From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Neuroscience

2022

The algorithmic and neural logic of perceptual decision-making

Gupta, Diksha

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01mc87pt47c

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

The algorithmic and neural logic of perceptual decision-making

Diksha Gupta

A Dissertation

Presented to the Faculty

of Princeton University

in Candidacy for the Degree

of Doctor of Philosophy

Recommended for Acceptance

by the

Princeton Neuroscience Institute

Adviser: Carlos D. Brody

November 2022

© Copyright by Diksha Gupta, 2022. All rights reserved.

Abstract

Decision-making based on noisy sensory stimuli is a fundamental part of everyday life. In this thesis, we present three studies that seek to advance the understanding of the algorithmic and neural logic underlying decision-making by reckoning with challenges posed by large-scale datasets and the complexity revealed by modern methodological techniques.

First, we consider the challenge posed by nonstationarities in a key decision variable - the criterion used to pick between two options - that is of particular concern in large longitudinal datasets. We characterize the pitfalls of a method proposed to accommodate for such nonstationarity and recommend a model-based alternative.

Second, we identify a unified mechanism that gives rise to two widely observed departures from the behavior of the optimal model - trial history biases and lapses. While these suboptimalities have traditionally been considered distinct, we demonstrate that normative decision-making under a misbelief about nonstationarity in the world gives rise to both history dependence and choices that appear to be evidence-independent or lapses. We test our model in choices of a large dataset of rats, and choices and reaction times of a novel reaction time task, and show that the constraints posited by the model are obeyed in these datasets.

Finally, we investigate the neural logic underlying perceptual decision-making. Decision-making is often conceptualized as a sequence of two sub-computations - gradual accumulation of evidence followed by thresholding to commit to a choice. In rats, the anterior dorsal striatum (ADS) and the frontal orienting fields (FOF) have been mapped onto these two theoretically defined computations. This provides an intriguing - yet untested - neural implementation of the decision process in which ADS and FOF form a functional feedforward hierarchy. We present results from simultaneous neural recordings and projection-specific inactivations that challenge this previously proposed mapping. Our results show that both ADS and FOF carry redundant task-relevant information, are involved throughout the accumulation process, and yet respond differentially to perturbations. We reconcile these conflicting observations from physiology and perturbations using a multi-region recurrent neural network model.

Acknowledgments

I would like to thank my advisor Carlos Brody who has been an incredible source of kindness and knowledge throughout my time as a graduate student. I am grateful that he enthusiastically encouraged me to pursue my ideas at key junctures of my PhD and prioritized my personal development and goals. I am thankful for his thoughtful guidance and optimistic approach that has helped me overcome several hurdles that surfaced along the way. I greatly admire and aspire to his clarity of thought, flair for articulating ideas and eye for rigor, and feel privileged to have received his mentorship.

I would also like to thank my thesis committee members, Jonathan Pillow and Ilana Witten. I always walked (zoomed?) out of my thesis committee meetings feeling more confident about my research and reassured about my progress. I am thankful for this positive prediction error. I am also grateful to Jonathan for generously sharing his time and knowledge inside and outside the classroom - his ideas have had a profound impact on my research.

I have been extremely fortunate to have had many wonderful collaborators and mentors who I wish to thank. I cannot imagine venturing into the brave new world of rodent training and experiments without the expert guidance of Chuck Kopec. I am deeply thankful to Chuck for his infinite patience, for teaching me how to solve problems creatively and for making the lab a fun (and organized!) environment. I would also like to thank Thomas Luo and Adrian Bondy for showing me the ropes of rodent electrophysiology. We all joined the Brody lab around the same time, and I have enjoyed their companionship immensely while

growing up as a researcher. Discussions with Thomas have helped me hone my critical thinking skills. I am inspired by his systematic and rigorous approach towards conducting science. I am grateful to Adrian for his mentorship and providing perspective when most needed. I have greatly benefited from his keen eye for the important issues and enthusiasm for the scientific endeavor. I would also like to thank Brian DePasquale for his willingness to help with all things technical and inspiring confidence in me. Early researchers are lucky to have Brian as their advocate. I am indebted to Marino Pagan for a most engaging and fun rotation experience which convinced me to join the Brody lab as a graduate student. Marino's enthusiasm for science is infectious and my conversations with him always spark new ideas. My thanks also to David Zoltowski for many engaging and helpful discussions about recurrent interactions in the brain over the past couple of years. I am thankful to Peter Bibawi, Jovanna Teran, Verity Elliott, Jessica Teran and the many Brody lab technicians, without whose help and support none of my experiments would have been possible.

My thanks to Alex Piet and Tyler Boyd-Meredith, the camaraderie I have shared with them has been one of the highlights of my graduate experience. I will always cherish our many hangouts and lunches. I also wish to thank Ahmed El-Hady, Alex Piet, Kevin Miller, Emily Dennis, Josh Julian, Angela Langdon, Manuel Schottdorf, Christine Constantinople, Athena Akrami, Ben Scott, Lucas Pinto, Edward Nieh, Esme Foncesca and Malavika Murugan for their feedback, invaluable mentorship and career advice. I want to thank the other graduate students in the Brody lab - Tim Kim, Sarah Jo Venditto, Jorge Yanar, Jess Breda and Wynne Stagnaro - with whom conversations are always fun and stimulating. I am thankful to Rebecca Khaitman, Alex Michaud, Ed Clayton for all that they did in support of this work. I am thankful to the ex-Brody lab members Bing Brunton, Jeff Erlich, Tim Hanks, Michael

Yartsev, my projects have been directly inspired by their work and I am constantly in awe of the deep thought that they put into designing their studies. Reading their findings has moulded my scientific thought process for the better. I feel fortunate to have been part of the incredible neuroscience community at Princeton, their intellect and brilliance has shaped my perspective on neuroscience. Over the years I have been lucky to have many wonderful friends, housemates, classmates, and cohort mates - I would like to thank them all for the fond memories and fun adventures.

I am thankful to my undergraduate mentor Amitabha Mukerjee who has hands down been the best teacher I have ever had. Even after nearly a decade, I can trace my every idea and project to an interest he inspired in me. I have been incredibly lucky to have him as a mentor and he is deeply missed!

No words are enough to express my gratitude towards my mom and dad. I am here today because of their love, unwavering support and the many sacrifices they have made and to them I dedicate this thesis. Thanks to Kapish, Minal and little Kridha for being my cheerleaders from a distance, knowing that they are rooting for me has kept me going. Thanks to my late grandparents, Tauji, Taiji, Meenu, Athiah and Mamiah for their unconditional support and love. Foremost, I would like to thank partner in life and science, Sashank Pisupati without whose love, care and patience none of this would have been possible. Thanks for believing in me when I lost faith in myself, thanks for making me your utmost priority, thanks for listening, talking through ideas with me, for proofreading my materials and for doing the housework (!).

To my parents

Contents

Abstract 3

1 Introduction 13
- 1.1 A primer on perceptual decision-making . . . 16
  - 1.1.1 Signal detection theory . . . 17
  - 1.1.2 Evidence accumulation . . . 18
- 1.2 Perceptual decision-making in rats . . . 21
  - 1.2.1 The Poisson Clicks task . . . 21
  - 1.2.2 Neural mechanisms . . . 24
- 1.3 Contributions and outline . . . 24

2 Limitations of a proposed correction for slow drifts in decision criterion 26
- 2.1 Introduction . . . 26
- 2.2 Results . . . 29
  - 2.2.1 Correction in action . . . 29
  - 2.2.2 Correction in misaction . . . 34
  - 2.2.3 A model-based solution . . . 39
  - 2.2.4 Comparing the model-free correction and the model-based approach 42
- 2.3 Discussion . . . 44
- 2.4 Methods . . . 46
  - 2.4.1 Simulation Details . . . 46
  - 2.4.2 History Bias and Correction . . . 48
  - 2.4.3 Model-based fitting . . . 49
  - 2.4.4 Dataset studied . . . 51
  - 2.4.5 Data and Code availability . . . 52

3 Trial history biases in evidence accumulation can give rise to apparent lapses **53**
3.1 Introduction . . . 53
3.2 Results . . . 56
3.2.1 A common mechanism produces history biases and apparent lapses 56
3.2.2 Rats display varying degrees of history-dependent threshold and lapse rate modulations . . . 62
3.2.3 Modeling history dependence in rat data . . . 64
3.2.4 History dependent initial points reproduce lapse rate modulations in data . . . 68
3.2.5 Reaction times support initial-point updating . . . 70
3.3 Discussion . . . 74
3.4 Methods . . . 78
3.4.1 Subjects . . . 78
3.4.2 Drift diffusion model of decision-making . . . 79
3.4.3 Models of initial point updating . . . 80
3.4.4 Psychometric curves . . . 81
3.4.5 Behavioral tasks . . . 82
3.4.6 Data modeling methods . . . 84
3.5 Supplementary materials . . . 90

4 Distributed evidence accumulation in frontal cortico-striatal circuitry **95**
4.1 Relevant background . . . 95
4.2 Introduction . . . 99
4.3 Results . . . 102
4.3.1 Task-relevant representations in FOF and ADS . . . 104
4.3.2 Cortico-striatal communication during evidence accumulation . . . 111
4.3.3 Perturbing FOF projections to ADS . . . 113
4.3.4 Bilateral FOF inactivations . . . 117
4.3.5 Modeling physiology and perturbations . . . 118

4.4 Discussion 126
4.5 Methods 133
4.5.1 Subjects and housing 133
4.5.2 Neuronal recordings and analysis 134
4.5.3 Optogenetic perturbations and analysis 140
4.5.4 Multi-region recurrent neural network model 143

5 Conclusions and outlook **149**

References **156**

Listing of figures

1.1 The Poisson Clicks task 22

2.1 The model-free correction for slow drifts in decision criterion 30
2.2 Limitations of the model-free correction 35
2.3 A model-based approach for accounting for slow drifts 40
2.4 Comparing the model-free correction and the model-based approach 43

3.1 Trial history-dependent initial points give rise to apparent lapses 57
3.2 History-dependent threshold and lapse rate modulations in a rat dataset . . 63
3.3 Models of rodent data including apparent and true lapses 65
3.4 Accumulator model reproduces lapse rate modulations in the data 67
3.5 Reaction time predictions from the model are borne out in data 71
3.1 Exponential filtering for initial point setting approximates Bayesian prior updates 90
3.2 Influence of within- and across-trial parameters on history-dependent modulation of biases and lapses in the psychometric function. 91
3.3 Performance measures across the rat dataset 92
3.4 Model fits to choices and reaction times of individual rats 93
3.5 Individual model comparison in the RT dataset 94

4.1 Proposed feedforward functional hierarchy in frontal cortico-striatal circuitry 101
4.2 Simultaneous recordings from FOF and ADS: experimental setup 103
4.3 Tuning of individual neurons to the accumulator variable: example neurons 105
4.4 Summary of neural tuning to the accumulator variable 106
4.5 Stimulus decoding from FOF and ADS 107
4.6 Choice and trial history decoding from FOF and ADS 108
4.7 Decoding decision variable from FOF and ADS population activity 109

4.8 Lead-lag relationship between decision-variables in FOF and ADS 110
4.9 Communication subspace between FOF and ADS 112
4.10 Silencing FOF axon terminals in ADS: experimental setup 113
4.11 Silencing FOF axon terminals in ADS: effect on choices 115
4.12 Silencing FOF axon terminals in ADS: effect on movement 116
4.13 Silencing FOF bilaterally 118
4.14 Architecture of the multi-region model 120
4.15 Training performance of the multi-region model 121
4.16 Average firing rates of units in the multi-region model 122
4.17 Decoding task-relevant variables from the multi-region model 123
4.18 Response of the multi-region model to novel perturbations 124
4.19 Response of multi-region model's units to perturbations 125

1

Introduction

The study of perceptual decision-making is perhaps one of the most impressive and exciting efforts geared towards understanding the neural basis of cognition. Linking perception to action, studies in this domain have carefully synthesized insights from statistics, physics, psychology, and neuroscience to meticulously describe with well reasoned process models, the choice that an agent might make given a perceptual stimulus, the time it might take them to make that choice and how different brain regions might participate during this process. Spanning research from zebra fish larvae to corvids to humans, fundamental hallmarks of perceptual decisions - that making hard decisions takes longer than easy decisions and that decisions depend upon other cognitive factors like expected costs, benefits, attention etc. - extend to other decision-making domains such as social, value-based, memory-based among others.

This raises the tantalizing possibility that understanding the neural mechanisms underlying this simple laboratory-tractable form of decisions, might shed light on decision-making in general.

This effort however, is not without its challenges. Early frameworks which were successful in describing decision processes of a few experimental subjects and carefully picked handfuls of neurons are now faced with large datasets - largely due to a burgeoning toolkit for large-scale neural measurements in tandem with the acceleration of perceptual decision-making research in rodents (Gold and Shadlen, 2007; Carandini and Churchland, 2013; Hanks and Summerfield, 2017). The first challenge is methodological: while large datasets from single subjects have been invaluable for precisely inferring decision-related parameters and informing long-standing debates about sources of noises (e.g. Brunton et al. 2013), recent forays into the longitudinal properties of such datasets have revealed that they are also more prone to carrying nonstationarities (Fründ et al., 2011; Dutilh et al., 2012; Schütt et al., 2016; Lak et al., 2020a; Roy et al., 2021; Ashwood et al., 2022) that can severely hinder such inferences unless explicitly accounted for.

The second challenge lies in characterizing and explaining the heterogeneity in decision-making behavior across individuals. Large populations lay bare the variability in, and multiplicity of decision-making strategies adopted by individuals, highlighting the ways in which they depart from our working models (Carandini and Churchland, 2013; Waskom et al., 2019; Musall et al., 2019). This demands updating decision-making models by identifying causes that could explain the various axes of individual variability in a parsimonious manner.

Perhaps the most significant challenge is that of mapping decision-making models onto their neural substrates and deducing their mechanistic logic (Gold and Shadlen, 2007; Brody

and Hanks, 2016; Hanks and Summerfield, 2017; Najafi and Churchland, 2018; Musall et al., 2019). The dominant narrative from early studies presented a promising and uncomplicated mapping in which individual brain areas formed feedforward modules that mirrored and neatly encapsulated the sequential algorithmic components of decision-making models that describe behavior. At the same time, these studies gravely noted that multiple other brain areas also carry similar representations and expressed hope that the study of simultaneous activity from these regions or their perturbations will help identify the few important players and their distinct roles . With the recent advent of tools for observing activity from populations of neurons across brain areas, and performing precise temporally-specific perturbations, this hope is being met with growing contrary and perplexing evidence (Katz et al., 2016; Yates et al., 2017; Hanks and Summerfield, 2017; Pinto et al., 2019; Koay et al., 2020). Hence there is an urgent need for the revision of old frameworks and the development of new ones that can entertain complex, distributed, and recurrent mappings.

This thesis seeks to address these challenges through three studies. We characterize pitfalls of a method proposed to accommodate non-stationarities during decision-making and offer a powerful alternative. We identify a unified mechanism that gives rise to two common empirically observed departures from the behavior of the optimal model, that have traditionally been considered distinct. We present results that challenge previous feedforward mappings of decision-making onto the cortico-striatal circuitry in rats and propose an alternative - a distributed and recurrent model that advocates for a more nuanced view of the mapping between representations, perturbations, and computations.

These studies rely heavily upon previous work on models of computations that underlie perceptual decision-making and an auditory decision-making task developed by the Brody

lab in rats. We begin by describing both of these in turn.

1.1 A primer on perceptual decision-making

In a party, how do we decide if a blurry face in the distance is a friend or a stranger - should we go greet them or carry on?

Perceptual decision-making is the process by which we shape our behavior using information derived from noisy sensory observations (like a blurry face). This involves gathering and evaluating sensory observations to infer the unknown state of the world and combining this estimate with our knowledge about the prior prevalence of different states and their relative benefits in order to make decisions that help meet our goals. Studies seeking to understand this process often focus on the role of sensory information in guiding behavior by manipulating the uncertainty or noise in sensory stimulus from trial to trial while keeping other pertinent variables (priors, benefits etc.) fixed. In our perceptual decision-making tasks, the uncertainty in sensory information arises from either the probabilistic nature of the source that generates sensory stimulus or from imperfect measurements of the stimulus by the observer. For instance, say if you are trying to read time from an LED watch, the first kind of uncertainty might be a display with a flickering led such that an 8 resembles 6 at times, and the second kind would be a dim display perhaps due to low battery which is barely above visual thresholds.

Two different formalisms - signal detection theory and evidence accumulation (also referred to as sequential sampling) are commonly used to study perceptual decisions. These formalisms prescribe what choice an observer should make ideally, given their sensory uncertainty and objectives. Signal detection theory assumes that the decision is based on a single

observation of noisy evidence whereas evidence accumulation accommodates multiple observations. These solutions are easiest to write and analyze when the number of possible choices are restricted to two i.e. for 2AFC or two alternative forced choice tasks and we restrict our discussion to this case.

1.1.1 Signal detection theory

Signal detection theory (SDT) specifies the process of making a decision based on a single observation of noisy evidence. It borrows its name from its early use by radar researchers to detect if an ambiguous stimulus was signal (e.g. a plane) or just noise (e.g. a bird). Similarly, within psychology it was applied to detection tasks where a subject is asked to report yes if they detect a stimulus and no otherwise. Unlike the classical psychophysics of the time, signal detection research incorporated rewards for hits and correct rejections, and penalties for misses and false alarms (Nevin, 1969), therefore combining psychophysics with operant conditioning and giving rise to the modern day decision-making research.

According to SDT, the subject occupies a true state of the world H_1 and H_2 (e.g. stimulus present vs absent), but its identity is hidden from the subject. It is assumed that the subject know the choice rules, i.e. it is optimal to choose L in H_1 and R in H_2. The subject obtains an observation ϵ of the state which is noisy, either due to external factors (states emit probabilistic cues) or internal factors (imperfect measurement or imperfect neural representation of the cues). This observation is often assumed to be Gaussian distributed around a mean decided by the state, and a variance reflecting the noise. The conditional distributions $P(\epsilon|H_1)$ and $P(\epsilon|H_2)$ describe the likelihoods of observing the evidence if the true state was H_1 and H_2 respectively.

SDT specifies the decision criterion β (or threshold) that a subject should apply to the likelihood ratio to make a choice for a variety of objectives.

$$\text{choose } H_1 \quad \text{if} \quad \frac{P(\epsilon|H_1)}{P(\epsilon|H_2)} \geq \beta$$

If the goal is accuracy and the prior probability favors one of the hypothesis, then $\beta = P(H_1)/P(H_2)$, from Bayes rule. Therefore, when the two states are equally likely $\beta = 1$. Similarly, a threshold on the likelihood ratio can be derived for maximizing value instead of accuracy (Dayan and Daw, 2008).

Hence, SDT is a powerful framework to make decisions based on priors, values and a single observation of sensory evidence to achieve a variety of goals. In Chapter 2, we work within this framework and analyze the effects of random drifts in decision criterion on choice.

1.1.2 Evidence accumulation

Evidence accumulation (EA) or sequential sampling is a natural extension of SDT to the case of multiple observations, where individual pieces of evidence about the true state of the world H_1 or H_2 arrive sequentially in time $\epsilon_1, \epsilon_2, ...\epsilon_t$ (assumed i.i.d.). As a result, at each timestep the decision-maker decides whether to stop the decision process and choose one of the two options L and R or continue the process and sample more evidence. Gathering more evidence helps reduce uncertainty about the hidden state of the world but also expends time and effort which are often costly - this gives rise to a speed-accuracy trade-off. EA specifies the decision rule to balance the benefits of making the decision early (i.e. avoiding sampling-related costs) against making the wrong decision and getting no reward.

We can compute the ratio of posterior beliefs about the true state at time t using Bayes

rule. The logarithm of this ratio, often called the *decision variable* is

$$\log \frac{P(H_1|\epsilon_{1..t})}{P(H_2|\epsilon_{1..t})} = \sum_{i=1}^{t} \log \frac{P(\epsilon_i|H_1)}{P(\epsilon_i|H_2)} + \log \frac{P(H_1)}{P(H_2)}$$

Like SDT, the decision rule in evidence accumulation under various objectives, can often be expressed as a pair of (sometimes time-varying) thresholds on the log posterior ratio, one for each choice. An in depth discussion of this can be found in Bogacz et al. 2006; Rao 2010 and Drugowitsch et al. 2012.

If the rewards for the two choices are equal and it is assumed that the reliability of the evidence is known *a priori* to the subject, and the subject wants to make the decision as soon as the estimate reaches a predetermined level of certainty C, then the thresholds are given by $\pm \log \frac{C}{1-C}$. The subject should continue the decision-making process until these thresholds are reached, and if the thresholds are crossed then the decision maker should choose the action corresponding to the inferred state.

$$\text{choose } H_1 \quad \log \frac{P(H_1|\epsilon_{1..t})}{P(H_2|\epsilon_{1..t})} \geq \log \frac{C}{1-C}$$

$$\text{choose } H_2 \quad \log \frac{P(H_1|\epsilon_{1..t})}{P(H_2|\epsilon_{1..t})} \leq \log \frac{1-C}{C}$$

Therefore, unlike SDT, EA models the time it takes to make a decision or the reaction time (RT) in addition to the choice. This algorithm has an interesting history and was used to break the enigma code in World War II, documented in Gold and Shadlen 2002.

It has long been shown that when the noise in sensory evidence is Gaussian with variance σ^2, this setting reduces to a drift diffusion model (DDM) in the continuum limit (Ratcliff

and Rouder, 1998) with the dynamics of decision variable given by the stochastic differential equation:

$$da = Adt + \sigma dW$$

where A represents the average increment rate of log likelihood ratio, W is a Wiener process and $a(0)$ is proportional to the log prior ratio of the two hypothesis. A bound on the value of a determines when the decision to commit to a choice is made. Perceptual neuroscience studies often go back and forth between the DDM and Bayesian representation of this computation. In Chapter 3, we will focus on the surprising behavioral consequences of trial-by-trial updates to the initial point $a(0)$.

EA offers several advantages over SDT for studying neural basis of decision-making: in EA the *decision variable* which determines the eventual choice differs from the instantaneous evidence, this dissociation helps differentiate neural representations related to decision-making from those involved in simply perception (Shadlen and Newsome, 1996; 2001; Huk and Shadlen, 2005; Hanks et al., 2015). Moreover, the temporally extended nature of the computation in EA helps delineate when a particular variable is represented or a computation is performed, aiding the discovery of neural correlates (Roitman and Shadlen, 2002; Huk and Shadlen, 2005; Peixoto et al., 2021).

Temporally extended perceptual decisions, as modeled by EA with timescales of seconds or longer have strong working memory demands. They not only require brains to represent the perceived evidence and compare its value to a threshold, but also to maintain the estimate of evidence accumulated so far in memory and integrate any new incoming evidence into it. How neurons with much smaller synaptic time constants coordinate their activity within and across brain areas to perform this feat is an active field of research and many theoretical

models have been proposed for this (reviewed in Wang 2008 and Goldman et al. 2009, also see Mante et al. 2013). Such neural mechanisms of evidence accumulation will be our focus in Chapter 4.

1.2 Perceptual decision-making in rats

The study of neural basis of evidence accumulation for perceptual decision-making was confined to non-human primates for the longest time. While many seminal and exciting discoveries have been and continue to be made in this model organism, rodents were introduced as a complementary model species to afford more access to genetic and molecular tools for large scale recordings and perturbations, and ameliorate the issue that arise due to low sample sizes typically used in monkey studies (Abbott, 2010).

Throughout this thesis, we study rats performing an auditory evidence accumulation task developed by Brunton et al. (2013), called the Poisson Clicks task.

1.2.1 The Poisson Clicks task

The Poisson Clicks task is a 2AFC task developed to train rats to make decisions based on accumulated evidence (Fig 1.1A). In this task, the rat is put in an operant chamber with three nose ports arranged on a curved wall. Each trial begins with the illumination of the LED on the center nose port, this indicates to a rat that it can start a trial by entering the center nose port. Upon trial initiation, after a variable delay ($0.5s - 1.3s$) auditory clicks play from two speakers placed to the left and right side of the rat. Each speaker plays auditory clicks sampled from independent Poisson processes with rates r_L and r_R. The clicks play for a variable duration on each trial ($0.2s - 1s$) such that the rat spends a total of $1.5s$ "fixating" in

the center port. If during this duration, the rat breaks the fixation for longer than $0.005s$ then the trial is considered a "violation" and is aborted. Once the clicks end, the center light turns off marking the end of fixation period and the rat is required to report which side played the greater number of clicks by poking its nose into one of the two side ports. Correct responses are rewarded with a $24\mu l$ drop of water at the correct side port and incorrect responses are punished with white noise.

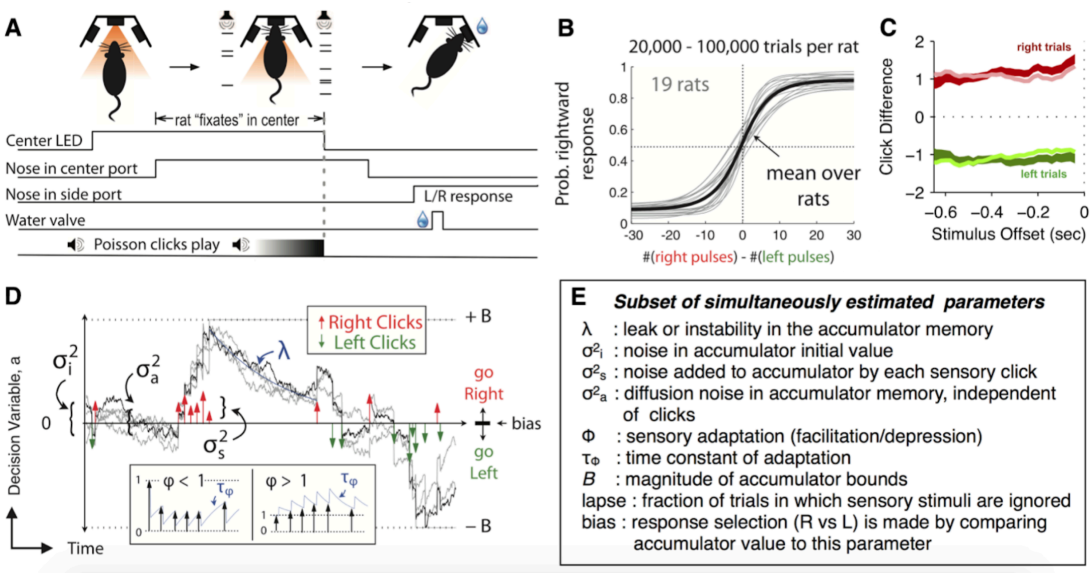


Figure 1.1: (A) Sequence of events in a trial (B) Psychometric curves of rats (C) Psychometric reverse correlation curves, showing that rats uniformly integrate the evidence throughout the stimulus period. (D) The Accumulator model: The accumulator variable (a, black trace) which represents the value of evidence accumulated so far, evolves stochastically through time as a function of model parameters and click times. At the end of the trial if $a >$ bias decision to go right is made else a left decision is made. (E) Description of parameters from D. Figure adapted from Brunton et al. (2013).

On each trial, the sum of expected click rates on the left and right sides is fixed $r_L + r_R =$ 40clicks/s, and the ratio $\gamma = \log(r_R/r_L)$ is varied to set trial difficulty. Greater $|\gamma|$ correspond to easier trials and smaller $|\gamma|$ correspond to harder trials. In this task, each click is a precisely timed quantum of evidence towards either a *left* or *right* decision. Due to this dis-

crete nature of the evidence, together with random timing of clicks and stimulus duration, the trials sample a large statistical space which enables fitting of complex behavioral models.

Brunton et al. (2013) showed that rats trained on this task achieve high accuracy (Fig 1.1B) and that clicks at each point in time influenced the left and right decisions similarly (Fig 1.1C). They also developed an accumulation to bound model (Fig 1.1D,E) modified to take into account the discrete nature of evidence. In the model, on a behavioral trial n, the evolution of accumulated evidence $a(t)$ in response to the left (δ_L) and right (δ_R) click trains is given by:

$$da = \begin{cases} 0, & \text{if } |a| \geq B \\ \lambda a dt + (\delta_{R,t} C_R(t)\eta_R - \delta_{L,T} C_L(t)\eta_L)dt + \sigma_a dW & \text{otherwise} \end{cases}$$
$$\text{where } \frac{dC}{dt} = \frac{1-C}{\tau_\phi} + (\phi - 1)C(\delta_{R,t} + \delta_{L,t})$$

where λ is the inverse time constant of the consistent drift in memory of $a(t)$. $C_R(t)$ and $C_L(t)$ are the magnitudes of each right and left click respectively after undergoing sensory adaptation (with adaptation strength ϕ and adaptation time constant τ_ϕ). The sensory noise that accompanies each click is represented by η_R, η_L which are Gaussian random variables with mean 1 and variance σ_s^2. The accumulation variable a also undergoes Brownian diffusion through the addition of a Wiener process (W) with variance σ_a^2. B represents the absorbing decision bound that prevents $a(t)$ from evolving further, if crossed. The initial value of the accumulator variable a is the Gaussian distribution $\mathcal{N}(0, \sigma_i)$. A choice is made by comparing the final value of the accumulator $a(T)$ to a side bias. A rightward choice is made if $a(T) >$ bias.

This model captures the rats' choices well and the best fitting parameters strongly support the view that the rats can accumulate evidence over long timescales and that the primary source of variability in their decisions is the new sensory evidence and not the accumulator's memory. In Chapter 3, we extend this model to capture trial-to-trial dependence in decisions. In Chapter 4, we leverage the moment-by-moment estimates of possible accumulator values derived from this model to analyze tuning properties of individual neurons.

1.2.2 Neural mechanisms

A short review of major advances in uncovering the neural mechanisms of decision-making based on accumulation of evidence is presented in Section 4.1.

1.3 Contributions and outline

This thesis describes three distinct studies all aimed at understanding the behavioral and neural principles that underlie perceptual decision-making. Each chapter can be read independently without any loss of clarity.

In Chapter 2 we discuss the impact of random drifts in decision-criterion on choice behavior and the many ways in which they induce patterns of history dependence that mimic signatures of active learning. We then investigate the pitfalls of a popular model-free correction from the literature proposed to account for this and show that the correction distorts one's inference away from the veridical strategies towards a narrow subset. To address these limitations, we propose a model-based approach for disambiguating systematic updates from random drifts, and demonstrate its success on real and synthetic datasets. This work is published in Gupta and Brody (2022).

In Chapter 3 we investigate whether history biases and lapses, two commonly observed suboptimalities could be linked in their origin. We explore the idea that history biases may reflect a misbelief about non-stationarity in the world, and demonstrate that normative decision-making under such beliefs gives rise to both history-dependence and choices that appear to be evidence-independent (or lapses). We show that the predictions made by this model are borne out in choices and reaction times of a large rat dataset during evidence accumulation. This work was awarded the Best Paper Award at the Reinforcement Learning and Decision Making Conference (2022).

In Chapter 4 we present results from projection-specific perturbations and analysis of task-relevant representations in FOF, ADS, and their communication subspace that challenge previous feedforward mappings of evidence accumulation onto cortico-striatal circuitry in rats performing the Poisson Clicks task. We show that both these regions are involved throughout the accumulation process, have redundant information and yet respond differentially to perturbations. We propose a distributed implementation of evidence accumulation that accounts for these puzzling observations, using a multi-region recurrent neural network model and discuss ongoing work studying its dynamics to gain mechanistic insight.

2

Limitations of a proposed correction for slow drifts in decision criterion

2.1 Introduction

Animals' choices in perceptual decision-making tasks often display a dependence on the recent history of stimuli, choices, and outcomes. This dependence is thought to arise from systematic updating of decision variables from trial to trial. These updates may reflect ongoing learning (Dayan and Daw, 2008), for instance an agent learning to perform a perceptual categorization task might update its beliefs about the prior probabilities of the different categories (Yu et al., 2009; Zhang et al., 2014), the category boundary separating them (Drugowitsch et al., 2019; Mendonça et al., 2020), or the values of the available actions (Lak

et al., 2020b;a; Pisupati et al., 2021). Alternatively, systematic updates may reflect heuristic strategies adopted by decision-makers due to resource constraints or mismatched objectives (Gigerenzer and Gaissmaier, 2011; Abrahamyan et al., 2016; Gardner, 2019). For example, a reward-seeking agent might be prone to repeating rewarded actions and avoid unrewarded ones, i.e. follow a win-stay lose-switch strategy, similarly high costs of motor switching may encourage an agent to repeat previously chosen actions, yielding stay biases. These learning processes and heuristic strategies are often identified by the distinct patterns of choices they predict.

In addition to these systematic updates, decision variables may drift randomly from trial to trial (Dutilh et al., 2012; Lak et al., 2020a; Ashwood et al., 2022; Cowley et al., 2020; Hennig et al., 2021; Roy et al., 2021; Lyamzin et al., 2021). These "unsystematic" drifts may arise from history-independent fluctuations in internal states such as attention, arousal and motivation, or from other unknown sources of noise (Renart and Machens, 2014). However, despite their history independence, unsystematic drifts may nevertheless produce correlations in choices that obscure the effect of systematic updates and thereby complicate their inference (Lak et al., 2020a; Sugawara and Katahira, 2021).

An important recent study (Lak et al., 2020a) considered the challenge posed by unsystematic drifts in one key decision variable, the decision criterion, i.e. the threshold for choosing one alternative over the other. The authors showed that a slow drift in criterion produces an apparent history dependence that mimics the signatures of updates from a systematic learning process. Hence, in order to remove the influence of these slow drifts, they proposed a model-free (MF) correction (Lak et al., 2020a), similar in spirit to an approach employed by (Mendonça et al., 2020). They reasoned that slow drifts would produce correlations in

choices across multiple successive trials. Thus behavior in a given trial would be correlated with both immediately previous and immediately future trials, giving the appearance that experience in the current trial would not just influence future choices (a causal process), but would also spuriously influence past choices (an acausal, therefore not biologically plausible process). Hence, they posited that the effect of slow drifts of decision criterion can be removed by *subtracting the influence of the current trial on past choice (the acausal component) from the influence of the current trial on future choice*. The simplicity of this correction is appealing and has already invited other authors to use it on their datasets. Indeed, recently the International Brain Laboratory reported that their dataset displayed win-stay lose-stay behavior, but when they accounted for possible slow drifts using the MF correction, they revealed instead a win-stay lose-switch strategy (The International Brain Laboratory et al., 2021).

While (Lak et al., 2020a) demonstrated convincingly that the MF correction can remove the influence of random slow drifts, they did so in the absence of any systematic updates of decision-variables. It remains to be shown if the MF correction is robust to the presence of learning or other heuristic strategies. If the MF correction indiscriminately removes correlations across choices, including those expected from systematic trial-by-trial updating of decision-relevant variables, that would considerably undermine its usefulness.

Here, using simulations we show that the proposed model-free (MF) correction fails in the presence of systematic trial-by-trial updating, yielding potentially misleading conclusions about the nature of underlying behavioral strategies. We show that the correction erroneously removes the effect of trial-by-trial updates that produce correlations across choices, thereby misidentifying a variety of update strategies as belonging to a narrow subspace of win-stay lose-switch or win-switch lose-stay biases. We advocate for an alternate, model-based ap-

proach for disambiguating systematic updates from random drifts, and demonstrate the success of this approach by fitting synthetic data. Finally, we apply the two approaches to real data from rodents (Brunton et al., 2013) and demonstrate that whereas the model-based approach proposed here preserves individual variability in systematic history biases across animals (while removing the influence of slow drifts), the previous model-free approach collapses this variability into a small subset of apparent strategies. Our results highlight the importance of evaluating assumptions underlying model-free corrections, and offer arguments for incorporating assumptions about the generative process directly into model fitting.

