered whether the drop in performance on sessions when the animal was using a misaligned kernel came about as a secondary effect, perhaps due to decreased motivation caused by frustration with a new task, or whether it could be simply predicted given the shape of the kernel. We investigated this by modifying the structure of the task on a subset of training sessions, such that the animal was confronted with more than a single set of discriminanda, which were randomly interleaved trial to trial. Only at the end of the trial, when the choice targets appeared, was the animal explicitly cued about the discriminanda for that trial.





**Figure 3.6. Observed effect of task strategy “misalignment” on performance.** **a.** Scatter plot of performance (as measured using the slope of the psychometric function) against the angle of kernel misalignment. Blue and red dots indicate sessions from monkey ‘lem’ and ‘ruf’, respectively (n=359 and 102). Average across both animals, smoothed with a Gaussian with 1° s.d., is shown in black, +/- 1 bootstrap s.e. Green dots refer to subsets of trials from an example session in monkey ‘lem’ (see Fig. 3.7). Marginal distribution of kernel misalignments, for both animals, is shown with the histogram (top). **b.** Example psychometric curves comparing performance on monkey ‘lem’ just before (green) and just after (red) the abrupt task switches shown in Fig. 3.5a. Just after task switches, when kernels were most misaligned, performance was sharply reduced, to chance or near-chance levels for the entire session (one session typically lasted about 3,000 trials).

The idea of using this interleaved task design was to minimize variability in cognitive factors like arousal, motivation, and the like that could differentiate performance across sets of orientations, such that the remaining variability would be explained simply by the kernel.



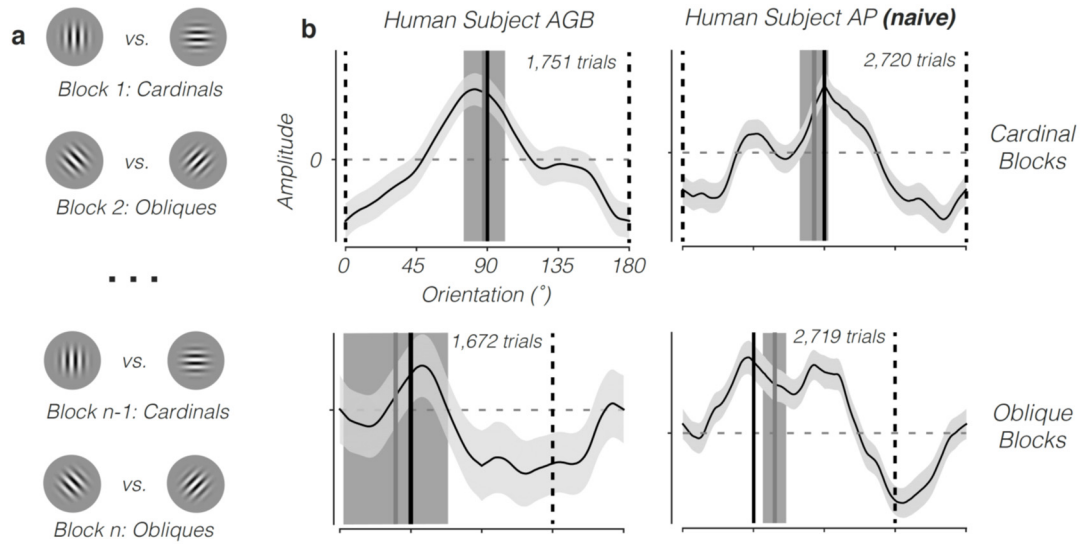
**Figure 3.7. Performance with interleaved orientation pairs.** **a.** Kernel  $\pm 1$  bootstrap s.e. observed on an example session when monkey ‘lem’ performed Task 1, with 3 pairs of near-oblique orientations presented in randomly interleaved trials. Vertical gray line indicates kernel circular mean, and vertical gray bar indicates  $\pm$  bootstrap s.e. Colored lines indicate the three pairs of interleaved orientations. Solid and dashed lines indicate the nominally “positive” and “negative” discriminandum orientations, respectively. **b.** Psychometric functions plotted separately for each pair of discriminanda. Performance is negatively correlated with the angle between the circular mean of the kernel and the positive discriminandum orientation. Error bars indicate 95%

bootstrap coverage intervals, and dashed lines are probit fits. **c.** Psychometric functions predicted by projecting stimuli onto the kernel, with Gaussian noise added to match overall percent correct. The match with the observed psychometric functions indicated that the differences in performance between the discriminanda can be explained simply as a result of a fixed task strategy. Error bars and axis labels are as in (b).

We calculated the animal's performance separately for each orthogonal pair of orientations and found significant differences in performance (Fig. 3.7a). These were entirely predictable given the kernel (Fig. 3.7b,c), and performance followed the same function of kernel misalignment as seen across days (see green dots in Fig. 3.6). This confirms that the effect of kernel misalignment of performance can be explained as a direct effect of the shape of the kernel.

#### *3.4.5 Comparison with Human Subjects*

The finding that the macaque subjects required many days to update their task strategy suggested to us a profound difference in human and non-human behavior. To directly test this, we trained two human subjects (one completely naïve, trained much like a monkey—see Methods) on Task 1. The human subjects were presented only with two tasks—oblique and cardinal discriminations—in alternating blocks of 100 trials (Fig. 3.8a). There was no cue indicating a block switch, although as with the monkeys, Gabor icons were presented after the stimulus presentation indicating the discriminanda. The humans reported their choice with a button press, and received feedback in the form of a full-field flash after incorrect choices. The human subjects achieved thresholds comparable to the macaques.



**Figure 3.8. Human subjects rapidly update task strategy.** **a.** Schematic illustrating the blockwise design of the human version of Task 1. Each block consisted of 100 trials. **b.** Kernels plotted individually by subject and block condition,  $\pm 1$  bootstrap s.e, combined across 5 sessions for each subject. Upright gray dashed lines and transparent gray regions indicate the circular mean  $\pm 1$  bootstrap s.e. Black bar indicates the nominally “positive” orientation. Subjects were able to adopt two distinct kernels for the two types of blocks.

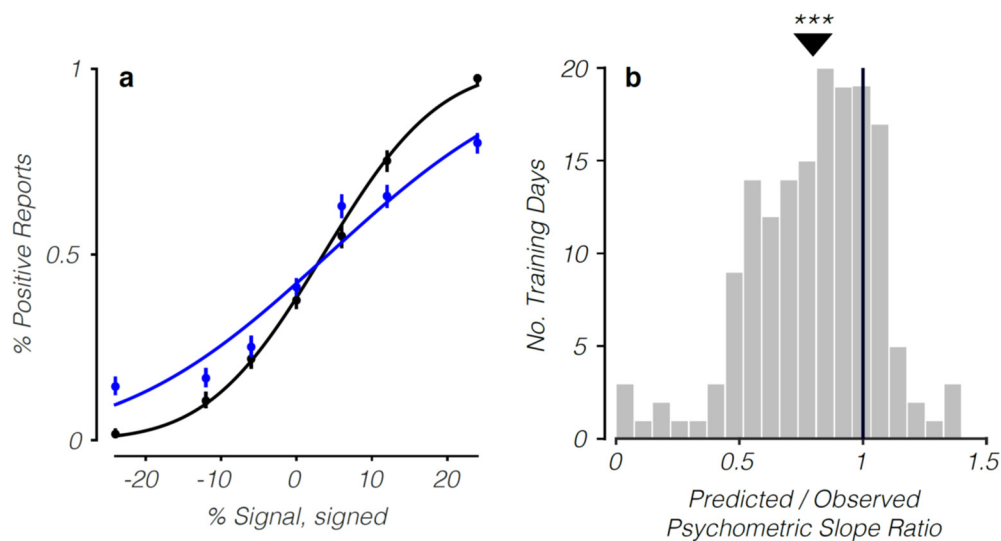
We measured kernels in the humans, separately according to the two block conditions, but combined across sessions (5 sessions per subject). The results were clear: subjects were able to adopt two different kernels, one appropriate for the cardinal discrimination blocks and one appropriate for the oblique discrimination blocks (Fig. 3.8b). Thus, they are capable of flexibly switching task strategies on the order of, at most, tens of trials, confirming our hypothesis of an interspecies difference. We discuss later our conclusions about what this may indicate about the relative cognitive abilities of humans and non-human primates and whether changes in task design may be used to accelerate the rate at which animal subjects update their task strategy.