2.2 Results

2.2.1 Correction in action

We begin by reproducing the results from (Lak et al., 2020a) that the proposed model-free (MF) correction removes spurious choice correlations that are introduced by slow drifts in decision criterion. We consider a generative model (Figure 1A) in which the stimulus affects the choice of an agent in accordance with signal detection theory (see Methods). On any given trial t, the agent compares a noisy perceptual sample, centered around the true stimulus s_t with logistic noise (scale β_S) to a decision criterion b_t, and makes a rightward choice if the sample exceeds the criterion and leftward choice otherwise. The probability of making a rightward choice is therefore given by the following logistic function:

$$p(c_t = 1|s_t, b_t) = \frac{1}{1 + e^{-(s_t - b_t)/\beta_S}}$$

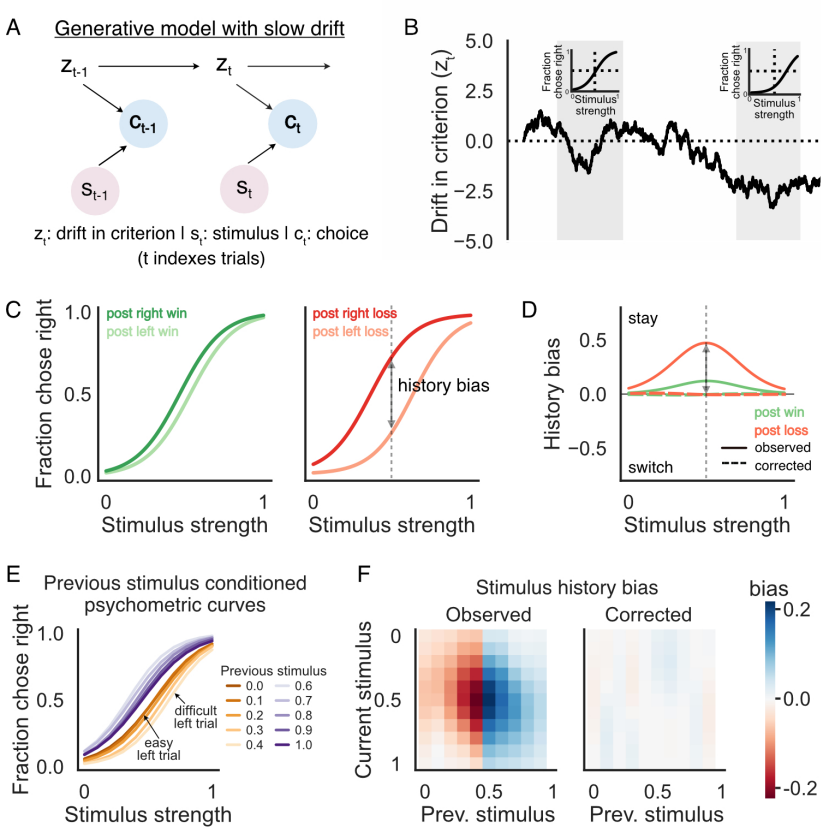


Figure 2.1: Correction for removing influences of slow drift in psychometric parameters on choice data (as proposed in Lak et al. 2020a) A. Generative model of choices based on signal detection theory: Each trial's choice c is made by comparing a noisy sample of that trial's stimulus s to a slowly drifting criterion (drift in the criterion represented by z, trials indexed by t) **B.** Trajectory of the slow drift in criterion over a period of 5000 trials revealing extended periods of consistent choice bias. Insets: psychometric curves for the trials highlighted in grey when choices are on average unbiased (left) and biased (right). Caption continued on next page.

The decision criterion b_t of this observer varies from trial-to-trial relative to a fixed baseline β_0 due to unsystematic variations arising from random drift z_t:

$$b_t = \beta_0 + z_t$$

Figure 2.1: (Previous page.) **C.** Psychometric curves (fraction of rightward choices as a function of the current trial's stimulus strength) conditioned on the previous trial's choice and outcome, displaying an apparent win-stay lose-stay bias. Colors indicate previous outcome (green - win, red - loss) and lightness indicates previous choice (dark - right, light - left) **D.** History bias, measured as the difference between conditioned psychometric curves following right and left wins (green) or losses (red). Solid lines indicate observed history bias, displaying an apparent small win-stay, larger lose-stay effect. Dashed lines indicate history bias after applying the MF correction, which removes the spurious choice correlation created by slow drift. **E.** Psychometric curves conditioned on the previous trial's stimulus when the previous trial was rewarded. Curves display an apparent dependence on past stimulus strength, and have bigger shifts following hard trials compared to easy trials. Colors indicate previous stimulus category (violet - rightward, amber - leftward) and lightness indicates strength of the stimulus (dark - high stimulus strength or easy trials, light - low stimulus strength or hard trials) **F.** (Left) Stimulus history bias, measured as the difference between psychometric curves conditioned on previous stimulus and average psychometric curve. Biases are larger following hard trials (color gradient along x-axis) and disproportionately affect the performance on hard trials (color gradient along y-axis). Color indicates the direction of induced bias (blue - towards right, red - towards left) (Right) Stimulus history biases after applying the model-free (MF) correction, which removes the spurious dependence due to slow drifts in decision criterion.

This drift z in decision criterion b (Figure 1B) was simulated according to the following autoregressive process (discrete time Ornstein-Uhlenbeck process):

$$z_t = (1 - \lambda)z_{t-1} + \sigma_d \epsilon_t$$

where ϵ_t is an i.i.d sequence of standard Gaussian random variables and σ_d sets the standard deviation of the Gaussian noise. The decay rate λ was fixed to a small value ($\lambda = 5e - 4$) to slightly dampen the pure Brownian motion thereby reducing the odds of the criterion drifting too far from the true boundary.

We assume that there is no learning or systematic trial-by-trial updating of any decision-related variables. In spite of that, the slowly drifting criterion produces correlations in the agent's choices across trials, such that when the psychometric curve is conditioned on previous trial's outcome and choice, it appears that the agent is following a small win-stay larger lose-stay strategy (Figure 1C)– in other words, it appears that subjects use the experience of a trial's outcome to affect future trials, when in fact reward outcomes do not affect the deci-

sion process at all. The psychometric curves following both correct and incorrect rightward choices are biased towards rightward choices and vice versa following leftward choices. We summarize these apparent history effects by taking the difference between conditioned psychometric curves following right and left wins/losses. Therefore, following correct trials the history bias is given according to:

$$\text{Bias}_{t|t-1}^{PC} = \frac{1}{2}\big(p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 1) - p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 0)\big)$$

where $o_{t-1} = 1$ if the trial $t-1$ is a win and is 0 otherwise (PC refers to post-correct). And to compute the bias following error trials (PE) we instead condition on trial $t-1$ being an incorrect trial or $o_{t-1} = 0$:

$$\text{Bias}_{t|t-1}^{PE} = \frac{1}{2}\big(p(c_t = 1|s_t, o_{t-1} = 0, c_{t-1} = 1) - p(c_t = 1|s_t, o_{t-1} = 0, c_{t-1} = 0)\big)$$

This metric is positive if the behavior has a stay bias and is negative when there is a tendency to switch. Even though the drift in criterion is independent of the previous trial's outcome, the stay bias appears previous outcome-dependent - its magnitude is smaller following correct choices than erroneous choices (Figure 1D, solid lines). This effect arises because the agent is more prone to errors and more likely to have a persistent choice bias when the criterion has drifted further away from the true boundary.

Further, we examine how recent sensory history affects future choices under such a generative model and observe that again, the slow drifts in criterion produce the semblance of systematic trial-by-trial updating (Figure 1E,F; replication of Figure 2, figure supplement 1 from (Lak et al., 2020a)). It appears as though the agent's choice on the current trial is modu-

lated by the previous trial's stimulus difficulty (Figure 1E). Following correct trials the agent has a higher propensity to repeat the correct response to the previous trial's category if the previous trial was difficult. In contrast if the past trial was an error the bias is higher following an easy trial (not shown). This apparent difficulty-dependent stimulus history bias (resembling confidence-guided updating) arises because of the statistics of occurrence of correct trials in the presence of drifts in criterion. Easy trials are resilient to drifts in criterion and are just as likely to be correct both when the drifts favor the correct option and when they are biasing the decision against it, whereas hard trials have a higher tendency of being correct when the drift in criterion is biased towards the correct option. As a result, trials following a correct hard trial are sampled from a distribution in which the median value of the criterion is biased, hence the conditioned psychometric curve has a large bias. Whereas, trials following an easy trial results from choices with relatively unbiased criterion values, hence the conditioned psychometric curve has a smaller bias.

We summarize these apparent stimulus history biases following rewarded trials for each pair of previous and current stimuli in Figure 1F (left panel). These biases are computed by subtracting the psychometric curve computed using all trials from the psychometric curves conditioned on the previous trial's stimulus strength (similar to (Lak et al., 2020a)):

$$\text{Bias}_{s_t|s_{t-1}} = \frac{1}{2}\big(p(c_t = 1|s_t, s_{t-1}, o_{t-1} = 1) - p(c_t = 1|s_t)\big)$$

Next, we apply the MF correction to this synthetic data. The MF correction exploits the idea that the effect of slow drifts should be similar for at least a small stretch of trials ($n >= 2$) and any influence of future rewards on past choices is acausal, likely due to slow drifts and needs to be removed. (Mendonça et al., 2020) proposed the variant of the correction for

choice-outcome biases and (Lak et al., 2020a) for stimulus history biases. We describe and test the efficacy of both these variants here. For post-correct (PC) history bias, the MF correction entails subtracting the (spurious) bias due to trial $t-1$ on the past trial $t-2$ from the bias due to trial $t-1$ on the future trial t:

$$\text{Corrected Bias}^{PC}_{t|t-1} = \text{Bias}^{PC}_{t|t-1} - \text{Bias}^{PC}_{t-2|t-1}$$

Post-error history bias is corrected for slow drifts in a similar manner. The MF correction for stimulus history effects stipulates that the effect of slow drifts in decision criterion can be removed by performing an equivalent operation on inferred stimulus history bias:

$$\text{Corrected Bias}_{s_t|s_{t-1}} = \text{Bias}_{s_t|s_{t-1}} - \text{Bias}_{s_{t-2}|s_{t-1}}$$

When applied to synthetic data, the MF correction removes these apparent choice history effects caused by the drifting criterion and successfully recovers the systematic component of past choice's influence on subsequent choice i.e. no influence (Figure 1D, dashed lines). Similarly, specious stimulus history effects could be successfully removed by applying the MF correction (Figure 1F, right panel).

2.2.2 Correction in misaction

In the past section, the generative model lacked active trial-by-trial updating and hence the correlations in choice were entirely due to the slow drift in decision criterion. Next we consider a more complex scenario, one in which the decision variables are systematically updated from trial-to-trial due to ongoing learning or other heuristic strategies in addition to the ran-

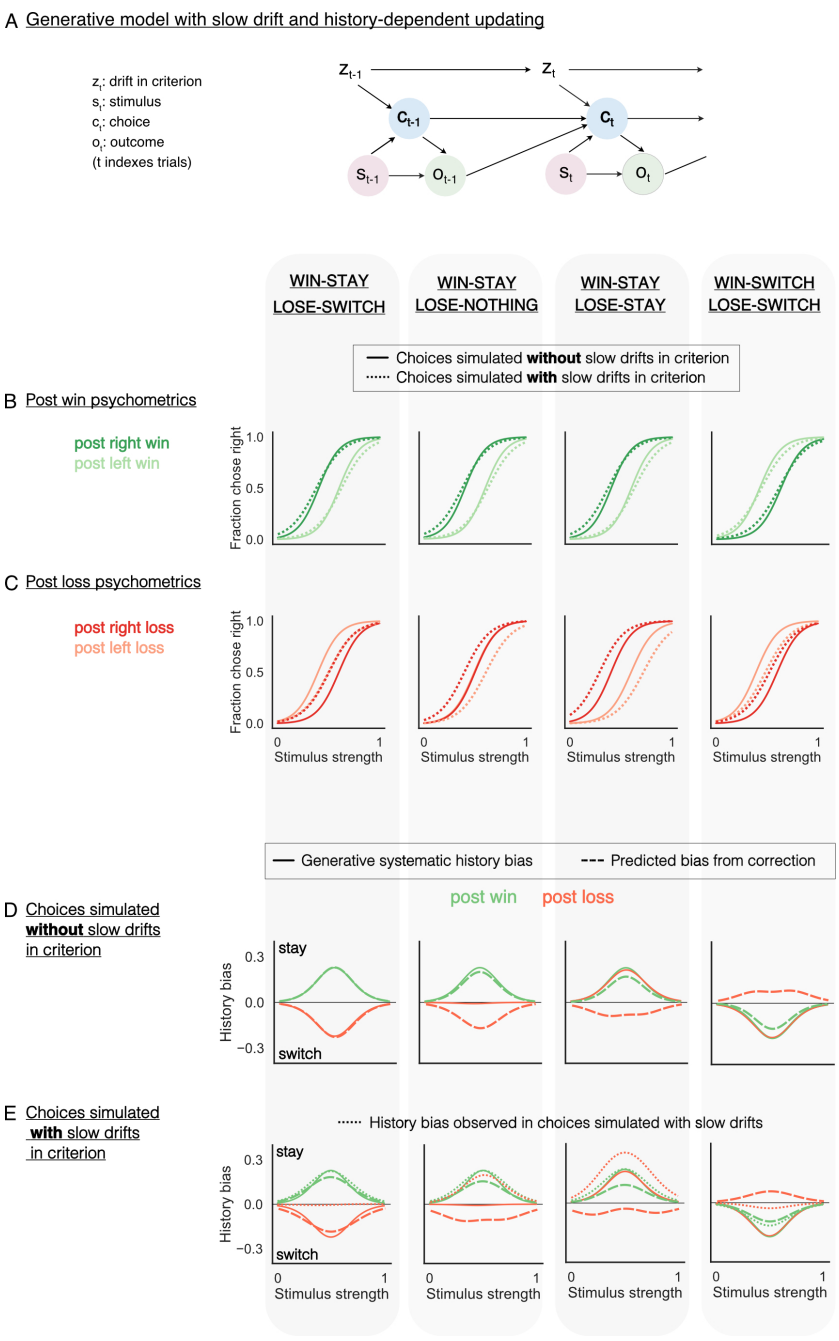


Figure 2.2: Correction fails to recover generative parameters when decision variables are actively updated from trial-to-trial Caption on next page.

Figure 2.2: (Previous page) **A.** Generative model in which the current trial's choice c_t is influenced by the previous choice c_{t-1} and previous outcome o_{t-1}, in addition to the current trial's stimulus s_t and a slow drift in criterion z_t . In some scenarios, the previous outcome does not influence the choice on the next trial e.g. win-stay lose-stay or win-switch lose-switch. **B.** Psychometric curves post rewarded trials, conditioned on previous trial's choice for a model with win-stay, lose-switch (column 1), win-stay lose-nothing (column 2), win-stay lose-stay (column 3) and win-switch lose-switch (column 4) biases. Solid lines represent psychometric curves without slow drift and dotted lines with slow drift. Addition of slow drift gives rise to worse perceptual sensitivity. **C.** Same as B but for trials following losses or unrewarded trials. Addition of slow drift masks lose-switch effects and overlays a stay bias across all updating strategies. **D.** History biases following wins (green) and losses (red) without slow drifts in criterion. Solid: generative systematic history biases, dashed: estimates from MF correction. MF correction nearly recovers the true generative post win history biases (dashed vs solid green line) across different updating strategies. For post loss biases (dashed vs solid red line), the MF correction has no effect for win-stay lose-switch bias (column 1), but produces spurious lose-switch effects for win-stay lose-nothing (column 2) and win-stay lose-stay (column 3) biases. For win-switch lose-switch (column 4) bias the correction returns a lose-stay bias. **E.** Same as D but for history biases following wins (green) and losses (red) with slow drifts in criterion. The correction fails to recover the generative history biases (dashed vs solid lines) in this case as well. Dotted lines depict the history bias observed in choices simulated with drift (corresponding to dotted lines in 2B,C)

dom drifts in decision criterion, and we determine how the MF correction performs under these circumstances.

Previous studies have shown that a large subset of possible learning strategies can be approximated by logistic regression models that directly represent the influence of reward and choice history on future choices (Katahira, 2015; Miller et al., 2019). For this reason, in this study we explore the space of possible choice and outcome history effects instead of examining the correction's effect in the presence of individual learning algorithms. We parametrically vary the bias induced by past choice (c) and outcome (o) and study the effect of MF correction both in the presence and absence of slow drifts (z) of decision criterion (Figure 2A). We find that while the MF correction can correctly recover a symmetric win-stay lose-switch history dependence, it fails to recover the generative history-dependence in other scenarios.

In order to simulate choices from this family of generative models, we update the randomly drifting decision criterion b_t every trial based on the recently observed choice and outcome

from the previous trial. Therefore, on any trial the decision criterion is given by:

$$b_t = z_t + \beta_0 + \beta_C(\mathbb{I}_{t-1}^{RC} - \mathbb{I}_{t-1}^{LC}) + \beta_E(\mathbb{I}_{t-1}^{RE} - \mathbb{I}_{t-1}^{LE})$$

where $\mathbb{I}_{t-1}^{RC}$, $\mathbb{I}_{t-1}^{LC}$, $\mathbb{I}_{t-1}^{RE}$ and $\mathbb{I}_{t-1}^{LE}$ are indicator variables denoting successes or failures on the previous trial ($t-1$) following rightward and leftward choices (RC refers to right correct, LC left correct, RE right error and LE left error). β_C and β_E determine how much previous trial's outcome modulates the criterion (β_C following correct trials and β_E following error trials). As before, β_0 is the fixed baseline and z_t denotes unsystematic variations arising from random drift.

First, we simulate data with a symmetric win-stay lose-switch rule, such that positive outcomes or correct trials bias the agent towards repeating the choice that led to them, and negative outcomes do the opposite i.e. $\beta_C = 1$, $\beta_E = -1$ and $z_t = 0 \; \forall \; t$ (Figure 2B,C, first column, solid lines). When, in addition to win-stay lose-switch, the decision criterion slowly drifts over trials ($z_t \neq 0$), the systematic lose-switch rule is obscured and the choice behavior seems to have no updating following error trials (Figure 2B,C first column, dotted lines). Moreover, the slope of the psychometric curve (also referred to as perceptual sensitivity) appears shallower.

We apply the MF correction to this data and examine the observed and corrected history biases with and without an underlying slow random drift in the decision criterion of the generative process. In the absence of slow drifts, as expected, the MF correction (dashed lines) does not change our estimate of history biases and we recover the generative win-stay lose-switch bias (Figure 2D, first column). But even in the presence of slow drifts, the MF correction successfully nullifies its contaminating influence on history bias and we approximately

recover the generative win-stay lose-switch bias Figure 2E first column).

Next, we consider a scenario in which the magnitude of biases following wins (i.e. rewarded trials) versus losses (unrewarded trials) are asymmetric. We eliminate the bias induced by negative outcomes and study win-stay lose-nothing ($\beta_C = 1, \beta_E = 0$; Figure 2B,C second column). Here the slow drift masks the true post-error biases and instead accentuates stay biases (Figure 2B,C dotted lines). In the absence of slow drift, we can directly observe the generative bias, therefore in principle the MF correction should not affect our estimate of history biases. But instead we observe that the MF correction distorts our estimate of history biases, skewing them to resemble a win-stay lose-switch updating strategy (Figure 2D second column).

This distortion arises from the interaction of average choice statistics cast by the win-stay lose-nothing strategy with the functional form of the correction. The MF correction dictates that we should be able to recover the systematic influence of a trial on the next trial by subtracting the acausal influence (influence of a trial on the previous trial) from the observed influence. In this case, the observed influence on error trials is zero (lose-nothing). The acausal influence however is non-zero despite the absence of slow drifts - this is because on average there are more correct trials than error, and hence more often than not the current choice is a repetition of past choice (due to win-stay), therefore the acausual influence is inferred to be a stay bias. So when we apply the correction and subtract this stay bias from zero we recover a spurious lose-switch bias on error trials. In the presence of slow drift too, the MF correction is unable to recover the generative systematic history bias (Figure 2E, second column solid vs dashed) and instead inaccurately estimates a stay bias following wins and a switch bias following losses.

When there is no dependence on previous outcomes, and the bias following both correct and error trials are towards the previous choice i.e. a win-stay lose-stay bias ($\beta_C = 1, \beta_E = 1$; Figure 2B,C third column), slow drifts in decision criterion exacerbate the stay effects following error trials. In synthetic data generated without any drifts in decision criterion but with stay bias in effect, the MF correction spuriously infers a win-stay lose-switch strategy (Figure 2D third column). A similar failure in recovering the generative parameters is observed in the presence of slow drifts (Figure 2E third column).

Next we examine the effect of the MF correction when the data is generated from an outcome independent switch rule i.e. win-switch lose-switch ($\beta_C = -1, \beta_E = -1$; Figure 2B,C last column). In this case too, slow drift obscures the true post-error effects and overlays a stay bias on it. Again, we find that the MF correction fails to recover generative biases both in the presence and absence of slow drifts - it instead returns a win-switch lose-stay like bias (Figure Figure 2D,E last column).

2.2.3 A model-based solution

The foregoing analysis demonstrates that the proposed MF correction produces inaccurate results when the systematic trial-to-trial updating of history bias deviates from a symmetric win-stay lose-switch strategy. The correction assumes that correlations in choice arise from processes unrelated to deterministic trial-by-trial updating - an assumption that is untrue for many learning or heuristic algorithms. This could be remedied by explicitly estimating the contributions of both slow drifts and systematic updates to choice behavior.

Here we describe one such approach for the simulations discussed in the previous section. We infer the parameters governing behavior by fitting the choices to the generative

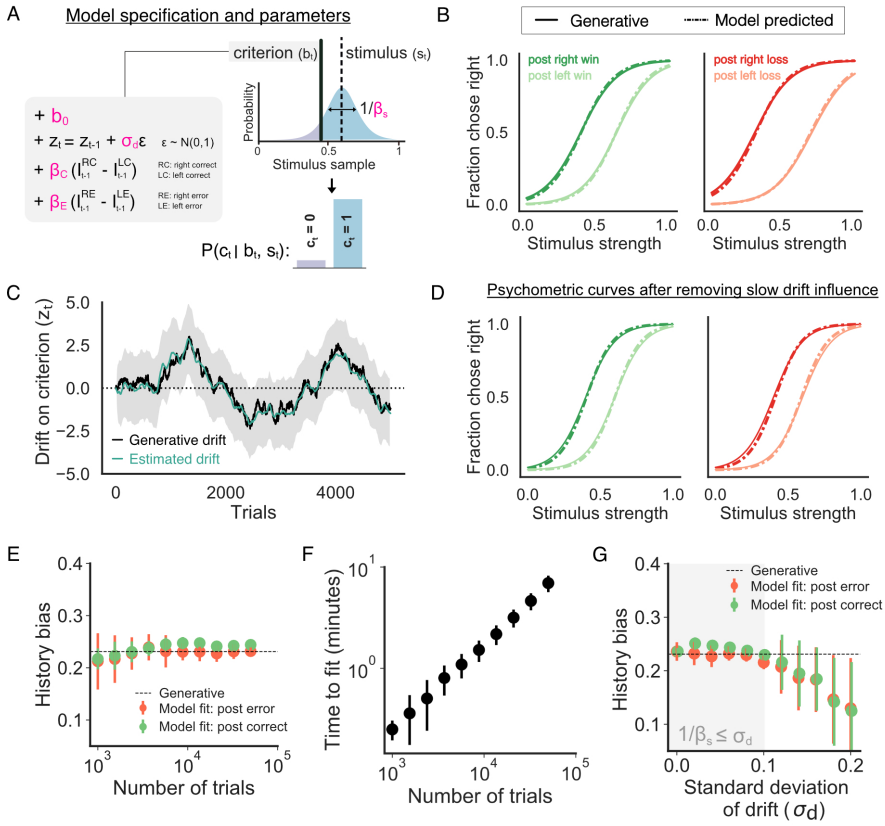


Figure 2.3: Fitting the model helps disambiguate systematic history biases from effects of slow drift A. Schematic of the model with parameters highlighted in pink. (Top right) Logistic distribution of noisy samples perceived by a subject for a given true stimulus s_t (dotted line). The probability of the current trial's choice c_t being rightward/leftward (shaded blue, bottom right) is given by the mass to the right/left of the criterion b_t (solid black line), which evolves according to the equations shown in the box (left - z_t is drift in criterion, $\mathbb{I}_{t-1}$ is an indicator variable for the previous trial's choice, outcome). **B.** Psychometric curves from simulated data (solid lines) shown alongside fits (dashed-dotted lines) to a model containing both history biases and slow drifts in criterion **C.** Inferred slow drifts from model fits (grey) compared to the true generative slow drift (black) showing good correspondence. Shaded error bars indicate two standard deviations of the posterior. **D.** Inferred history biases from model fits (dashed-dotted lines) compared to the true history biases (solid lines) showing good correspondence. Note: these are hypothetical psychometric curves that would have been observed in the absence of slow drift, and hence not directly observable from the data as in B) Caption continued on next page.

model at hand: a signal detection theory observer with logistic noise, systematic trial-history dependent biases and drifts in criterion (Figure 3A, see Methods). We fit the parameters $\Theta = \{\sigma_d, b_0, \beta_S, \beta_C, \beta_E\}$ of the model to choices using the Expectation Maximization al-

Figure 2.3: (Previous page.) **E.** Mean inferred systematic history bias (at the point of indifference) following correct (green) and error (red) trials as a function of the number of trials (in log-scale). Dotted line represents the generative systematic bias (symmetric win-stay lose-stay). Error bars denote standard deviation across 20 independent simulated datasets. **F.** Time taken to fit the model as a function of the umber of trials (both axes in log-scale). Error bars denote standard deviation across 20 independent simulated datasets. **G.** Mean inferred systematic history bias (at the point of indifference) following correct (green) and error (red) trials as a function of standard deviation of the drift. Dotted line represents the generative systematic bias (symmetric win-stay lose-stay). Error bars denote standard deviation across 20 independent simulated datasets.

gorithm with a Laplace approximation of the posterior over the latent state (drifting variable; (Macke, 2011) as implemented in the ssm library (Linderman et al., 2020).

We simulate data from an agent that follows a win-stay lose-stay strategy and has random drifts in its decision criterion (similar to Figure 2, third column). The MF correction is ineffective in recovering the generative parameters in this case. The model successfully captures trends in the data (Figure 3B: Fits (dashed-dotted lines) plotted alongside observed choice and outcome conditioned psychometric curves from the simulated data (solid lines)). We next examine if the model is able to tease apart the relative influences of slow drift and trial history bias on choice behavior and indeed, the inferred most likely trajectory of the latent drifting variable (Figure 3C, grey line) closely tracks the true simulated drift (Figure 3C, black line). Furthermore, the model correctly infers the true parameters governing the systematic component of the behavior. We demonstrate this by simulating choices in the absence of slow drifts with both the generative parameters and estimated parameters. The previous choice and outcome conditioned psychometrics from the generative and estimated parameters show good correspondence (Figure 3D). Additionally, this method helps recover the true perceptual sensitivity, which is otherwise confounded by the drifting biases (as in Figure 2B,C).

Further, we explore the performance of this model-based approach in low trial-count and high drift-noise regimes. We simulated datasets with systematic win-stay lose-stay biases (same

as Figure 3B), with varying trial counts ranging from 1000 to 50000. We then fit these datasets with the model, and examined the estimated systematic component of history bias (measured at the point of indifference, i.e. the stimulus strength at which the probability of making rightward and leftward choices is equal) as well as the time to fit. Even for trial counts as low as 1000, the model accurately infers the win-stay lose-stay bias with reasonable precision (Figure 3E), with fitting time increasing linearly with trial count (Figure 3F). We also simulated datasets with varying amounts of noise in the drift, with standard deviations ranging from 0 to 0.2. The model accurately recovered the systematic history bias with good precision as long as the noise in random drift was lower than sensory noise ($\sigma_d \leq 1/\beta_s$, grey box). Beyond this regime, even though the model correctly fits observed choices (not shown) and correctly infers the direction of systematic updates, its estimates are biased and variable due to the high degree of indeterminacy in the data.

2.2.4 Comparing the model-free correction and the model-based approach

Next, we compare the history biases estimated by the two approaches with synthetic as well as real data. We simulate choices for a variety of possible post-correct and post-error biases (Figure 4A-C, squares, see Table 1 in Methods) in presence of random drifts in decision criteria and evaluate the results obtained by applying the MF correction and by model-fitting. We summarize these history effects by plotting the choice history bias observed following correct trials against that observed following error trials, measured at the point of indifference i.e. the stimulus strength at which the fraction of rightward choices is 0.5. First, we show that consistent with our previous simulations, the presence of slow drifts exaggerates stay biases in the data (Figure 4A, dots) especially following error trials, and obscures the true underlying

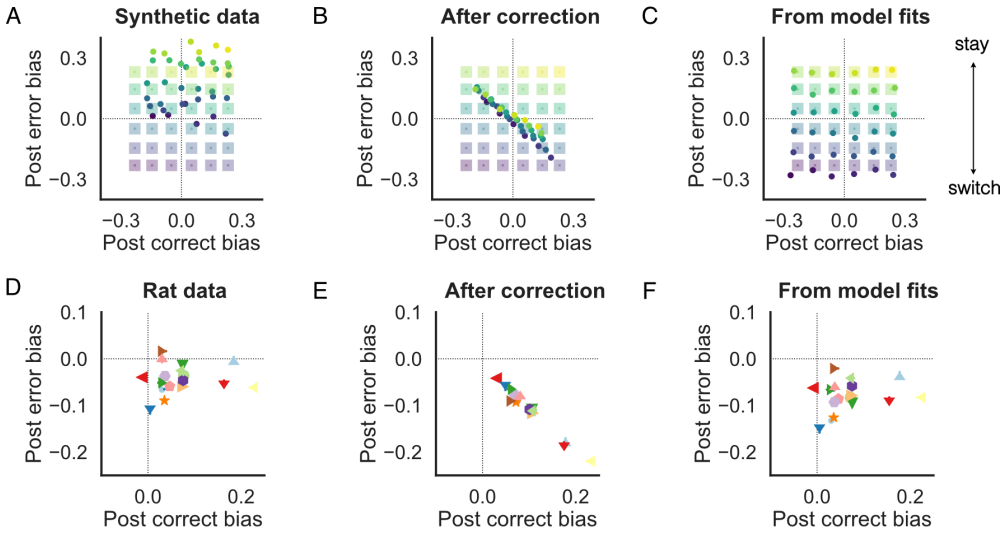


Figure 2.4: The model-based approach but not MF correction preserves variability in history updating in synthetic and real datasets A. History biases following left and right wins (x-axis) or losses (y-axis) at the point of indifference Squares denote the biases caused by systematic trial by trial updates, and dots denote the observed biases when slow drifts are present. Each color represents a unique setting of β_C and β_E parameters, which span a spectrum of stay (positive values) and switch (negative values) biases. **B.** Distribution of biases after applying MF correction **C.** Distribution of inferred history biases from fits to a model that includes slow drift and trial-by-trial updating. **D.** Distribution of observed history biases from 19 rats (dataset from (Brunton et al., 2013)). Symbols denote individual rats. **E.** Distribution of biases after applying correction from (Lak et al., 2020a). MF correction distorts the distribution, uniformly producing equal degrees of win-stay and lose-switch biases. **F.** Distribution of inferred history biases from fits to a model that includes slow drift and trial-by-trial updating. The distribution is shifted relative to A), but individual rats retain their relative position and asymmetries between post-correct and post-error updating.

update strategy (Figure 4A, squares). Application of the MF correction does not selectively remove the influence of slow drifts, rather it warps all the history effects to lie along the diagonal represented by win-switch lose-stay and win-stay lose-switch strategies (Figure 4B, dots). In contrast, the model-based inference of systematic updates successfully removes the influence of slow drifts and reveals the underlying strategy for all considered settings (Figure 4C, dots vs squares).

Next, we analyze choice data from 19 rats participating in a two-alternative auditory evidence-accumulation task (data from (Brunton et al., 2013); see Methods for more details). In this

study, even though successive trials were independent of each other, the rat subjects showed varying extents of win-stay and lose-switch biases (Figure 4D). The application of MF correction to this dataset, squashes the variability in history biases across all rats and returns a homogenous win-stay lose-switch bias (Figure 4E). In contrast, the model fits retain the individual variability (Figure 4F) and infer a higher post error switch bias compared to that observed in raw data by accounting for stay tendency that is induced by slow drifts (Figure 4A).

2.3 Discussion

Nonstationarities in decision-relevant variables, if overlooked, can bias one's estimates of psychophysical parameters and obscure strategies that underlie behavior (Fründ et al., 2011). Previous work (Lak et al., 2020a) has shown that unaccounted for slow drifts in decision criterion could obscure trial-by-trial updates produced by active learning, and proposed a correction to remove its influence (Figure 1). Here we investigate the performance of this correction under a range of different generative models, and demonstrate that it fails to selectively remove the influence of slow drifts in the presence of systematic trial-by-trial updates of decision variables (Figure 2). Moreover, applying the correction corrupts the estimates of trial-history influences, biasing them towards a small subset of possible strategies. To mitigate these shortcomings, we propose an alternate approach of explicitly modeling slow fluctuations in the decision-making process (Figure 3). We demonstrate that this model-based approach can successfully disambiguate systematic updates from non-specific drifts, hence preserving the structure of individual variability in behavioral strategies (Figure 4).

Non-specific drifts in decision variables may arise from fluctuations in internal states such

as attention, arousal and motivation, or from other internal sources of noise (Renart and Machens, 2014). Hence, some authors have used neural measurements to gauge their dynamics (Ridderinkhof et al., 2003; Eichele et al., 2008; Cowley et al., 2020). Of particular interest is the work of Cowley and colleagues who leverage slow drifts in stimulus encoding in visual and prefrontal cortex to explain fluctuations in behavior and pinpoint the affected decision-variable.

Even in the absence of such detailed measurements, it may be possible to account for unsystematic influences by modelling the noisy dynamics of unobserved variables, as we have done in this manuscript for the decision criterion (see (Smith et al., 2004) for formative work). This approach has also been used to identify systematic and random contributions to action value learning (Findling et al., 2019), allowing for the decomposition of errors into noise-driven and choice-driven components.

In general, it is likely that multiple decision variables drift over time, or that systematic updates occur over longer timescales, in which case it is important to incorporate those assumptions into the model. While the model we suggest can be readily extended to estimate additional forms of dynamics, other authors have put forward alternate approaches. Previous studies have proposed regression models to infer the trajectory of psychometric parameters over trials, regularizing their estimates with empirically determined priors (Bak et al., 2016; Kattner et al., 2017; Bak and Pillow, 2018; Roy et al., 2021). In another study authors consider a generative model in which the latent state which governs the setting of psychometric parameters undergoes discrete change (Ashwood et al., 2022). In parallel, some generative models might allow for the use of filtering methods to recover signal from the observations. Indeed, previous studies have shown that low-pass filtering the sequence of choices using a

moving average filter can adequately help estimate the slow drifts in bias (Cowley et al., 2020; Mochol et al., 2021) without inducing the kind of biases induced by the MF correction. The relative utility of these different approaches (state space models, regression, filtering, corrections) would depend on the desired level of explanatory power, number of data samples and efficiency of the inference algorithms among other factors. In any case, studies would benefit from explicit comparison between different hypothesized generative models.

Outside the realm of trial-based behavior many exciting advances have been made in inferring the dynamics underlying naturalistic behaviour (Wiltschko et al., 2015; Eyjolfsdottir et al., 2016; Sharma et al., 2018; Calhoun et al., 2019). In the future, it would be interesting to bring these advances into the study of learnt perceptual behaviour and develop inference strategies for more sophisticated generative models. Exploring the origins and logic of slow fluctuations inferred with such models might help advance our understanding of the principles that underlie behaviour and learning.

2.4 Methods

2.4.1 Simulation Details

We simulated a signal detection theory observer that compares a noisy sample of the stimulus (range 0 to 1, corrupted by logistic noise) to a decision criterion, making a rightward choice if the sample exceeds the criterion and leftward choice otherwise.

The decision criterion b_t of this observer varies from trial to trial relative to a fixed baseline β_0, due to unsystematic variations arising from random drift z_t and systematic variations based on the choice and outcome of the previous trial c_{t-1}, o_{t-1}. The drift z in decision criterion b over trials t was simulated according to the following autoregressive process (discrete

time Ornstein-Uhlenbeck process):

$$z_t = (1 - \lambda)z_{t-1} + \sigma_d \epsilon_t$$

where ϵ_t is an i.i.d sequence of standard Gaussian random variables. Throughout the study the decay rate was fixed to a small value ($\lambda = 5e - 4$) to prevent the criterion from drifting too far from the true boundary and the standard deviation of the Gaussian noise, σ_d was set at 0.05 when applicable. The systematic variations based on previous choice and outcome took the form of an additive bias to the decision criterion. The bias was $\beta_C / -\beta_C$ following rightward/leftward choices and positive outcomes, and $\beta_E / -\beta_E$ following rightward /leftward choices and negative outcomes. Therefore, the decision criterion on any trial was given by the following equation:

$$\begin{aligned} b_t(c_{t-1}, o_{t-1}, z_t) = \beta_0 + z_t \\ + \beta_C \mathbb{I}(c_{t-1} = 1, o_{t-1} = 1) - \beta_C \mathbb{I}(c_{t-1} = 0, o_{t-1} = 1) \\ + \beta_E \mathbb{I}(c_{t-1} = 1, o_{t-1} = 0) - \beta_E \mathbb{I}(c_{t-1} = 0, o_{t-1} = 0) \end{aligned}$$

where $\mathbb{I}(c_{t-1}, o_{t-1})$ denotes an indicator variable for a particular combination of previous choice c_{t-1} and outcome o_{t-1}. The baseline value β_0 was set to produce an equal proportion of leftward and rightward choices when stimulus carried no information i.e. stimulus strength of 0.5.

Finally, choice c on trial t depended upon the stimulus strength on the current trial s_t, and the decision criterion b_t. The probability of making a rightward choice was given by the

logistic function:

$$p(c_t = 1|s_t, b_t) = \frac{1}{1 + e^{-(s_t - b_t)/\beta_S}}$$

The psychometric functions were fit using the logistic regression function from the scikit-learn module in Python (Pedregosa et al., 2011).

	Number of trials	β_S	b_t	β_C	β_E	σ_d
Fig 1C-D	40,000	10	-5	0	0	0.05
Fig 1E-F	80,000	10	-5	0	0	0.10
Fig 2B-E solid (left to right)	80,000	10	-5	1,1,1,-1	-1,0,1,-1	0.00
Fig 2B-E dotted (left to right)	80,000	10	-5	1,1,1,-1	-1,0,1,-1	0.05
Fig 3B-D	40,000	10	-5	1	1	0.05
Fig 3E-F	see x-axis	10	-5	1	1	0.05
Fig 3G	20,000	10	-5	1	1	see x-axis
Fig 4A-C	40,000	10	-5	$\pm 1, \pm 0.6, \pm 0.2$	$\pm 1, \pm 0.6, \pm 0.2$	0.05

TABLE 1 : Parameters used in simulations

2.4.2 History Bias and Correction

History bias post-correct/error at each stimulus strength is computed as the difference between the probability of going right following rightward and leftward correct/error trials. Therefore, the history bias following correct trials is given by

$$bias^{PC}_{t|t-1} = \frac{1}{2}\left(p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 1) - p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 0)\right)$$

And similarly, following error trials history bias is computed as

$$\frac{1}{2}\left(p(c_t = 1|s_t, o_{t-1} = 0, c_{t-1} = 1) - p(c_t = 1|s_t, o_{t-1} = 0, c_{t-1} = 0)\right)$$

To correct for the influence of slow drifts, we subtract the acausal bias due to the past trial

on the choice preceding it from the bias due to the past trial on current choice $bias_{t-2|t-1}$ (Lak et al., 2020a; Mendonça et al., 2020), i.e. following a rewarded trial the corrected bias for each stimulus strength was computed according to:

$$bias^{PC}_{t|t-1} - bias^{PC}_{t-2|t-1} = \frac{1}{2}\big(p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 1) - p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 0)\big) - \frac{1}{2}\big(p(c_{t-2} = 1|s_{t=2}, o_{t-1} = 1, c_{t-1} = 1) - p(c_{t-2} = 1|s_{t-2}, o_{t-1} = 1, c_{t-1} = 0)\big)$$

The corrected bias following error trials was similarly computed, but by conditioning on previous error trials i.e. $o_{t-1} = 0$.

2.4.3 Model-based fitting

We denote the choice on trial $t \in 1...T$ as c_t such that $c_t = 1$ if the choice is towards right and 0 otherwise. The choice on any given trial is formed by comparing the value of a noisy perceptual sample (centered around the true stimulus s_t with logistic noise) to the category boundary or criterion b_t, which we assume is time-varying and is modulated by trial-history. If the value of the sample exceeds the criterion, a rightward choice is made. Therefore, conditioned on the drifting criterion and the stimulus the choice of an agent on any given trial is given by a Bernoulli distribution with mean $\mathbb{E}[c_t|b_t, s_t] = \frac{1}{1+\exp(-(\beta_S s_t - b_t))}$ where β_S determines the sensitivity of an agent to the perceptual stimulus s_t.

We assume that the drift z_t in criterion b_t evolves according to a random walk with step size

σ_d, we assume that the mean and variance of the initial state are known ($z_0 = 0, \sigma_0^2 = 0.01$):

$$z_1 \sim \mathcal{N}(z_0, \sigma_0^2)$$
$$z_{t+1}|z_t \sim \mathcal{N}(z_t, \sigma_d^2)$$

The decision criterion b_t is the summation of a fixed baseline β_0, the drift, and trial-history influences:

$$b_t = z_t + \beta_0 + \beta_C(\mathbb{I}_{t-1}^{RC} - \mathbb{I}_{t-1}^{LC}) + \beta_E(\mathbb{I}_{t-1}^{RE} - \mathbb{I}_{t-1}^{LE})$$

where $\mathbb{I}_{t-1}^{RC}$, $\mathbb{I}_{t-1}^{LC}$, $\mathbb{I}_{t-1}^{RE}$ and $\mathbb{I}_{t-1}^{LE}$ are indicator variables denoting successes or failures on the previous trial $(t-1)$ following rightward and leftward choices. β_C and β_E determine how much previous trial's choice and outcome modulate the criterion.

This model for choice behavior is essentially a linear dynamical system with Bernoulli emissions. Therefore, to infer the parameters $\Theta = \{\sigma_d, b_0, \beta_S, \beta_C, \beta_E\}$ of this latent variable model we use an expectation-maximization algorithm, similar to those described before (Smith and Brown 2003, Macke et al. 2011, Macke et al. 2015). For the E-step, we require the log of posterior distribution $P(\mathbf{z}|\mathbf{c}, \Theta)$ over the latent drift given the observed pattern of choices and our current estimate of the parameters Θ:

$$\log P(\mathbf{z}|\mathbf{c}, \Theta) = \sum_{t=1}^{T} \left(c_t h_t + \log\left(1 - \frac{1}{1+e^{-h_t}}\right)\right) - \frac{(z_1 - z_0)^2}{2\sigma_0^2} - \frac{1}{2\sigma_d^2}\sum_{t=1}^{T-1}(z_{t+1} - z_t)^2 + \text{const}$$

where $h_t = \beta_S s_t - b_0 - z_t - \beta_C \mathbb{I}_t^C - \beta_E \mathbb{I}_t^E$. Owing to Bernoulli observations this distribution is not available in closed-form but is concave, so we approximate it with a multi-

variate Gaussian distribution (Laplace approximation) $P(\mathbf{z}|\mathbf{c},\Theta) \sim \mathbb{N}(\mu^*,\Sigma^*)$ where $\mu*$ is a vector of size T and is set to maximize the posterior $\mu^* = \arg\max_{\mathbf{z}} P(\mathbf{z}|\mathbf{c},\Theta)$ and Σ^* is of size $T \times T$ and is set to the inverse Hessian of the log posterior evaluated at μ^* i.e. $\Sigma^* = -[\nabla^2_{\mathbf{z}} \log P(\mathbf{z}|\mathbf{c},\Theta)|_{\mathbf{z}=\mu^*}]^{-1}$.

The M-step updates parameters by optimizing the expected total data log-likelihood with respect to the parameters

$$\Theta^* = \arg\max_{\theta} \int [\log(\mathbf{c}|\mathbf{z},\Theta) + \log P(\mathbf{z}|\Theta)]\mathbb{N}(\mathbf{z}|\mu^*,\Sigma^*)d\mathbf{z}$$

We update the parameters using L-BFGS except σ_d which we update analytically. We use the *ssm package* (Linderman et al., 2020) in Python for this inference.

2.4.4 Dataset studied

We analyze the trial-history effects of nineteen rats (50223 ± 21915 trials per rat) performing an evidence accumulation task, originally published in (Brunton et al., 2013). In this task, the subjects were presented with two simultaneous streams of randomly-timed discrete pulses of evidence (clicks), one from a speaker to their left and the other to their right, for a predetermined duration. The subjects were required to maintain fixation throughout the entire stimulus, failure to do so led to a violation trial. At the end of the stimulus, the subjects had to orient towards the side which played the greater number of clicks to obtain a water reward. The discrimination difficulty was controlled on each trial by varying the ratio of the generative Poisson rates of the two click streams. Successive trial's difficulty and rewarded side were independently sampled. Since the animals neither made a choice nor received an outcome on violation trials, we ignore them while computing trial-history effects.

2.4.5 Data and Code availability

The code and data associated with this manuscript is available here: https://doi.org/10.24433/CO.8821874.v1 as a standalone executable.

acknowledgements

We thank Armin Lak, Tim Kim, Thomas Luo and Adrian Bondy for helpful discussion and comments on this manuscript.

This work was done in collaboration with Charles Kopec and Brian DePasquale.

3

Trial history biases in evidence accumulation can give rise to apparent lapses

3.1 Introduction

Experienced perceptual decision makers have long been known to deviate from the predictions of optimal decision-theory, displaying several suboptimalities in their decision-making. Among the most pervasive of these is the dependence of behavior on the recent history of observed stimuli, performed actions or received outcomes, despite it being disadvantageous and leading to worse performance (Cho et al., 2002; Gold et al., 2008; Busse et al., 2011; Carandini and Churchland, 2013; Zhang et al., 2014; Fründ et al., 2014; Scott et al., 2015; Abrahamyan et al., 2016; Odoemene et al., 2018; Akrami et al., 2018; Pinto et al., 2018; Urai et al., 2019;

Hermoso-Mendizabal et al., 2020; Mendonça et al., 2020; Lak et al., 2020a; Mochol et al., 2021; Roy et al., 2021; The International Brain Laboratory et al., 2021). History biases may arise due to a strategy that is optimized for naturalistic settings, where continual learning of priors, action-values or other decision variables helps agents adapt to changing environments but is maladaptive in experimental settings where the statistics of the environment are held stationary (Yu and Cohen, 2009; Molano-Mazon et al., 2021). To date, decision-theoretic accounts have accommodated history biases by modeling them as a modulating factor on the perceptual evidence that drives decisions about choices (Nevin, 1969; Ratcliff and Rouder, 1998; Bogacz et al., 2006; Busse et al., 2011; Goldfarb et al., 2012; Gardner, 2019; Urai et al., 2019; Hermoso-Mendizabal et al., 2020). In these accounts, the influence of history biases can be overcome with sufficient perceptual evidence.

A second suboptimality that is less well studied but nevertheless widely prevalent in decision makers, is the presence of lapses, or (asymptotic) errors that are immune to high levels of evidence (Wichmann and Hill, 2001; Law and Gold, 2009; Busse et al., 2011; Gold and Ding, 2013; Brunton et al., 2013; Carandini and Churchland, 2013; Wang et al., 2018; Pinto et al., 2018; Pisupati et al., 2021; Shushruth and Shadlen, 2021). Due to this evidence-independence, lapses are assumed to arise from nuisance mechanisms that are separate from the perceptual decision-making process and are often imputed to ad-hoc sources of noise including inattention, motor errors etc.

However, several recent results suggest that these two suboptimalities may be linked in their origin. In primates, Gold and colleagues have observed that learning reduces dependence on recent trial history (Gold et al., 2008) as well as lapse probabilities (Law and Gold, 2009). Intriguingly, mice trained on a visual detection task have been reported to show higher

levels of history dependence on sessions with higher lapse probabilities (Busse et al., 2011). Moreover, lapses have been shown to display Markov dependencies, rather than the traditionally assumed independent probabilities of occurrence across trials (Ashwood et al., 2022). Furthermore, lapses have been proposed to reflect forms of exploration (Pisupati et al., 2021) that are sensitive to trial-by-trial updates of action value. Likewise, neural perturbations of secondary motor cortex and striatum in rodents have been shown to substantially impact both lapses (Erlich et al., 2011; 2015; Yartsev et al., 2018; Guo et al., 2019; Pisupati et al., 2021; Sindreu et al., 2021) and trial-history influences on decisions (Siniscalchi et al., 2019; Sindreu et al., 2021). Together, these observations discredit the assumption that history biases and lapses have independent causes and raise the possibility that perhaps some of the variance ascribed to lapses emerges from history dependence.

What cognitive process could underlie the links between history biases and lapses? Here we explore the idea that history biases may reflect a misbelief about non-stationarity in the world, and demonstrate that normative decision-making under such beliefs gives rise to both history-dependence and choices that appear to be evidence-independent. This corresponds to an accumulation to bound process with history dependent initial state updates. We fit this model to a dataset of 152 rats trained on an auditory decision-making task. Despite the heterogeneity in history biases and lapse rates in this population, we show that the a large fraction of lapses can be explained by presence of history dependence during evidence accumulation. Further, our model makes predictions about the time it takes to make decisions. We test these predictions in a rat dataset with reaction time reports, and show that it captures patterns of choices, reaction times and their history dependence in this dataset. This model improves our ability to predict the precise dynamics of decision variables within and across

trials in perceptual decision making tasks and helps parses variance attributed to lapse into history-dependent and independent components.

3.2 Results

3.2.1 A common mechanism produces history biases and apparent lapses

Well trained subjects in 2AFC tasks are often assumed to have faithfully learnt the likelihood function and priors that determine the structure of the task (Bogacz et al., 2006; Gold and Shadlen, 2007). Given this knowledge, the optimal decision making strategy entails combining any knowledge about prior prevalence of the available options with incoming pieces of evidence, continuing to gather more evidence until a desired threshold of confidence is reached in favor of one of the options* (Gold and Shadlen, 2007; Dayan and Daw, 2008; Drugowitsch et al., 2012) (Fig 3.1B top). This strategy for making decisions converges to a drift diffusion model (DDM) when evidence is sampled continuously (Bogacz et al., 2006). In a DDM, one's belief about the correct option maps onto a diffusing particle that drifts between two boundaries (Fig 3.1B). The first boundary that the particle crosses determines the decision. Correspondingly, the initial state of this particle encodes the influence of priors and, the drift rate is set by the likelihood of incoming evidence.

However in general, subjects may not have the knowledge that the task structure is stationary, and may instead assume that it is constantly evolving (Yu and Cohen, 2009). In this case, even experienced subjects would not converge to a static estimate of these quantities, but would instead incessantly update them from trial to trial. Here we consider choice behavior

*In tasks where the reliability of incoming evidence (stimulus difficulty) varies from one trial to next, it has been shown that ideal observers should have time-varying bounds (Drugowitsch et al., 2012), however for simplicity here we assume that the bounds are stationary during and across trials.

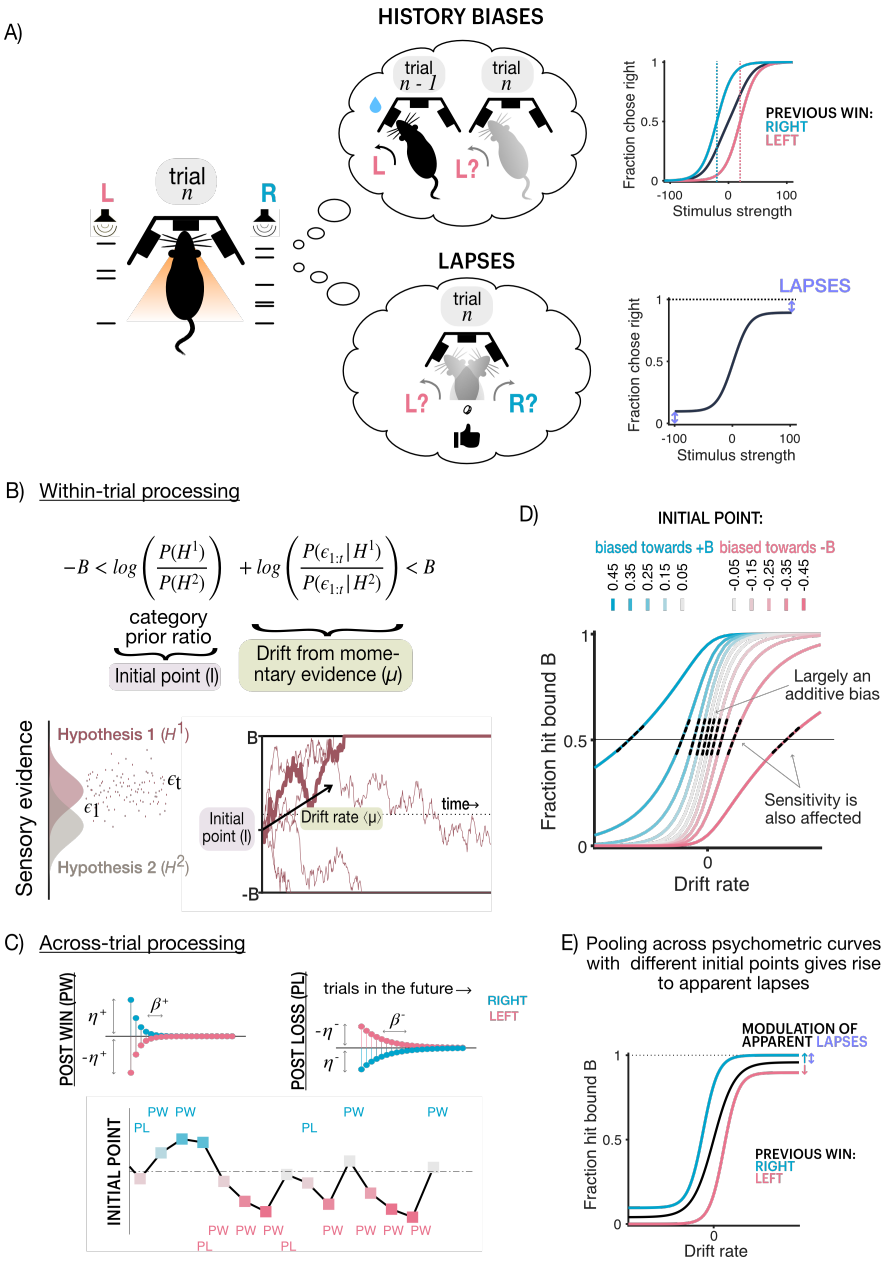


Figure 3.1: Trial history-dependent initial points give rise to apparent lapses (A) Schematic of two commonly observed suboptimalities in decision-making: history biases (top) and lapses (bottom). (Left): Rat performing a perceptual decision making task, where it has to make one of two decisions (left, right) based on accumulated sensory evidence (auditory clicks on either side). Caption continued on next page.

Figure 3.1: (Previous page.) (Top left): History biases i.e. a tendency for the decision on the current trial (n) to be inappropriately influenced by what happened on the previous trial (n-1) in addition to the accumulated sensory evidence. In this example, a previously rewarding leftward decision is likely to be repeated. (Top right): Typically assumed effect of history bias on the psychometric curve, which is the proportion of rightward decisions as a function of the stimulus strength. History biases are thought to most strongly affect decisions when the sensory evidence is weak i.e. around the inflection point of the curve (threshold parameter), shifting it horizontally. (Bottom left): Lapses i.e. a tendency to make seemingly random choices on some trials, irrespective of the accumulated sensory evidence. (Bottom right): Typically assumed effect of lapses on the psychometric curve, of vertically scaling the endpoints or asymptotes of the curve. **(B)** Standard normative model of within-trial processing during evidence accumulation. (Top) Decision rule that produces the most accurate decisions in the shortest amount of time, in which a decision is made when the summed log-ratios of category priors and accumulated evidence exceeds one of two decision bounds. This corresponds to a drift-diffusion process where the prior term sets the initial point (I) and the rate of accumulating evidence sets the drift rate (mu). (Bottom left): Schematic of the generative model, where one of two hypotheses (H1, H2) produce noisy samples of evidence over time (epsilon). (Bottom right): Schematic of the aforementioned drift-diffusion process, showing a sample trajectory based on noisy evidence (bold line) that leads to a rightward decision when the positive bound is hit. Thin lines depict alternate trajectories based on different noisy instantiations of the same drift rate (black arrow). **(C)** Model of across-trial updating of initial points. Model of across trial processing that can accommodate several forms of prior updates. Past choices and outcomes can additively affect the initial point with different magnitudes (η) and exponentially decaying timescales (β) depending on whether they were wins (top left) or losses (top right). (Bottom): Example sequence of trials, labelled by whether they follow a previous win (PW) or previous loss (PL) showing the cumulative effect of trial history on initial point updates. Colors denote different initial point biases, same as (C). **(D)** Effect of initial point values on psychometric functions. Colors denote different initial point levels, towards the positive (blue) or negative (pink) decision bounds. Small deviations from 0 in the initial point (grey) lead to largely additive, horizontal biases in the psychometric curve whereas larger deviations (saturated colors) have more complex effects, additionally reducing its effective slope (dotted black lines) or "sensitivity" to the stimulus. **(E)** Pooling different initial point biases gives rise to apparent lapses. Psychometric function (black) pooled across trials with different initial point biases (due to updating) has apparent lapses (purple arrow), moreover conditioning the psychometric curve on whether the previous trial was a rightward (blue) or leftward (pink) win reveals a modulation of these apparent lapses by trial history.

that results from non-stationary beliefs about priors, which result in trial-to-trial updates to the initial point[†]. We assume that initial points are set based on an exponential filter over previous choices and outcomes where past wins and losses affect initial points with different magnitudes and timescales (Fig 3.1C, top). So the initial point I at trial $n+1$ is given by:

$$I(n+1) = i^{RC}(n+1) + i^{LC}(n+1) + i^{RE}(n+1) + i^{LE}(n+1)$$

[†]For a treatment of non-stationary likelihood functions which yield variability in drift rate, see (Drugowitsch et al., 2019; Mendonça et al., 2020)

where each filter i^h decays by a factor of β^h, and is incremented by a factor of η^h following the observation of that particular choice-outcome pair, i.e

$$i^h(n+1) = \eta^h 1^h(o_n) + \beta^h i^h(n) \quad \text{where} \quad h = \{RC, LC, RE, LE\}$$

o_n is the choice-outcome pair observed on trial n and $1^h(o_n)$ is an indicator function that is 1 when $o_n = h$ and is 0 otherwise. Such a filter can approximate optimal updating strategies under a variety of non-stationary beliefs. As an example, we consider initial point updates during Bayesian learning of priors under the belief that the prior probabilities of the two hypotheses can undergo unsignaled jumps (Yu and Cohen 2009; Zhang et al. 2014; Supp Fig.3.1A). We show that the exponential filter successfully approximate these updates (Supp Fig.3.1B). Nevertheless, we use this more flexible parameterization to allow for asymmetric learning from different choices and outcomes, which could be beneficial under generative models where one believes that one category persists for longer than another (requiring different decay rates), or correct and incorrect outcomes are not equally informative (requiring different update magnitudes). For instance, in a prior-tracking experiment where previous correct choices had a cumulative effect, but errors had a resetting effect (Hermoso-Mendizabal et al., 2020), this could be captured in the exponential filter by faster decay rates for errors.

What are the consequences of such trial-by-trial updating of initial points for choice behavior? In a DDM, for a given initial point I and drift rate μ, the probability of choosing the option corresponding to bound $B+$ is given by:

$$P(B+) = \frac{1 - e^{-2\mu(B+I)/\sigma^2}}{1 - e^{-4\mu B/\sigma^2}}$$

where B is the magnitude of the bound and σ^2 is the squared diffusion coefficient. The resultant psychometric curves for different values of initial points are plotted in Fig 3.1D. This expression reduces to a logistic function of $\mu B/\sigma^2$ only when $I = 0$. Small deviations in the initial state manifest as largely additive biases to the evidence, shifting psychometric curves horizontally towards the option favored by the initial point, as if modulating the psychometric threshold parameter (Fig 3.1C lighter colors). Interestingly, large deviations in the initial point produce qualitatively different effects on choices (Fig 3.1C darker colors). They not only bias the choices towards the option consistent with the initial point but additionally reduce the effective sensitivity to evidence. This can be seen as reduction in slopes at the point of indifference (Fig 3.1C dashed lines). Therefore, trial to trial deviations in the initial state produce history-biased choices which have differently diminished dependence on the evidence.

The average choice behavior obtained by pooling choices with different history-biased initial-points is a mixture of psychometric curves with varying thresholds and sensitivity to perceptual evidence. Such a psychometric curve is heavy-tailed (Shen and Ma, 2019; Nguyen et al., 2019) and appears to have asymptotic errors or lapses (Fig 3.1E, black curve). These asymptotic errors are not truly random decisions or *true lapses*, rather they are *"apparent lapses"* arising from deterministic history-based updates. In such a setting, the psychometric curves obtained by conditioning on past trial's choice and outcome or *history conditioned* psychometric curves are both horizontally and vertically shifted i.e. they show history-dependent modulations in both threshold and lapse rate parameters of the psychometric curve (Fig 3.1E, Supp Fig. 3.2B). Lapse rates that are modulated by trial-history are uniquely produced by history-biased initial points in this formulation (and therefore reflect apparent lapses), this

is in contrast to lapse rates observed in the unconditioned psychometric curve which might have additional extraneous causes (Wichmann and Hill, 2001; Ashwood et al., 2022; Pisupati et al., 2021) (and therefore reflect both apparent and true lapses).

In this model, since history modulations of psychometric thresholds and lapse rates arise from one unified process, they are not allowed to vary independently of the decision-making process or of each other, rather their relative magnitudes are intimately coupled with and constrained by accumulation variables. For instance, increased magnitudes or timescales of initial point updating produce large fluctuations in the initial point across trials. This in turn reduces the effective sensitivity of the accumulation process to evidence, giving rise to more apparent lapses and history biases (Supp Fig. 3.2A). Similarly, changes in within-trial parameters of accumulation can dramatically influence these history modulations (Supp Fig. 3.2C). Decisions made with smaller accumulator bounds are more sensitive to initial point modulations, and therefore give rise to higher apparent lapses and threshold modulations. Higher levels of sensory noise have a similar effect, yielding more apparent lapses - this is consistent with recent reports of lapse rates being modulated by sensory uncertainty (Pisupati et al., 2021). Finally, more impulsive integration exaggerates the influence of initial points, hence producing more apparent lapses and history biases.

Some definitions:

- *Lapse rate*: Lapse rates capture the difference between perfect performance and observed performance at the asymptotes, measured through sigmoidal fits to the psychometric curves.

- *True lapse*: A true lapse is an independent cognitive process through which agents generate stochastic evidence-independent choices.
- *Apparent lapses*: Apparent lapses are deterministic evidence-dependent choices, that nonetheless contribute to lapse rates.

3.2.2 Rats display varying degrees of history-dependent threshold and lapse rate modulations

Next, we sought to test if history dependence contributes to psychometric lapse rates by modulating the initial state of the accumulator in empirical data. To this end, we analyze data from 152 rats (37522 ± 22090 trials per rat; Supp Fig.3.3A) trained on a previously developed auditory task that requires accumulation of pulsatile evidence over time (Brunton et al., 2013). In this task, the subject is presented with two simultaneous streams of randomly-timed discrete pulses of evidence, one from a speaker to their left and the other to their right (Fig 3.2A). The subject must maintain fixation throughout the entire stimulus, and subsequently orient towards the side which played the greater number of clicks to receive a water reward. The trial difficulty, duration of stimulus and the correct answer were set independently on successive trials. Since this task delivers sensory evidence through pulses that have random but precisely known timing it provides high statistical power to characterize decision variables that give rise to the choice behavior.

Rats showed high accuracy in the data that we analyzed (accuracy stats, Supp Fig.3.3B). This performance was stable, with little to no change in accuracy over time (average accuracy change of, Supp Fig.3.3C) hence reflecting the asymptotic behavior rather than task acquisition. Rats in this dataset showed history dependence in their choices, largely tending towards

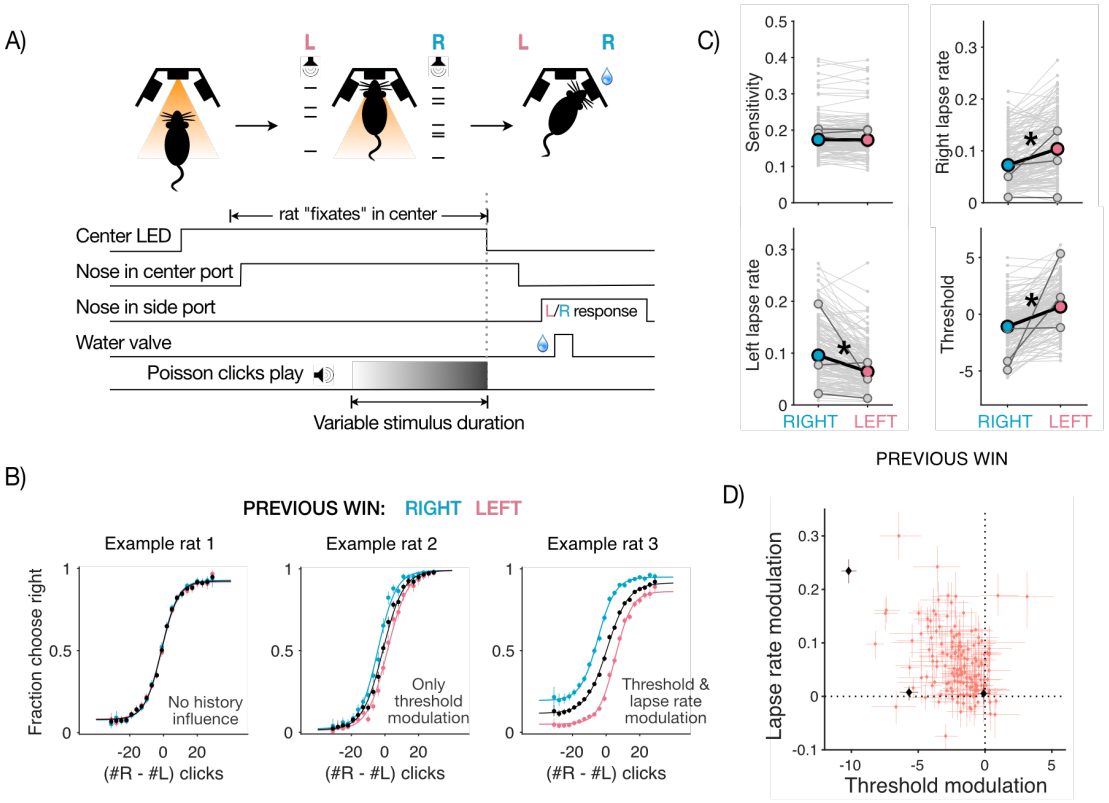


Figure 3.2: History-dependent threshold and lapse rate modulations in a large-scale rat dataset (A) Schematic of evidence accumulation task in rats: (Top): Phases of the "Poisson clicks" task, including trial initiation in center port (left), evidence accumulation based on two streams of Poisson-distributed auditory clicks (middle) and choice report in one of two side ports followed by water reward for correct choices (right). (Bottom): Time-course of trial events in a typical trial. **(B)** Individual differences in history-dependence: Psychometric functions of three example rats from a large-scale dataset, displaying different kinds of history modulation. Choices are plotted conditioned on previous left (blue), right (pink) or all wins (black). (Left): Example rat with no history-dependence in choices, resembling the ideal observer. (Middle): Example rat with modulations of the threshold parameter alone, resembling the dominant conceptualization of history bias. (Right): Example rat with history-dependent modulation of both threshold and lapse rate parameter, similar to the majority of the population. **(C)** Dataset displays significant modulations of both threshold and lapse rate parameters: Scatters showing parameters of psychometric functions following leftward wins (post left, blue) or rightward wins (post right, pink). Each pair of connected gray points represents an individual animal, solid colored dots represent average parameter values across animals. Trial history does not significantly affect the sensitivity parameter (first column) but significantly affects threshold (second column) and left, right lapse rate parameters (third, fourth column).**(D)** Scatter comparing threshold and lapse rate modulations in the entire population. Each dot is an individual animal, error bars are bootstrap CIs. Black points represent example rats. The majority of the population lies in the top left quadrant, showing co-modulations of both threshold and lapse rate parameters by history (stats).

a "win-stay, lose-switch" dependence (Supp Fig.3.3E). We found substantial individual variability in the dependence of rats' choices on history in the dataset. Some rats are barely influenced by history, confirming to the optimal strategy expected in such a stationary task (Fig 3.2B left). Some others show dependence resembling the dominant conceptualization of history biases - a history-dependent modulation of the psychometric threshold parameter (Fig 3.2B middle). Still others showed history-dependent modulation of both threshold and lapse rate parameters (Fig 3.2B right). The population as a whole most closely resembles the latter-most rat, with both threshold and lapse rate parameters being significantly different following left and right wins while sensitivity is not affected ($p = 8e-1$ for sensitivity, $3e-17$ for bias, $8e-8$ for left lapse, $6e-7$ for right lapse, calculated using Mann-Whitney U-test, Fig 3.2C). Promisingly, the consistency in direction of modulation of threshold and left/right lapse rates - towards rightward (leftward) choices following right (left) wins - matches the model behavior (Fig 3.1E). Moreover, the vast majority of rats show co-modulations of both these parameters by history (Fig 3.1D). Across rats, on average 0.17 ± 0.12 fraction of total lapse rates are modulated by trial history and therefore could potentially reflect apparent rather than true lapses (Supp Fig.3.3D).

3.2.3 Modeling history dependence in rat data

In order to test whether the history-modulations in lapses in the rat dataset arise from history-dependent initial points, we extend an accumulator model (Brunton et al., 2013) to incorporate trial-by-trial initial point updating (Fig 3.1D). This accumulator model leverages our knowledge of when each pulse of auditory evidence is played on a given trial, to model the sensory adaptation process as well as to set the noise and drift of the accumulator variable

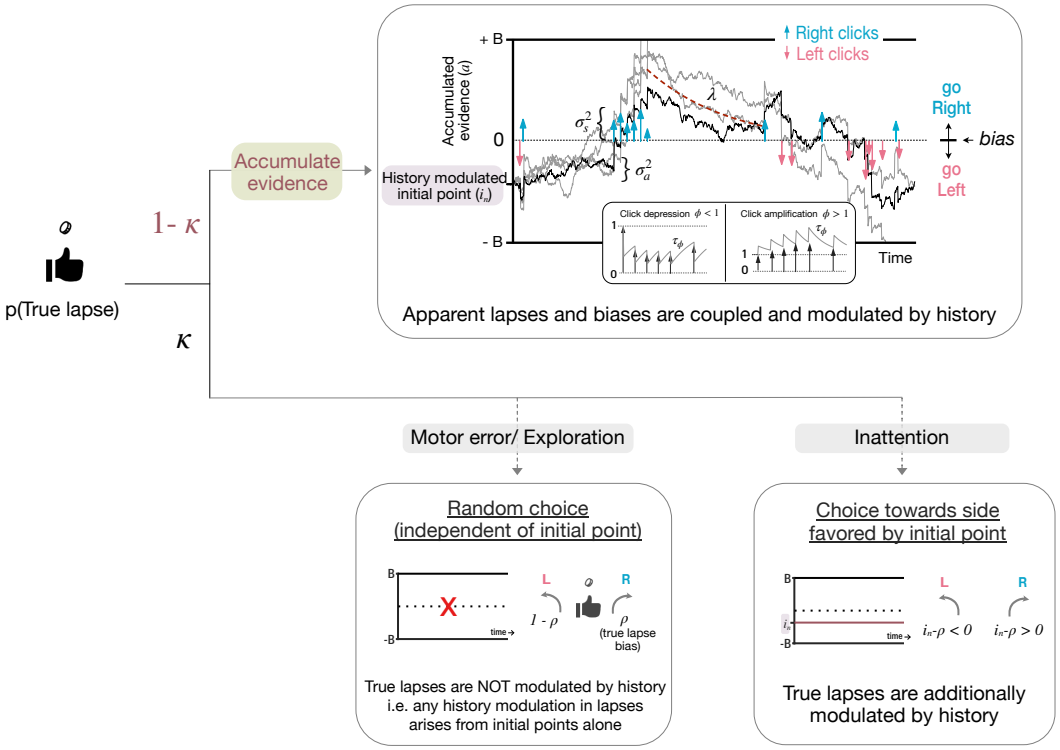


Figure 3.3: Models of rodent data including apparent and true lapses Schematic of model used to fit rodent data in the Poisson Clicks task: **(Top left)**: Probability of making true lapses (κ), in addition to the apparent lapses produced by history-dependent initial points. **(Top right)**: Schematic of accumulation-to-bound model adapted to the discrete Poisson Clicks task, whose initial points are modulated by trial history (purple) producing history-dependent apparent lapses and biases. The model consists of sensory noise (σ_s^2) in click magnitudes, adaptation of successive click magnitudes based on an adaptation scale (ϕ) and timescale (τ_ϕ), accumulator noise (σ_a^2) added at each timestep, leak in the accumulator (λ), and decision bounds +/-B. **(Bottom)**: Two different possible models of true lapses: (Left): Motor error/exploration model, that occasionally chooses a random action irrespective of the initial point, with some bias (ρ) reflecting an error in motor execution or random exploration. This model produces true lapses that are not modulated by history, such that any history modulations arise from the initial points alone. (Center, Right): Inattention model, that occasionally fail to attend to the stimulus, and relies on the initial point to make an informed decision. In this model, true lapses are also modulated by history in addition to apparent lapses. A deterministic decision is made based on the difference between the initial point and a bias (ρ)

(Fig 3.3 top right, Methods). The model also includes a feedback parameter that controls whether integration is leaky, perfect or impulsive.

Further, we augment the accumulator model with two models of true lapses in order to explicitly capture and disambiguate them from any apparent lapses arising from history-modulated initial points (Fig 3.3 bottom). The first variant models true lapses due to exploration or motor error, both of these phenomenons produce lapses that are not modulated by history. Exploration (for instance, choosing randomly on a small proportion of trials i.e. an epsilon-greedy strategy) has adaptive value in the presence of non-stationary beliefs about action value, as it helps quench the resultant uncertainty in action values. However, it is indistinguishable from motor errors, which lead to random choices due to failures in motor execution. This formulation of lapses is similar to that used in (Brunton et al., 2013). The other variant models lapses due to inattention, i.e. failures to attend to the stimulus on a small fraction of trials. We consider the optimal behavior given lack of sensory evidence due to inattention. In this strategy the choice is made deterministically towards the option favored by the initial point. Therefore, in this inattentional variant in contrast to the motor error/exploration variant, true lapses themselves are modulated by trial-history via initial points. We consider decisions arising from both these processes to be true lapses since they are truly independent of the evidence, unlike apparent lapses which still retain some evidence-dependence.