### *3.4.6 Are Subjects Really at Threshold?*

A common goal of experimental paradigms in which animal subjects make perceptual judgments is to reduce the signal-to-noise ratio of the stimuli until animals are forced to fully exploit the sensory evidence available to them. When subjects are at this threshold level of performance, changes in nominal signal level are, by definition, impossible to distinguish from noise. Subjects must learn that a procedure that feels like educated guessing in fact yields reward rates that are considerably higher than chance. If subjects are not at threshold, they can adopt a strategy of waiting for high-signal trials and then simply choose randomly on trials they detect as having little or no signal, a potentially serious behavioral confound. We demonstrate a novel means of detecting such behavior, and show that monkey subjects do, to some extent, employ it, even though typical methods of ensuring threshold performance were used.

The principle behind our approach was to compare estimates of the animal's sensitivity to stimulus fluctuations, using self-consistency measures, and their sensitivity to changes in nominal signal level, given by the shape of the psychometric function. To do this, we used the measured kernel as a means of modeling subject's choices, as discussed above. This time, we added Gaussian noise to the projections of the stimuli onto the kernel prior to thresholding. The s.d. of this noise term was constrained so that the self-consistency of the resulting choices matched the observed self-consistency of the subject. If the subject is truly at threshold, this should provide an accurate model of the subject's choices. If instead, this generates a psychometric function whose slope is lower than the one observed, it means the subject performs better in response to changes in nominal signal level than would be predicted based on their sensitivity to fluctuations in

the stimulus. This suggests that subjects are in fact able to distinguish stimulus fluctuations from changes in nominal signal level and are at their most sensitive when they detect the presence of nominal signal. This would mean they are not at threshold. In principle, this model can also overestimate the slope. This would mean that subjects are systematically more sensitive on zero-signal trials. It seems unlikely that subjects would systematically adopt such a strategy.



**Figure 3.9. Self-consistency underpredicts overall performance. a.** The psychometric function observed for an example session, monkey ‘*lem*’ (black). The psychometric function predicted using the observed kernel and decision noise constrained by the subject’s observed self-consistency in response to repeated trials (blue). Error bars are 95% bootstrap coverage intervals, and solid curves are probit fits. **b.** The distribution of slopes ratios of the psychometric functions (predicted/observed) across sessions. Only sessions for which the estimated kernel provided an adequate model of the subject’s behavior (that is, prediction accuracy matched observed self-consistency) were used (n=161 sessions, both subjects). The average slope ratio (0.8) was significantly less than 1 ( $p < 10^{-5}$ ), indicating that subjects are generally more sensitive to changes in nominal signal level than would be predicted by their sensitivity to variability in the stimulus.

For the majority of sessions (77%), we found that the slope reconstructed using this method was indeed smaller than the slope of the observed psychometric function, indicating that animals are less sensitive to stimulus fluctuations than to changes in

nominal signal level (Fig. 3.9). Across all sessions, the average slope ratio (predicted/observed) was 0.8. We observed considerable variability in the slope ratio across sessions, although it is difficult to say how much of this may be due to noise in the measurements that constrain the predictions. We conclude that, in general, subjects are able to distinguish stimulus fluctuations from changes in nominal signal level and are at their most sensitive when they detect the presence of nominal signal, violating a necessary condition of threshold performance (although this effect is modest).

### **3.5 Discussion**

In this chapter, I discussed the use of psychophysical reverse correlation to probe the decision strategies of macaque subjects performing a 2AFC orientation discrimination task. Developed as a tool for human psychophysics, PRC has been less often used in animal studies. However, it is particularly useful for studying animal behavior, because our understanding of the decision strategies they employ is inherently limited, yet crucial for interpreting the data.

The most striking result we obtained was that macaque subjects are remarkably slow to update their decision strategies when task instruction changes. We identified a significant influence of prior training history on the shape of the kernels that persisted for days or even weeks. This hysteresis impaired, in some cases devastatingly so, the subject's ability to receive reward. By contrast, human subjects performing the same task were able to rapidly update their task strategy in a blockwise fashion within single sessions.

We considered whether the macaque subjects update their task strategies so slowly because they are incapable of doing otherwise. Alternatively, this may be a clever strategy lets the subjects achieve a desired rate of reward without having to change their kernel. A definite answer to this question would require a better understanding of the unknown objective functions that monkey subjects attempt to optimize in experimental settings. While it may be correct to presume that human subjects seek to maximize percent correct, monkey subjects likely assign positive weight to the rate of reward, negative weight to effort and attention expended, as well as weights on other factors that are not well understood. This makes it hard to use behavioral measures to assess their true capabilities.

Nonetheless, several pieces of evidence point towards the view that our results reveal a novel and fundamental limitation. First, if the monkeys are in fact able to more quickly update their kernels, the fact that they did not do so suggests a considerable degree of effort that is traded off in some way against the increased rate of reward it would have entailed. Second, the slow updating of their kernels leads, in some cases, to severe and prolonged decreases in reward rate that persists for days. This was particularly striking on the occasions when we deliberately switched the task such that the subject's kernel should dramatically impair performance. These abrupt task switches caused the subject's performance to drop to near-chance levels, strong evidence that subjects find great difficulty in efficiently updating their kernels, even when it is necessary to obtain reward. This suggests a degree of inflexibility in the sensorimotor association underlying their performance that was unexpected. Nonetheless, it seems likely that changes in the



design of the task and training protocol may ameliorate this deficit. We discuss possible examples in Chapter 4.

Whether or not these can mitigate the slow kernel updating we observed, the task design and training methods we used are similar to those that have been widely used in other studies involving animal subjects. This strongly suggests that animal subjects in many prior studies were not performing the task as instructed. As we have argued, measures of performance are not necessarily useful in guarding against this possibility. While we did observe some training sessions where the subject's kernel misalignment dramatically impaired performance, these were typically the result of changed task instruction designed to produce this behavior based on prior knowledge of the subject's kernel. Without this prior knowledge, these profound deficits are unlikely to be observed frequently. When they have been observed, experimenters have quite reasonably removed these sessions from their analysis, since good behavior is a prerequisite. The typical regime is the one in which the kernel *is* misaligned, but not so much as to produce a performance deficit that cannot simply be attributed to day-to-day variability in performance. Given that a systematic feature of measures that depend on animal subject's performance (CPs, neurometric/psychometric threshold ratios, and others) is their pronounced variability (Britten et al., 1992; Yu, Dickman, DeAngelis, & Angelaki, 2015), we speculate that some of this may have been due to the variable location of subject's kernels, a factor which has not previously been considered.

These results also suggest the possibility of an even more serious behavioral confound for neuroscience studies in which experimenters seek to understand an effect on neuronal activity of dynamically changing task instruction. Such studies may have

inadvertently compared data across conditions associated with identical, or at least very similar, decision strategies, particularly if the timescale of dynamic changes in the task is short (e.g. Cohen & Newsome, 2008). Including objective measures of the task strategies subjects actually employ will be an extremely useful means of avoiding such confounds in the future (see our use of PRC in Chapter 2).

### 3.6 Methods

#### *Modeling A Misaligned Task Strategy*

To explore the impact on performance of a misaligned task strategy, we modeled the behavior of a subject performing Task 1. The model subject performed the task simply by comparing the relative power at two orthogonal orientations (that is, the task strategy only has support at two orientations, where it assigns values of equal magnitude and opposite sign). By construction of the stimulus (see Chapter 2 Methods), fluctuations in orientation power are normal, and it seems reasonable to model internal noise as an additional Gaussian term. Thus we can model the subject's performance using the discriminability of two Gaussians with equal variance (given by the sum of the two noise terms) whose means are determined by the amplitude of the orientation filter applied to the stimulus at the two orientations used in the decision. Proportion correct is then given analytically:  $0.5 + 0.5 * \tanh \frac{d'}{\sqrt{2}}$ , where

$$d' = \frac{\eta_W(\theta, \sigma_{ori}) - \eta_W(90 + \theta, \sigma_{ori})}{\sigma_{noise}} \quad (3.1)$$

$\eta_W(\Phi, \sigma)$  is the wrapped normal distribution with 180°-period, mean 0°, standard deviation  $\sigma$ , evaluated at angle  $\Phi$  in degrees.  $\sigma_{ori}$  is the standard deviation of the

orientation filter applied to the stimulus, and  $\sigma_{noise}$  is the standard deviation of the noise term.  $\theta$  is the angle between the decision orientations and the discriminanda. We set  $\sigma_{noise}$  to a fixed value to match the average percent correct of the animal subjects across a typical range of orientation bandwidths.