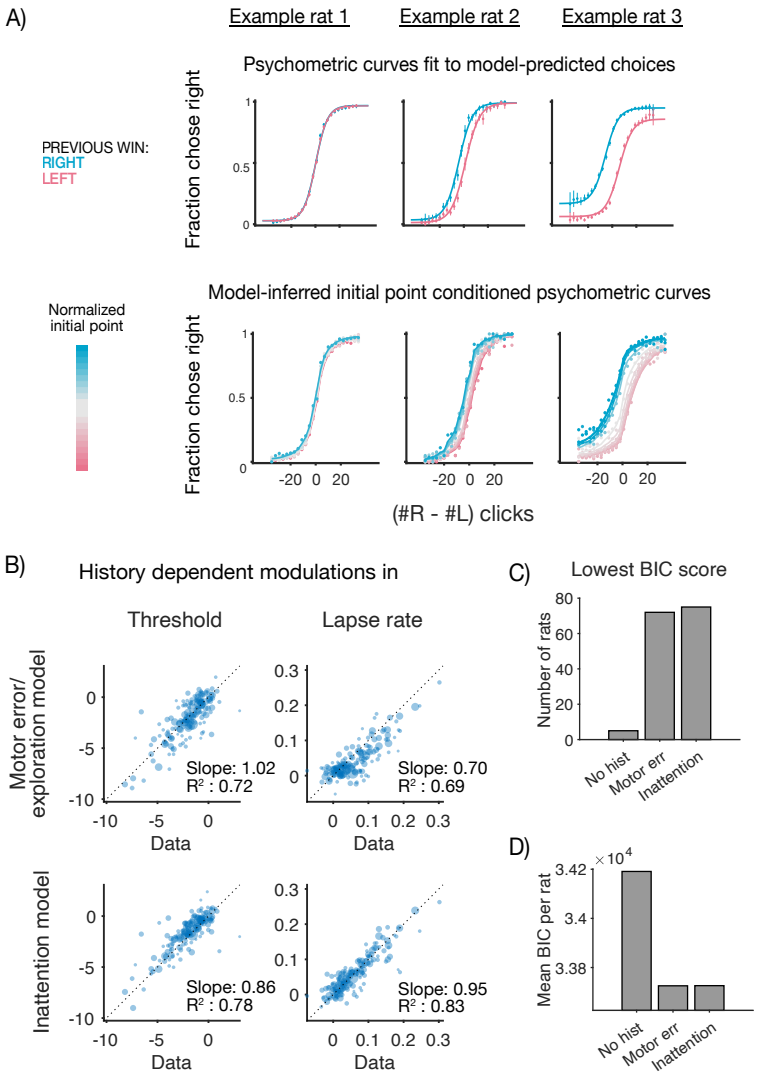


Figure 3.4: Accumulator model with history dependent initial points reproduces lapse rate modulations in the data (A) Model fits to individual rats: (Top): Psychometric data from 3 example rats (points) conditioned on previous rightward (blue) or leftward (wins), overlaid on model-predicted psychometric curves (dashed line) from the motor error/exploration model. The three example rats were chosen to illustrate the diversity of history effects in the dataset, ranging from no history effects (left) to history effects that largely created horizontal biases (center) and history effects that additionally affected lapse rates (right). (Bottom): Psychometric curves (solid line) from the same example rats conditioned on model-inferred initial points (colors from red to blue), showing a similar pattern to analytical predictions in Fig 3.1D.Caption continued on next page.

Figure 3.4: (Previous page.) **(B)** Individual differences in history effects captured by different models: History modulations of threshold (left) and lapse rate (right) parameters measured from psychometric fits to the raw data (x-axis) v.s. model predictions (y-axis). (Top): Motor error/exploration model captures a majority of the variability in both threshold and lapse rate modulations, as evidenced by the points lying close to the unity line, suggesting that these modulations arise from the coupled biases and apparent lapses produced by initial point-biased accumulation of evidence. (Bottom): same as Top but for Inattention model.**(C)** Distribution of best fitting models for individual rats: Overall bar height for each model denotes the total number of rats for which that model scored the lowest BIC score. **D**

Model comparison using BIC by pooling data across rats and computing mean BIC score per rat. Lower scores indicate better fits. Scores favor motor-error/exploration and inattention variants comparably.

3.2.4 History dependent initial points reproduce lapse rate modulations in data

We perform trial-by-trial fitting of choices from each individual rat to the accumulator model with history-dependent initial points in combination with different variants of true lapses, using maximum likelihood estimation (Methods).

In order to assess to what extent history modulations in lapse rates can be accounted for by apparent lapses produced by history-dependent initial points, we first focus on the exploration/motor error model, since in this model true lapses are not modulated by history. We find that the exploration/motor-error model is able to capture both threshold and lapse rate modulations well across different regimes of rat behavior, as evident from fits to example rats (Fig 3.4A top). Moreover conditioning rats' psychometric curves on model-inferred initial point values reveals that the initial points capture a substantial amount of variance in choice probabilities (Fig 3.4A bottom), resembling theoretical predictions (Fig 3.1C).

Next, we compare the threshold and lapse rate modulations produced by this model to the modulations in the data measured through conditioned psychometrics (Fig 3.4B). Once again, we find that across individuals, the motor error/exploration model captures a sub-

stantial amount of variance [$R^2 = 0.72$ (threshold parameter), $R^2 = 0.69$ (lapse rate parameter)] and shows good correspondence to the empirical modulations in data [slope= 1.02 (threshold parameter), slope $= 0.70$ (lapse rate parameter)]. The inattention model, which additionally allows true lapses to depend on history (Fig 3.3), captures more variance [$R^2 = 0.86$ (threshold parameter), $R^2 = 0.83$ (lapse rate parameter)] however it shows higher correspondence with lapse rate modulations at the expense of threshold modulations [slope= 0.86 (threshold parameter), slope $= 0.95$ (lapse rate parameter)]. This minor improvement in accounting for history modulations in lapse rates suggests that apparent lapses produced by history-dependent initial points are a major driver of this effect in the dataset.

Further, we use Bayes Information Criterion (BIC) to evaluate which of our model variants is the most parsimonious description for each rat in the dataset (Fig 3.4C-D). Most rats were best fit either by the motor error/exploration (72/152 rats) or the inattention (75/152 rats) models. For a small number of rats (5/152), the BIC score recommended that history modulation of neither the initial points nor true lapses is warranted (Fig 3.4C). Next we examined which model best captured the choices in the population as a whole, so we computed the mean BIC across rats (Fig 3.4D). Unsurprisingly, models with history modulations had lower mean BIC than the model without. Between the motor error/exploration and inattention models, both models had comparable mean BIC scores, although the motor error model won out marginally (Mean BIC score for motor error/exploration: 33725.64, inattention: 33726.25).

Therefore, in most rats, apparent lapses caused by history-biased initial points are the predominant cause of history-modulations of lapse rates and in about half the population, might be the only one. Altogether, these results support the hypothesis that history-biased initial

points contribute to (apparent) lapses in this dataset, which sometimes form up to about 60% of the total lapse rate (Fig 3.2E).

3.2.5 Reaction times support initial-point updating

In the accumulation-to-bound framework the time it takes for the drifting particle to hit the bound determines how long a subject deliberates before committing to a decision, therefore in addition to choices this framework makes clear predictions about their reaction times (RTs). So far, we have gathered evidence that the choice patterns that emerge from history-modulated initial points can describe rats' choices, next we examine if this mechanism can also account for the time it takes to make those choices.

We trained rats ($n = 6$) on a new variant of the auditory evidence accumulation task, modified to include reaction time reports (Fig 3.5A). In this task, stimulus is played as long as the rat subject maintains their nose in the center port and stops immediately when the subject withdraws the nose. The subject has to correctly report which speaker's auditory click train is sampled from a higher Poisson rate to receive a water reward. Rats perform this task with high accuracy (Fig 3.5B, left, average accuracy: 0.75 ± 0.02, number of trials 37205 ± 14247) albeit with dependence on trial-history despite it being detrimental to their performance (Fig 3.5B, right). Modulations of psychometric parameters by trial history (Fig 3.5C) resemble that of non-reaction time task (Fig 3.2C) i.e. following left and right wins, bias and lapse parameters are significantly different across the population of rats while sensitivity is not affected ($p = 0.69$ for sensitivity, 0.004 for threshold, 0.02 for left lapse rate, 0.02 for right lapse rate, calculated using Mann-Whitney U-test).

Reaction times (RTs) of these subjects display several signatures consistent with history-

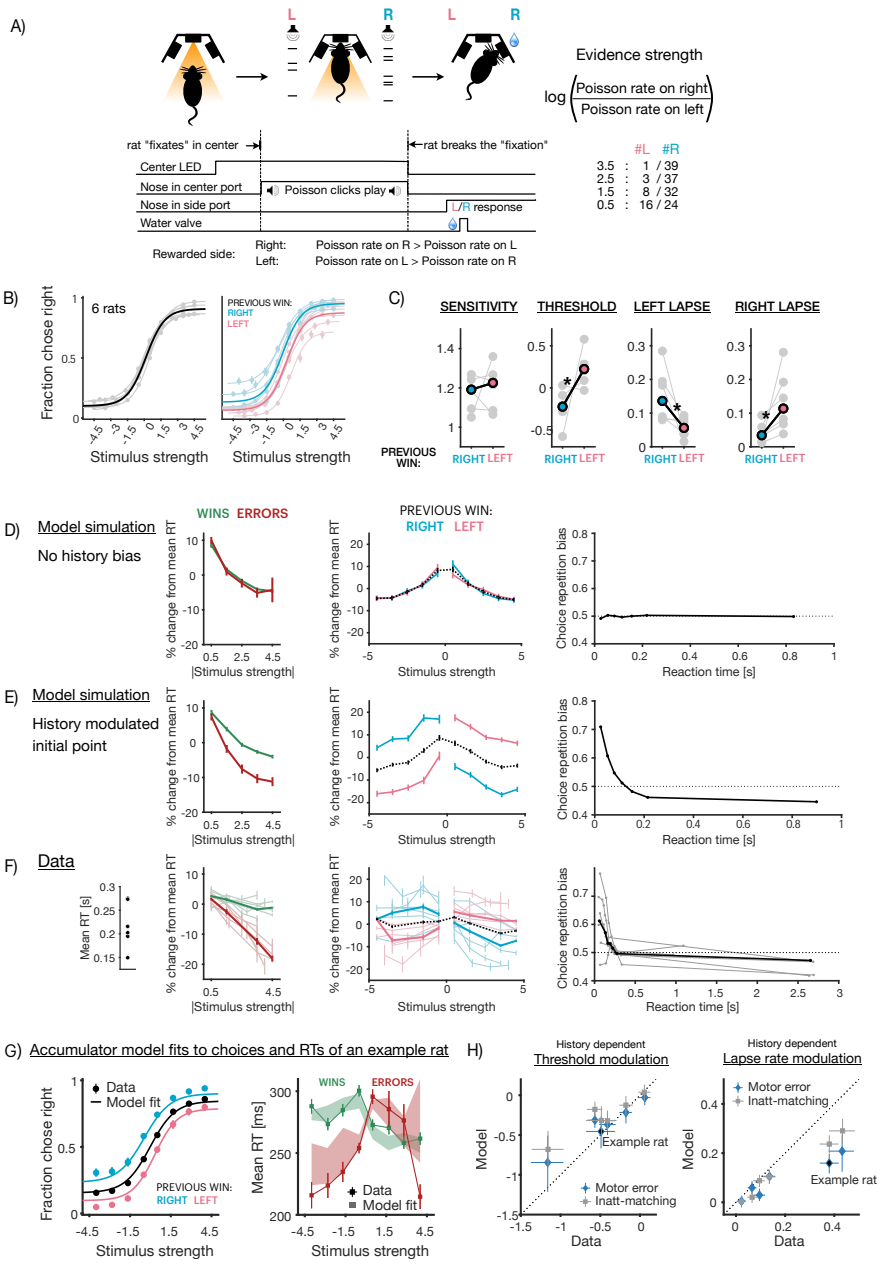


Figure 3.5: Reaction time predictions are borne out in data (A) Schematic of reaction time task in rats, with similar structure to (Fig 3.2A), with two modifications: rats are allowed to break "fixation" anytime during the trial and make a choice, and are rewarded for choosing the side with the higher Poisson rate, encouraging longer sampling for more accurate estimates.

Figure 3.5: (Previous page.) **(B)** Average choice behavior on all trials (left) and following previous right or left wins (right) of 6 rats on the reaction time task (solid line), overlaid on individual rat behavior (translucent lines) Caption continued on next page. **(C)** Average parameters (solid points) of history-conditioned psychometric curves, overlaid on individual parameters (translucent points) showing significant history modulations in threshold and lapse rate parameters ($p<0.05$, Mann-Whitney U-test) **(D-F)** Reaction time signatures D) expected from models with no history bias, E) expected from models with history modulated initial points and F) observed in data. First, error reaction times are expected to be shorter if initial points are history modulated, as seen in data (Left column, red curves are below green curves in E,F). Second, post right win reaction times are expected to be lower on rightward stimuli, and similarly following left wins (Middle column, blue (pink) curves on the right (left) are lower than dotted lines in E,F). Finally, repetition biases in choices are expected to occur more frequently for short reaction times, when the effect of initial points is strong (Right column, curves are above dotted line for smaller RTs in E,F). **(G)** Joint fits of accumulator model (with motor-error true lapses) to choices (left) and reaction times (right) of an example rat show good correspondence to data. Data represented by points (circles: choices, squares: reaction times) and model fits represented by lines (choices) or shaded bars (reaction times, thickness represents 95% bootstrap prediction intervals). Reaction times (right) are split by wins (green) or errors (red). **(H)** Scatter plot showing correspondence between history modulations in threshold (left) or lapse rate (right) parameters derived from data (x-axis) and model fits (y-axis). Individual points represent rats, with colors indicating fits from motor error (blue) or inattention-with-matching models. (grey).

modulation of initial points (Fig 3.5D-F). First, trial-to-trial variability in initial state of the accumulator is expected to give rise to shorter RTs on error trials compared to correct trials (Fig 3.5E, left) (Ratcliff and Rouder, 1998). This is unlike a standard DDM with no trial-to-trial variability in parameters, where RTs for correct and error trials are of similar magnitudes (Fig 3.5D, left). Indeed in the rat dataset, error RTs are consistently shorter than correct RTs across rats (Fig 3.5F, left). Interestingly, error RTs have been often reported to be longer than correct RTs in monkeys and humans (Roitman and Shadlen, 2002; Shevinsky and Reinagel, 2019), which has been attributed to either presence of collapsing bounds or trial-to-trial variability in drift rate (Ditterich, 2006a;b; Drugowitsch et al., 2012; Nguyen and Reinagel, 2020). Second, history modulations in initial points are expected to produced shorter RTs when the current stimulus matches the previous stimulus following rewarded choices (Fig 3.5E, middle; Yu and Cohen 2009; Goldfarb et al. 2012). We find that this signature is also present in the dataset across rats (Fig 3.5F, middle). Finally, variability in the

initial point is expected to be most influential early in the decision process, largely affecting the leading edge of the RT distribution, corresponding to a precipitous decrease in repetition bias as a function of RT (Fig 3.5E, right; Urai et al. 2019). Indeed, the data displays this pattern, with repetition bias being most prominent for short RTs, disappearing and turning into a weak alternation bias for long RTs (Fig 3.5F, right). Taken together, these three signatures offer promising evidence for the prevalence of history dependent initial point modulations in the dataset.

We further test this hypothesis by performing simultaneous trial-by-trial fitting of choices and RTs to an accumulator model with history-dependent initial points (see Methods). We fit this data with both motor error/exploration and a flexible variant of the inattention model in which on true lapse trials subject chooses stochastically, with a probability given by a sigmoidal function of the initial point. Such a strategy is often called probability matching and although suboptimal, has found empirical support in many perceptual tasks (Mamassian et al., 2002). We do not enforce an exact matching but instead allow this model to flexibly under- or over-match the bias due to initial point, depending on the slope parameter of the sigmoid which is fit to the data. When this slope parameter goes to infinity, this model picks deterministically based on the initial point, similar to the previous inattention model. On the other hand when the slope goes to zero, choices on true lapse trials are no longer dependent on initial points, reducing to the motor error/exploration model, with intermediate parameters interpolating between these two extremes. We refer to this as inattention-matching model. We find that the joint patterns of choices, reaction times and their history modulations in the data are captured by both models (Fig 3.5G - example rat showing fits from motor-error/inattention model, Supp Fig. 3.4 - all rats). The motor error model ac-

counts for substantial variance in history-dependent threshold and lapse rate modulations (Fig 3.5H). The inattention model shows a slight improvement upon this and is the better model overall based on BIC (Supp Fig. 3.5), with some residual variance not captured by either of the models. Nonetheless, accounting for history-dependent updating of initial points can help render a sizable fraction of decisions - that would have been otherwise attributed to lapses - predictable on a trial-by-trial basis.

3.3 Discussion

History biases and lapses have both long been known to plague decision-making studies across species, however they have largely been assumed to be distinct from each other, despite their frequent co-occurence and co-modulation. Here, we propose that normative accumulation under misbeliefs of non-stationarity can produce both history biases and apparent lapses, offering an explanatory link between the two suboptimalities. We show that history-dependent updates to the initial point of accumulation produce choices with varying biases in psychometric thresholds as well as varying sensitivities to evidence, yielding apparent, history-modulated lapse rates when choices are averaged across trials (Fig 3.1). We then analyzed a large rat dataset that had history modulations in both psychometric threshold and lapse rate parameters (Fig 3.2). We fit this dataset with an accumulator model that included apparent lapses due to history-dependent initial points, as well as true lapses (Fig 3.3), and found that history-dependent initial points accounted for a substantial amount of variance in both threshold and lapse rate modulations in the dataset, capturing a wide range of regimes in the dataset (Fig 3.4). Finally, we tested the reaction time predictions of the model on a novel task in rats, and confirmed that the dataset showed signatures of initial point updating, and the

model could successfully capture choices, reaction times and history modulations in psychometric thresholds and lapse rates (Fig 3.5). Altogether, our results suggest that history biases and a substantial amount of variance attributed to lapses may reflect a common mechanistic process, whose evolution can be precisely tracked both within and across trials.

History biases in perceptual decision making tasks have been modeled using initial point updates to DDMs in humans and non-human primates (Gold et al., 2008; Goldfarb et al., 2012; Zhang et al., 2014). These studies tended to have relatively small magnitudes of history bias, and miniscule lapse rates, hence being well captured by small deviations in the initial point of a DDM, which largely yield horizontal shifts in the psychometric function. This regime of initial point updates is well approximated by a logistic function with additive biases, which is the dominant descriptive model used to characterize history-dependent psychometric curves (Busse et al., 2011; Carandini and Churchland, 2013; Fründ et al., 2014; Abrahamyan et al., 2016; Gardner, 2019; Pinto et al., 2018; Odoemene et al., 2018; Urai et al., 2019; Hermoso-Mendizabal et al., 2020; Roy et al., 2021; Ashwood et al., 2022; Bolkan et al., 2022). However, as we demonstrate, when deviations in the initial point are large, this logistic approximation breaks down. This fact has been overlooked in much of the literature and hence even in datasets with large history biases and lapses the logistic formulation continues to be favored (Odoemene et al., 2018; The International Brain Laboratory et al., 2021; Roy et al., 2021; Ashwood et al., 2022), albeit requiring additional components. Our demonstration predicts that the full range of initial point effects should resemble concurrent, trial-by-trial changes in both threshold and sensitivity parameters of the logistic function. Indeed, (Ashwood et al., 2022) found that apparent lapses in several rodent datasets can be better captured by runs of trials with such concurrent modulations, yielding biased "disengaged"

states.

In our treatment, we only considered history-dependent updates to the initial point of a DDM. Such a mechanism is normative under non-stationary beliefs about the prior [‡], which is our favored interpretation as it aligns with other studies of history biases (Gold et al., 2008; Yu and Cohen, 2009; Summerfield and Koechlin, 2010; Goldfarb et al., 2012; Mulder et al., 2012; Abrahamyan et al., 2016; Molano-Mazon et al., 2021). Nevertheless, these updates may also reflect other heuristic strategies (Gigerenzer and Gaissmaier, 2011) which we accommodate using our flexible parameterization of initial point updates. Animals may entertain non-stationary beliefs about other elements of the decision process, such as the rewards or likelihoods (Dayan and Daw, 2008; Mendonça et al., 2020; Lak et al., 2020a; Pisupati et al., 2021). Normative updating in such situations still reduces to initial point updates in simple settings (for e.g. non-stationary rewards for a single difficulty) (Simen et al., 2009; Rorie et al., 2010), but in more complex ones it *additionally* affects drift rates (Palmer et al., 2005; Eckhoff et al., 2008; Hanks et al., 2011; Drugowitsch et al., 2019; Fan et al., 2018; Urai et al., 2019; Mendonça et al., 2020). While we don't see reaction time signatures of drift rate variability in our dataset (error RTs are shorter than correct RTs, instead of being longer), such updates may be another potential mechanism by which history-modulated apparent lapses could arise.

Lapse rates are often considered to be a mixed bag comprising several different noise processes, yet most studies so far have focused on one or more of these component processes in isolation (Pisupati et al., 2021; Ashwood et al., 2022). In this work, we have attempted a more expansive approach of considering multiple processes at once, in an attempt to partition lapse

[‡]Note that this is the case if the agent assumes that a shift in the prior over categories maps onto an overall shift in the prior over difficulties - see (Drugowitsch et al., 2019) for a detailed treatment

rate variance into mixtures of deterministic and stochastic components. We distinguished apparent lapses that interact with sensory evidence from three models of "true" lapses that are evidence independent - inattention with maximizing policies, inattention with matching policies and epsilon-greedy exploration (which is identical to motor error). We find that the behavior of our rats is best described by a mixture of apparent lapses and the latter two true lapse variants, with about equal evidence for each. A previous study proposed an evidence-dependent model of true lapses – uncertainty-guided exploration (Pisupati et al., 2021), in order to account for the scaling of lapse rates with sensory noise. While we don't explicitly consider this model, our model of apparent lapses already displays this property, with higher levels of sensory noise leading to more frequent apparent lapses.

Our model predicts that an increased reliance on history (i.e. bigger initial points) should produce more apparent lapses, all other parameters being held constant. Indeed, this could provide an explanation that links disparate sets of observations from previous studies: while some studies have reported that perturbations of secondary motor cortex and striatum give rise to higher lapse rates (Erlich et al., 2015; Yartsev et al., 2018; Guo et al., 2019; Sindreu et al., 2021; Pisupati et al., 2021), other have shown that the effects of perturbing these regions seems to resemble an increased history-dependence (Sindreu et al., 2021; Luo et al., 2021). Interpreting these results through the lens of our model, we'd conclude that these regions play a crucial role in the interaction of history-dependent initial points with sensory evidence. Indeed, Luo et al. 2021 find that this increased history dependence upon perturbation is mediated by increased bias in initial value of the neurally derived accumulator variable. Our model could also help explain why Busse et al. 2011 found that mice with higher lapse probabilities showed higher history dependence. Or results from The International Brain

Laboratory et al. 2021 who observed a modulation in lapse rates in addition to horizontal biases upon explicit manipulation of category priors.

One interesting future line of investigation is to probe the precise nature of the model of non-stationarity over priors assumed by animals in such tasks. The range of values inferred by our flexible parameterization could offer a useful starting point for this line of investigation - for instance, Dynamic Belief Models (Yu and Cohen, 2009; Ryali et al., 2018), a popular class of generative models over priors, correspond to a narrowly constrained set of parameter values in our model. Such an understanding would not only afford more reliable control of behavior and more accurate interpretation of neural correlates in stationary tasks, but could also yield insight into the inductive biases that allow animals to learn quickly and efficiently in non-stationary, naturalistic settings.

3.4 Methods

3.4.1 Subjects

Animal use procedures were approved by the Princeton University Institutional Animal Care and Use Committee (IACUC #1853). All subjects were adult male Long Evans rats, typically housed in pairs. Rats that trained during the day were housed in a reverse light cycle room. Rats had free access to food but in order to to motivate them to work for water reward, they were placed on a controlled water schedule: 2-4 hours per day during task training, usually 7 days a week and between 0 and 1 hour ad lib following training.

3.4.2 Drift diffusion model of decision-making

We use a standard formulation of sequential decision-making based on (Bogacz et al., 2006; Drugowitsch et al., 2012), in which an agent is faced with a stream of noisy sensory evidence $\epsilon_{1:t}$ coming from one of two hypotheses H_1 and H_2. The agent has to decide between sampling for longer or choosing one of two actions L,R (reaction time regime) or has to choose one of two actions after a fixed amount of evidence (fixed duration regime). Such a problem can be formulated as one of finding an optimal policy π_t in a partially-observable markov decision process (Rao, 2010; Drugowitsch et al., 2012), whose solution can be written as a pair of thresholds on the log-posterior ratio $\log(\frac{g(t)}{1-g(t)})$, where $g(t) = p(H_1|\epsilon_{1:t})$:

$$\pi_t = \begin{cases} \text{choose L}, & -B \geq \log(\frac{g(t)}{1-g(t)}) \\ \text{sample}, & -B < \log(\frac{g(t)}{1-g(t)}) < B \\ \text{choose R}, & \log(\frac{g(t)}{1-g(t)}) \geq B \end{cases}$$

The log posterior ratio can be further broken down into a sum of log prior ratios and log likelihood ratios, using Bayes rule:

$$\log \frac{p(H_1|\epsilon_{1:t})}{p(H_2|\epsilon_{1:t})} = \log \frac{p(H_1)}{p(H_2)} + \log \frac{p(\epsilon_{1:t}|H_1)}{p(\epsilon_{1:t}|H_2)}$$

The optimal policy can equivalently be expressed in terms of the prior and sum of momentary sensory evidence $x(t) = \sum_t \epsilon_t$, which are sufficient statistics of the posterior (Drugowitsch et al., 2012; Piet et al., 2018). In the continuous time limit, when the average rate of evidence increments or drift rate is μ, and the standard deviation of sensory noise is σ, this corresponds

to a drift diffusion model, whose initial point I is proportional to the log prior ratio, and that terminates when it reaches one of two bounds (Bogacz et al., 2006):

$$dx = \mu dt + \sigma dW, \quad x(0) = I = k \cdot \log \frac{p(H_1)}{p(H_2)}$$

In this case, the probability of choosing rightward actions, i.e. hitting the upper bound can be written analytically as follows (derived from Palmer et al. 2005):

$$P(B+) = \frac{1 - e^{-2\mu(B+I)/\sigma^2}}{1 - e^{-4\mu B/\sigma^2}}$$

In cases where trial difficulties (and hence drift rates) vary from trial to trial the optimal policy includes time-dependent, collapsing bounds on the posterior. However, under certain circumstances, constant bounds on $X_t = \sum_t \epsilon_t$ implement close-to-optimal collapsing bounds on the posterior (Denève, 2012; Drugowitsch et al., 2012), which is the regime we assume for our analysis.

3.4.3 Models of initial point updating

We model initial point updating as a sum of exponential filters over past choice-outcome pairs (RC: right-correct, LC: left-correct, RE: right-error, LE: left-error). So the initial point I at trial $n+1$ is given by:

$$I(n+1) = i^{RC}(n+1) + i^{LC}(n+1) + i^{RE}(n+1) + i^{LE}(n+1)$$

where each filter i^h decays by a factor of β^h, and is incremented by a factor of η^h following the observation of that particular choice-outcome pair, i.e

$$i^h(n+1) = \eta^h 1^h(o_n) + \beta^h i^h(n) \quad \text{where} \quad h = \{RC, LC, RE, LE\}$$

o_n is the choice-outcome pair observed on trial n and $1^h(o_n)$ is an indicator function that is 1 when $o_n = h$ and is 0 otherwise.

When β^h and η^h are the same $\forall h$, this rule reduces to an approximation of the Bayesian update for the Dynamic Belief Model (Yu and Cohen, 2009), which tracks a prior that undergoes discrete unsignaled switches at a fixed rate. We compared this unconstrained model to models with various constraints on the decay and magnitude parameters (same parameters for corrects v.s. errors, left v.s. right etc). While model comparison revealed that not every rat required all parameters to be different, the unconstrained model is the most general form that best captures behavior across rats.

3.4.4 Psychometric curves

Psychometric curves model the probability of a subject choosing one of the options (e.g. right) as a function of stimulus strength. We parametrize the psychometric curve as a 4-parameter logistic function:

$$P(\text{choose Right}) = \kappa_0 + \frac{\kappa_1}{1 + e^{-b(x - x_0)}}$$

where x_0 is the threshold parameter that additively biases the stimulus x, b measures sensitivity to the stimulus, κ_0 is the left asymptote or left lapse rate and κ_1 scales the logistic function.

Therefore, the right asymptote is given by $\kappa_0 + \kappa_1$ and the right lapse rate itself is given by $1 - (\kappa_0 + \kappa_1)$. We fit all four of these parameters $\{\kappa_0, \kappa_1, x_0, b\}$ to choices generated by either the DDM (Fig 3.1), rats (Fig 3.2, 3.4, 3.5), or accumulator models adapted to the tasks (Fig 3.4, 3.5) using a gradient-descent algorithm (interior-point) to maximize the (Binomial) log likelihood of choices using MATLAB's constrained optimization function *fmincon*. κ_0 and κ_1 were both constrained to lie within the interval $[0, 1]$. 95% confidence intervals on these parameters were generated using bootstrapping.

History modulation of psychometric parameters: To summarize the effects of trial history on psychometric parameters we fit independent psychometric curves to choices conditioned on 1-trial back choice-outcome history i.e. following rightward wins (RW) and leftward wins (LW). Modulation of the threshold parameter by history was then computed as $x_0^{RW} - x_0^{LW}$. To quantify the modulation of lapse rate parameter by history we first computed the difference in the left and right asymptotes following rightward and leftward wins: $\kappa_0^{RW} - \kappa_0^{LW}$ and $(\kappa_0^{RW} + \kappa_1^{RW}) - (\kappa_0^{LW} + \kappa_1^{LW})$ respectively. The net modulation of lapse rates with trial history is given by the sum of these differences: $2(\kappa_0^{RW} - \kappa_0^{LW}) + (\kappa_1^{RW} - \kappa_1^{LW})$.

3.4.5 Behavioral tasks

Auditory evidence accumulation task: Rats were trained with previously established protocol (Brunton et al., 2013; Hanks et al., 2015; Erlich et al., 2015; Yartsev et al., 2018) using the BControl system. Briefly, rats were put in an operant chamber with three nose ports. They were trained to begin a trial by poking their nose into the middle port. This initiated two simultaneous streams of randomly-timed discrete auditory clicks for a prede-

termined duration after a variable delay (0.5-1.3s), one from a speaker to their left and the other to their right. Rats were required to maintain "fixation" throughout the entire stimulus (1.5s), failure to do so led to a violation trial. At the end of the stimulus, rats had to poke towards the side which played the greater number of clicks to obtain a water reward. Stimulus difficulty was varied from trial-to-trial by changing the ratio of the generative Poisson rates of the two click streams. Trial difficulty and rewarded side were independently sampled on each trial.

We analyzed rats which performed greater than 30,000 trials, at 70% or more accuracy. Sessions with less than 300 trials or less than 60% accuracy for either of the choices were excluded. Since rats typically perform this task for many months after having passed the final training stage, to minimize nonstationarities in the data (due to break in training because of holiday closures etc.) and ensure that we are analyzing asymptotic performance, we identified temporally contiguous sessions with stable accuracy by performing change-point detection on smoothed trial hit rate using MATLAB's *findchangepts* function. The partition with most number of trials was included in the analysis. Since the animals neither made a choice nor received an outcome on violation trials, we ignore them while computing trial-history effects.

Auditory evidence accumulation task with reaction time reports To measure rats' reaction times in addition to choices we modified the auditory evidence accumulation task in two ways. First, we relaxed the "fixation" requirement and instead allowed rats to sample the stimulus for as long as they want. As soon as rats broke fixation by removing their nose from the center port, the stimulus stopped and the rats were required to report their decision by poking into one of the side ports. For any given trial, the time that the rat spent sampling the stimulus was its reaction time. Second, we rewarded rats if they correctly

reported the side which had greater underlying Poisson rate rather than the side which played the greater number of clicks. This helped eliminate the trivial strategy of culminating a decision after the first click and having perfect accuracy by simply reporting the side of that click without any need for evidence accumulation.

In practice, we followed the same training protocol as the interrogation task (Brunton et al., 2013) but with the modified reward rule. Once the rats were fully trained on the interrogation protocol we gradually reduced the duration of delay between stimulus onset and trial initiation as well as the fixation period. Most rats maintained high accuracy (>70%) upon this manipulation, if rats performance did not meet this criterion even after a week of training they were excluded. Rats tended to have worse accuracy early in the session, so we omitted the first 50 trials from our analysis. After the first 50 trials, we confirmed that the accuracy in the first and second halves of the session was comparable.

3.4.6 Data modeling methods

Accumulator model To models subjects choices and RTs, we used accumulation to bound model modified to take into account the discrete nature of evidence in our behavioral tasks (Brunton et al., 2013). In the model, on a behavioral trial n, the evolution of accumu-

lated evidence $a(t)$ in response to the left (δ_L) and right (δ_R) click trains is given by:

$$da = \begin{cases} 0, & \text{if } |a| \geq B \\ \lambda a dt + (\delta_{R,t} C_R(t) \eta_R - \delta_{L,T} C_L(t) \eta_L) dt + \sigma_a dW & \text{otherwise} \end{cases}$$

$$\text{where } \frac{dC}{dt} = \frac{1 - C}{\tau_\phi} + (\phi - 1) C(\delta_{R,t} + \delta_{L,t}) \quad \text{and}$$

$$a(t = 0) = I(n)$$

where λ is the inverse time constant of the consistent drift in memory of $a(t)$. $C_R(t)$ and $C_L(t)$ are the magnitudes of each right and left click respectively after undergoing sensory adaptation (with adaptation strength ϕ and adaptation time constant τ_ϕ). The sensory noise that accompanies each click is represented by η_R, η_L which are Gaussian random variables with mean 1 and variance σ_s^2. The accumulation variable a also undergoes Brownian diffusion through the addition of a Wiener process (W) with variance σ_a^2. B represents the absorbing decision bound that prevents $a(t)$ from evolving further, if crossed. The initial value of the accumulator variable a varies from trial-to-trial and is set based on exponentially filtered history of previous choices and outcomes (see Methods section on Models of initial point updating). A choice is made by comparing the final value of the accumulator $a(T)$ to a side bias. A rightward choice is made if $a(T) >$ bias.

Since the model quantifies noise sources on each trial, it requires estimating the evolution of a noise-induced probability distribution $P(a(t))$. We compute $P(a(t))$ by solving the Fokker-Planck equations that correspond to model dynamics (see (Brunton et al., 2013; DePasquale et al., 2021) for numerical methods). The probability of making a rightward choice

at the end time-point T of a trial, given accumulation model parameters θ^{acc} is:

$$P(\text{choose R}|\delta_R, \delta_L, \theta^{acc}) = \int_{a=\text{bias}}^{\infty} daP(a(T)|\delta_R, \delta_L, \theta^{acc})$$

MODELS OF TRUE LAPSES We assume that some fraction of choices κ arise from processes extraneous to evidence accumulation such as motor error/exploration, inattention with maximizing policy and inattention with probability matching. We parameterize these processes with θ^{lapse} and refer to them as "true lapses":

- In the motor error/exploration variant, the probability of making a choice towards the right - when lapsing - is given by ρ.

$$P(\text{choose R}|\theta^{lapse}) = \rho$$

- In the inattention with maximizing variant, the subject lapses towards the side favored by the initial point relative to a bias ρ. So the probability of a rightward choice due to inattention on trial n is:

$$P(\text{choose R}|\theta^{lapse}) = \begin{cases} 1 & \text{if } i(n) - \rho > 0 \\ 0.5 & \text{if } i(n) - \rho = 0 \\ 0 & \text{if } i(n) - \rho < 0 \end{cases}$$

- In the inattention with matching variant, the probability of lapsing towards right depends on the initial point through a sigmoidal function whose slope m (or matching

constant) as well as bias ρ is a free parameter:

$$P(\text{choose R}|\theta^{lapse}) = \frac{1}{1 + e^{-m(i(n)-\rho)}}$$

Hence the total probability of making a rightward choice due to accumulation and true lapses is:

$$P(\text{choose R}|\Theta) = (1-\kappa)P(\text{choose R}|\delta_R, \delta_L, \theta^{acc}) + \kappa P(\text{choose R}|\theta^{lapse})$$

where $\Theta = \{\theta^{acc}, \theta^{lapse}, \kappa\}$.