This equation also has the form to describe the performance of a model subject whose stimulus filters are broad, but where the distribution of orientation in the stimulus is discrete, such that signal is varied by changing the relative orientation energy at two orthogonal orientations. In this case  $\theta$  describes the angle between the discriminanda and the center of a wrapped normal distribution describing the subject's task strategy. We expect, but do not explicitly show, that the form of the solution would be qualitatively similar if both the task strategy and the distribution of stimulus energy were allowed to be broad.

### *Predicting Kernel Peaks from Training History*

We used a multilinear regression model to predict the observed circular mean of a kernel on a given session  $n$  ( $K_n$ ) given the set of nominally positive discriminanda orientations presented on  $m$  prior training sessions ( $O_{n-m}, O_{n-m+1}, \dots, O_{n-1}$ ) and on the current session ( $O_n$ ). Because both the independent and explanatory variables in the model are circular, we used a circular-circular regression method that treats the problem linearly in terms of trigonometric functions of the original variables (Sarma & Jammalamadaka, 1993). The regression equations were given by:

$$\sin K_n = \beta_{sin} + \sum_{i=0}^m \beta_i \sin O_{n-i} + \varepsilon \quad (3.2)$$

and

$$\cos K_n = \beta_{cos} + \sum_{i=0}^m \beta_i \cos O_{n-i} + \varepsilon. \quad (3.3)$$

$\beta_i$  is the coefficient reflecting the weight of the discriminandum orientation  $i$  sessions back, and  $\beta_{sin}$  and  $\beta_{cos}$  are model constants. Circular variables were doubled, to account for the 180-degree period of orientation. The regression was solved with Tikhonov-regularized least-squares. For the regularization, we used a 2<sup>nd</sup>-order finite difference operator (corresponding to convolution with the vector [1 -2 1]) as the Tikhonov matrix, in order to confer temporal smoothness on the coefficients. The optimal strength of regularization was determined empirically under 50-fold cross-validation. We also picked the training history length ( $m$ ) empirically in the same manner. For monkey ‘*lem*’, we used  $m=33$  days back. For monkey ‘*ruf*’, we used  $m=12$  days back.

We experimented with an extended version of this model that allowed cross-terms (that is, terms relating the sine of the explanatory variables to the cosine of the independent variable) but found that these additional terms did not improve model performance.

#### *Calculating the Slope of the Psychometric Function*

We used a probit model to fit the psychometric functions. The fitted dispersion parameter was used to quantify the slope of the fitted curve. (This is equivalent to the inverse of the s.d. of best-fitting cumulative Gaussian, and is proportional to the peak in the first derivative of the fitted curve.) Because % signal was not directly comparable for the two task variants and due to inter-subject differences, slopes are shown after normalizing by the average across sessions, separately for each task variant and subject.

### *Human Psychophysics*

Human subjects performed a version of Task 1 (described in detail in Chapter 2). The stimuli were displayed in much the same way, with a few minor differences. The monitors subtended a smaller visual angle ( $21.8^\circ \times 17.4^\circ$ ), and subjects maintained head-fixation using a chin rest. Subjects were instructed to fixate a central square, but eye position was not monitored. Stimuli were displayed for a duration of 600ms at 100 Hz, after which time the subject was free to make a choice by pressing one of two buttons. Incorrect choices were followed by a brief full-field flash. The stimulus was always displayed along the lower vertical meridian, with a peak spatial frequency of 1.1 cpd. For the two subjects AP and AGB, median envelope s.d.'s were  $0.7^\circ$  and  $0.4^\circ$ , and median eccentricities were  $3.4^\circ$  and  $2.8^\circ$ , respectively. The choice targets appeared to the left and right of the fixation point. Discriminanda within the range  $0-90^\circ$  corresponded to a rightward choice target and a right button press. The naïve subject was told only that left and right button presses indicated choices, and that the feedback indicated an incorrect choice. No other instructions were given, so that this subject had to deduce the task requirements from feedback alone, as did the monkeys. Block switches were uncued.

## **Chapter 4:**

### **Summary and Future Directions**

#### **4.1 Chapter 2 Summary and Future Directions**

Variability in the activity of sensory neurons in response to a fixed sensory stimulus is typically described as “noise” in the sensory representation. If this is correct, this variability places important constraints on the ability of downstream processes to correctly infer the sensory input. These constraints have been explored in a number of influential theoretical studies, with two notable conclusions about the consequences of structured noise correlations on 2AFC discrimination. First, these affect the reliability with which the stimulus can be reconstructed, and second, they affect the magnitude of correlations between variability in single neurons and perceptual choice.

In Chapter 2, I presented the results of an experiment designed to test whether structured noise correlations are fixed or change dynamically with a subject’s task. The strategy was to record from populations of V1 neurons in behaving macaques using multi-electrode arrays, while the subjects performed an orientation discrimination task using orthogonal discriminanda. By changing the set of discriminanda from session to session, we could identify fixed and task-dependent components of noise correlation structure.

This experiment revealed that the structure of noise correlations changed systematically with the subject’s task:  $r_{sc}$  was high for pairs of neurons preferring the same discriminandum orientation and low for pairs preferring opposite discriminanda. A

multilinear regression model could not identify a significant portion of structured noise correlations that remained fixed.

These data are consistent with the presence of a variable feedback signal that increases the firing rate of neurons contributing to one decision pool while decreasing the firing rate of neurons contributing to the other pool. Under the assumption that the noise correlations are naively read out using linear pooling, we found they would generate CP that quantitatively matched the data. Thus, the data may be reconciled with the interpretation of CP as a causal effect of sensory neuronal variability on choice, under the following assumptions: 1) the effects of feedback are read out as though they were sensory evidence, and 2) the feedback itself is not influenced by or in some other way correlated with the decision. Importantly, this is the first experimental evidence supporting the predictions of linear pooling relating CP and structured noise correlations, because we are the first to look at the entire correlation matrix.

However, the necessity of the assumptions above presents problems with this interpretation, given other aspects of the data presented in Chapter 2 and in other studies. One such problem is that the structure of the noise correlations introduced by feedback are “differential”: that is, they mimic variability in the stimulus along the task axis. If these are naively read out as sensory evidence, this would contaminate the sensory evidence and reduce the subject’s performance. While quantitative studies would be needed to better understand the magnitude of this effect, one intriguing possibility is that downstream computations discount the influence of feedback on the sensory representation so as to avoid this, effectively segregating sensory and non-sensory components of variability. This would attenuate the feedforward effect of the structured

noise correlations on choice, necessitating the interpretation that some of observed CP is an effect of feedback.

Next, several studies have found that CP plateaus or increases consistently (see §2.7.3; Hendrikje Nienborg & Cumming, 2009; Shadlen et al., 1996)) over the course of a fixed-duration trial. This cannot be reconciled with the reduced influence of stimulus fluctuations on choices, as measured with PRC (Kiani et al., 2008; Hendrikje Nienborg & Cumming, 2009). This suggests that CP partially reflects an effect of the decision process itself.

Perhaps the strongest evidence comes from studies in which subjects discriminate stimuli that are bistable (Dodd et al., 2001; see §1.3.1). This perceptual property implies a weak influence of sensory noise on the perceptual choice. The strong CPs observed in these studies therefore supports the alternative view that the variability in the subject's perceptual state causally influences the firing rate of sensory neurons via feedback.

Given these data, we believe our results support the view that the task-dependent structured noise correlation we observe in V1 reflects variability in a feedback signal related to the decision variable itself. Consequently, we conclude that signals related to perceptual inference/decision-making are present even at the earliest cortical processing stage for vision. This presents a fundamentally new view of sensory neuronal variability and suggests a complex relationship between sensory and non-sensory brain areas that cannot be captured by traditional models that divide perceptual processes into “representational” and “inferential” components (Johnson, 1980; Tolhurst et al., 1983). Because most theoretical studies have not included analysis of choice-related feedback



signals, it is not currently possible to make quantitative statements about the downstream effects of this feedback on the decision process.

These results point to two directions for future research. The first is to investigate the unknown neural circuitry underlying the novel feedback that we observed as correlated variability in V1. One approach would be to causally manipulate candidate source areas while measuring any changes in the structure of correlated noise in V1. One potential candidate for the (at least proximate) source of the effects we observe in V1 is the dense cortico-cortical feedback from area V2 (Perkel, Bullier, & Kennedy, 1986). To address precisely this possibility, Rick Born's group is currently undertaking a replication of our experiment with and without inactivation of area V2 using implanted cryoloops (Ponce, Lomber, & Born, 2008; Camille Gomez-Laberge and Rick Born, *personal communication*). If the task-dependent noise correlation structure is abolished or otherwise perturbed under V2 inactivation, this would add causal evidence to the argument we have made in favor of a top-down origin and would shed insight on its anatomical origins. A similar approach would be to use simultaneous recordings in V1 and V2 or another candidate area to predict, in the trial-to-trial manner, the state of correlated fluctuations in V1 based on neuronal activity in the candidate area.

Apart from corticocortical feedback, inputs from subcortical areas are also interesting potential sources of choice-related feedback to sensory cortices. For instance, area V1 receives significant extra-retinal inputs from the LGN (McAlonan, Cavanaugh, & Wurtz, 2008) and the pulvinar (Purushothaman, Marion, Li, & Casagrande, 2012). The pulvinar in particular serves as an important site of convergent inputs from diverse brain regions (Berman & Wurtz, 2011; Wurtz, McAlonan, Cavanaugh, & Berman, 2011)

including the basal ganglia (indirectly through the superior colliculus), a set of structures increasingly thought to play an important role in perceptual decision making (Ding & Gold, 2013; Krauzlis, Bollimunta, Arcizet, & Wang, 2014). The pulvinar projects broadly throughout visual cortex, making it potentially useful as a means of broadcasting aspects of the decision process to diverse visual areas. Future investigations into the role these and other structures play in perceptual decision making and in modulating the choice-related activity of sensory neurons may shed light on the circuit architecture underlying our results.

The finding that feedback to V1 introduces “differential” correlations also suggests serious problems for experimental attempts to estimate the “information capacity” of sensory neuronal populations using simultaneous recordings, an active area of current research (Cohen & Maunsell, 2009; Graf, Kohn, Jazayeri, & Movshon, 2011; Moreno-Bote et al., 2014; Zohary et al., 1994). The underlying assumption of these studies is that all correlated variability is equivalent to noise in the sensory representation. If a component is not, the results may not apply. More specifically, without knowledge of the state of centrally-generated signals originating elsewhere in the brain, the estimates of information capacity will necessarily be downwardly biased, and by an unknown amount. Therefore, such studies cannot provide meaningful results until the state of internal variables that influence correlations, and their role in the decision process, is better understood.

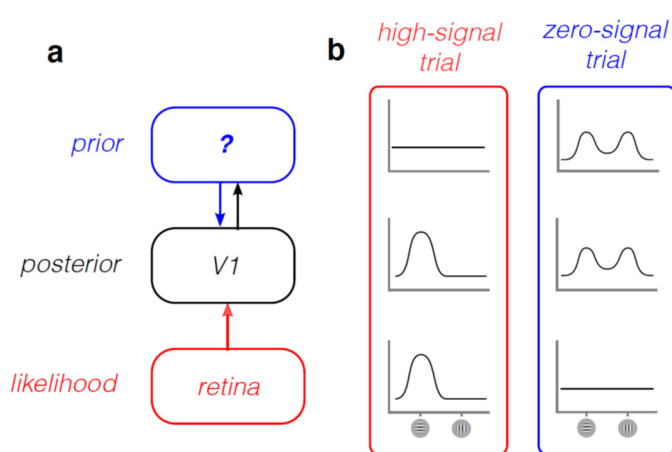
A second promising avenue for future research is computational modeling. Currently, our understanding of the role of sensory neuronal variability in perceptual decision making is largely based on theoretical models, particularly because we lack the

experimental tools to measure variability at the population level and to manipulate it in ways that may reveal its causal influence on decisions. Given the presence of choice-related feedback in sensory neurons, purely feedforward models are unlikely to provide a complete description of the processes underlying perceptual decision making.

A significant problem with incorporating feedback signals into these models is that they introduce additional degrees of freedom without sufficient constraints to make the questions well posed. One framework that incorporates feedback in a constrained, principled way has been proposed by Ralf Haefner and colleagues (Haefner et al., 2014). It describes perceptual decision making through the lens of Bayesian inference, assigning the algorithmic components of Bayesian inference (likelihoods, priors, and posteriors) to neuronal substrates of visual perception (Fig. 4.1a). (Currently the model has been formulated to describe data from the visual system, although there is no reason its results are not applicable to other modalities).

In this framework, the likelihood consists of the sensory evidence transmitted in a feedforward manner from the retina. The prior consists of expectations about the structure of the visual world gained through past experience. This is represented as internally-generated feedback from downstream brain areas. These downstream areas are also responsible for the decision itself, but in the absence of sensory evidence their activity is dominated by the influence of the prior. For instance, during performance of the orthogonal orientation discrimination task we used, this prior would contain two peaks corresponding to the discriminanda orientations (Fig. 4.1b). Lastly, the posterior is represented in the population activity of visual cortical neurons themselves. This is the biggest departure from the pooling model. In the pooling model, activity in these neurons

represents a stochastic encoding of the visual input. In the Haefner model, it represents a probabilistic representation of the subject's *belief* about the presence of features in the visual world, which includes both the feedforward influence of sensory evidence and feedback influence of the prior. Importantly, neurons represent all these variables probabilistically, creating trial-to-trial fluctuations that allow for predictions about CP and noise correlations. Finally, sensory neuronal activity causally influences the downstream areas, creating a self-reinforcing loop which generates dynamical structure over the course of single trials. (For this reason, the Haefner model falls into the category of models illustrated in Fig. 2.10b.)



**Figure 4.1. The Haefner model.**

**a.** As part of the generative model, sensory neurons (here, in V1) represent the posterior likelihood of the presence of visual features in the world. They are influenced both by a *prior* encoding context, beliefs, and expectations, and a *likelihood* encoding retinal input. **B.** On high-signal trials (trial with horizontal orientation shown) the strong likelihood dominates the posterior via feedforward inputs.

On low-signal trials, the structure of the prior (induced by the expectation implicit in the task that one of the two discriminanda is present) becomes apparent, and affects the posterior through feedback. All curves represent the probability distribution of the encoded variable over orientation. (This figure is adapted from one that was generously shared by Richard Lange and Ralf Haefner).

This model can reproduce several key findings in the perceptual decision making literature. First, it reproduces the finding that early evidence is more strongly weighted in a fixed-duration discrimination task. The reason for this is the positive feedback loop that results from feedforward and feedback connections. Early evidence biases the

decision variable through feedforward connections, and the resulting feedback attenuates the influence of sensory input presented later in the trial.

Second, the model predicts task-dependent structured noise correlations during a 2AFC orientation discrimination task very similar to what we observed. The reason for this is that the subject's prior contains peaks corresponding to the discriminanda (Fig. 4.1b). The resulting feedback signal is sampled from this bimodal distribution (i.e. alternates between the two discriminanda across time) generating a lattice-like pattern of noise correlations similar to what we observed (Fig. 2.4e).

Lastly, the model predicts a dynamic timecourse of CP that qualitatively matches our data (see §2.7.3) and those of past studies (Nienborg & Cumming, 2009). Early in the trial, CP primarily reflects the feedforward pathway because no coherent top-down belief has formed yet (that is, feedback alternates between both alternatives). As the trial progresses, the top-down belief becomes stronger (that is, more peaked around one orientation) due to the self-reinforcing feedforward and feedback connections in the model. This generates CP of higher magnitude later in the trial.