MODEL FITTING The model parameters were fit to individual rats by maximizing the log likelihood of the observed choices of the rat $\mathbf{c_{obs}}$, i.e. by maximizing

$$\ln \mathcal{L}(\mathbf{c_{obs}}|\delta_{\mathbf{R}}, \delta_{\mathbf{L}}, \Theta) = \Sigma_n \ln P(c_{obs,n}|\delta_{R,n}, \delta_{L,n}, \Theta)$$

where n indexes trials. Constrained optimization was performed in Julia using Optim package. We computed gradients for parameter optimization using a forward-mode automatic differentiation package. The reported maximum likelihood parameters and likelihood values (used for model comparison) are from model fits to the entire dataset. We fit a random subset of 10 rats using 5-fold cross-validation (85% training dataset, 15% test dataset) but this yielded very similar maximum likelihood parameters and virtually identical test and training log-likelihoods. Hence, to save on computing time we fit the different model variants to each rat's entire dataset. This agreement between test and training likelihoods is likely due to the

large number of trials in our dataset and the modest number of parameters in our model.

Simultaneous modeling of choices and RTs In decision-making tasks, observed reaction times (RTs) are often thought of as comprised of stimulus sampling or decision times (DTs, the time it takes for the subject's accumulated evidence to hit the bound) and non-decision related processing times (NDTs). In our datasets we observed that reaction times tended to be slower following incorrect trials and that they grew longer over the course of a session. These effects could be isolated just to RTs and were not observed in choice behavior. To model these trends we conceptualize non-decision times as arising from a separate drift diffusion process whose drift ν is additionally modulated by current trial number n and previous trial's outcome. These non-decision time drift-diffusion processes terminate when the bound ω is hit. We assume that the non-decision times for each choice $k \in \{L, R\}$ have independent bounds (ω_k) and drifts (ν_k). So the non-decision times for a trial n are samples from the following Wald or Inverse Gaussian (IG) distribution:

$$\tau_n^{NDT} \sim IG\left(\frac{\omega_k}{\nu_k - \alpha n + \gamma_o 1^-_{(n-1)}}, \omega_k^2\right)$$

where $k \in \{L, R\}$ and $1^-_{(n-1)}$ is an indicator function which is 1 if the previous trial was incorrect and is 0 otherwise. α parameterizes the impact of trial number on NDTs and γ_o parameterizes the impact of previous trial's outcome on current trial's NDT.

We fit the model by maximizing the joint log likelihood of the observed choices and RTs. For any given trial, we can compute the likelihood of observing a particular reaction time RT_{obs} and choice c_{obs} due to accumulation by marginalizing over possible decision or bound

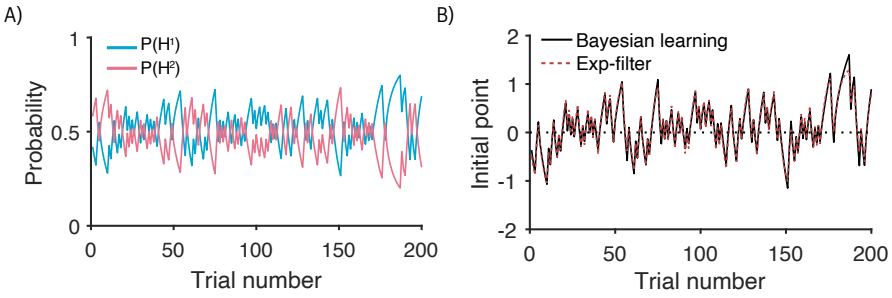
hitting times $\tau_{c_{obs}}$ for the observed choice:

$$P(c_{obs}, RT_{obs}|\delta_R, \delta_L, \theta^{acc}, \theta^{NDT}) = \int_0^{RT_{obs}} P(\tau_{c_{obs}}|\delta_R, \delta_L, \theta^{acc}) P(c_{obs}, RT_{obs}|\theta^{NDT}, \tau_{c_{obs}}) d\tau_{c_{obs}}$$

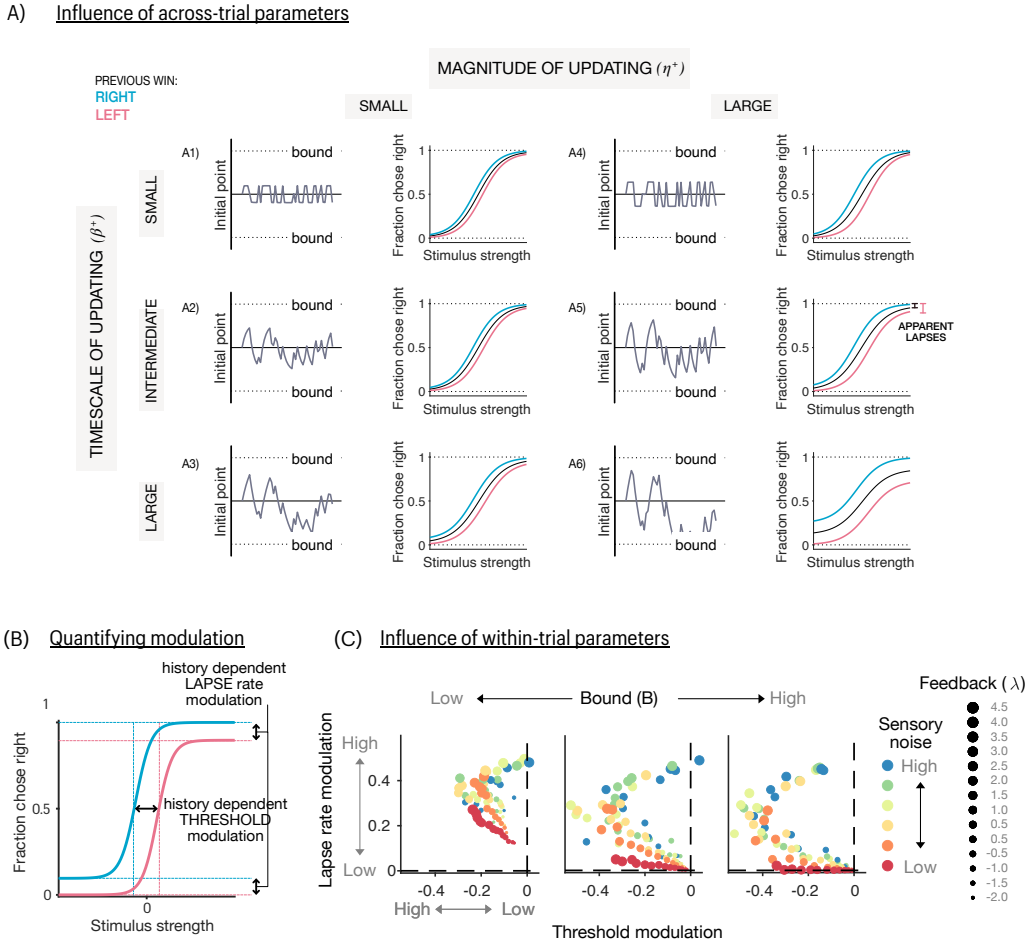
On true lapse trials, RTs were assumed to arise from NDTs alone and therefore the joint likelihood due to accumulation and true lapses is given by:

$$\mathcal{L}(c_{obs}, RT_{obs}|\delta_R, \delta_L, \Theta) = (1-\kappa)P(c_{obs}, RT_{obs}|\delta_R, \delta_L, \theta^{acc}, \theta^{NDT}) + \kappa P(c_{obs}, RT_{obs}|\theta^{lapse}, \theta^{NDT})$$

where $\Theta = \{\theta^{acc}, \theta^{NDT}, \theta^{lapse}, \kappa\}$.

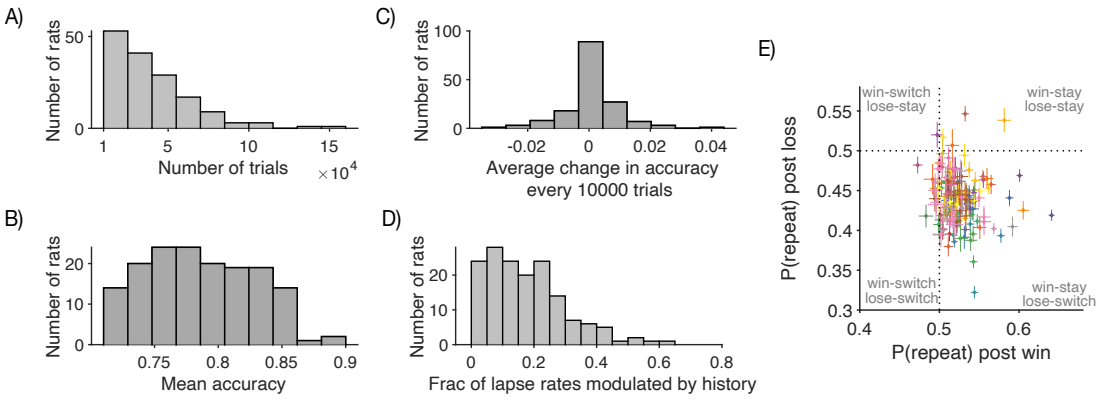


Supplementary Figure 3.1: Exponential filtering for initial point setting approximates Bayesian prior updates under assumptions of non-stationarity A: Example of a mis-belief in a non-stationary prior. Traces represent belief about prior probability of two hypotheses H^1 and H^2 inferred from a random sequence of trials drawn from a stationary symmetric prior, under the misbelief that the prior is occasionally undergoing unsignalled jumps. Such an assumed generative model is often referred to as the Dynamic Belief Model (DBM; Yu and Cohen 2009). **B**: Initial point updates corresponding exactly to the fluctuating prior beliefs in (A) that emerge from Bayesian learning (black line), plotted against approximate initial points derived from exponential filtering (dotted red line) of past choices and outcomes. The exponential filter provides a good approximation of exact Bayesian updates, while being more expressive and flexible to capture the possibility of other generative models and corresponding update rules.

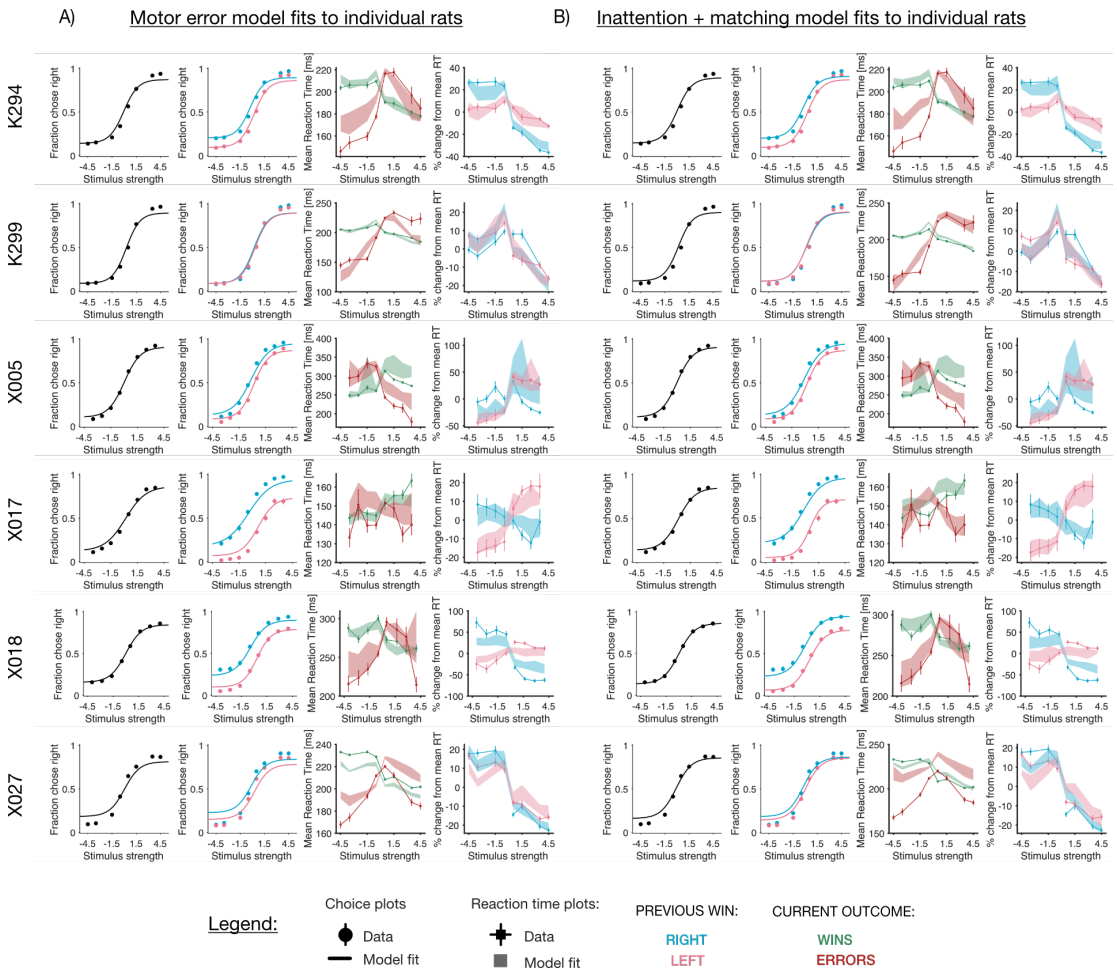


Supplementary Figure 3.2: Influence of within- and across-trial parameters on history-dependent modulation of biases and lapses in the psychometric function. (A) Influence of across-trial parameters on history-dependent modulation: Effects of varying magnitude of trial-by-trial updating (η - columns) and timescale of updating (β - rows) on initial point trajectories (gray lines) and psychometric curves (black - conditioned on all previous wins, blue - previous right wins, pink -previous left wins). (Top row) Small timescales of updating lead to fast fluctuations in initial points, and mostly horizontal shifts in psychometric curves with trial history, for both small and large magnitudes of updating (A1 and A4). (Middle row) Intermediate timescales of updating lead to slower fluctuations in initial point that have a cumulative effect across trials. For large magnitudes of updating (A5) these can give rise to apparent lapses (black intervals) as well as history-dependent modulation of these lapses (pink intervals). (Bottom row) Long timescales of updating lead to stronger cumulative initial point biases across trials, yielding apparent lapses and lapse rate modulations even for small magnitudes (A3). When combined with large magnitudes of updating (A6) lead to initial points that sometimes exceed the bounds, leading to a combination of apparent lapses (initial points within bounds) and deterministic, stimulus-independent decisions (initial points outside bounds). Caption continued on next page.

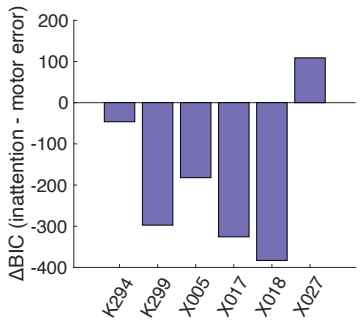
Supplementary Figure 3.2: (Previous page.) **(B)** Quantifying modulation of psychometric properties: The difference between psychometric curves conditioned on previous wins (blue) or losses (pink) can be quantified using two metrics - the horizontal distance between the midpoints of psychometric curves ("threshold modulation") and the vertical distance between its asymptotes ("lapse rate modulation") **(C)** Effects of varying the parameters of the within-trial drift diffusion model (DDM) on history-dependent threshold (x-axis) and lapse rate modulations (y-axis). Colors denote levels of sensory noise, size of dots denote values of the feedback parameter of the DDM. The feedback parameter determines if the accumulation is leaky ($\lambda < 0$, ignores early evidence), perfect ($\lambda = 0$, uses all evidence) or impulsive ($\lambda > 0$, ignores later evidence). Plots from left to right are for low, intermediate and high values of bound respectively. High bounds predominantly give rise to threshold modulations, however high positive values of feedback and higher levels of sensory noise additionally produce lapse rate modulations. Lapse rate modulations are dramatically increased by lower bounds for many different values of feedback and by higher values of sensory noise.



Supplementary Figure 3.3: Performance measures across the rat dataset (A) Histogram of trial counts for all rats in the population. Most rats completed on the order of 1e4 trials. **(B)** Histogram of mean accuracy showing that rats showed good performance on the task (mean accuracy $\pm$ std dev: 0.79 ± 0.04). **(C)** Average change in mean accuracy every 1e4 trials, showing that rats' performance was stable over time, reflecting asymptotic behavior rather than task acquisition. **(D)** Histogram of history-modulated lapse rates as a fraction of total lapse rates. A sizeable portion of the population had non-zero fractions, suggesting that history-dependence could potentially account for some lapse rate variance. **(E)** Scatter comparing repetition bias following wins and losses. Each point is a rat, error bars are Wilson binomial CIs. Most of the population occupied the bottom right quadrant, showing a "win-stay, lose-switch" bias i.e. repetitions following wins and alternations following losses.



Supplementary Figure 3.4: Model fits to choices and reaction times of individual rats A Fits of an accumulator model with history-dependent initial points, and additional true lapses arising from motor error. Each row shows fits to an individual rat, and each column depicts a different behavioral measure summarizing choices (first column, psychometric curve), reaction times (third column, win RTs in green and error RTs in red), and history modulations in choices/reaction time (second/fourth column, psychometric curves/RTs conditioned on previous right wins (blue) or left wins (pink)). Data represented by points (circles: choices, squares: reaction times) and model fits represented by lines (choices) or shaded bars (reaction times, thickness represents 95% bootstrap prediction intervals). **B** Same rats as in (A), with fits of an accumulator model with history-dependent initial points and true lapses arising from inattention with matching.



Supplementary Figure 3.5: Individual model comparison in the RT dataset Difference in BIC scores between the inattention and motor error models, for individual rats in the reaction time dataset. Negative values indicate that inattention was the winning model, as is the case for most rats.

This work was done in collaboration with Charles Kopec, Adrian Bondy, Thomas Luo and Verity Elliott.

4

Distributed evidence accumulation in frontal cortico-striatal circuitry

4.1 Relevant background

In our daily lives, we constantly have to makes decisions based on noisy stimulus. One example of this, which has overshadowed the lives of most during the COVID-19 pandemic is deciding whether or not it is safe to attend public gatherings in indoor spaces. This decision depends on many factors, what are the case numbers like? are folks vaccinated? are they going to be masking? Each of these individual pieces of evidence only partially informs the decision but accumulating evidence from many factors helps reduce overall uncertainty.

Models which formalize the temporal dynamics that underlie such decisions based on

noisy or ambiguous evidence are often referred to as accumulation-to-bound models. These models have strong links to the Bayesian decision theory and help decompose the decision-making process into its computational components. According to these models, a subject forms a decision by gradually accumulating the evidence for or against the possible choices - this accumulation over time helps the subject average over different noisy samples and improve their estimate of the underlying signal. The gradual evidence accumulation continues until the integrated value reaches a predetermined threshold (or bound), one for each category or until the stimulus stream expires. The choice favored by the final value of the accumulated evidence then drives the decision. Over the past few decades decision-making studies in humans, non-human primates, and rodents have revealed that decisions in a wide range of decision-making domains including perceptual, value-based, memory-based, social, among others, are well described by the accumulation-to-bound models.

Within perceptual decision-making, studies often probe this fundamental process by asking subjects to infer and pick which of the two categories an ambiguous stimulus is drawn from. A seminal example of such a paradigm is the random dots motion (RDM) task. In this task, subjects decide between two possible, opposite directions of motion based on a stimulus that consists of a field of dots. On any given trial, some fraction of the dots move coherently in one direction and this fraction is manipulated from trial-to-trial to vary the decision difficulty. In this and similar paradigms, the accumulation-to-bound models predict that the decision accuracy should improve as a function of stimulus duration (roughly with $\sqrt{t}$), even as the average evidence available per unit time stays constant. These models also predict that the subjects should take longer to make decisions based on weak evidence (or on hard trials) than strong evidence (or on easy trials), since on strong evidence trials the accumulated value

quickly ramps up and hits the bounds, driving the decision. Both these behavioral signatures of evidence accumulation have been observed widely in many perceptual studies. Importantly, detailed fitting of the accumulation model to behavior and comparison with other decision-making strategies has further corroborated that accumulation-to-bound models are useful descriptors of the decision-making process.

The search for neural basis of evidence accumulation has benefited from conceptualization of the decision-making process as a feedforward chain of computations that together constitute an accumulator model, i.e. gradual accumulation of momentary sensory evidence followed its thresholding towards the end of the decision process to make a categorical choice (Schall, 2001; Gold and Shadlen, 2007). Among the pioneering studies in this line of research were experiments in rhesus monkeys performing the RDM task. Early set of electrophysiological recordings had reported that middle temporal (MT) and medial superior temporal (MST) areas of the visual cortex encode the momentary sensory evidence about motion direction of the dot stimulus during the RDM task. Downstream of these sensory areas, trial-averaged firing rates of select neurons in the lateral intraparietal (LIP) subregion of the posterior parietal cortex (PPC) buildup in a manner that resemble the ramping temporal dynamics of an accumulator model (Shadlen and Newsome, 1996; 2001; Roitman and Shadlen, 2002). On easy trials, the firing rates ramp rapidly whereas on hard trials they increase rather slowly, resembling the movement of accumulated evidence in an accumulator model on high and low coherence trials respectively. In further resemblance, in versions of the task where monkeys control when to make a decision, the trial-averaged firing rates reach a common level prior to the choice irrespective of the trial difficulty as if the decision is being

triggered by hitting of an evidence bound (Roitman and Shadlen, 2002). These were among the first neural recordings that found compelling evidence for a straightforward one-to-one mapping of the behavioral accumulator model onto the neural substrate (Shadlen and Kiani, 2013), spurring much interest in LIP and its putative rodent analogue PPC.

Further work probing the representations, dynamics, and causal contributions of LIP and its rodent analogue using modern statistical and experimental methods, has painted a more complicated picture of its role in evidence accumulation (Brody and Hanks, 2016; Huk et al., 2017). Ramping activity in LIP has been traditionally interpreted as reflecting the accumulation of motion evidence from MT. However, statistical analysis of single-trial coupling between simultaneous activity from the two areas suggests that ramping activity in LIP can not be sufficiently captured by the accumulation of motion evidence from MT alone, and requires the addition of a ramping, choice-related regressor (Yates et al., 2017). To further complicate matters, there have been contradictory reports about whether LIP neurons show ramping dynamics on single trials at all, with some studies showing that noise properties of LIP population responses show signatures consistent with accumulation dynamics (Steinemann et al., 2022; Stine et al., 2022), and other studies finding that neural responses are better described by discrete "stepping" dynamics (Latimer et al., 2015; Zoltowski et al., 2019). Finally, studies attempting to study the causal requirement of LIP during evidence accumulation have reported no decision-making impairment (Katz et al., 2016), little impairment (Zhou and Freedman, 2019) or rapid compensation (Jeurissen et al., 2022) in response to reversible inactivation of LIP neurons. Mirroring these sets of observations, even though the rodent PPC has been shown to carry graded representation of accumulated evidence during the Poisson Clicks task and closely related task variants (Hanks et al., 2015; Scott et al., 2017;

Morcos and Harvey, 2016) its inactivations have failed to produce impairments in evidence accumulation (Erlich et al., 2015; Licata et al., 2017).

The fact that animals' decision making stays intact even when activity in these parietal areas is silenced, strongly suggests that LIP/PPC participate in evidence accumulation as part of a wider distributed network, if at all. Indeed, other fronto-cortical and subcortical areas have been identified in both monkeys (Horwitz and Newsome, 1999; Kim and Shadlen, 1999; Ding and Gold, 2010; Purcell et al., 2010; Ding and Gold, 2012; 2013; Mante et al., 2013; Peixoto et al., 2021; Jun et al., 2021; Stine et al., 2022) and rodents (Znamenskiy and Zador, 2013; Erlich et al., 2015; Hanks et al., 2015; Scott et al., 2017; Yartsev et al., 2018; Pinto et al., 2019; Deverett et al., 2019; Orsolic et al., 2021; Pisupati et al., 2021; Koay et al., 2020; Musall et al., 2021; Bolkan et al., 2022) that show signatures consistent with the accumulator model and causally impair the decision-making process.

In this chapter, we review the puzzling evidence about the roles of two such promising areas in rats, the Frontal Orienting Fields (FOF, putative analogue of monkey FEF) and the anterior dorsal striatum (ADS, analogue of monkey caudate) and use projection-specific inactivations, simultaneous recording and modeling to further our understanding of neural basis of evidence accumulation.

4.2 Introduction

Decision-making based on noisy perceptual evidence is a core cognitive process that is often conceptualized as a sequence of two sub-computations - gradual accumulation of evidence followed by thresholding to commit to a choice. In rats, the anterior dorsal striatum (ADS) and the frontal orienting fields (FOF) have been mapped onto these two theoretically defined

computations. ADS represents the graded evidence and is causally required throughout the accumulation process (Yartsev et al., 2018), whereas FOF represents the categorical choice and is causally necessary only at the very end of accumulation period when the graded evidence has to be thresholded (Erlich et al., 2015; Hanks et al., 2015; Piet et al., 2017). These observations provide an intriguing - yet untested - neural implementation of the decision process, one in which ADS and FOF form a functional feedforward hierarchy such that evidence accumulated in ADS is thresholded in FOF.

While appealing, this proposal is surprising in many ways. This proposed implementation maps onto a multisynaptic pathway in which signals from ADS must be relayed through other basal ganglia nuclei, thalamus and/or superior colliculus before they reach FOF. This is despite the fact that FOF projects monosynaptically to ADS and is one of its major cortical inputs. Anatomically connected brain areas can selectively communicate information to their downstream targets (Kaufman et al., 2014; Semedo et al., 2019), so it is certainly an interestingly possibility that the information flow during evidence accumulation renders the monosynaptic projection from FOF to ADS non-functional and only involves the multisynaptic pathway. However, the substantial heterogeneity in neural encoding within FOF, with neurons representing both graded evidence and categorical choice (Hanks et al., 2015), along with FOF's anatomical placement, strongly evokes the possibility that FOF might have a role in gradual accumulation as well.

Indeed, some recent studies pose challenges to the interpretation of FOF's unihemispheric optogenetic perturbations, that had provided compelling evidence to confine FOF's role to the thresholding epoch. It was recently observed that unlike bihemispheric inactivations, transient unihemispheric inactivations of cortex (similar to the ones that helped map FOF

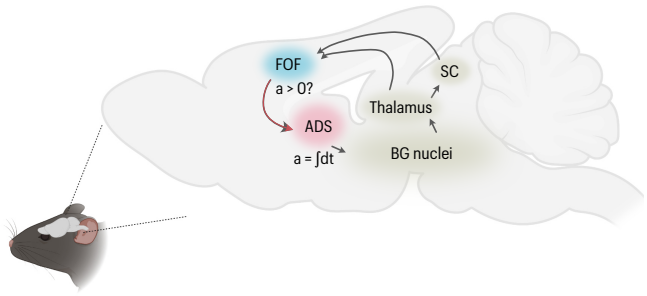


Figure 4.1: Proposed feedforward functional hierarchy underlying evidence accumulation in the rat brain Schematic showing proposed implementation of evidence accumulation (Hanks et al., 2015; Yartsev et al., 2018; Brody and Hanks, 2016). The anterior dorsal part of striatum (ADS, pink) is thought to be involved in gradual accumulation of sensory evidence and represents this accumulated value "a" in a graded fashion (Yartsev et al., 2018). The accumulated evidence is thought to be thresholded, and represented in categorical form in the frontal orienting fields (FOF, blue), a part of the rat secondary motor cortex (Hanks et al., 2015; Erlich et al., 2015; Piet et al., 2017). The pathway from ADS to FOF involves a multi-synaptic pathway (grey arrows through olive-green regions) through basal ganglia (BG) nuclei, the thalamus and the superior colliculus (SC) that could potentially support this transformation. Such a feedforward functional hierarchy also implies that corticostriatal projections from FOF to ADS (red arrow) play no role in this computation.

onto thresholding) can be quite robust to perturbations and recover through information from the other hemisphere (Li et al., 2016). Since the network has enough time to recover early in the accumulation period but not towards the end, interhemispheric compensation might underlie the lack of impairment observed during FOF's inactivations early in the decision process, obscuring its role. Moreover, there is mounting evidence showing that whole-region non-specific perturbations may conceal effects that can be revealed by more specific perturbations that target a particular projection (Warden et al., 2012; Otis et al., 2017; Kim et al., 2017).

Given these facets, it is all the more important to directly evaluate the proposed transformations between ADS and FOF. Testing this proposal would require recording activity from both areas simultaneously and examining if the expected representations, transformations, and their latencies hold on single-trials and when inter-individual variability is accounted

for. This is also important since slight differences in cognitive strategies (e.g. accumulation with recency or primacy bias) or training histories between individuals can map onto different neural representations and dynamics (Gilad et al. 2018; Latimer and Freedman 2021; Arlt et al. 2022; Bolkan et al. 2022, Pagan et al. *in prep*).

Here, using simultaneous multiarea population recordings and projection-specific perturbations in rats performing an evidence accumulation task, we provide evidence against the modular, feedforward hypothesis in which evidence accumulated in ADS is thresholded in FOF. We show that ADS does not have privileged information about evidence, rather both regions carry similar amounts of information that evolves at comparable timescales. Further, we reveal an unexpected role for FOF during gradual accumulation of evidence using perturbations. We show that unlike inactivation of FOF neurons, selective silencing of FOF inputs to striatum impairs decision-making throughout the accumulation process and not just during choice formation. We revise the feedforward hypothesis and advocate for an alternate distributed implementation of evidence computation for decision-making in the cortico-striatal circuitry. We instantiate this proposal using a multiregion recurrent neural network which reproduces our results from perturbation and recording experiments.

4.3 Results

In order to test if the hypothesized transformation between ADS and FOF holds on single trials, we recorded simultaneous population activity from ADS and FOF of rats, trained to perform a previously developed decision-making task (Brunton et al., 2013). This task requires accumulation of auditory evidence over hundreds of milliseconds for good performance. On any given trial, rats are presented with two streams of randomly timed auditory clicks, one

played from a speaker to their left and the other from a speaker to their right (Fig 1.1A). At the end of the stimulus, rats are rewarded with a drop of water for correctly reporting the side which played the greater number of clicks.

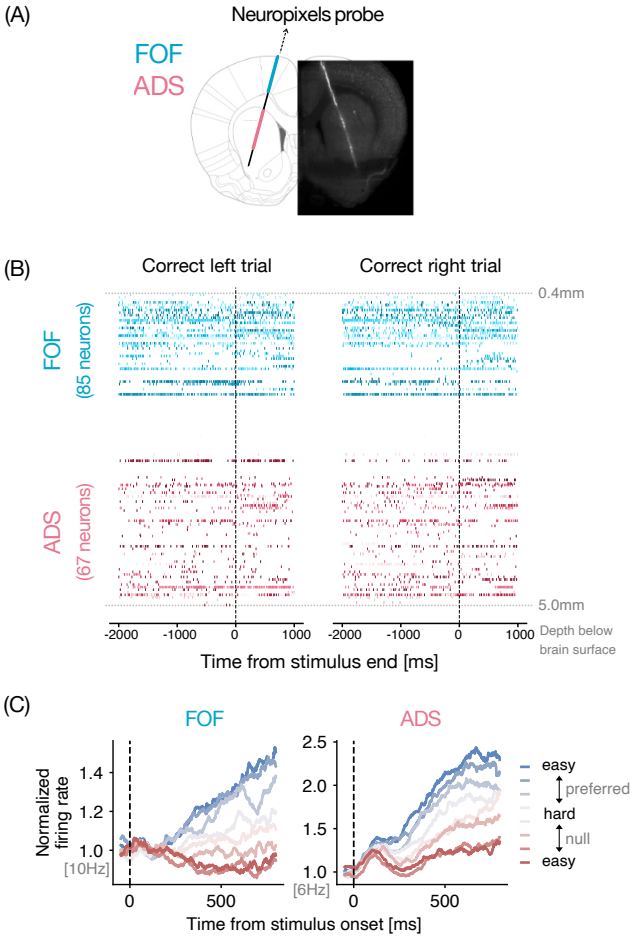


Figure 4.2: Experimental setup for recording simultaneously from FOF and ADS (A) Schematic showing neuropixels targeting (AP 1.9, ML 1.15 at a 15 degrees angle in the coronal plane) to record simultaneously from FOF and ADS, overlaid on histology from an example rat. Brain atlas was adapted from Paxinos and Watson, 2006 (B) Spike rasters showing simultaneous neural activity recorded with neuropixels probes from two example correct trials (left and right) from one session. Neurons are arranged in order of their estimated depth from the brain surface. Aqua blue shades represent FOF neurons and pink shades represent ADS neurons (C) Average FOF (left) and ADS (right) population responses during stimulus period for different strengths of sensory evidence. Blue and red colors correspond to stimuli from preferred and non-preferred sides of the neurons, respectively. Darker hues correspond to easier trials and greyish hues correspond to harder trials. Both areas show stimulus strength dependent ramping activity. (FOF n = 222 neurons, ADS n = 239 active side-selective neurons from 5 rats).

While rats (n = 5) performed this task, we used Neuropixels probes (Jun et al., 2017; Luo et al., 2020) to record neural activity from FOF and ADS simultaneously (Fig 4.2A-B). We recorded from 788 FOF and 1756 ADS neurons across 12 sessions. Cells with mean firing rate of at least 1Hz during the stimulus period were considered active and included for further analysis (468 FOF, 561 ADS neurons). In both FOF and ADS, we found that neurons

that significantly modulated their activity in a side-selective manner (determined using ROC analysis, see Methods 4.5.2) on average exhibited firing rates that ramped upwards for stimuli favoring their preferred side with a slope proportional to the stimulus strength (Fig 4.2C). These trial-averaged stimulus-dependent ramping responses have long been interpreted as a neural correlate of the evidence accumulation process (Gold and Shadlen, 2007) and are consistent with previous reports in rats from FOF and ADS (Hanks et al. 2015; Yartsev et al. 2018), and analogous regions in primates (FEF and caudate: Ding and Gold 2012; 2010). However, both graded and categorical representations of accumulated evidence can give rise to such ramping responses (Latimer et al., 2015; Hanks et al., 2015), hence to distinguish between these possibilities we characterized the encoding and decoding profiles of task-relevant variables in the two regions with more directed analysis.

4.3.1 Task-relevant representations in FOF and ADS

The feedforward functional hypothesis posits that ADS has more graded encoding of evidence than FOF and therefore predicts that evidence can be decoded from ADS earlier and with higher accuracy compared to FOF. We sought to compare the differences in the extent of graded encoding between FOF and ADS by using the method developed in Hanks et al. (2015). This method computes the 'tuning curves' or maps of how firing rates of neurons vary as a function of the accumulator value. To do so, it leverages the model from Brunton et al. 2013 (Fig 1.1D-E) to obtain trial-by-trial, moment-by-moment estimates of the value of subject's accumulator and relates it to the simultaneously observed firing rates of neurons. Previous application of this method had found that tuning curves in FOF were more step-like or categorical, with accumulator values to one side of the decision boundary giving rise

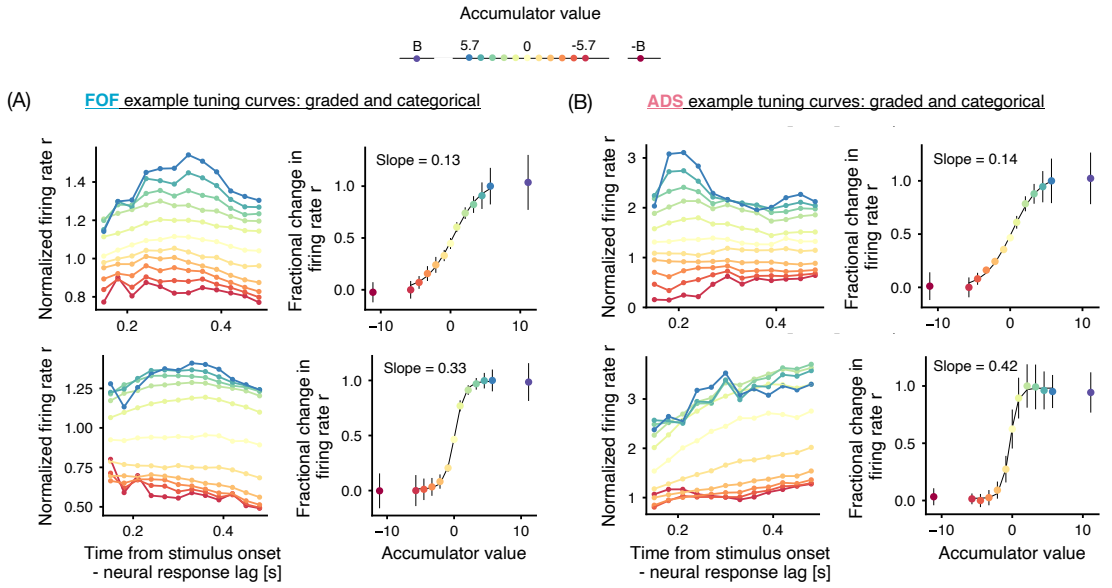


Figure 4.3: Example neurons from FOF and ADS exhibiting graded and categorical tuning to the accumulator variable (A) Two example neurons from FOF exhibiting graded (top) and categorical (bottom) tuning to the behaviorally inferred accumulator variable. (Left) Neuron's firing rate as a function of time from the stimulus onset time minus the neural response lag of 100ms. Firing rates are grouped by binned values of the accumulator variable (spectral colors; $\pm$B represent the sticky bounds of the behavioral accumulator model). (Right) Dots - firing rate as a function of the accumulator variable (mean $\pm$ sem), obtained by averaging the firing rate map from (left) over time. To account for differences in dynamic range across neurons, responses have been scaled to span the range from 0 to 1. Black curves show the tuning curves or sigmoids fit to the data. Slopes of these sigmoids at the transition between negative and positive accumulator values is used to summarized the nature of tuning. The neuron in the top panel exhibits a graded response to the accumulator variable, signified by the shallow slope of its tuning curve. The neuron in the bottom panel exhibits a sharp dependence on the sign of the accumulator variable, reflected in the steeper slope of its tuning curve. (B) Same as A but showing two example neurons with graded (top) and categorical (bottom) responses from ADS.

to one cluster of firing rates, and the values on the other side of the decision boundary giving rise to another cluster (Hanks et al., 2015). On the other hand, previous application in ADS had found prevalence of neurons whose firing rates varied smoothly with the graded value of the accumulator (Yartsev et al., 2018).

When applying this analysis to side-selective neurons in our data, we found examples of both graded and categorically tuned neurons in FOF as well as ADS (Fig 4.3). Surprisingly, we found that the median slope was slightly higher for ADS (0.25 ± 0.13, mean $\pm$ std) com-

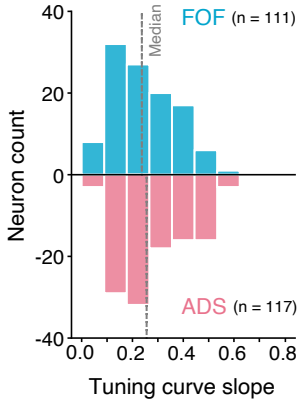


Figure 4.4: Distribution of tuning curve slopes in FOF and ADS Histogram of tuning curve slopes in FOF (aqua blue; n = 111, 3 rats) and ADS (pink; n = 117, 3 rats), inferred by fitting a sigmoid to the relationship between firing rate and accumulator value (as in Fig 4.3). Grey lines indicate the median slopes for the two populations (FOF = 0.23 $\pm$ 0.12; ADS = 0.25 $\pm$ 0.13). The two populations have similar distribution of tuning to the accumulator variable, with neurons representing the accumulator value in both graded (smaller slopes) and categorical (larger slopes) fashion. The medians of the two populations were not significantly different (P > 0.05, Mann Whitney U test).

pared to FOF (0.23 ± 0.12) but with no significant differences between FOF and ADS (Fig 4.4, P = 0.09 Mann-Whitney U test). Instead, both areas carried large numbers of graded neurons, with a sizeable fraction of categorical neurons. These results demonstrate that the two regions do not differ in their encoding of the accumulator variable, when variability across trials, sessions and animals is controlled for. This is contrary to expectations from previous studies that were performed in separate cohorts of animals, suggesting that the two regions' contributions to the decision-making process may not be as distinct as previously hypothesized.

We also compared the decodability of information about evidence and other task variables between FOF and ADS populations. Despite the similarity in trial-averaged tuning to the accumulator variable, it is still possible that this information is decodable earlier in ADS, and/or with higher accuracy. For each region, we trained separate linear decoders to predict the cumulative difference in the number of right and left clicks at each time point during the trial (Fig 4.5A). Interestingly, both regions showed comparable decoding performance over time, as measured by the correlation between the true and predicted cumulative stimulus on held out time points, with no significant differences between the regions (Fig 4.5B; two

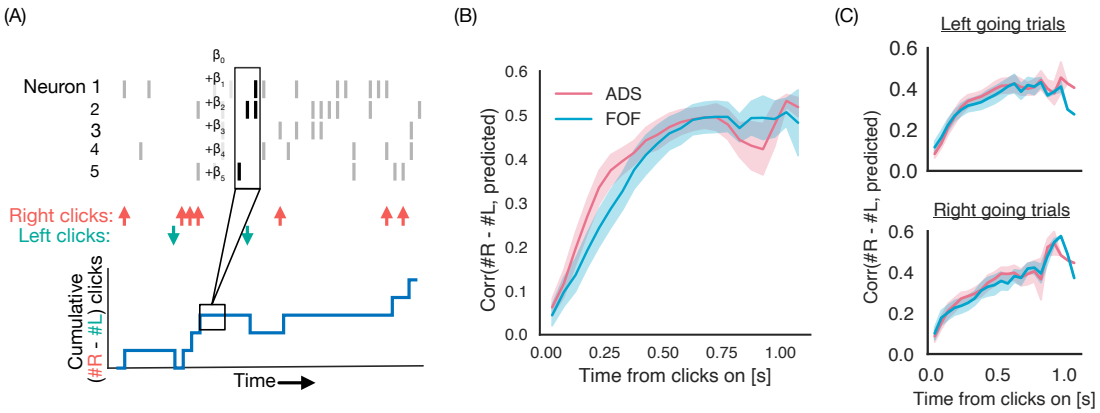


Figure 4.5: Similar strength and timecourse of stimulus decoding in FOF and ADS (A) Schematic showing the linear decoder designed to predict the cumulative click difference in the number of right and left clicks at any given time during a trial using a linear combination of neural activity from FOF or ADS. We assumed a neural encoding lag of 100ms and regularized the linear weights on neurons with L1 penalty using cross-validation. (B) Stimulus decoding performance (mean $\pm$ sem) of the linear decoder on held-out time points as a function of time from stimulus onset (n = 5 rats). Linear decoders were fit independently to the two populations (FOF: aqua blue; ADS: pink) while controlling for the number of neurons from the two regions. Performance is measured as the correlation between the actual trajectory of difference in number of clicks and the prediction from linear decoding of neural responses. Decoding performance and its timecourse did not significantly differ between FOF and ADS (P>0.05, two-way RM ANOVA) (C) Same as B but performed separately for left (top) and right (bottom) choice trials to control for choice. The two areas show high decoding performance which evolves with a similar timecourse (P>0.05, two-way RM ANOVA).

way repeated measures ANOVA, P = 0.23 for region and P = 0.13 for region x time). These results did not change when choice information was controlled for, by separately decoding the stimulus on left and right choice trials (Fig 4.5C). The results were similar also when instead of the veridical stimulus, the mean accumulated evidence inferred with the behavioral model (as in Fig 4.3) was decoded from the two populations (not shown).

Additionally, we trained logistic decoders to predict other binary task-relevant variables such as choice and trial-history from the two regions. The two regions showed high decoding of choice (Fig. 4.6A) but once again with no significant differences in decoding performance over time (two way repeated measures ANOVA, P = 0.98 for region and P = 0.12 for region x time). Past trial's choice could be decoded well into the next trial with substantial

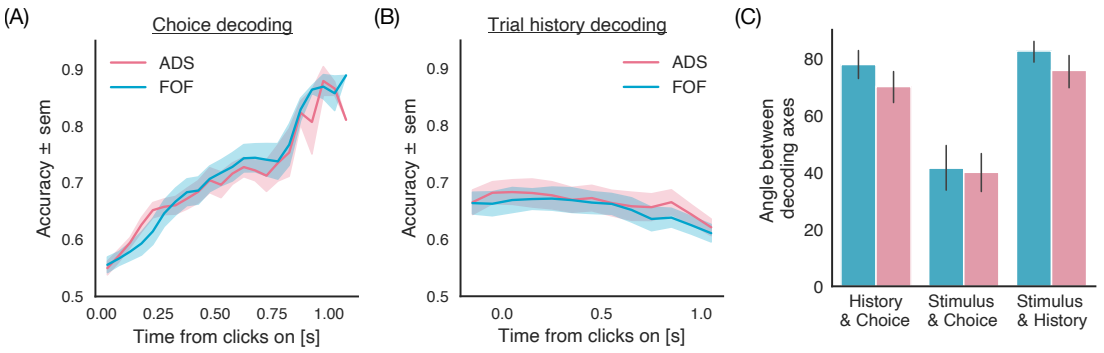


Figure 4.6: Similar choice and trial-history decoding performance in FOF and ADS (A) Cross-validated choice decoding accuracy as a function of time from stimulus onset for FOF (aqua blue) and ADS (pink) neural activity (n = 5 rats). Time courses and strength of choice decoding from the two populations did not differ significantly (P>0.05 two-way RM ANOVA). Decoding was performed with an L2 regularized logistic decoder, while controlling for number of neurons from the two areas. (B) Same as A but decoding for choice from past trial (n = 5 rats). No significant differences were observed between the two populations (P > 0.05, two-way RM ANOVA) (C) Pairwise angles (mean $\pm$ sem) between choice (panel A), trial history (panel B) and stimulus (Fig. 4.5B) decoding axes i.e. vector of linear weights on neurons inferred for decoding. The two areas have similar geometry of decoding axes, with choice and evidence axes being highly aligned, and trial-history axis almost perpendicular to them both.

accuracy from both regions, however there were no significant differences between the two regions (two way repeated measures ANOVA, P = 0.32 for region and P = 0.63 for region x time). Finally, we compared the relative geometry of the three different decoding axes (stimulus, history and choice) in the two regions, by measuring the pairwise angles between them. Evidence and choice decoding axes were more aligned with each other compared to trial history decoding, but with no significant differences between the geometry of the two regions (Wilcoxon signed-rank test, P = 0.09 for choice & history, P = 0.79 for choice & stimulus, P = 0.05 for stimulus & history). Altogether, these results show that the two regions have indistinguishable representations of task relevant variables, leading us to believe that their involvement in decision-making may be strongly related.

Our conclusions about simultaneous evolution of evidence-related representations in the two regions are also supported by an independent analysis which decodes the decision vari-

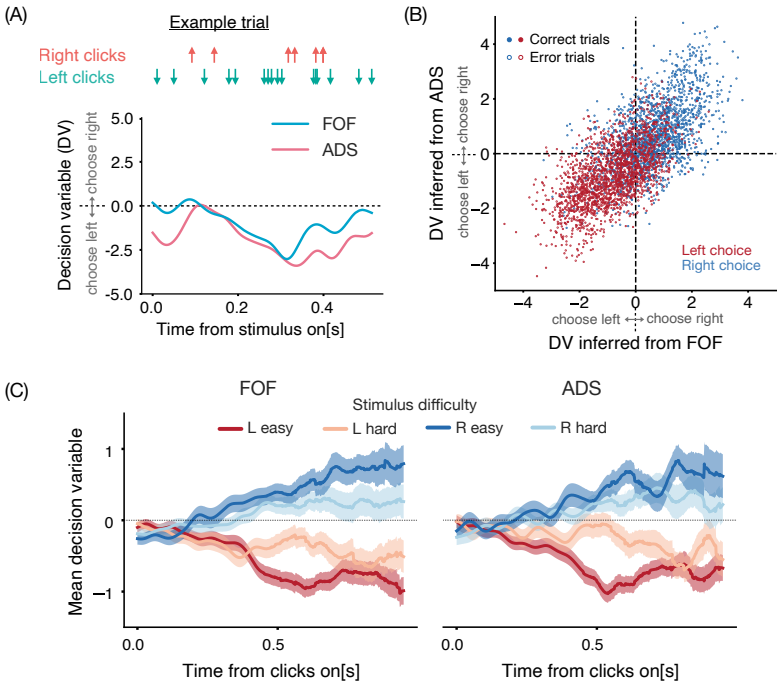


Figure 4.7: Similarity in trajectories of decision variable decoded from FOF and ADS population activity (A) Example trial showing the trajectory of the decision variable (DV) decoded from the simultaneous population activity of FOF and ADS (blue, pink line respectively), using the logistic decoder (see Methods, Fig 4.6). The DV represents the log odds ratio of making one choice over another, with higher positive (negative) values indicating a stronger likelihood of rightward (leftward) choices. Both regions have comparable DV time-courses, and seem to be influenced by the sequence of right and left clicks (coral, green arrows respectively). (B) Scatter plot comparing DV values inferred from FOF (x-axis) and ADS (y-axis) activity at the same time-points from an example session. Dots represent time-points across trials, with filled (unfilled) circles indicating correct (incorrect) trials. Points are color-coded according to the actual choice of the animal, with blue (red) indicating rightward (leftward) choices. The DV inferred from both areas shows good correspondence with each other and the true choice. (C) Mean DV trajectories from FOF (left) and ADS (right) over time from an example session and separated by stimulus difficulty (red - easy leftward stimulus, light red - hard leftward, blue - easy rightward, light blue - hard rightward). Shaded regions indicate S.E.M. Both regions show similar DV time-courses and similar dependence on stimulus difficulty.

able (see Chapter 1) from the neural population activity (Kiani et al., 2014; Li et al., 2016; Peixoto et al., 2021). A logistic decoder trained to predict choice from the neural activity finds the hyperplane in the neural population space (with firing rates of individual neurons as axes), that best separates the population responses associated with the two choices. Projection of the population firing rates at any given time (i.e. a point in this population space)

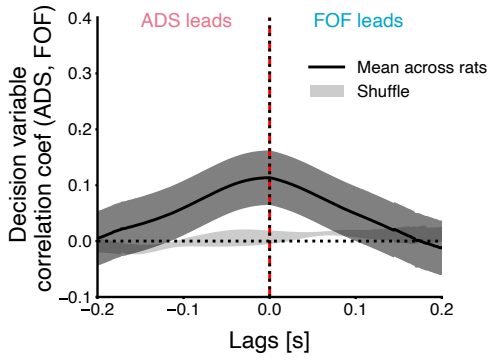


Figure 4.8: No discernible lead-lag relationship between decision-variables in FOF and ADS. Time-lagged cross-correlogram between decision-variable trajectories decoded from FOF and ADS, showing the correlation coefficient (y-axis) at different time lags (x-axis: positive (negative) values represent FOF trajectories leading (lagging) ADS). Solid black line indicates mean across sessions and rats, with dark-gray shaded region representing S.E.M. Light gray region indicates a shuffled control, computed by cross-correlating FOF activity with ADS activity on trials with shuffled (rather than matched) identities. The cross-correlogram is symmetric with a wide peak at zero, hence showing no strong evidence of any lead-lag relationships.

onto this hyperplane reflects the model's prediction confidence in units of log-odds ratio of the two choices, with positive values indicating confidence towards rightward choices and negative values for leftward ones. This projection, often referred to as the decision variable or DV, estimates the animal's internal decision variable (Kiani et al., 2014) and has been successfully leveraged to infer changes of mind in real-time and subsequently for biasing animal's choices (Peixoto et al., 2021). This neurally inferred, time-varying DV can capture not just the influence of sensory evidence on choice but also other known (e.g. trial history) and unknown determinants of the animal's choice. This approach therefore complements our previous analyses as it relaxes the assumptions made by our behavioral model that describes the evolution of the accumulator variable (as in Fig 4.3, see Methods) and does not assume a fixed stimulus encoding lag in the two regions (as in Fig 4.5).

We sought to examine if these neurally inferred DVs from the two regions exhibit any systematic lead-lag relationships in their evolution, in order to infer the dominant direction of task-related information flow within the frontal cortico-striatal circuitry. To this end, we used the logistic decoder trained to predict the eventual choice (from Fig 4.6) and estimated

the DV on single trials from the two regions (Fig 4.7A). DVs from FOF and ADS showed good correspondence with each other and with the eventual true choice of the animal on both correct and error trials (Fig 4.7B). Averaged across trials of different difficulties, DV values (unsigned) increased over the trial, with the rate of increase dependent on the strength of sensory evidence – exhibiting the classic ramping signature expected from evidence accumulation (Fig 4.7C). Consistent with our previous analyses, we did not find any lead-lag relationship in the evolution of DV in the two regions (Fig 4.8). While the DVs showed significant correlations at short latencies on average, the cross-correlogram of these DVs averaged over rats and sessions had a peak at a latency of 0ms with no asymmetries for positive or negative lags that would have indicated a dominant feedforward or feedback relationship (Semedo et al., 2022).

4.3.2 Cortico-striatal communication during evidence accumulation

Next, we probed the content of interaction between FOF and ADS during decision-making more generally, rather than restricting ourselves to activity in pre-specified task-relevant subspaces, such as along the decision hyperplane (Fig 4.7). To measure such interactions we used canonical correlation analysis (CCA; Hotelling 1936) to search for dimensions in the population activity spaces of the two brain regions, along which neural activity is maximally correlated. We inferred these dimensions by analyzing 'noise' correlations (Fig 4.9A) or correlations in trial-by-trial fluctuations in neural activity after removing the expected mean firing rates for each trial (Cohen and Kohn, 2011). Noise correlations are commonly thought to reflect co-fluctuations in neural activity arising from either direct or indirect communication between brain regions. Since, in our behavioral task we do not have repeats of the same stim-

ulus, we fit a general linear model (GLM) with temporally-extended kernels for left and right clicks, and other major task events to each neuron's response and used predictions from this model as estimates of the mean expected response (Fig 4.9A; Methods). We performed CCA on the residual population activity from the two regions and defined a set of canonical axes, one for each region, often referred to as communication subspace (Kohn et al. 2020).

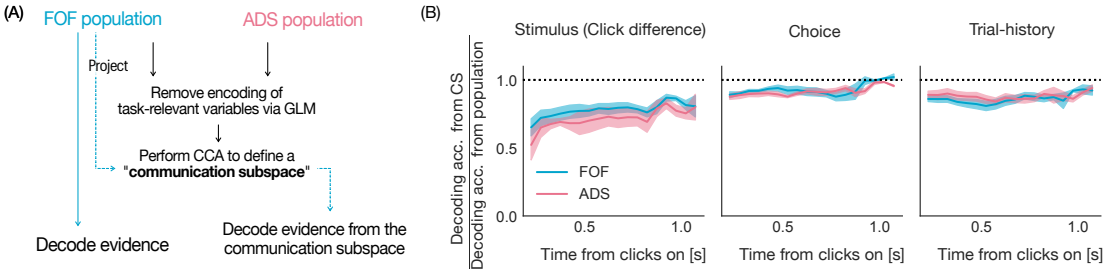


Figure 4.9: Communication subspace analysis using CCA reveals shared information between FOF and ADS (A) Schematic depicting pipeline used to define a communication subspace between FOF and ADS, using Canonical Correlation Analysis (CCA). Activity from both regions is first passed through a GLM to extract single-neuron encoding of any task-relevant variables. This encoding is subtracted from each neuron's activity to remove any signal correlations between neurons, and the residual activity of the two populations are passed into CCA to extract communication subspaces for both regions. Finally, decoding of task variables is performed on both the raw population activity from each region, as well as from the activity projected into that region's communication subspace. (B) Time-course of information about task-relevant variables present in the communication subspace (CS) of FOF (blue) and ADS (pink), plotted as a fraction of the total information about that variable in the entire region's population activity. A fraction of 1 indicates that all the information about that variable is present in the CS. Stimulus information (left) is present in both regions' CS and increases over the course of the trial, with FOF sharing a slightly higher fraction of its information in its CS. Choice (center) and trial history (right) information is present throughout the trial in both regions' CS at a comparable level.

We compared the information content of activity in FOF and ADS's communication subspaces (CS) to the whole population activity space by measuring the decoding accuracy of stimulus, choice and trial history from the two (Fig 4.9B). Remarkably, we found that the majority of stimulus information in the two regions could also be decoded from both regions' CS and this fraction increased over the course of the trial, with FOF sharing a slightly higher fraction of its information in its CS. Information about choice and trial history could

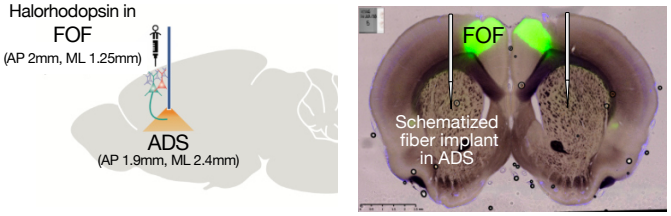


Figure 4.10: Silencing FOF axon terminals in ADS: experimental setup (Left) Schematic showing viral delivery of Halorhodopsin (eNpHR3.0) to FOF and sharp optical fiber implant in ADS. Virus was injected bilaterally and 594mm wavelength laser (25-33mW) was delivered unilaterally through the fiber optic cable on 25% of trials to silence FOF axon terminals in ADS. (Right) Example histology from a rat expressing eNpHR3.0 bilaterally in FOF.

be decoded throughout the trial in both regions' CS at a comparable level and almost equalled the accuracy from the full population space. Altogether, these results indicate that the representations of these task-relevant variables in the two regions reside in the space that captures the shared trial-by-trial noise fluctuations and therefore are likely communicated between the two.

4.3.3 Perturbing FOF projections to ADS

From the analysis of simultaneous activity of FOF and ADS we observed that the two regions have remarkably similar encoding and decoding of decision-relevant variables, and that there is no discernible lead-lag relationship between the representations of the evolving decision variable in the two regions. While this redundancy could reflect that these signals are inherited by both FOF and ADS from a third (unknown) brain region, given the causal involvement of these regions during decision-making - a more interesting and likely possibility is that this redundancy reflects ongoing recurrent interactions (Semedo et al., 2022) between FOF and ADS. Are these interactions necessary for making decisions? Establishing importance of these interactions would put strong constraints on the network mechanisms that underlie decision-making.

Therefore, we sought to probe this question in the experimentally tractable monosynaptic projection from FOF to ADS using unilateral optogenetic inactivations. We injected the inhibitory opsin, halorhodopsin (AAV5-CaMKIIa-eNpHR3.0-eYFP) bilaterally in FOF and delivered light (25-33mW, 594nm) via a sharp fiber optic implanted in ADS to silence the activity of the FOF-ADS projection neurons' axon terminals (Fig 4.10). FOF to ADS projections (or cortico-striatal projections in general) are composed of two major cell types: intratelencephalic (IT) neurons, projecting bilaterally to other cortical structures and striatum, and pyramidal tract (PT) neurons, projecting unilaterally to subcortical structures including striatum (Dudman and Gerfen, 2015). Our question does not require differentiating between these cell types so our unilateral inactivations essentially inhibit all inputs from FOF (IT and PT, both hemispheres) to the illuminated ADS. Axon terminal silencing has high specificity not just for the cell type but also for the projection target and therefore has minimal off-target effects, however optogenetic tools for this technique are still not perfected (Wiegert et al., 2017; Rost et al., 2022). For our experiments, we use halorhodopsin because it remains one of the best opsins for sub-second terminal silencing (Mahn et al., 2016; Wiegert et al., 2017; Rost et al., 2022) and has had demonstrable success in behavioral experiments (e.g. Tye et al. 2011; Stuber et al. 2011; Hwang et al. 2019).

First, to test whether this projection has any role in the decision-making process, we silenced the activity of FOF-ADS terminals in either hemisphere during the whole trial (2s period trial initiation until 500ms after stimulus end; Fig 4.11A) on a subset of trials (25%). This inactivation significantly increased the proportion of choices ipsilateral to the laser side in rats injected with eNpHR3.0 (Fig 4.11B; $P < 0.001$, nonparametric bootstrap test), with no significant effects on the controls injected with eYFP (Fig 4.11B; $P = 0.78$, nonparametric

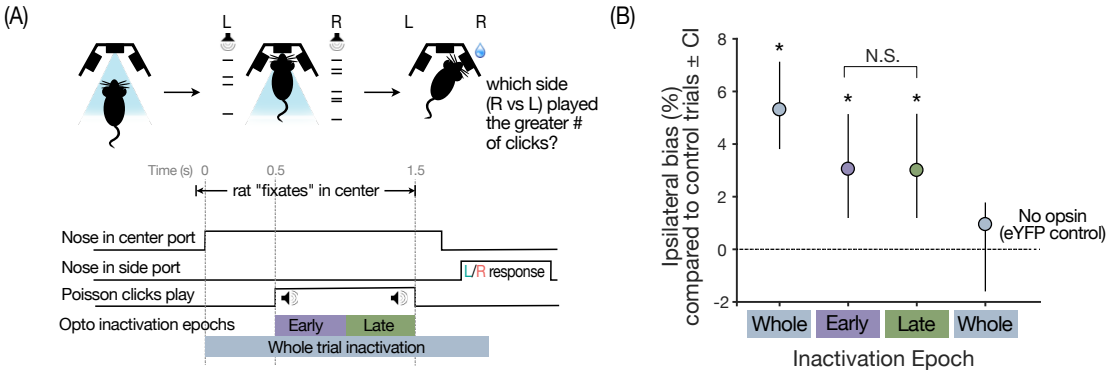


Figure 4.11: Silencing FOF axon terminals in ADS: effect on choices (A) (Top) **Evidence accumulation task in rats:** Schematic of the events in the task (adapted from Brunton et al. 2013). From left to right: rats initiate a trial by poking their nose into the center port, two streams of randomly timed auditory clicks play from speakers on their left and right side, rats are rewarded for choosing the side with the greater number of clicks. (Bottom) **Timeline of task events showing perturbation epochs:** FOF axon terminals in ADS were perturbed during one of three epochs: *Whole-trial* (grey bar) started as soon as the rat entered the center port and terminated after 500ms from stimulus end. *Early* (purple) spanned the first half of a 1s long stimulus and *late* (green) spanned the second half of a 1s long stimulus. It is predicted that if the FOF-ADS projection is involved in conversion of the graded evidence to a categorical choice, then the inactivation should have an effect on behavior only at the end of the stimulus presentation i.e. during the late epoch. Whereas, if the projection is involved in evidence accumulation, a process that occurs throughout the stimulus presentation, inactivations during both early and late epochs should have an effect. (B) **FOF to ADS projections are involved throughout the accumulation process:** The ipsilateral bias, or the difference in percentage of trials the rat went to the side ipsilateral to the laser on inactivation v.s. control trials plotted for different inactivation epochs. Circles represent the mean ipsilateral bias and error bars represent 95% bootstrap confidence intervals. Whole-trial inactivations (6rats, 10 hemispheres) in rats injected with NpHR produced a significant impairment in behavior ($P<0.001$, non-parametric bootstrap test) and no significant effects in the eYFP controls ($P=0.78$) (2rats, 3 hemispheres). Early and late inactivations (5 rats, 8 hemispheres) also produced significant impairments ($P<0.005$ non-parametric bootstrap test) with no significant differences between the effects of the two epochs ($P=0.49$) suggesting that FOF-ADS projections are involved throughout the accumulation process.

boostrap test), suggesting that the projection is involved in the decision process.

Next, we assessed the temporal extent of involvement of this projection, by restricting the inactivation to either early (first half of a 1s long stimulus) or late (second half of a 1s long stimulus) epochs of the evidence accumulation or stimulus period (Fig 4.11A; similar to Hanks et al. 2015; Yartsev et al. 2018). With this design, it is predicted that if the FOF-ADS projection is involved in conversion of the graded evidence to a categorical choice, then the inactivation should have an effect on behavior only at the end of the stimulus presentation

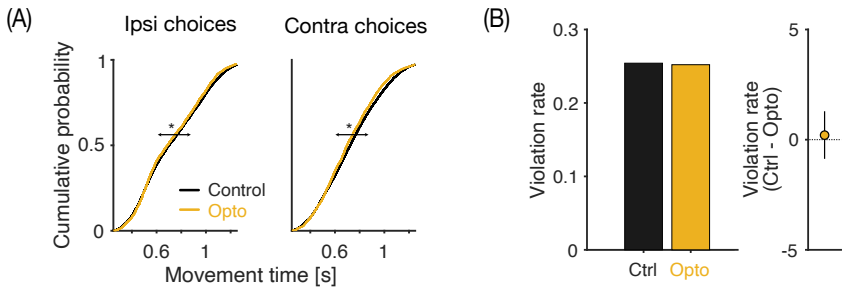


Figure 4.12: Silencing FOF axon terminals in ADS: effect on movement (A) **Silencing reduces movement times** Cumulative distribution of movement times across (n = 6) rats on control (black) and whole-trial inactivation (orange) trials. Movement time is defined as the time taken by the rats to leave the center port and enter one of the two side ports to report their choice. Silencing of FOF axon terminals in ADS significantly reduced the mean movement times for both choices ipsiversive (left, P = 0.05 Mann-Whitney U test) and contraversive (right, P = 0.001 Mann-Whitney U test) to the laser. (B) **Silencing doesn't affect rate of trial completion** (left) Mean violation rates observed across rats (n=6) on control (black) and whole-trial inactivation (orange) trials. Rats are expected to "fixate" at the center port for 1.5s from trial initiation, failure to do so aborts the trial. Violation rates measure the rate of premature exits from the center port. An increased violation rate might reflect an inability to complete trials due to gross motor/cognitive impairment induced by laser. (right) Mean violation rates between control and opto trial were not different (P = 0.58, non-parametric bootstrap test).

i.e. during the late epoch. Whereas, if the projection is involved in evidence accumulation, a process that occurs throughout the stimulus presentation, inactivations during both early and late epochs should have an effect. Much to our surprise, when we restricted inactivations of the FOF-ADS projection to the early and late epochs, we found that the inactivations led to an ipsilateral bias in both epochs (Fig 4.11B; early, P = 0.002; late, P = 0.007, nonparamteric bootstrap test) with no significant differences in the degree of impairment (P = 0.49). This is in stark contrast to the effect of whole region FOF inactivations (Hanks et al., 2015) which caused an ipsilateral bias only during the late epoch. Instead, these results resemble ADS inactivations (Yartsev et al., 2018) which caused an ipsilateral bias in both early and late epochs. Altogether, these results implicate the FOF-ADS projection in the gradual accumulation process, revealing a previously obscured role for FOF neurons. Consistent with this, the involvement of cortico-striatal neurons throughout the stimulus period was also recently

reported in mice performing a similar task using somatic inhibition (Musall et al., 2021).

Finally, we also examined the effect of inactivations on movements (Fig 4.12). Terminal silencing slightly but significantly reduced mean movement times for choices both ipsiversive ($P = 0.05$ Mann-Whitney U test) and contraversive ($P = 0.001$ Mann-Whitney U test) to the laser. However, silencing did not increase rates of fixation violations ($P = 0.58$, nonparametric bootstrap test). This indicates that the laser did not induce any gross motor impairments affecting animals' ability to successfully complete trials.

4.3.4 Bilateral FOF inactivations

Optogenetic inhibition of FOF inputs to ADS impairs decisions both in early and late stimulus epochs (Fig 4.11), indicating a role for FOF in the gradual accumulation process, much like ADS. This is consistent with the highly similar neural representations observed in the two regions (Figs 4.4, 4.5, 4.6) and suggests that rather than being incidental, evidence representations in FOF might be important for decision-making behavior. Interestingly, this result also raises the question: why don't the unilateral whole-region FOF inactivations disrupt decisions during the accumulation period? In mouse M2 interhemispheric recovery was shown to be the reason for lack of impairments observed during transient unilateral inactivations in early task periods (Li et al., 2016; Chen et al., 2021). So a good candidate explanation might be that FOFs in the two hemispheres are coupled and can support each other's recovery when aberrant inputs from the other hemisphere are detected. The localized nature of terminal inhibition targeting FOF inputs just to ADS, and not to the other FOF might bypass such a recovery.

We tested this hypothesis by carrying out bilateral optogenetic inactivations in a separate

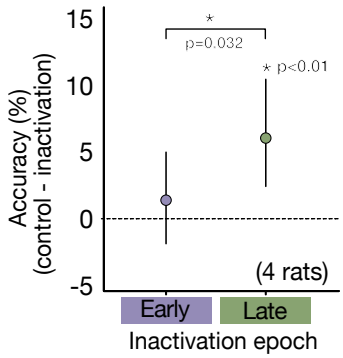


Figure 4.13: Silencing FOF bilaterally reduces performance only in late stimulus epochs. Difference in accuracy between control and inactivation trials, plotted for temporally specific bilateral inactivations of FOF in early (purple) or late (green) stimulus epochs. Circles represent mean accuracy difference across 4 rats, error bars represent 95% bootstrap confidence intervals. Only inactivations in the late epoch produce significant reductions in accuracy (p<0.01, nonparametric bootstrap test), differing significantly from inactivations in early epochs (p<0.032, nonparametric bootstrap test) in which there is no significant impairment.

cohort of rats (Fig 4.13). We silenced FOF activity by delivering laser bilaterally either during the first half of the stimulus (early inactivation epoch), or the second (late inactivation epoch; same as Fig 4.11A). We observed a significant decrement in performance in only the late epoch ($P < 0.01$, nonparametric bootstrap test), inactivation during the early accumulation epoch did not significantly hamper the performance ($P > 0.05$, nonparametric bootstrap test). This implies that interhemispheric compensation does not underlie the lack of impairment observed on inactivating FOF early during evidence accumulation. So the question of why projection specific inactivations impair decisions early during accumulation but whole region FOF inactivations don't, remains outstanding.

4.3.5 Modeling physiology and perturbations

What network properties could confer the FOF-ADS circuitry with the robustness that we observed while it engages in computations important for decision-making? To approach this challenge, we turn to computational modeling. Our experimental findings are hard to reconcile with a feedforward orgnanisation of these brain regions, wherein regions have localized functions (as previously hypothesized). They might however be concordant with a circuit that computes in a collaborative and distributed fashion through recurrence and feedback,

which allows for adaptive rescue of computations upon perturbation (e.g. Kopec et al. 2015; Li et al. 2016; Duan et al. 2021a; Bredenberg et al. 2021; Mejías and Wang 2022). Therefore, we translated our findings into a mechanistic multi-region network model that instantiates a recurrent circuit between FOF and ADS, inspired by anatomical properties and captures decision-making behavior. We sought to analyze the behavior of this network in unperturbed and perturbed states to identify mechanism(s) through which it can be sensitive to projection specific inactivations but resilient to whole region FOF inactivations during the accumulation period. This model let us perform virtual experiments to assay the consequences of perturbations with arbitrary temporal and cell-type specificity - a prohibitive prospect *in vivo* due to time, costs, tools, and number of animals involved. We consider this exploration an important step towards generating intuitions and hypotheses about the function of complex multi-region networks, given the predominance of single region models within the decision-making literature.

We used recurrent neural networks (RNN) in our modelling, since several RNN-based models have been successful at capturing animal behavior as well as neural representations during perceptual decision-making (Rigotti et al., 2010; Mante et al., 2013; Carnevale et al., 2015; Remington et al., 2018). Moreover, the tools for reverse-engineering such networks have been effective at revealing their dynamical systems properties (Mante et al., 2013; Remington et al., 2018; Maheswaranathan et al., 2019). We imposed biological constraints on the RNN, by initializing its weights with possible anatomical connections while following Dale's law and maintained these constraints on the architecture throughout training (following methods developed in Song et al. 2016). The model's architecture is schematized in Fig 4.14. Briefly, the network included two modules for the two brain hemispheres and each

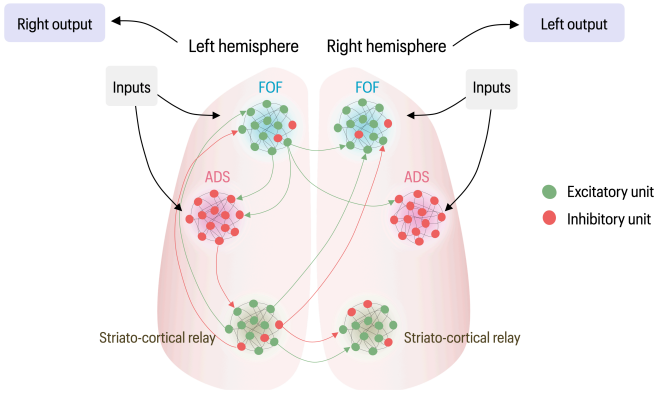


Figure 4.14: Architecture of the multi-region RNN model Schematized model architecture. The model has two modules that represent the two brain hemispheres. Within each module are three submodules that correspond to FOF, ADS and the striato-cortical relay. For clarity, we show long range projections between submodules from just the left module. FOF has two kinds of projection patterns, modeled after IT and PT neurons. The first kind, projects bilaterally to striatum and the other hemisphere, whereas the other kind projects ipsilaterally to striatum. ADS sends inhibitory ipsilateral projections to the striato-cortical relay and the relay projects bilaterally to FOFs and the other hemisphere's striato-cortical relay. Inhibitory projections and units are depicted in red, and excitatory projections and units are depicted in green. All units in the model followed Dale's law. The inputs were sent to all excitatory units in FOF and ADS and the outputs were read out from all units in the two hemispheres.

module in turn consisted of three submodules with 50 all-to-all connected neural units - representing FOF, ADS and the multi-synaptic relay between ADS and FOF (i.e. other basal ganglia nuclei, thalamus, SC, etc.). In the FOF submodules 20% of the neurons were inhibitory whereas in ADSs' 100% of the neurons were inhibitory, approximating the known distribution of E-I neurons in these brain regions. We modeled long-range projection patterns after empirical observations. The FOF submodule's projection patterns followed that of cortical long-range pyramidal neurons (IT and PT neurons), ADS projected unilaterally to the relay module and the relay module projected bilaterally to the two FOFs, completing the recurrent loop. The sensory inputs were sent to both FOF and ADS submodules, since it is known that the auditory cortex projects diffusely to both FOF and ADS in the rat brain. Outputs were read out from all neural units in the FOF and ADS submodules.