Apart from its good fit to neuronal data, this model is also appealing because it provides a normative framework for understanding what otherwise appears to be a curious feature of our data: choice-related feedback “contaminates” the sensory representation. Under the rationale of the Haefner model, this contamination is a means of implementing prior beliefs. In the strict confines of psychophysical tasks, this may lead to suboptimal performance, since it leads to confusion about the causes of sensory neuronal firing. However, in naturalistic contexts, combining prior beliefs with sensory information, particularly when the sensory input is weak or ambiguous, may be adaptive.

Implementing this combination at early stages of sensory processing (and thus eliminating any unbiased stage of sensory representation) may have an additional advantage of promoting stability and consistency in any parallel downstream computations that rely on the sensory information (Luu & Stocker, 2016).

The Haefner model has succeeded in providing a parsimonious account of a great deal of neuronal data that has been difficult to reconcile otherwise. However it requires further experimental validation. To this end, it makes several untested predictions. These emerge from the fact that the feedback relates directly to a subject having learned the implicit structure of the task. This predicts that CP and structured noise correlations will emerge over the course of learning and be correlated positively with performance<sup>6</sup>. Our data provide provisional experimental support for the second prediction. We found that one animal had consistently lower behavioral thresholds, and we also observed stronger CPs and structured noise correlations in this animal (data not shown). The experiments mentioned previously that are being undertaken by Rick Born's group are also designed to explicitly examine the evolution of CP and structured noise correlations during learning. They plan to begin neuronal recordings in naïve monkey subjects, and continue recording as subjects becomes expert on the task.

Another prediction is that the structure of noise correlations in a sensory area should reflect the structure of the task. For instance, in the context of 3AFC discrimination, the correlation matrix should have three peaks (corresponding to within-

---

<sup>6</sup> Increases in CP with learning have been previously observed (Law & Gold, 2008) and can be due to a variety of factors, such as decreases in pooling noise, making this a poor test for the Haefner model. However, a learning-associated increase in the amplitude of noise correlation structure is a unique prediction.

pool pairs) and three troughs (corresponding to between-pool pairs). This is a novel prediction which awaits experimental testing.

The Haefner model makes a strong commitment to the view of perception as Bayesian inference, which may not be necessary for generating the model predictions. It will be important to know if other models of perceptual decision making that include choice-related feedback can explain the data without an explicitly Bayesian explanation.

#### **4.2 Chapter 3 Summary and Future Directions**

In Chapter 2, we identified a source of task-dependent feedback that manifested as changes in  $r_{sc}$  structure in V1. A crucial assumption of our experimental design was that we could successfully manipulate subjects' allocation of internal resources by changing the set of orientations they needed to discriminate to receive reward. Unlike prior studies, we objectively measured subjects' task strategies using psychophysical reverse correlation, to confirm the validity of this assumption. We showed that they did update their psychophysical kernels to reflect task instruction.

In Chapter 3, we explored aspects of the subjects' psychophysical performance in more detail. Our main finding was that the monkey subjects slowly "rotate" their psychophysical kernels over a timescale of tens of thousands of trials after task instruction changes. Because task instruction was typically changed more frequently than this, the subjects usually used a task strategy which was somewhat misaligned. Occasionally, we chose a task to maximize difficulty given the subject's kernel location. During these sessions, we found that the subjects' performance dropped to chance levels for extended periods of time (>1,000 trials). By contrast, human subjects performing the

same task were able to rapidly adopt an appropriate kernel within tens of trials after a change in the discriminanda orientations.

We found that we could predict almost half (39%) of the variability in the monkey subjects' performance across sessions only with knowledge of the angle between the kernel peak orientation and the task orientation. Despite this clear relationship, subjects were able to maintain reasonably good performance with a modest degree of misalignment. Given other sources of behavioral variability, we conclude that the modest impact of kernel misalignment would be difficult to detect without directly measuring it.

Our results suggest serious behavioral confounds in studies involving monkey psychophysics. These are likely to have gone unnoticed in prior studies. This is particularly concerning for studies in which the discriminanda are updated more frequently than subjects can update their kernels. Given that many such studies frequently change the discriminanda to match the preferences of recorded neurons, this is likely to be a widespread effect. Studies combining electrophysiology and psychophysics often report considerable variability in subject's psychophysical thresholds (Britten et al., 1992; Yu, Dickman, DeAngelis, & Angelaki, 2015). The slow strategy updating characterized here may be an important contributing factor. This would have introduced a source of variability in measures that contain a behavioral component that was not accounted for.

Studies of the neuronal effects of *dynamic* changes in task context may be particularly subject to confounds of this sort (Cohen & Newsome, 2008). In these studies, subjects may have been able to perform well across multiple task contexts while actually performing a fixed internal task strategy. Future studies will benefit from the use of PRC to identify and avoid confounds of this sort.



The slow kernel updating we observed in the monkey subjects is likely due to a combination of two factors: 1) a previously uncharacterized inflexibility in their sensorimotor contingencies, and 2) the use of training procedures that inadvertently incentivized them to adopt inflexible task strategies. Identifying the contribution of these two components will require further experimentation in training procedures and task designs. We discuss several changes that could be employed to encourage subjects to more rapidly update their task strategies.

First, increasing the cost of incorrect choices during training would effectively increase the penalty for kernel misalignment and potentially incentivize subjects to adopt a more flexible strategy. Although we did use a reward structure that penalized random guessing (see Chapter 2 Methods), the use of task designs in which there are more than two alternatives would much more strongly penalize this strategy (by reducing the reward rate from  $1/2$  to  $1/n$ , given  $n$  alternatives). However,  $n$ -choice tasks (for  $n > 2$ ) create significant analysis hurdles: for instance, more sophisticated methods of measuring decision strategies than those employed here would be required. A useful middle ground may be to use a task designed to encourage flexible kernels during early training and one designed for efficient analysis during later sessions that contribute to the final dataset.

Second, our subjects may also have been unintentionally encouraged to slowly update their strategy because we often kept the discriminanda the same for multiple training sessions in a row. If we had frequently changed the discriminanda at an early point in their training, this may have caused them to adopt more flexible kernels from the outset. Lastly, the fact that the choice target locations were fixed given the discriminanda (see Chapter 2 Methods) means that the subjects were not overtly incentivized to make

use of the oriented Gabor icons to inform their choices. In other words, they may have used a strategy that involved a low-level sensorimotor association rather than an abstract, categorical decision. Trial-by-trial changes in the choice target locations may have encouraged the subjects to make better use of the orientation cues to form a categorical decision. Further work will be needed to clarify whether these or other methods can be employed to circumvent the slow kernel updating our subjects employed.

For several reasons, we believe it is unlikely that the slow task strategy updating characteristic of the monkey subjects in our experiments is particular to discriminations in the orientation domain. First, it seems implausible that the behavioral limitation does not relate, at least in part, to constraints on domain-general processes downstream from visual cortex, rather than the specific way orientation is represented. Another argument relates to the widespread, albeit often anecdotal, observation that monkey subjects are poor to generalize across even slight changes in a task configuration. We speculate that the slow kernel updating we report may underlie these frequent reports of poor generalization. Finally, the most direct evidence comes from preliminary PRC data from monkeys performing a shape discrimination task (Long Sha and Roozbeh Kiani, *personal communication*). In this task, possible stimuli come from a set of complex shapes spanning a subspace defined by a set of predetermined features (for instance, one “dimension” might be indexed by the length of a triangle protruding from the right side of the object). The experimenters then arbitrarily chose hyperplanes in this high-dimensional parameter space as the discrimination boundary defining the task. Preliminary data suggest that, after a change in the discrimination boundary, monkey subjects require multiple days of retraining during which time their internal decision

boundary slowly becomes closer and closer to the new optimal boundary. These results appear qualitatively similar to what we have reported in the orientation domain, despite the fact that this task involves a highly abstracted feature space that is presumably represented in a population of neurons anterior to V1.