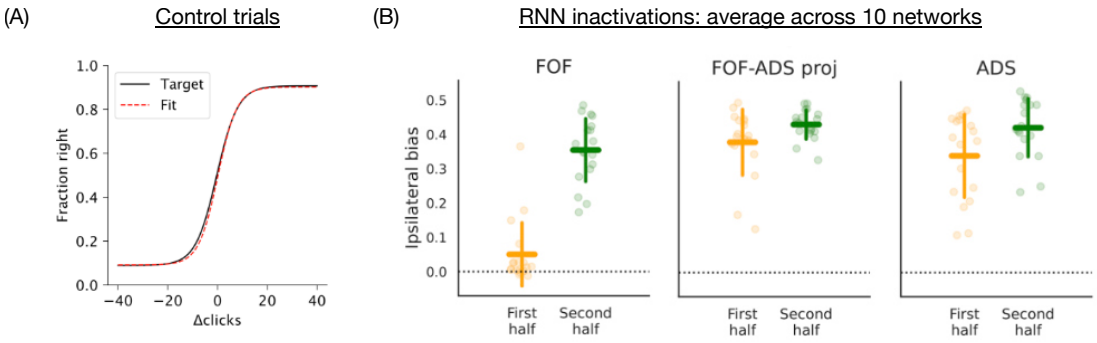


Figure 4.15: Model successfully recapitulates behavior on training and inactivation trials (A) Psychometric curve for model-produced choices (dotted red) overlayed on target choices produced by the Brunton model (black). The model successfully produces target (rat-like) behavior including suboptimalities such as lapses (asymptotes of the psychometric curve). (B) Average bias towards the ipsilateral choices exhibited upon unilateral inactivations of FOF, FOF-ADS projections and ADS (left to right) during either the first or second half of the stimulus. Bars represent the mean $\pm$ CI across 10 trained networks, each dot represents results from inactivations in one of the hemispheres of the trained networks. The network recapitulates empirical data and shows ipsilateral bias only upon inactivation of FOF in the second half. FOF-ADS and ADS inactivations produce ipsilateral bias both in the first and second halves, resembling rat data.

In addition to this multi-region modular structure, our use of the RNN model differs from previous studies in two ways. First, we trained the RNN weights (using backpropagation through time) to perform the evidence accumulation task in a way that matches animal behavior, rather than like an ideal observer. To achieve this, we generated training data by sampling trials from the behavioral model of Brunton et al. (2013). This imparted the network with suboptimalities such as lapses that are commonly observed in rat behavior (see Chapter 3). Second, on a quarter of trials during training, we turned down the gain on weights either within an FOF submodule, ADS submodule or FOF inputs to an ADS submodule during either the first or the second half of the stimulus, essentially mimicking our unilateral inactivation experiments. On these inactivation trials we trained the network to produce an ipsilateral choice, if upon inactivation such a bias was observed empirically. Our use of perturbation data to introduce structure and specialization within a multi-region

RNN model is novel, and complements previous approaches that have relied either on differences in neural representations or just the architecture to lend RNNs area-specific identities (Yang and Molano-Mazón 2021, also see Cowley et al. 2022).

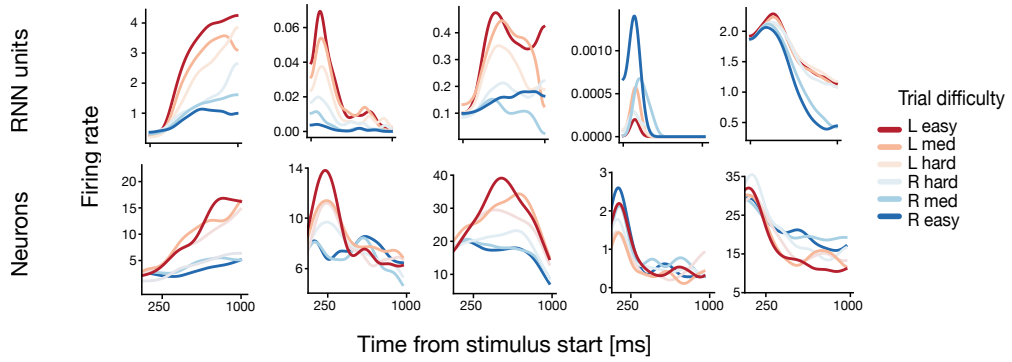


Figure 4.16: Average firing rates of model units bear resemblance to recorded neurons PSTHs (aligned to stimulus onset, conditioned on stimulus difficulty) for example units from the model (top) and recorded data from FOF and ADS (bottom). Both model and real units show variable timecourses and diversity in when the strength of evidence is encoded - transiently or throughout the stimulus.

We trained 10 such networks with different initializations and studied their behavior. Once trained, the model successfully reproduced choice behavior on control (Fig 4.15A) and inactivation trials (Fig 4.15B). In the model - much like the experimental data - unilateral FOF inactivations introduced an ipsilateral bias only in the second half of the stimulus, but FOF to ADS projections and ADS caused a bias throughout the stimulus.

The model responses exhibited several features that resembled neural representations, despite not being explicitly trained to do so. The average firing rate of model units, just like neurons from FOF and ADS had heterogeneous temporally patterned responses (Fig 4.16). Some neurons encoded the evidence strength transiently whereas other encoded it persistently, capturing the diversity observed in recorded neurons from FOF and ADS. Moreover, the timecourse of stimulus and choice decoding were similar across model FOF and ADS, as

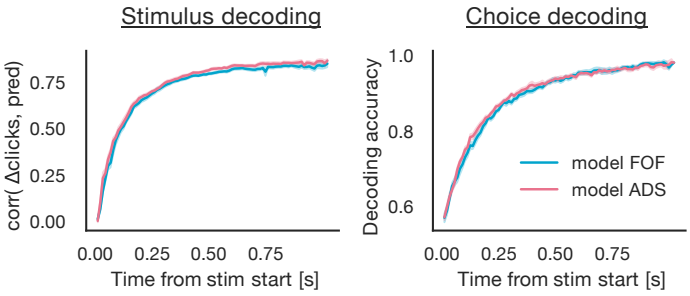


Figure 4.17: Model FOF and ADS have similar timecourse of stimulus and choice decoding (Left) Correlation between the true cumulative difference in number of right and left clicks at any time point during the trial and that predicted by a linear decoder trained on either model FOF activity (aqua blue) or model ADS activity (pink), as a function of time from stimulus onset. Both model areas show similar strength and timecourse of stimulus decoding, resembling our observations from analysis of simultaneous FOF and ADS activity. (Right) Accuracy of a logistic decoder trained to predict the eventual choice from responses of either model FOF (aqua blue) or ADS (pink) during the stimulus. Again, as in the data, the choice decoding accuracy evolves similarly in the two regions, with similar strengths.

we had observed in our population recordings (Fig 4.17, compare to Figs 4.5 and 4.6).

In addition to these similarities in responses, the model was also successful at predicting effects of a perturbation it wasn't trained on. During training, the network was exposed to unilateral inactivations only, so bilateral inactivations of FOF formed a good test of the model's robustness properties. Prior modeling work comparing the robustness of different architectures (Li et al., 2016) has shown that modular and symmetric networks tend to be robust to unilateral inactivations but not bilateral ones. Therefore, we expected the model to deviate from our experimental findings and show disrupted performance during bilateral inactivations of FOF in the first half of stimulus - unless our novel training protocol with perturbation data was successful at instilling some robustness properties of the real FOF-ADS circuitry into the network. To our surprise, much like the data, the model only showed worse performance in the second half of bilateral FOF inactivations, and not in the first half (Fig 4.18, for rat data see Fig 4.13).

We examined the responses of neural units during different inactivations by projecting

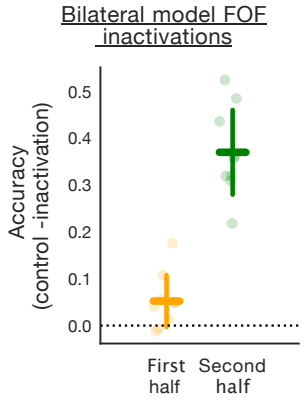


Figure 4.18: Model successfully predicts response to novel perturbations Difference in accuracy between model's choices on control trials and trials in which model FOF was inactivated bilaterally either during the first (orange) or the second (green) half of the stimulus. Bars represent the mean $\pm$ CI across 10 trained networks, dots represent the difference for individual networks. Despite not being trained on bilateral inactivations, the network reproduces the responses observed for these inactivations *in vivo*: network's decision-making behavior is impaired greatly when FOF is inactivated in the second half and minimally in the first half.

population responses from all submodules onto the first three principal components (which capture >85% variance; Fig 4.19). Examination of neural responses on trials with unilateral inactivation of FOF in the first half revealed that after an initial perturbation, the network activity recovers during the second half (Fig 4.19A), facilitating normal decision-making behavior. Selective removal of the three different inputs to an FOF submodule: stimulus, input from the other FOF submodule and inputs from the relay - showed that only the inputs from the relay are important for this recovery (not shown). Such a recovery was absent during inactivations of FOF projections to an ADS submodule (Fig 4.19B), instead these inactivations sent the population responses along the path that normally drives ipsilateral choices and subsequently gave rise to choices biased towards the inactivated side.

In ongoing work we are probing the dynamical structure of these trained models, in unperturbed and perturbed states using the reverse-engineering approach from Sussillo and Barak 2013. We minimized the squared speed of the system's dynamics to find the network's population response patterns (or points in the neural state space) which are more or less stable, such that if the network is allowed to evolve from these patterns with no external input,

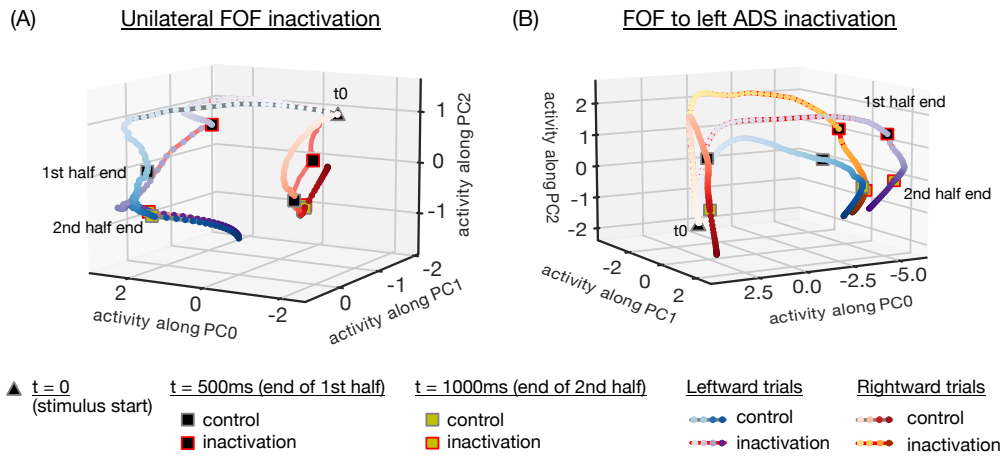


Figure 4.19: Model recovers in response to FOF but not projection specific inactivations Projection of responses of model units onto the first three principal components, showing population trajectories during control leftward (grey lines, blue dots), control rightward (grey lines, purple dots), inactivation leftward (red lines, red dots) and inactivation rightward (red lines, orange dots) trials. Colored squares denote key timepoints during the trial: end of the 1st half (black filled squares), and end of the 2nd half (green filled squares) on control (grey outlines) and inactivation (red outlines) trials. (A) Unilateral left FOF inactivation in the 1st half makes population trajectories deviate from control trajectories at the 1st half end, but trajectories recover to the same endpoint by the 2nd half end. (B) Inactivation of FOF inputs to the left ADS make population trajectories deviate from control at the 1st half end, but these trajectories do not recover correctly and instead end up on the incorrect side at the 2nd half end.

the drift in the network's response is negligible on the timescale of normal operation. Following past work, we refer to these slow points — that reflect the long-term memory of the network — as fixed points. During unperturbed states, the network as a whole possessed a continuum of fixed points that approximated a line attractor (Seung 1996; similar to previous reports Mante et al. 2013; Maheswaranathan et al. 2019), i.e. the fixed points had one stable direction that pointed towards the neighbouring fixed points, and responses along other dimensions rapidly collapsed back onto the fixed points. The approximate line attractor was oriented with respect to the readout such that fixed points at one end of the line attractor strongly drove rightward choices, on the other end strongly drove leftward choices and the remaining fixed points in between interpolated between those two extremes. So, the evidence

was integrated through movement of the population response's projection along the approximate line attractor, with clicks to the right pushing the network along one direction of the line attractor, and clicks to the left in the opposite direction. We are currently investigating how this dynamical structure is changed upon different inactivations, and which structural properties are crucial for the differences.

4.4 Discussion

The two distinct computations required for making decisions based on noisy evidence - gradual accumulation of evidence followed by its categorization to reach a decision - have been respectively mapped onto two distinct brain regions, ADS and FOF (Yartsev et al., 2018; Hanks et al., 2015) in rats. However this implied feedforward functional hierarchy going from ADS to FOF has never been directly evaluated. We sought to fill this important gap by studying simultaneous population recordings from these two regions and the effects of selective silencing of FOF inputs to ADS. The feedforward functional hypothesis posits that ADS has more graded encoding of accumulator value than FOF and therefore predicts that evidence can be decoded from ADS earlier and with higher accuracy compared to FOF. Contrary to this expectation, we found that the two regions do not differ in their encoding of the accumulator value when variability across trials, sessions and animals is controlled for (Figure 4.4). Instead, both regions showed substantial heterogeneity in the tuning of individual neurons to the estimated value of accumulator, with neurons encoding both graded and categorized value of the accumulator (Figure 4.3). In further disagreement with the feedforward hypothesis, we showed that evidence can be decoded from neural populations in both regions with similar accuracy and that this information evolves at comparable timescales in the

two regions (Figure 4.5). Investigating further, through three lines of evidence we demonstrated that these accumulation-related signals are shared between the two regions and are involved in the decision-making behavior. First, we showed that neurally inferred accumulator estimates from the two regions show symmetric cross-correlation with no discernible lead-lag relationship, consistent with a recurrent interaction (Figure 4.8). Second, we found that evidence could be decoded with high accuracy from the functional communication subspace that captures trial-by-trial co-fluctuations between the two regions (Figure 4.9). Finally, we optogenetically silenced FOF axon terminals in ADS during time periods that differentially overlap with the two sub-computations and found that such a manipulation impairs decision-making during all accumulation periods (Figure 4.11). Together, these results indicate that the frontal cortico-striatal circuitry participates in evidence accumulation in a distributed manner, unlike the previously hypothesized compartmentalized feedforward fashion.

The similarities we observed in the encoding of task-relevant features in FOF and its downstream target ADS during decision formation are largely consistent with recent reports in mice in simple decision-making tasks, which found that choice signals in secondary motor cortex (that FOF is a part of) and striatum emerge at indistinguishable time scales (Steinmetz et al., 2019) and that striatum reflects summed activity of its cortical inputs (Peters et al., 2021). Moreover, our results about encoding of sensory evidence in FOF are in agreement with a recent study of dorsal cortex's role during decision-making in mice which reported that secondary motor cortex prominently encodes choice independent evidence-related signals (Orsolic et al., 2021). In the random dots motion task (Roitman and Shadlen, 2002), the analogues to rat FOF and ADS, monkey FEF and caudate respectively have been reported to

carry similar accumulation-related signals early during the stimulus presentation with comparable prevalence of trial-difficulty and choice modulated signals during the stimulus epoch (FEF: Kim and Shadlen 1999, caudate: Ding and Gold 2010, FEF and caudate: Ding and Gold 2012; Ding 2015; Fan et al. 2020). Curiously, these studies reported an earlier emergence of choice modulation in firing rates of individual neurons in caudate compared to FEF (Ding and Gold, 2012). We took the approach of investigating when such a selectivity emerges in the population as a whole, by decoding elapsed evidence from the population firing rates and did not observe differences between the two areas. Our conclusion is also supported by comparing regression kernels of a general linear model with regressors for choice, stimulus pulses etc. fit to individual neurons (not discussed; similar to Park et al. 2014). A number of reasons could underlie these differences: foremost, these recordings were not made simultaneously which leaves open the possibility that concurrent recordings in monkeys will reveal similarities in the timescales with which choice information evolves in the two regions. It could also be that there are indeed species-specific differences, or the fact that our recordings were performed in a variable duration version of the task and not a reaction time version induces subtle differences in strategies and underlying neural mechanisms. Nevertheless, in stark agreement across species, this past work had noted that caudate activity does not converge to a common level of activity just before saccade onset, whereas FEF's does; our recordings from FOF and ADS also show this pattern. When aligned to the end of the stimulus instead of the beginning, we observe a higher convergence to common firing rates for the two choices across trial-difficulties in FOF compared to ADS – ADS neurons continue to show modulation with trial strength (not shown). This similarity in evidence representation and functional role of cortex (FOF) and striatum has been predicted and is therefore

consistent with previous theoretical accounts that have sought to embed the decision-making computation into the cortico-basal ganglia network (Bogacz and Gurney, 2007; Rao, 2010). However similarities with Lo and Wang 2006 are unclear, which predicts transient rather than sustained responses in striatum and models striatum's role as the setting of evidence thresholds. Some recent forays into monkey FEF and caudate (Ding, 2015; Fan et al., 2020; Doi et al., 2020) have proposed differences in representations and roles of these areas when when the environment is nonstationary and the subject has to constantly adapt to changing priors or reward schedules, in other recent work differences in roles of brain regions have been revealed when the subject is not yet proficient and is still acquiring the task (Zhong et al., 2019; Wolff et al., 2022). Therefore, manipulations of these environmental variables and examining these areas during learning might be fruitful endeavors in the future to further understand the role of FOF and ADS in learning and execution of perceptual decisions.

To understand if the evidence representations in FOF and ADS are related, we used CCA to identify dimensions in the neural population space that captured co-fluctuation in the trial-to-trial activity of FOF and ADS over and above those introduced by external task events and defined a functional communication subspace (Hotelling, 1936; Semedo et al., 2020; Kohn et al., 2020; Kang and Druckmann, 2020; Keeley et al., 2020; Semedo et al., 2022). We found that during the stimulus period, the communication subspace carries stimulus-related information such that evidence can be decoded fairly well from just the activity present in the communication subspace. This did not have to be the case, as using similar approaches recent work has identified patterns of activity or information that are shared between brain areas as well as those that are kept private within an area. For instance, using reduced rank regression, a close variant to our approach, it has been reported that majority of the variance in V1 is kept

private and that V1 only selectively shares information related to a few of its activity patterns with V2 (Semedo et al., 2019). In rat M1 and M2 it has been shown that shared dynamics across areas identified using CCA become more related to task performance with learning (Veuthey et al., 2020). Similarly in mice, CCA was recently used to show that the content of communication between A1 and mPFC changes during different contexts based on a control signal from mPFC, this helps select revelant sensory information during contextual decision making (Barbosa et al., 2022). Such communication subspaces defined based on noise correlations do not distinguish between correlations due to direct communication from those induced by communication arriving from common inputs to the two regions. While our projection-specific inactivations support a causal role for the interaction from FOF to ADS, our conclusions about communication of evidence between the two regions can be further strengthened if decrements in evidence-related information are observed upon perturbation of one of the areas, in paired perturbation and recording experiments.

Through specifically targeting the cortico-striatal pathway and silencing FOF axon terminals in ADS, we revealed a role for FOF throughout the gradual accumulation process. This finding might seem at odds with the past work that had concluded using unilateral silencing of FOF somas, that FOF's causal role is restricted to the decision commitment period and does not extend to the accumulation period (Hanks et al., 2015) – however it is in line with recent exciting work highlighting how different projection pathways from the same cortical region can carry different information and differentially impair behavior upon perturbation (e.g. Warden et al. 2012; Znamenskiy and Zador 2013; Li et al. 2015; Murugan et al. 2017; Hwang et al. 2019; Hirokawa et al. 2019; Musall et al. 2021; Tervo et al. 2021; Duan et al. 2021b; Wolff et al. 2022; Park et al. 2022; also see Cox and Witten 2019; Jazayeri and Ostojic

2021 for discussion) producing effects that could be otherwise obscured by whole region perturbations. A notable example of this comes from an auditory frequency discrimination task in rats (Znamenskiy and Zador, 2013), where perturbations of auditory cortex neurons with high (low) frequency tuning did not bias animals' choices towards high (low) responses, but selective activation of striatum projecting auditory cortex neurons with high (low) frequency tuning biased the animals' choices as predicted. Also of note is a recent study in mice (Musall et al., 2021), which has also proposed a role for corticostriatal projections from frontal cortex throughout gradual accumulation of evidence. Our results add an important piece of evidence in favor of the emerging principle that projection-specific perturbations can unmask effects not visible in whole-region perturbations (Warden et al., 2012; Znamenskiy and Zador, 2013; Kim et al., 2017). Given recent controversies about the causal roles of different regions in evidence accumulation (Raposo et al., 2014; Erlich et al., 2015; Katz et al., 2016; Licata et al., 2017; Jeurissen et al., 2022), our approach might offer a way to further clarify the roles of these brain regions and help with identifying the key communication channels operant during decision-making.

These results however do raise the question of why it is that unilateral perturbations of FOF neurons are unsuccessful at disrupting decisions during the accumulation period. We experimentally tested one of the leading explanations from the literature – that the other cortical hemisphere can help with recovery and confer robustness to unilateral cortical perturbations (Li et al., 2016; Chen et al., 2021). However, surprisingly we found that effects of bilateral inactivations of FOF resembled those of unilateral inactivations and did not extend to the early accumulation period (Figure 4.13). Given the vast space of possibilities we turned to computational modeling to gain mechanistic insight into how these differences in

perturbation could arise while FOF and ADS accumulate evidence in a distributed fashion. Therefore, we developed a multi-region RNN model that is constrained by anatomical connectivity and obeys Dale's law (Figure 4.14). We successfully trained it to capture behavior on control and perturbation trials and showed that the model produces evidence decoding profiles that match simultaneous recording data, and replicates novel perturbation effects that it wasn't trained on (Figures 4.15-4.18). The model's dynamics indeed show recovery during unilateral and bilateral perturbations of FOF, but not during perturbations of its projections to ADS. Ongoing work is examining the model's dynamics to understand how accumulation unfolds in perturbed and unperturbed states using methods for reverse-engineering networks (Barak et al., 2013; Mante et al., 2013; Maheswaranathan and Sussillo, 2020; Driscoll et al., 2022).

While the encoding/decoding properties of units in our network resemble those of the empirical data, a complementary approach would be to train the RNN such that individual units in the network additionally match observed neural firing rates (Fisher et al., 2013; Rajan et al., 2016; Andalman et al., 2019; Duncker et al., 2020; Perich et al., 2021). It remains an exciting empirical question of how these different training objectives, algorithms, and constraints alter the dynamics of the network and which objectives make better predictions. Past studies seeking to model the complex web of multi-region interactions that give rise to cognition (Murray et al. 2017; Kleinman et al. 2019; Pinto et al. 2019; Bredenberg et al. 2021; Márton et al. 2020; Mejías and Wang 2022, reviewed in Perich and Rajan 2020; Yang and Molano-Mazón 2021) by training RNNs have relied either on differences in neural representations or just the architecture to lend RNNs area-specific identities. We believe our use of projection-specific perturbation data to introduce structure is novel and is likely to

tighten the correspondence between the models and data (also see Cowley et al. 2022). We are hopeful that this approach of using multi-region RNNs in tandem with projection-specific perturbations and simultaneous recording will offer a powerful framework for generating hypotheses and intuitions about how recurrently connected brain regions perform decision-making in a distributed fashion.

4.5 Methods

4.5.1 Subjects and housing

A total of 19 male Long-evans rats (*Rattus norvegicus* from Hilltop, PA) were used for this study. Of these, 5 animals were used for neural recordings and 14 for optogenetic experiments. Investigators were not blinded to experimental groups during data collection or analyses. Animal use procedures were approved by the Princeton University Institutional Animal Care and Use Committee (IACUC #1853) and were carried out in accordance with National Institute of Health standards.

Rats were housed in Techniplast cages and all training and testing were conducted during the dark cycle. Animals that trained during the day were housed in a 12 hr reverse light cycle room. Animals were pair housed whenever possible, but were always single housed after optic fiber or Neuropixels implantation to prevent damage to the implant. Rats had free access to food but were placed on a controlled water schedule in order to motivate them to participate in behavioral task for water rewards. They obtained water rewards during behavioral training sessions (1-5hrs). Following behavioral training rats received either an unlimited water supplement in an ad lib access period of up to 1hr or a controlled supplement so that their total access to water within a period of 24hrs equalled 3% of their body weight.

4.5.2 Neuronal recordings and analysis

Acquisition and pre-processing

We followed the general surgery and Neuropixels probe implantation methods for recording chronically in freely moving animals from Luo et al. (2020) to record unihemispherically from FOF and ADS of 5 rats. The Neuropixels probes were assembled in a compact 3D printed casing, printed in-house using Formlabs SLA printer. This assembly allowed the probe to be stereotaxically manipulated, provided electromagnetic shielding, prevented it from contacting biological fluids or other adhesive materials applied during surgery and imparted mechanical protection to the implant for robust chronic tethered recordings. We targeted FOF and ADS with one penetration, by implanting the probe at AP 1.9mm and ML $\pm$1.15mm from bregma, at an angle of 15° in the coronal plane. Before implantation, the probe shank was dipped for several seconds in a dye (DiI, 1-2 mg/ml in isopropyl alcohol) to allow for histological reconstruction of probe location. We inserted at least 6mm of the probe shank into the brain; the actual depth of implantation varied between animals.

Electrophysiological recordings were made using either commercially available Neuropixels 1.0 acquisition hardware (Putzeys et al., 2019) or the earliest test-phase IMEC acquisition hardware. We used SpikeGLX software to acquire the data. On every session we recorded from 384 sites arranged in a checkerboard pattern spanning banks 0 and 1 of the probe. The data were automatically spike sorted offline using Kilosort2 with default parameters and then manually created using JRClust's GUI (Jun et al., 2017). During manual curation, each unit detected by Kilosort2 was inspected, if the events comprising the unit had near zero amplitudes or non-physiological waveforms then the entire unit was discarded. If the unit was

judged to be comprised of multiple distinct waveform then it was split into two or more units and finally spatially neighboring units were compared using cross-correlograms, drift patterns and waveform similarity to determine if they should be merged. After following these steps, all units with $< 5\%$ refractory violations were considered as reflecting spikes from a single neuron.

Analysis of neural recordings

For all analyses, we analyzed neurons that had an average firing rate >1Hz during the stimulus period, neural responses were aligned to stimulus onset and included only up to the time of movement onset for each trial. To compute the average population response to different stimulus strengths, we selected neurons that had large differences in their firing rates during the stimulus period for trials that subsequently resulted in rightward versus leftward choices ($|\text{auc} - 0.5| > 0.05, P < 0.05$, Receiver operating characteristic). The choice that produced the larger response was defined as the preferred side. For each neuron, firing rates for individual trials were calculated by smoothing the spike trains with a causal half-Gaussian filter with 0.05s standard deviation, and then normalizing by neuron's mean response at stimulus onset.

Neural tuning curves: Following Hanks et al. (2015), we fit the trial-by-trial behavioral model developed in Brunton et al. (2013) separately to each rats' choices. In order to best describe the behavior on recording sessions and account for nonstationarities, we first fit the model to choices from sessions (30-90) prior to recording. We then used the best-fit parameters and their estimated confidence intervals to impose Gaussian priors on the parameters and

refit the model to behavior from just the recording days. The mean of the Gaussian prior was set to the maximum likelihood estimate of the parameters, and the standard deviation was set to 2 times the estimated confidence interval. Once fit, we obtained the trial-by-trial, moment-by-moment estimate of the likely accumulator variable trajectories while taking into account the choice made by the rat from the model. This is computed by selectively propagating accumulator values that are consistent with rat's choice - at the end of the stimulus - backwards in time.

We collated across trials and used these estimates of the accumulator value to compute the joint probability distribution between accumulator value, time and observed firing rates for each neuron. Following past work, the neural time was assumed to lag behind the time in the model by 100ms. Our results are invariant to small deviations ($50 - 100$ms) from this assumed lag. Firing rates were computed by smoothing the spike trains with a Gaussian filter with 0.75s standard deviation. We marginalized the joint probability distribution over time to extract the response of the neuron given the value of the accumulator. We analysed the time period between 0.15 and 0.5s from stimulus onset and responses of side-selective neurons ($|\text{auc} - 0.5| > 0.075, P < 0.05$). These settings yielded firing rate maps that were largely stable in time for different accumulator values. We accounted for differences in the dynamic ranges of the neural response, by scaling the responses to span the range from 0 to 1. We then fit the estimated relationship between scaled neural responses r and the accumulator value a with a four-parameter sigmoid:

$$r = k_1 + \frac{k_2}{1 + e^{-k_3(a - k_4)}}$$

At zero-crossing, the slope of this equation is given by $k_2 k_3 / 4$. We compared this slope pa-

rameter for individual neurons across the two regions and used the Mann-Whitney U test to determine whether the medians of each region's population distribution were significantly different.

Decoding analyses: We trained linear decoders to predict the cumulative difference in number of right and left clicks (Δclicks) from neural population activity. To be able to compare decoding performance from the two regions, we controlled for differences in the number of recorded neurons by sub-sampling neurons from the region with higher number, while maintaining the overall distribution of side selectivity. We considered time points aligned to stimulus onset, and assumed a neural encoding lag of 100ms. Δclicks during the trial and neural responses were both binned into 50ms bins, neural responses were further smoothed with a Gaussian kernel with 75ms standard deviation. Neural responses were then assembled into a design matrix X with dimensions $T \times N$, where T is equal to the total number of time points (time points from stimulus onset to termination $\times$ number of trials), and N is the number of neurons. Similarly, Δclicks were concatenated across trials to yield Y with dimensions $T \times 1$. We fit one set of regression weights β_{st} for all time points such that $Y = X\beta_{st}$ and imposed L1 penalty. We performed 10-fold cross-validation to assess the decoder's performance (reported in Fig 4.5). Within each fold the strength of regularization was estimated independently using 10-fold cross-validation. Fitting was performed using the sklearn package in Python (Pedregosa et al., 2011). To determine if the decoding performance or its timecourse was different between the two regions, we performed a two-way repeated measures ANOVA with time (time bins from stimulus onset) and area (FOF, ADS) as factors using Pingouin package in Python (Vallat, 2018).

To decode binary variables, choice and past trial's choice we used logistic decoders. We assembled the design matrix similarly, however without any neural lags since the target variables are constant for a trial. We duplicated the eventual or past choice for all time bins from a trial to construct the Y variable. We fit the decoders with L2 penalty, once again with 10-fold cross-validation. Again, we used sklearn package in Python for fitting (Pedregosa et al., 2011). While estimating the decision variables or DVs using the decoders trained to predict choice, we binned the neural activity into 1ms bins and convolved them with a Gaussian kernel with 25ms standard deviation before projecting onto the decision hyperplane. These parameters were picked based on recovery analysis on simulated synthetic datasets. However, our results are not sensitive to small variations in these parameters.

We measured the similarity between the decoding axes or vector of inferred regression weights β_i where $i \in \{ch, st, h\}$ by first computing the corresponding unit vectors $\hat{\beta}_{ch}, \hat{\beta}_{st}$ and $\hat{\beta}_h$ and then calculating the pairwise angles between them in degrees:

$$\theta_{ij} = \frac{180}{\pi} \cos^{-1} |\hat{\beta}_i^\intercal \hat{\beta}_j|$$

In this formulation, an angle of 0° indicates that the axes are perfectly aligned, whereas an angle of 90° indicates that the axes are orthogonal or maximally unrelated. We determined if the angles between decoding axes were different between the two regions using Wilcoxon signed rank test.

Canonical correlation analysis (CCA): We used CCA to define dimensions in the neural population space of FOF and ADS which reflect their shared dynamics. In a paired multivariate dataset such as this one, CCA finds a pair of dimensions – one for each set of ob-

servations – such that linear projections of observations onto these dimensions are maximally correlated across the two datasets (Hotelling, 1936). The linearly projected observations are often referred to as canonical variables and the corresponding correlations as canonical correlations. Much like principal component analysis (PCA), CCA can be used to define multiple such dimensions, such that each subsequent pair of canonical variables is orthogonal and less correlated than the previous. So for paired datasets, $\mathbf{X_a} \in \mathbb{R}^{t \times n_a}$ and $\mathbf{X_b} \in \mathbb{R}^{t \times n_b}$ where t represents the number of data points, and n_a, n_b denote the number of observed dimensions in each dataset, CCA finds $r \leq \min(n_x, n_y)$ linear dimensions, $\mathbf{w}_a^{i*} \in \mathbb{R}^{n_x}$ and $\mathbf{w}_b^{i*} \in \mathbb{R}^{n_y}$ for $i \in \{1, 2, \cdots, r\}$ by maximizing the objective function:

$$\mathbf{w}_a^{i*}, \mathbf{w}_b^{i*} = \underset{\mathbf{w}_a, \mathbf{w}_b}{\operatorname{argmax}} \operatorname{corr}(\mathbf{X}_a \mathbf{w}_a^i, \mathbf{X}_b \mathbf{w}_b^i)$$

$$(\mathbf{X}_a \mathbf{w}_a^i)^\top (\mathbf{X}_a \mathbf{w}_a^j) = 0, \quad (\mathbf{X}_b \mathbf{w}_b^i)^\top (\mathbf{X}_b \mathbf{w}_b^j) = 0 \quad \forall j \neq i : i, j \in \{1, 2, \cdots, r\}$$

subject to the additional constraints that the canonical variables should have unit length, i.e. $\|\mathbf{X}_a \mathbf{w}_a^i\|_2 = 1$ and $\|\mathbf{X}_b \mathbf{w}_b^i\|_2 = 1$. Multiple solutions have been proposed to solve for the projections $\mathbf{w}_a$ and $\mathbf{w}_b$, we used the solution proposed by Hotelling 1936 which solves for the dimensions by performing singular value decomposition on a correlation matrix.

We used CCA on the residuals from a general linear model (GLM) fit individually to each neuron's firing rates in order to remove signal-related correlations. The GLM described neuronal firing rates during a period of 2.25s aligned to the initiation of a trial, as a combination of temporal kernels aligned to the time of each task event. We included kernels for trial initiation, left and right clicks, previous trial's choice, current trial's choice, reward and punishment. Kernel for choice was allowed to be acausal, whereas all other kernels were causal in

their timing. Following past work (Park et al., 2014; Bondy et al., 2020), to impose smoothness onto the temporal kernels, we parameterized each of these kernels as made of raised cosine bases and fit the weights with ridge penalty using cross-validation with sklearn in Python (Pedregosa et al., 2011). We fixed the hyper-parameters, such as number and duration of these kernels by measuring cross-validation performance on a small grid. We binned the neural activity into 50ms bins, smoothed with a Gaussian kernel of 100ms standard deviation for this analysis. GLM captured the firing rates well with a mean cross-validated R^2 of 0.53 ± 0.04. We used residuals only during the stimulus period to define the communication subspace using CCA. To determine the dimensionality of the communication subspace, we first computed the correlation for each axis on held out test datasets using 10-fold cross validation and picked the number of dimensions which showed significant correlations (> 95 percentile) compared to a distribution of correlations estimated from CCA performed on trial-shuffled datasets (10^3 shuffles). All datasets had $1 - 8$ significant dimensions. For decoding from the communication subspace, procedures identical to full population space (section above) were followed.

4.5.3 Optogenetic perturbations and analysis

Optical fiber chemical sharpening

We followed methods similar to (Hanks et al., 2015) and used standard off the shelf 50-125μm FC-FC duplex fiber optic cables (FiberCables.com). The metal casing of the connector and the outer protective layers of the cable were removed, yielding a 1.5cm of fiber optic cable with inner plastic coating (typically clear) intact. To etch the fiber tip, 2-2.5mm of the fiber tip was submerged in a 5ml Eppendorf tube containing 48% hydrofluoric acid,

topped with mineral oil. Over the course of 17mins, the fiber tip was slowly pulled out of the acid using a motor (Narishige) producing a taper in the inner plastic coating. Then the speed of the motor was increased and the motor was run until the tip was entirely removed from the acid (typically 13mins), producing a sharp fiber tip, usually with uniform and broad light scatter. We measured the laser beam transmission efficiency of the etched fibers using an integrating sphere photodiode power sensor (Thor Labs). Fibers that did not produce sufficiently broad or uniform scatter or had efficiency $< 85\%$ were discarded.