This shape discrimination task, while different from the task we used in a number of respects, also shares an important feature: changes in task instruction require subjects to reweight sensory evidence within a fixed feature space, and therefore likely contained within a single population of neurons. Another class of discrimination tasks involve stimuli that vary in multiple modalities encoded in physically distinct neuronal populations. For instance, Mante, Sussillo, Shenoy, & Newsome (2013) trained monkeys to perform a discrimination task involving a stimulus that could vary both in direction of motion and color. Importantly, motion and color are known to be represented in largely non-overlapping regions of visual cortex (DeYoe & Van Essen, 1988; Livingstone & Hubel, 1988; Zeki, 1978). Subjects were cued in a blockwise fashion which of the two features to discriminate to receive reward. Behavioral data confirmed that the subjects were able to rapidly adjust their task strategy from one using color to one using motion on the time scale of only tens of trials (three orders of magnitude faster than our monkeys updated their task strategies). We speculate that this ability may rely on the fact that subjects did not necessarily need to adjust any read out weights to do the task. Instead they only needed to switch modalities, which may be subserved by a switch between two fixed sets of readout weights applied to distinct sets of neurons.

## 4.3 Summary of Contributions

### *Summary of Contributions: Chapter 2*

- We found that the structure of spike-count correlations in macaque V1 is not fixed, but rather changes systematically with the task being performed. We excluded a number of possible stimulus-driven factors that could explain this, strongly implying that the source of these dynamic changes is centrally generated and is directly caused by subjects' engagement in the task.
- Using a multilinear regression model, we separated the noise correlation structure into two components: one that was fixed and one that changed dynamically with the task. The vast majority of the variance was explained by the dynamic component.
- The lattice-like correlation structure is consistent with the observed CPs, assuming linear pooling. However, this would require that the V1 spikes generated by feedback be naively read out as sensory evidence, because the fixed component of noise correlation structure was insufficient to generate CP of the appropriate magnitude. Alternatively, the source of feedback may be correlated with the ongoing decision process itself, implying that CP reflects an effect of choice on sensory neurons.
- The lattice-like correlation structure closely resembles “differential” correlations. This implies that feedback contaminates the sensory representation, unless downstream computations can segregate self-generated and external sources of V1 activity.

*Summary of Contributions: Chapter 3*

- Psychophysical reverse correlation demonstrates that monkey subjects' psychophysical kernels frequently diverge from those suggested by task instruction in an orthogonal orientation discrimination task. Underlying this is a strong hysteresis in their task strategies—after a change in task instruction, subjects slowly update their kernels on the order of tens of thousands of trials.
- We were able to predict 39% of the day-to-day variability in subjects' psychophysical performance only with knowledge of the angle of kernel misalignment, identifying a source of behavioral variability in monkey psychophysical performance that has not previously been well documented.
- By placing the discriminanda specifically to challenge subjects' kernels, we could reduce their performance to chance for entire sessions (>1000 trials), demonstrating a profound deficit in the flexibility of their sensorimotor contingencies that strongly impairs their ability to receive reward.
- Human subjects were able to rapidly update their psychophysical kernels to match task instruction even when the discriminanda were alternated in a blockwise fashion within single sessions, showing that the deficit is specific to animal subjects.
- We found that monkey subjects were slightly more sensitive to changes in nominal signal level than predicted based on their sensitivity to random stimulus fluctuations. This suggests they are able to distinguish these and were thus not fully at their psychophysical threshold.

## Bibliography

- Abbott, L. F., & Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Computation*, *11*(1), 91–101. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9950724>
- Ahumada Jr, A. J. (1996). Perceptual classification images from Vernier acuity masked by noise. In *Perception ECVF abstract* (Vol. 25, p. 0). Pion Ltd.
- Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews. Neuroscience*, *7*(5), 358–66. <http://doi.org/10.1038/nrn1888>
- Averbeck, B. B., & Lee, D. (2006). Effects of noise correlations on information encoding and decoding. *Journal of Neurophysiology*, *95*(6), 3633–3644. <http://doi.org/10.1152/jn.00919.2005>
- Bair, W., Zohary, E., & Newsome, W. T. (2001). Correlated firing in macaque visual area MT: time scales and relationship to behavior. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(5), 1676–97. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11222658>
- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, *29*(4), 1165–1188. <http://doi.org/10.1214/aos/1013699998>
- Berman, R. A., & Wurtz, R. H. (2011). Signals conveyed in the pulvinar pathway from superior colliculus to cortical area MT. *Journal of Neuroscience*, *31*(2), 373–384. <http://doi.org/10.1523/JNEUROSCI.4738-10.2011>
- 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(01), 87.

<http://doi.org/10.1017/S095252380000715X>

Britten, K., Shadlen, M., Newsome, W. T., & Movshon, J. (1992). The Analysis of Visual Motion : A Comparison of Neuronal and Psychophysical Performance. *The Journal of ...*, 12(December), 4745–4765. Retrieved from

<http://www.jneurosci.org/content/12/12/4745.short>

Callaway, E. M. (2004). Feedforward, feedback and inhibitory connections in primate visual cortex. *Neural Networks : The Official Journal of the International Neural Network Society*, 17(5-6), 625–32. <http://doi.org/10.1016/j.neunet.2004.04.004>

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51(13), 1484–1525. <http://doi.org/10.1016/j.visres.2011.04.012>

Celebrini, S., & Newsome, W. T. (1994). Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *The Journal of Neuroscience*, 14(7), 4109–4124. Retrieved from

<http://www.jneurosci.org/content/14/7/4109>  
<http://www.jneurosci.org/content/14/7/4109.full.pdf>  
<http://www.ncbi.nlm.nih.gov/pubmed/8027765>

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., ... Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience*, 13(3), 369–78.

<http://doi.org/10.1038/nn.2501>

Cohen, M. R., & Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nature Neuroscience*, 14(7), 811–9. <http://doi.org/10.1038/nn.2842>

- Cohen, M. R., & Maunsell, J. H. R. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*, *12*(12), 1594–600.  
<http://doi.org/10.1038/nn.2439>
- Cohen, M. R., & Newsome, W. T. (2008). Context-dependent changes in functional circuitry in visual area MT. *Neuron*, *60*(1), 162–73.  
<http://doi.org/10.1016/j.neuron.2008.08.007>
- Cook, E. P., & Maunsell, J. H. R. (2002). Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nature Neuroscience*, *5*(10), 985–94.  
<http://doi.org/10.1038/nn924>
- Crapse, T. B., & Basso, M. A. (2015). Insights into decision making using choice probability. *Journal of Neurophysiology*, *114*(6), 3039–3049.  
<http://doi.org/10.1152/jn.00335.2015>
- Cumming, B. G., & Nienborg, H. (2016). Feedforward and feedback sources of choice probability in neural population responses. *Current Opinion in Neurobiology*, *37*, 126–132. <http://doi.org/10.1016/j.conb.2016.01.009>
- Cumming, B. G., & Parker, A. J. (1999). Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(13), 5602–18. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10377367>
- de Lafuente, V., & Romo, R. (2005). Neuronal correlates of subjective sensory experience. *Nature Neuroscience*, *8*(12), 1698–1703. <http://doi.org/10.1038/nn1587>
- Dean, A. F. (1981). The variability of discharge of simple cells in the cat striate cortex. *Experimental Brain Research*, *44*(4), 437–440. <http://doi.org/10.1007/BF00238837>



- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective neurons in the macaque. *Journal of Neuroscience*, *4*(8), 2051–2062.
- DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, *11*(5), 219–226.  
[http://doi.org/10.1016/0166-2236\(88\)90130-0](http://doi.org/10.1016/0166-2236(88)90130-0)
- Ding, L., & Gold, J. (2013, August). The basal ganglia's contributions to perceptual decision making. *Neuron*. Elsevier Inc. <http://doi.org/10.1016/j.neuron.2013.07.042>
- Dodd, J. V, Krug, K., Cumming, B. G., & Parker, A. J. (2001). Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *The Journal of Neuroscience*, *21*(13), 4809–21. <http://doi.org/21/13/4809> [pii]
- Ecker, A. S., Berens, P., Cotton, R. J., Subramaniyan, M., Denfield, G. H., Cadwell, C. R., ... Tolias, A. S. (2014). State Dependence of Noise Correlations in Macaque Primary Visual Cortex. *Neuron*, *82*(1), 235–248.  
<http://doi.org/10.1016/j.neuron.2014.02.006>
- Ecker, A. S., Berens, P., Tolias, A. S., & Bethge, M. (2011). The effect of noise correlations in populations of diversely tuned neurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(40), 14272–83. <http://doi.org/10.1523/JNEUROSCI.2539-11.2011>
- Ecker, A. S., Denfield, G. H., Bethge, M., & Tolias, A. S. (2016). On the Structure of Neuronal Population Activity under Fluctuations in Attentional State. *Journal of Neuroscience*, *36*(5), 1775–1789. <http://doi.org/10.1523/JNEUROSCI.2044-15.2016>
- Eckstein, M. P., & Ahumada, A. J. (2002). Classification images: a tool to analyze visual strategies. *Journal of Vision*, *2*(November), 1x.

- Goard, M., & Dan, Y. (2009). Basal forebrain activation enhances cortical coding of natural scenes. *Nature Neuroscience*, *12*(11), 1444–9. <http://doi.org/10.1038/nn.2402>
- Goris, R. L. T., Movshon, J. A., & Simoncelli, E. P. (2014). Partitioning neuronal variability. *Nature Neuroscience*, *17*(6), 858–865. <http://doi.org/10.1038/nn.3711>
- Graf, A. B. a, Kohn, A., Jazayeri, M., & Movshon, J. A. (2011). Decoding the activity of neuronal populations in macaque primary visual cortex. *Nature Neuroscience*, *14*(2), 239–45. <http://doi.org/10.1038/nn.2733>
- Gu, Y., Angelaki, D. E., & DeAngelis, G. C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nature Neuroscience*, *11*(10), 1201–1210. <http://doi.org/10.1038/nn.2191>
- Haefner, R. M., Berkes, P., & Fiser, J. (2014). The implications of perception as probabilistic inference for correlated neural variability during behavior, 26. *Neurons and Cognition*. Retrieved from <http://arxiv.org/abs/1409.0257>
- Haefner, R. M., Gerwinn, S., Macke, J. H., & Bethge, M. (2013). Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nature Neuroscience*, *16*(2), 235–42. <http://doi.org/10.1038/nn.3309>
- Henry, G. H., Bishop, P. O., Tupper, R. M., & Dreher, B. (1973). Orientation specificity and response variability of cells in the striate cortex. *Vision Research*, *13*(9), 1771–1779. [http://doi.org/10.1016/0042-6989\(73\)90094-1](http://doi.org/10.1016/0042-6989(73)90094-1)
- Herrero, J. L., Roberts, M. J., Delicato, L. S., Gieselmann, M. A., Dayan, P., & Thiele, A. (2008). Acetylcholine contributes through muscarinic receptors to attentional modulation in V1. *Nature*, *454*(7208), 1110–4. <http://doi.org/10.1038/nature07141>

- Higham, N. J. (1988). Computing a nearest symmetric positive semidefinite matrix. *Linear Algebra and Its Applications*, *103*, 103–118. [http://doi.org/10.1016/0024-3795\(88\)90223-6](http://doi.org/10.1016/0024-3795(88)90223-6)
- Hruba, L., Ott, T., Nieder, A., Pourriahi, P., & Nienborg, H. (2015). Gain modulation by serotonin in the macaque primary visual cortex. In *Society for Neuroscience Abstracts*.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, *148*, 574–591. <http://doi.org/10.1113/jphysiol.2009.174151>
- Johnson, K. O. (1980). Sensory discrimination: decision process. *Journal of Neurophysiology*, *43*(6), 1771–1792.
- 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. [http://doi.org/10.1016/0042-6989\(80\)90128-5](http://doi.org/10.1016/0042-6989(80)90128-5)
- Kang, I., & Maunsell, J. H. R. (2012). Potential confounds in estimating trial-to-trial correlations between neuronal response and behavior using choice probabilities. *Journal of Neurophysiology*, *108*(12), 3403–15. <http://doi.org/10.1152/jn.00471.2012>
- Kiani, R., Hanks, T. D., & Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J Neurosci*, *28*(12), 3017–3029. <http://doi.org/10.1523/JNEUROSCI.4761-07.2008>
- Kohn, A., & Smith, M. a. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *The Journal of Neuroscience : The Official Journal of*

- the Society for Neuroscience*, 25(14), 3661–73.  
<http://doi.org/10.1523/JNEUROSCI.5106-04.2005>
- Krauzlis, R. J., Bollimunta, A., Arcizet, F., & Wang, L. (2014). Attention as an effect not a cause. *Trends in Cognitive Sciences*, 1–8. <http://doi.org/10.1016/j.tics.2014.05.008>
- Krug, K., Cumming, B. G., & Parker, a J. (2004). Comparing perceptual signals of single V5/MT neurons in two binocular depth tasks. *Journal of Neurophysiology*, 92(3), 1586–1596. <http://doi.org/10.1152/jn.00851.2003>
- Law, C.-T., & Gold, J. I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature Neuroscience*, 11(4), 505–513.  
<http://doi.org/10.1038/nn2070>
- Liu, J., & Newsome, W. T. (2005). Correlation between speed perception and neural activity in the middle temporal visual area. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 25(3), 711–22.  
<http://doi.org/10.1523/JNEUROSCI.4034-04.2005>
- Livingstone, M., & Hubel, D. (1988). Segregation of Depth: Form, Anatomy, Color, Physiology, and Movement, and Perception. *Science. New Series*, 240(4853), 740–749. <http://doi.org/10.1126/science.3283936>
- Luu, L., & Stocker, A. A. (2016). Choice-induced biases in perception. *bioRxiv*.  
<http://doi.org/10.1101/043224>
- Maier, A., Aura, C. J., & Leopold, D. a. (2011). Infragranular sources of sustained local field potential responses in macaque primary visual cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 31(6), 1971–80. <http://doi.org/10.1523/JNEUROSCI.5300-09.2011>

- Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*(7474), 78–84.  
<http://doi.org/10.1038/nature12742>
- Maunsell, J. H., & Cook, E. P. (2002). The role of attention in visual processing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *357*(1424), 1063–1072.  
<http://doi.org/10.1098/rstb.2002.1107>
- Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. <http://doi.org/10.1016/j.tins.2006.04.001>
- McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(1), 431–441.  
Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9870971>
- McAlonan, K., Cavanaugh, J., & Wurtz, R. H. (2008). Guarding the gateway to cortex with attention in visual thalamus. *Nature*, *456*(7220), 391–400.  
<http://doi.org/10.1038/nature07382>
- Mehta, a D., Ulbert, I., & Schroeder, C. E. (2000). Intermodal selective attention in monkeys. I: distribution and timing of effects across visual areas. *Cerebral Cortex (New York, N.Y. : 1991)*, *10*(4), 343–358. <http://doi.org/10.1093/cercor/10.4.343>
- Mitchell, J. F., Sundberg, K. a, & Reynolds, J. H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron*, *63*(6), 879–888.  
<http://doi.org/10.1016/j.neuron.2009.09.013>
- Mitzdorf, U., & Singer, W. (1979). Excitatory synaptic ensemble properties in the visual cortex of the macaque monkey: a current source density analysis of electrically

- evoked potentials. *The Journal of Comparative Neurology*, 187(1), 71–83.  
<http://doi.org/10.1002/cne.901870105>
- Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., & Pouget, A. (2014). Information-limiting correlations. *Nature Neuroscience*, 17(10), 1410–1417.  
<http://doi.org/10.1038/nn.3807>
- Neri, P., & Levi, D. M. (2006). Receptive versus perceptive fields from the reverse-correlation viewpoint. *Vision Research*, 46(16), 2465–2474.  
<http://doi.org/10.1016/j.visres.2006.02.002>
- Newsome, W. T., Britten, K. H., Movshon, J. A., & Shadlen, M. N. (1989). Single Neurons and the Perception of Visual Motion. In *Neural Mechanism of Visual Perception* (pp. 171–197).
- Newsome, W. T., & Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*.  
<http://doi.org/http://www.ncbi.nlm.nih.gov/pubmed/3385495>
- Niell, C. M., & Stryker, M. P. (2010). Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron*, 65(4), 472–479.  
<http://doi.org/10.1016/j.neuron.2010.01.033>
- Nienborg, H., Cohen, M. R., & Cumming, B. G. (2012). Decision-Related Activity in Sensory Neurons: Correlations Among Neurons and with Behavior. *Annual Review of Neuroscience*, 35(1), 463–483. <http://doi.org/10.1146/annurev-neuro-062111-150403>
- Nienborg, H., & Cumming, B. G. (2007). Psychophysically measured task strategy for

- disparity discrimination is reflected in V2 neurons. *Nature Neuroscience*, 10(12), 1608–1614. <http://doi.org/10.1038/nn1991>
- Nienborg, H., & Cumming, B. G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature*, 459(7243), 89–92. <http://doi.org/10.1038/nature07821>
- Nienborg, H., & Cumming, B. G. (2010). Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality? *Current Opinion in Neurobiology*, 20(3), 376–381. <http://doi.org/10.1016/j.conb.2010.05.002>
- Nienborg, H., & Cumming, B. G. (2014). Decision-Related Activity in Sensory Neurons May Depend on the Columnar Architecture of Cerebral Cortex. *Journal of Neuroscience*, 34(10), 3579–3585. <http://doi.org/10.1523/JNEUROSCI.2340-13.2014>
- Olshausen, B. a, & Field, D. J. (2005). How close are we to understanding v1? *Neural Computation*, 17(8), 1665–99. <http://doi.org/10.1162/0899766054026639>
- Parker, A. J., Krug, K., & Cumming, B. G. (2002). Neuronal activity and its links with the perception of multi-stable figures. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 357(1424), 1053–1062. <http://doi.org/10.1098/rstb.2002.1112>
- Paukert, M., Agarwal, A., Cha, J., Doze, V. A., Kang, J. U., & Bergles, D. E. (2014). Norepinephrine controls astroglial responsiveness to local circuit activity. *Neuron*, 82(6), 1263–1270. <http://doi.org/10.1016/j.neuron.2014.04.038>
- Perkel, D. J., Bullier, J., & Kennedy, H. (1986). Topography of the afferent connectivity of area 17 in the macaque monkey: a double-labelling study. *The Journal of*

- Comparative Neurology*, 253(3), 374–402. <http://doi.org/10.1002/cne.902530307>
- Pitkow, X., Liu, S., Angelaki, D. E., DeAngelis, G. C., & Pouget, A. (2015a). How Can Single Sensory Neurons Predict Behavior? *Neuron*, 87(2), 411–423. <http://doi.org/10.1016/j.neuron.2015.06.033>
- Pitkow, X., Liu, S., Angelaki, D. E., DeAngelis, G. C., & Pouget, A. (2015b). How Can Single Sensory Neurons Predict Behavior? *Neuron*, 87(2), 411–423. <http://doi.org/10.1016/j.neuron.2015.06.033>
- Ponce, C. R., Lomber, S. G., & Born, R. T. (2008). Integrating motion and depth via parallel pathways. *Nature Neuroscience*, 11(2), 216–223. <http://doi.org/10.1038/nn2039>
- Price, N. S. C., & Born, R. T. (2010). Timescales of sensory- and decision-related activity in the middle temporal and medial superior temporal areas. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 30(42), 14036–14045. <http://doi.org/10.1523/JNEUROSCI.2336-10.2010>
- Purushothaman, G., Marion, R., Li, K., & Casagrande, V. a. (2012). Gating and control of primary visual cortex by pulvinar. *Nature Neuroscience*, 15(6), 905–12. <http://doi.org/10.1038/nn.3106>
- Rabinowitz, N. C., Goris, R. L. T., Cohen, M. R., & Simoncelli, E. P. (2015). Attention stabilizes the shared gain of V4 populations. *eLife*, (November). <http://doi.org/http://dx.doi.org/10.7554/eLife.08998>
- Reynolds, J. H., & Chelazzi, L. (2004). Attentional Modulation of Visual Processing. *Annual Review of Neuroscience*, 27(1), 611–647. <http://doi.org/10.1146/annurev.neuro.26.041002.131039>



- Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*(7451), 585–90. <http://doi.org/10.1038/nature12160>
- Ringach, D., G. S., & Shapley, R. (1997). A subspace reverse correlation technique for the study of visual neurons. *Vision Research*, *37*(17), 2455–2464.
- Romo, R., Hernández, A., Zainos, A., Brody, C., & Salinas, E. (2002). Exploring the cortical evidence of a sensory-discrimination process. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *357*(1424), 1039–1051. <http://doi.org/10.1098/rstb.2002.1100>
- Ruff, D. A., & Cohen, M. R. (2014). Attention can either increase or decrease spike count correlations in visual cortex. *Nature Neuroscience*, *17*(11), 1591–7. <http://doi.org/10.1038/nn.3835>
- Sarma, Y., & Jammalamadaka, S. R. (1993). Circular Regression. *Statistical Sciences and Data Analysis*.
- Schattschneider, D. (1978). The Plane Symmetry Groups: Their Recognition and Notation. *The American Mathematical Monthly*, *85*(6), 439. <http://doi.org/10.2307/2320063>
- Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *16*(4), 1486–510. Retrieved from <http://www.jneurosci.org/content/16/4/1486.short>
- Shuler, M. G., & Bear, M. F. (2006). Reward Timing in the Primary Visual Cortex.

- Science*, 311(5767), 1606–1609. <http://doi.org/10.1126/science.1123513>
- Sillito, A. M., Cudeiro, J., & Jones, H. E. (2006). Always returning: feedback and sensory processing in visual cortex and thalamus. *Trends in Neurosciences*, 29(6), 307–16. <http://doi.org/10.1016/j.tins.2006.05.001>
- Smith, M. A., & Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 28(48), 12591–603. <http://doi.org/10.1523/JNEUROSCI.2929-08.2008>
- Smolyanskaya, A., Haefner, R. M., Lomber, S. G., & Born, R. T. (2015). A Modality-Specific Feedforward Component of Choice-Related Activity in MT. *Neuron*, 87(1), 208–219. <http://doi.org/10.1016/j.neuron.2015.06.018>
- Snippe, H., & Koenderink, J. (1992). Information in channel-coded systems: correlated receivers. *Biological Cybernetics*, 190, 183–190. Retrieved from <http://link.springer.com/article/10.1007/BF00201025>
- Sompolinsky, H., Yoon, H., Kang, K., & Shamir, M. (2001). Population coding in neuronal systems with correlated noise. *Physical Review E*, 64(5), 051904. <http://doi.org/10.1103/PhysRevE.64.051904>
- Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research*, 23(8), 775–785. [http://doi.org/10.1016/0042-6989\(83\)90200-6](http://doi.org/10.1016/0042-6989(83)90200-6)
- Tomko, G. J., & Crapper, D. R. (1974). Neuronal variability: non-stationary responses to identical visual stimuli. *Brain Research*, 79(3), 405–418. [http://doi.org/10.1016/0006-8993\(74\)90438-7](http://doi.org/10.1016/0006-8993(74)90438-7)

- Treue, S., & Martínez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*(6736), 575–9.  
<http://doi.org/10.1038/21176>
- Uka, T., & DeAngelis, G. C. (2004). Contribution of area MT to stereoscopic depth perception: Choice-related response modulations reflect task strategy. *Neuron*, *42*(2), 297–310. [http://doi.org/10.1016/S0896-6273\(04\)00186-2](http://doi.org/10.1016/S0896-6273(04)00186-2)
- Welker, C. (1976). Receptive fields of barrels in the somatosensory neocortex of the rat. *The Journal of Comparative Neurology*, *166*(2), 173–89.  
<http://doi.org/10.1002/cne.901660205>
- Wurtz, R. H., McAlonan, K., Cavanaugh, J., & Berman, R. A. (2011). Thalamic pathways for active vision. *Trends in Cognitive Sciences*, *15*(4), 177–184.  
<http://doi.org/10.1016/j.tics.2011.02.004>
- Yu, X., Dickman, J. D., DeAngelis, G. C., & Angelaki, D. E. (2015). Neuronal thresholds and choice-related activity of otolith afferent fibers during heading perception. *Proceedings of the National Academy of Sciences*, (14), 201507402.  
<http://doi.org/10.1073/pnas.1507402112>
- Zeki, S. M. (1978). Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex. *Journal of Physiology*, (277), 273–290.
- Zénon, A., & Krauzlis, R. J. (2012). Attention deficits without cortical neuronal deficits. *Nature*, *489*(7416), 434–437. <http://doi.org/10.1038/nature11497>
- Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, *370*(6485), 140–143. <http://doi.org/10.1038/370140a0>