Optogenetic virus injection and fiber implantation

We followed previously described procedures for general surgery and virus injections (Erlich et al., 2011; Hanks et al., 2015); here we describe procedures specific to this study. We bilaterally injected adeno-associated virus, either AAV5-CaMKIIa-eYFP-eNpHR3.0 or AAV2/5-CamMKIIa-EYFP-WPRE-hGHpA using a Nanoject (Drummond Scientific) in FOF. Five closely spaced injection tracts were made in each FOF's craniotomy. The tracts were typically arranged $500\mu m$ apart at the 4 vertices and the center of a cross spanning AP, ML axes, centered at AP 2mm ML $\pm$1.3mm. In each tract, 16 injections of $14.1nL$ were made every $100\mu m$ in depth, starting at $200\mu m$ below brain surface i.e. from 0.2-1.8mm. Virus was expelled at $20nL/s$ and the injections were made once every $10s$. At the final injection in a tract the pipette was left in place for at least 2min before removal. In each craniotomy a total of $1.128\mu l$ of virus was injected over a 30min period consisting of 80 separate injections.

For bilateral FOF silencing, chemically sharpened optical fibers were then lowered down the central injection tracts to a depth of 1mm. For inhibition of FOF terminals in ADS we made additional craniotomies centered over ADS (AP 1.9mm, ML $\pm$2.4mm) and lowered

etched fibers to a depth of 3.5-3.9mm. The craniotomies were filled with kwik-sil (World Precision Instruments) and the fibers were secured to the skull with viscous dental composite (Absolute Dentin, Parkell). Optical fibers were then enclosed in a custom 3D printed casing and the implant was filled with dental acrylic, allowing only the FC connectors to protrude. Before behavioral experiments, viral constructs were allowed to develop for 6-8 weeks for FOF silencing, and for 10-12 weeks for terminal silencing. Accurate viral injection targeting and expression was verified histologically.

Optogenetic perturbation

For optogenetically perturbing neural activity during behavior, a fiber rotary joint (Princetel) was mounted on the ceiling of the behavioral chamber. In the beginning of a session, the FC connectors in the rat's implant were connected to the patch cables attached to the fiber rotary joint. The rotary joint was connected to the laser beam source. For FOF inactivations, the laser beam from a 200mW, 532nm laser (OEM Laser Systems) was used. To deliver illumination bilaterally, the laser beam was split into two beams of roughly equal power ($\sim$ 25mW) using a beam splitter (Doric). An OBIS 594nm (Coherent) laser beam of power 25-33mW was used for silencing of FOF axon terminals in either left or right ADS (unilateral silencing). Illumination was delivered continuously on a subset of trials (25%) by opening and closing a shutter with a 5V TTL pulse.

In the inactivated trials, light was delivered during one of three different epochs. The first epoch was 2000ms long, starting on trial initiation and ending 500ms after the termination of the stimulus. We call this "whole-trial" inactivation and this type of stimulation was delivered to a cohort of 6 rats (10 hemispheres). The second and third epochs were 500ms long and

spanned the first and second halves respectively of a 1s long auditory stimulus. We refer to them as "early" and "late" inactivation epochs respectively. Early and late inactivations were done in sessions separate from the whole trial inactivations in 5 rats (8 hemispheres; 3 of these rats also belonged to whole-trial inactivation cohort). In sessions with early and late inactivations, 50% of the trials had 1s long stimulus. Duration of the stimulus on rest of the trials was uniformly distributed between 0.2-1s. On half of these 1s long stimulus trials, light was delivered either during early or late epoch, and the other half served as control trials.

We measured behavioral bias resulting from optogenetic inactivation by following methods used in Hanks et al. (2015) and Yartsev et al. (2018). First, we binned the trials on the basis of stimulus strength and for each of the 10 binned stimulus strengths, we computed the percentage of trials during which the rats chose the side ipsilateral to the optically stimulated side on control and inactivation trials. The mean difference across stimulus strengths in this 'go ipsi' rate between inactivation and control trials measured the bias caused by optogenetic inactivation. A positive bias represents an increase in ipsilateral choices upon laser stimulation. Confidence intervals and statistical comparisons for this bias parameter were calculated using nonparametric bootstrap procedures.

4.5.4 Multi-region recurrent neural network model

We trained an RNN with biological constraints on connectivity to perform the evidence accumulation task from Brunton et al. 2013. The RNN is composed of standard firing-rate units ($N = 300$) with a rectifying nonlinearity and $\tau = 20ms$ as the time constant of decay

of network units. RNN dynamics are governed by the equations:

$$\tau\dot{\mathbf{x}} = -\mathbf{x} + W^{rec}\mathbf{r} + W^{in}\mathbf{u} + \sqrt{2\tau\sigma^2_{rec}}\xi \qquad \text{where} \quad \mathbf{r} = [\mathbf{x}]_+$$

where the variable $\mathbf{x}(t)$ is an N dimensional vector containing the activation of neural units in the network. These activations are mapped to the corresponding firing rates $\mathbf{r}(t)$ by passing $\mathbf{x}$ through the threshold linear nonlinearity which maps individual activations to positive firing rates: x_i if $x_i > 0$ and 0 otherwise for $i = 1, ..., N$. Recurrent connection weights in the network are given by the $N \times N$ matrix W^{rec} and the connection weights from the inputs to network units are given by the $N \times N^{in}$ matrix W^{in}. The network receives 4-dimensional ($N^{in} = 4$) time-varying inputs, $\mathbf{u}(t) = [s_L(t) \quad s_R(t) \quad h(t) \quad l(t)]^\top$, which represent the left clicks, right clicks, trial-history and inattention inputs respectively (see section below for details). Noise intrinsic to the network is represented by $N \times 1$ dimensional $\xi(t)$, independent white noise with zero mean and unit variance. The network output $z(t)$ is read out linearly, as a weighted sum of the firing rates of network units, with weights W^{out} and a bias b:

$$z = W^{out}\mathbf{r} + b$$

Network inputs and targets

The network received 4 inputs $\mathbf{u}(t) = [s_L(t) \quad s_R(t) \quad h(t) \quad l(t)]^\top$ at every time step. The first two inputs — $s_L(t)$ and $s_R(t)$, correspond to the randomly timed discrete auditory evidence that is presented to the rats during the task, from speakers to their left and right sides respectively. These inputs are zero except when a left/right click occurrs. The click times were sampled from Poisson processes, with Poisson rates for left/right inputs sampled inde-

pendently on each trial, from the same distribution as the one used for rats. At the times when left clicks occurred the average input was equal to -1 , however we added independent Gaussian noise η with zero mean and variance $\sigma_S^2 = 2$ to each click, hence the input for any given left click was $-(1+\eta)$. σ_S^2 represents the variance of sensory noise accompanying each sensory input, and was set to match the average sensory noise variance measured in rats performing this task (Brunton et al., 2013). Similarly, when a right click occurred, the magnitude of input s_R was set to $1+\eta$. The overall timing of these stimulus inputs mirrored that of the real task. On each trial, the stimulus duration i.e. the time during which clicks played (in s) was sampled uniformly from the interval $[0.2, 1.0]$ and was preceded by a variable delay such that the stimulus always ended at 1.5s from the trial start.

The other two inputs are related to the internal state of the animal and are included to introduce suboptimalities in the network's behavior. The third input $h(t)$ represents trial history. Its value was set to either $+2$ or -2, for the first 5 time steps of a trial and was 0 everywhere else. The positive (negative) value indicated that the bias due to trial history is towards the right (left) choice. We included this input because trial history influences animal's behavior (Chapter 3) and was strongly encoded in both FOF and ADS (Fig 4.6). For the results presented in this chapter currently, this input was turned off. The fourth input $l(t)$ indicated if the network's choices were to be informed by the stimulus or made randomly (inattentive lapses). On trials when the network was to choose randomly this input's value was set to 1 throughout the trial, and was zero otherwise.

Choice targets for each trial were obtained by passing stimulus, history and lapsing inputs to the behavioral model from (Brunton et al. 2013, also see Methods section in Chapter 3). If the target choice was towards right (left), then the target was set to $+1$ (-1) in the 200ms

following termination of stimulus, i.e. from 1.5 to 1.7s. On a quarter of trials, we turned down the gain of certain recurrent weights in W^{rec} by a factor of 0.1 so as to simulate unilateral inactivations of FOF, ADS or FOF inputs to ADS during either the first or second half of a 1s long stimulus. On these trials, if an ipsilateral bias was observed experimentally then the choice targets were not obtained from the model but instead always set to the ipsilateral side.

Network connectivity and training

We engineered the recurrent weight matrix W^{rec} so as to capture the major anatomical features of the FOF-ADS circuitry. The network included two modules for the two brain hemispheres (150 units each) and each module in turn consisted of three submodules with 50 all-to-all connected neural units - representing FOF, ADS and the multi-synaptic relay between ADS and FOF (i.e. other basal ganglia nuclei, thalamus, SC, etc.). All units in the network followed Dale's law. In other words, all recurrent projections from a unit were either positive or negative but never both. We refer to the units which send out positive recurrent projections as excitatory units and the ones with negative projections as inhibitory units. Within each submodule, we set the ratio of excitatory to inhibitory units to match the known distribution of E-I neurons in these brain regions. Therefore, FOF submodules had 20% inhibitory units and ADS had 100%. For the relay submodule we set the ratio to 30%.

In FOF submodules, inhibitory units projected locally within the submodule, and excitatory units projected to other submodules in a manner aiming to mimic the long-range projections of cortical pyramidal neurons (Dudman and Gerfen, 2015). About 30% of the excitatory neurons from an FOF submodule followed the pattern of projections observed in

cortical IT neurons - they sent projections to the other FOF submodule and the two ADS submodules. Another 30% followed projection patters of cortical PT neurons - they sent projects to the ipsilateral ADS submodule. The remaining FOF excitatory units only had local recurrent connections. The ADS submodules projected ipsilaterally to the relay module, since we are not aware of any cross-hemipshere projections between the two hemispheres. We do not differentiate units in ADS as belonging to D1/D2 subtypes etc. Given the abstract nature of the relay submodule, we put minimal constraints on its recurrent connections and it was allowed to project bilaterally to the two FOFs and to the other relay submodule with both excitatory and inhibitory projections.

To initialize W^{rec}, we first sampled a matrix of values from a normal distribution with zero mean and variance $1/N$ while enforcing the connectivity patterns as detailed above. Then we balanced the average weights of excitatory and inhibitory inputs into each unit and set the spectral radius to 1.3.

All units in the FOF and ADS submodules received all 4 inputs i.e. W^{in} projected only to FOF and ADS units, and outputs were read out from all units in the network. Input and outputs weight were initialized using samples from Glorot normal distributions (Glorot and Bengio, 2010).

While training, to enforce the Dale constraints on connectivity we parameterized the recurrent weight matrix W^{rec} as a product of a non-negative matrix $W^{rec,+}$ and a diagonal matrix D of 1s and -1s identifying excitatory and inhibitory units i.e. $W^{rec} = W^{rec,+}D$. We maintained the sparse connectivity across submodules by further performing element-wise multiplication with a connectivity mask M^{rec}, such that $W^{rec} = M^{rec} \odot (W^{rec,+}D)$. We simulated the network's dynamics using Euler updates with $\Delta t = 2$ms and generated

$240,000$ trials for this, with minibatches of 64 trials each. We optimized W^{in}, W^{rec}, W^{out}, b and $x(0)$ to minimize the mean-squared error (MSE) between the network output and targets, using Adam (Kingma and Ba, 2017) in TensorFlow (Abadi et al., 2015). The methods used for analysis of activity and perturbation response of the multi-region models were identical to that used for rat data.

5

Conclusions and outlook

A major quest for neuroscience is to understand the computations and neural mechanisms that give rise to cognitive behavior, of which making well-considered decisions based on sensory evidence is a fundamental building block. Through cleverly designed studies and theoretical model development tremendous progress has been made towards understanding the many decision-making strategies, their merits and objectives as well as the neural correlates and mechanisms that underlie them (Gold and Shadlen, 2007; Hanks and Summerfield, 2017). Excitingly, new methodologies such as large-scale behavioral and neural measurements, and precise neural perturbations have revealed unexpected complexity at multiple levels (Najafi and Churchland, 2018) – these range from nonstationarities in decision-variables, to variability and heterogeneity in choice behavior across individuals, to redundancy in func-

tional role and compensation between brain regions. In this thesis, we sought to address these emerging challenges through three studies and we summarize our findings below.

For decades, decision-making studies have used patterns of history dependence in choices to infer ongoing learning and heuristic strategies, and consequently argued for or against optimality of behavior (Yu and Cohen, 2009; Gardner, 2019; Rahnev et al., 2020). Recent work however, has raised the concern that such inference may be confounded by slow, random drifts in decision variables, leading to potentially erroneous conclusions (Fründ et al., 2011; Lak et al., 2020a; Roy et al., 2021). This is particularly relevant to large behavioral datasets collected over long timescales, where the effects of these drifts become pronounced. A recent study, Lak et al. 2020a considered the challenge posed by drifts in one key decision variable, the criterion, i.e. the threshold for choosing one alternative over the other. They showed that a drifting criterion produces an apparent history dependence that mimics the signatures of updates from an active learning process. Hence, they proposed a correction to remove the influence of such drifts, an approach that is being widely adopted by multiple groups (The International Brain Laboratory et al., 2021; Mendonça et al., 2020; Masset et al., 2020).

In Chapter 2, we show that this correction fails in the presence of a wide range of commonly assumed updating strategies. It fails to selectively remove the influence of slow drifts, thereby misidentifying a variety of updating strategies as belonging to a narrow subset. In contrast to this "model-independent" correction, we advocate for an approach that explicitly models the dynamics of both slow drifts and active updating of decision variables. We demonstrate the success of this model-based approach on real and synthetic datasets. Our method offers many advantages: it is accurate for a large spectrum of updating strategies, it

yields a faithful estimate of the latent trajectory of drifts which could be leveraged for investigating neural correlates or sources of these fluctuations, and the general framework is flexible and can accommodate additional assumptions. In conclusion, our work urges caution in applying the correction proposed by Lak et al. 2020a and offers a powerful alternative approach to account for random drifts in decision criterion. This work should be of interest to human and animal researchers studying learning, decision-making and fluctuating internal states. Future work could explore the use of such model-based approaches to better track neural correlates of decision variables in the presence of slow fluctuations, or perhaps even investigate the neural correlates and origins of these fluctuations themselves.

Even in simple tasks, subjects with significant experience often deviate from the optimal policy. A prominent example of this shows up in perceptual decision-making tasks where subjects often assume dependencies between successive trials even if they are independent, and consequently show history biases in their decisions (e.g. Busse et al. 2011; Fründ et al. 2011). In addition, on a sizeable fraction of trials subjects seem to choose stochastically, independent of the evidence - such errors are called lapses (Wichmann and Hill, 2001; Carandini and Churchland, 2013). These suboptimalities have traditionally been considered distinct however several studies have noted that history biases and lapses covary in prevalence during learning and at asymptotic performance, and their proposed neural substrates overlap. What cognitive process could underlie the links between these suboptimalities? In Chapter 3, we explored the idea that history biases may reflect a misbelief about non-stationarity in the world (Yu and Cohen, 2009), and demonstrated that normative decision-making under such beliefs gives rise to both history-dependence and choices that appear to be evidence-

independent. This corresponds to an accumulation to bound process with history dependent initial state updates.

We fit this model to a large dataset of rats trained on an auditory decision-making task. Despite the heterogeneity in history biases and lapse rates in this population, we show that the constraints posited by the model are obeyed in this dataset. Further, our model makes predictions about the time it takes to make decisions. We test these predictions in a rat dataset with reaction time reports, and show that it captures patterns of choices, reaction times and their history dependence. Given the widespread prevalence of lapses and history biases in animal behavioral data in both healthy and perturbed brain states, this model will be of broad interest to researchers looking for principled explanations for these suboptimalities, ways to intervene on the behavior to manipulate them and methods to improve estimation of decision variables in their presence. Moreover, since this model produces apparent lapses through deterministic trial-by-trial updating, it may provide a dynamical account of the processes underlying quasi-stochastic behavior, such as exploration.

A fruitful future endeavor could involve jointly modeling behavioral and neural variability to further increase predictability - for instance, using neural measurements from decision-making areas to constrain and improve our estimates of initial points, accumulator trajectories and ultimately different kinds of errors in choices. Such procedures to jointly fit neural and behavioral data are being developed by BrianDePasquale (DePasquale et al., 2021) and Thomas Luo in the Brody lab. Future work could also achieve stronger comparisons between different mechanistic models of these suboptimalities by carefully manipulating key task features, such as trial sequences and rewards, in ways that differentially affect different processes. Such manipulations could help partition out inattentive, exploratory and appar-

ent lapses with higher confidence, as well as distinguish variability due to initial point updates from other forms of updating (such as updates of drift rates or decision bounds).

Decision-making is often conceptualized as a sequence of sub-computations, that are then mapped onto specialized modules within neural circuitry, chained together in an effectively feedforward fashion. Such a conceptualization has been successful in describing the transformations in the primate visual system (Serre et al., 2007; Yamins et al., 2013) and has been extended to the circuitry underlying perceptual decision-making in the face of noisy sensory evidence. In Chapter 4 we test such a proposal in rats, where the anterior dorsal striatum (ADS) and the frontal orienting fields (FOF) have been respectively mapped onto two theoretically defined sub-computations – gradual accumulation of noisy evidence followed by thresholding to commit to a choice. ADS is known to represent the graded evidence and be causally required throughout the accumulation process (Yartsev et al., 2018), whereas FOF is known to represent the categorical choice and be causally necessary only at the very end of accumulation period during decision commitment (Erlich et al., 2015; Hanks et al., 2015; Piet et al., 2017). These observations therefore provide an intriguing neural implementation of the decision process, one in which ADS and FOF form a functional feedforward hierarchy such that evidence is accumulated in ADS and later thresholded in FOF.

Using simultaneous multi-region population recordings and projection-specific perturbations we provide evidence against this modular, feedforward hypothesis. We show that ADS does not have privileged information about evidence, rather both regions carry similar amounts of information that evolves at comparable timescales. By analyzing the co-fluctuations in simultaneous single trial activity in ADS and FOF, we demonstrate that these accumulation-

related signals are shared between the two regions. Further, we reveal an unexpected role for FOF-ADS interactions during gradual accumulation of evidence using projection-specific perturbations. We show that unlike inactivations of FOF neurons, selective silencing of FOF projections to striatum impairs decision-making throughout the accumulation process and not just during choice formation. We revise the feedforward hypothesis and advocate for an alternate, distributed and recurrent implementation of evidence accumulation in the cortico-striatal circuitry. We instantiate this proposal using a multi-region recurrent neural network which reproduces our results from perturbation and recording experiments.

Future studies could gain even more fine-grained insight into the interactions between regions by modeling the low-dimensional latent dynamics in these regions and their influence on each other (Zoltowski et al., 2020; Glaser et al., 2020). Indeed, we are pursuing this approach in collaboration with David Zoltowski and Jonathan Pillow. Another avenue for future work could be to test the predictions of the RNN model, by simultaneously recording from regions (such as FOF) and their downstream targets (such as ADS) during precise optogenetic perturbations using opto-ephys approaches. Such experiments could help confirm the expected effects of recovery and/or the expected loss of information downstream, and help further revise or constrain the model if these expectations are not met. Another potential use case of the RNN in future endeavors, would be to estimate uncertainty about different parameters (such as connection weights; Fortunato et al. 2019) and use this to guide follow-up experiments, prioritizing experiments that might be most informative in reducing this uncertainty. The RNN modeling framework could also be fruitfully extended to include other nodes in the circuitry that are known to be important for decision-making. In rats, these include anterior parts of medial prefrontal cortex (aM2; Luo et al. 2021), posterior parts

of striatum (pStr; Bondy et al. 2020; Pisupati et al. 2021) and other cortical regions such as the posterior parietal cortex (PPC; Raposo et al. 2014; Erlich et al. 2015; Hanks et al. 2015). All of these regions have been known to encode evidence, choice or trial history, RNN modeling in combination with projection-specific approaches could offer insights into how these nodes interact to produce decision-making behavior. At least two other subcortical structures - the superior colliculus (SC) and thalamus - are known to be involved in this decision-making network, and there are hints in the literature that both of these regions might support robustness or recovery in frontal cortex in related tasks (Kopec et al., 2015; Guo et al., 2017) making them promising candidates to explore in evidence accumulation tasks using this approach. Finally, the RNN could be modified to include more biologically plausible forms of training (e.g. Murray 2019) allowing us to extend our modeling approach to learning stages of the task. This could help reveal differences between regions, their projections and perturbations that are visible during learning but not during expert performance (Zhong et al., 2019; Wolff et al., 2022). The expressive nature of such structured neural network models offers us one possible model class that can match the complexity of modern systems neuroscience - of both its experimental manipulations as well as the dynamics of the underlying neural substrate - while retaining tractability in fitting experimental data. Ultimately, we hope that such approaches will yield insight into the nuanced relationship between decision-making computations and the algorithmic and neural logic that implements them in the brain.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X. (2015). TensorFlow: A system for large-scale machine learning. page 21.

Abbott, A. (2010). Neuroscience: The rat pack. *Nature*, 465(7296):282–283. Number: 7296 Publisher: Nature Publishing Group.

Abrahamyan, A., Silva, L. L., Dakin, S. C., Carandini, M., and Gardner, J. L. (2016). Adaptable history biases in human perceptual decisions. *Proceedings of the National Academy of Sciences*, 113(25):E3548–E3557. Publisher: National Academy of Sciences Section: PNAS Plus.

Akrami, A., Kopec, C. D., Diamond, M. E., and Brody, C. D. (2018). Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature*, 554(7692):368–372. Number: 7692 Publisher: Nature Publishing Group.

Andalman, A. S., Burns, V. M., Lovett-Barron, M., Broxton, M., Poole, B., Yang, S. J., Grosenick, L., Lerner, T. N., Chen, R., Benster, T., Mourrain, P., Levoy, M., Rajan, K., and Deisseroth, K. (2019). Neuronal Dynamics Regulating Brain and Behavioral State Transitions. *Cell*, 177(4):970–985.e20.

Arlt, C., Barroso-Luque, R., Kira, S., Bruno, C. A., Xia, N., Chettih, S. N., Soares, S., Pettit, N. L., and Harvey, C. D. (2022). Cognitive experience alters cortical involvement in goal-directed navigation. *eLife*, 11:e76051. Publisher: eLife Sciences Publications, Ltd.

Ashwood, Z. C., Roy, N. A., Stone, I. R., Urai, A. E., Churchland, A. K., Pouget, A., and Pillow, J. W. (2022). Mice alternate between discrete strategies during perceptual decision-making. *Nature Neuroscience*, 25(2):201–212. Number: 2 Publisher: Nature Publishing Group.

Bak, J. H., Choi, J. Y., Akrami, A., Witten, I., and Pillow, J. W. (2016). Adaptive optimal training of animal behavior. *Advances in Neural Information Processing Systems*, 29:1947–1955.

Bak, J. H. and Pillow, J. W. (2018). Adaptive stimulus selection for multi-alternative psychometric functions with lapses. *Journal of Vision*, 18(12):4–4. Publisher: The Association for Research in Vision and Ophthalmology.

Barak, O., Sussillo, D., Romo, R., Tsodyks, M., and Abbott, L. F. (2013). From fixed points to chaos: Three models of delayed discrimination. *Progress in Neurobiology*, 103:214–222.

Barbosa, J., Proville, R., Rodgers, C. C., Ostojic, S., and Boubenec, Y. (2022). Flexible selection of task-relevant features through across-area population gating. Pages: 2022.07.21.500962 Section: New Results.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4):700–765. Place: US Publisher: American Psychological Association.

Bogacz, R. and Gurney, K. (2007). The Basal Ganglia and Cortex Implement Optimal Decision Making Between Alternative Actions. *Neural Computation*, 19(2):442–477.

Bolkan, S. S., Stone, I. R., Pinto, L., Ashwood, Z. C., Iravedra Garcia, J. M., Herman, A. L., Singh, P., Bandi, A., Cox, J., Zimmerman, C. A., Cho, J. R., Engelhard, B., Pillow, J. W., and Witten, I. B. (2022). Opponent control of behavior by dorsomedial striatal pathways depends on task demands and internal state. *Nature Neuroscience*, 25(3):345–357. Number: 3 Publisher: Nature Publishing Group.

Bondy, A. G., Luo, T. Z., Gupta, D., and Brody, C. D. (2020). Anterior and posterior striatum play distinct roles in evidence accumulation. Denver, CO.

Bredenberg, C. J., Savin, C., and Kiani, R. (2021). Recurrent neural circuits overcome partial inactivation by compensation and re-learning. Technical report. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results Type: article.

Brody, C. D. and Hanks, T. D. (2016). Neural underpinnings of the evidence accumulator. *Current Opinion in Neurobiology*, 37:149–157.

Brunton, B. W., Botvinick, M. M., and Brody, C. D. (2013). Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science*, 340(6128):95–98. Publisher: American Association for the Advancement of Science Section: Report.

Busse, L., Ayaz, A., Dhruv, N. T., Katzner, S., Saleem, A. B., Schölvinck, M. L., Zaharia, A. D., and Carandini, M. (2011). The Detection of Visual Contrast in the Behaving Mouse. *Journal of Neuroscience*, 31(31):11351–11361. Publisher: Society for Neuroscience Section: Articles.

Calhoun, A. J., Pillow, J. W., and Murthy, M. (2019). Unsupervised identification of the internal states that shape natural behavior. *Nature Neuroscience*, 22(12):2040–2049. Number: 12 Publisher: Nature Publishing Group.

Carandini, M. and Churchland, A. K. (2013). Probing perceptual decisions in rodents. *Nature Neuroscience*, 16(7):824–831. Number: 7 Publisher: Nature Publishing Group.

Carnevale, F., de Lafuente, V., Romo, R., Barak, O., and Parga, N. (2015). Dynamic Control of Response Criterion in Premotor Cortex during Perceptual Detection under Temporal Uncertainty. *Neuron*, 86(4):1067–1077.

Chen, G., Kang, B., Lindsey, J., Druckmann, S., and Li, N. (2021). Modularity and robustness of frontal cortical networks. *Cell*, 184(14):3717–3730.e24.

Cho, R. Y., Nystrom, L. E., Brown, E. T., Jones, A. D., Braver, T. S., Holmes, P. J., and Cohen, J. D. (2002). Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. *Cognitive, Affective, & Behavioral Neuroscience*, 2(4):283–299.

Cohen, M. R. and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nature Neuroscience*, 14(7):811–819. Number: 7 Publisher: Nature Publishing Group.

Cowley, B. R., Calhoun, A. J., Rangarajan, N., Pillow, J. W., and Murthy, M. (2022). One-to-one mapping between deep network units and real neurons uncovers a visual population code for social behavior. Pages: 2022.07.18.500505 Section: New Results.

Cowley, B. R., Snyder, A. C., Acar, K., Williamson, R. C., Yu, B. M., and Smith, M. A. (2020). Slow Drift of Neural Activity as a Signature of Impulsivity in Macaque Visual and Prefrontal Cortex. *Neuron*, 108(3):551–567.e8.

Cox, J. and Witten, I. B. (2019). Striatal circuits for reward learning and decision-making. *Nature Reviews Neuroscience*, 20(8):482–494. Number: 8 Publisher: Nature Publishing Group.

Dayan, P. and Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4):429–453.

Denève, S. (2012). Making Decisions with Unknown Sensory Reliability. *Frontiers in Neuroscience*, 6.

DePasquale, B., Pillow, J. W., and Brody, C. (2021). Neural population dynamics underlying evidence accumulation in multiple rat brain regions. Technical report. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results Type: article.

Deverett, B., Kislin, M., Tank, D. W., and Wang, S. S.-H. (2019). Cerebellar disruption impairs working memory during evidence accumulation. *Nature Communications*, 10(1):3128. Number: 1 Publisher: Nature Publishing Group.

Ding, L. (2015). Distinct dynamics of ramping activity in the frontal cortex and caudate nucleus in monkeys. *Journal of Neurophysiology*, 114(3):1850–1861. Publisher: American Physiological Society.

Ding, L. and Gold, J. I. (2010). Caudate Encodes Multiple Computations for Perceptual Decisions. *Journal of Neuroscience*, 30(47):15747–15759. Publisher: Society for Neuroscience Section: Articles.

Ding, L. and Gold, J. I. (2012). Neural Correlates of Perceptual Decision Making before, during, and after Decision Commitment in Monkey Frontal Eye Field. *Cerebral Cortex*, 22(5):1052–1067.

Ding, L. and Gold, J. I. (2013). The Basal Ganglia's Contributions to Perceptual Decision Making. *Neuron*, 79(4):640–649.

Ditterich, J. (2006a). Evidence for time-variant decision making. *European Journal of Neuroscience*, 24(12):3628–3641. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-9568.2006.05221.x.

Ditterich, J. (2006b). Stochastic models of decisions about motion direction: Behavior and physiology. *Neural Networks*, 19(8):981–1012.

Doi, T., Fan, Y., Gold, J. I., and Ding, L. (2020). The caudate nucleus contributes causally to decisions that balance reward and uncertain visual information. *eLife*, 9:e56694. Publisher: eLife Sciences Publications, Ltd.

Driscoll, L., Shenoy, K., and Sussillo, D. (2022). Flexible multitask computation in recurrent networks utilizes shared dynamical motifs. Pages: 2022.08.15.503870 Section: New Results.

Drugowitsch, J., Mendonça, A. G., Mainen, Z. F., and Pouget, A. (2019). Learning optimal decisions with confidence. *Proceedings of the National Academy of Sciences*, 116(49):24872–24880. Publisher: National Academy of Sciences Section: PNAS Plus.

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., and Pouget, A. (2012). The Cost of Accumulating Evidence in Perceptual Decision Making. *Journal of Neuroscience*, 32(11):3612–3628.

Duan, C. A., Pagan, M., Piet, A. T., Kopec, C. D., Akrami, A., Riordan, A. J., Erlich, J. C., and Brody, C. D. (2021a). Collicular circuits for flexible sensorimotor routing. *Nature Neuroscience*, 24(8):1110–1120. Number: 8 Publisher: Nature Publishing Group.

Duan, C. A., Pan, Y., Ma, G., Zhou, T., Zhang, S., and Xu, N.-l. (2021b). A cortico-collicular pathway for motor planning in a memory-dependent perceptual decision task. *Nature Communications*, 12(1):2727. Number: 1 Publisher: Nature Publishing Group.

Dudman, J. T. and Gerfen, C. R. (2015). Chapter 17 - The Basal Ganglia. In Paxinos, G., editor, *The Rat Nervous System (Fourth Edition)*, pages 391–440. Academic Press, San Diego.

Duncker, L., O'Shea, D. J., Shenoy, K. V., and Sahani, M. (2020). A dynamical model with E/I balance explains robustness to optogenetic stimulation in motor cortex. In *Computational and Systems Neuroscience (COSYNE)*, Denver, CO.

Dutilh, G., van Ravenzwaaij, D., Nieuwenhuis, S., van der Maas, H. L. J., Forstmann, B. U., and Wagenmakers, E.-J. (2012). How to measure post-error slowing: A confound and a simple solution. *Journal of Mathematical Psychology*, 56(3):208–216.

Eckhoff, P., Holmes, P., Law, C., Connolly, P. M., and Gold, J. I. (2008). On diffusion processes with variable drift rates as models for decision making during learning. *New Journal of Physics*, 10(1):015006. Publisher: IOP Publishing.

Eichele, T., Debener, S., Calhoun, V. D., Specht, K., Engel, A. K., Hugdahl, K., Cramon, D. Y. v., and Ullsperger, M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proceedings of the National Academy of Sciences*, 105(16):6173–6178. ISBN: 9780708965108 Publisher: National Academy of Sciences Section: Biological Sciences.

Erlich, J. C., Bialek, M., and Brody, C. D. (2011). A Cortical Substrate for Memory-Guided Orienting in the Rat. *Neuron*, 72(2):330–343.

Erlich, J. C., Brunton, B. W., Duan, C. A., Hanks, T. D., and Brody, C. D. (2015). Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *eLife*, 4:8166.

Eyjolfsdottir, E., Branson, K., Yue, Y., and Perona, P. (2016). Learning recurrent representations for hierarchical behavior modeling. *arXiv:1611.00094 [cs]*. arXiv: 1611.00094.

Fan, Y., Gold, J. I., and Ding, L. (2018). Ongoing, rational calibration of reward-driven perceptual biases. *eLife*, 7:e36018. Publisher: eLife Sciences Publications, Ltd.

Fan, Y., Gold, J. I., and Ding, L. (2020). Frontal eye field and caudate neurons make different contributions to reward-biased perceptual decisions. *eLife*, 9:e60535. Publisher: eLife Sciences Publications, Ltd.

Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., and Wyart, V. (2019). Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nature Neuroscience*, 22(12):2066–2077. Number: 12 Publisher: Nature Publishing Group.

Fisher, D., Olasagasti, I., Tank, D. W., Aksay, E. R. F., and Goldman, M. S. (2013). A Modeling Framework for Deriving the Structural and Functional Architecture of a Short-Term Memory Microcircuit. *Neuron*, 79(5):987–1000. Publisher: Elsevier.

Fortunato, M., Blundell, C., and Vinyals, O. (2019). Bayesian Recurrent Neural Networks. arXiv:1704.02798 [cs, stat].

Fründ, I., Haenel, N. V., and Wichmann, F. A. (2011). Inference for psychometric functions in the presence of nonstationary behavior. *Journal of Vision*, 11(6):16–16. Publisher: The Association for Research in Vision and Ophthalmology.

Fründ, I., Wichmann, F. A., and Macke, J. H. (2014). Quantifying the effect of intertrial dependence on perceptual decisions. *J Vis*, 14(7):9–9.

Gardner, J. L. (2019). Optimality and heuristics in perceptual neuroscience. *Nature Neuroscience*, 22(4):514–523. Number: 4 Publisher: Nature Publishing Group.

Gigerenzer, G. and Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, 62(1):451–482. _eprint: https://doi.org/10.1146/annurev-psych-120709-145346.

Gilad, A., Gallero-Salas, Y., Groos, D., and Helmchen, F. (2018). Behavioral Strategy Determines Frontal or Posterior Location of Short-Term Memory in Neocortex. *Neuron*, 99(4):814–828.e7.

Glaser, J. I., Whiteway, M., Cunningham, J. P., Paninski, L., and Linderman, S. W. (2020). Recurrent Switching Dynamical Systems Models for Multiple Interacting Neural Populations. Pages: 2020.10.21.349282 Section: New Results.

Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pages 249–256. JMLR Workshop and Conference Proceedings. ISSN: 1938-7228.

Gold, J. I. and Ding, L. (2013). How mechanisms of perceptual decision-making affect the psychometric function. *Progress in Neurobiology*, 103:98–114.

Gold, J. I., Law, C.-T., Connolly, P., and Bennur, S. (2008). The Relative Influences of Priors and Sensory Evidence on an Oculomotor Decision Variable During Perceptual Learning. *Journal of Neurophysiology*, 100(5):2653–2668. Publisher: American Physiological Society.

Gold, J. I. and Shadlen, M. N. (2002). Banburismus and the Brain: Decoding the Relationship between Sensory Stimuli, Decisions, and Reward. *Neuron*, 36(2):299–308.

Gold, J. I. and Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, 30(1):535–574. _eprint: https://doi.org/10.1146/annurev.neuro.29.051605.113038.

Goldfarb, S., Wong-Lin, K., Schwemmer, M., Leonard, N. E., and Holmes, P. (2012). Can Post-Error Dynamics Explain Sequential Reaction Time Patterns? *Frontiers in Psychology*, 3.

Goldman, M., Compte, A., and Wang, X.-J. (2009). Neural Integrator Models. In *Encyclopedia of Neuroscience*, pages 165–178. Elsevier.

Guo, L., Weems, J. T., Walker, W. I., Levichev, A., and Jaramillo, S. (2019). Choice-Selective Neurons in the Auditory Cortex and in Its Striatal Target Encode Reward Expectation. *Journal of Neuroscience*, 39(19):3687–3697. Publisher: Society for Neuroscience Section: Research Articles.

Guo, Z. V., Inagaki, H. K., Daie, K., Druckmann, S., Gerfen, C. R., and Svoboda, K. (2017). Maintenance of persistent activity in a frontal thalamocortical loop. *Nature*, 545(7653):181–186. Number: 7653 Publisher: Nature Publishing Group.

Gupta, D. and Brody, C. D. (2022). Limitations of a proposed correction for slow drifts in decision criterion. *Neurons, Behavior, Data analysis, and Theory*, pages 1–19. Publisher: The neurons, behavior, data analysis and theory collective.

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., and Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, 520(7546):220–223.

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., and Shadlen, M. N. (2011). Elapsed Decision Time Affects the Weighting of Prior Probability in a Perceptual Decision Task. *Journal of Neuroscience*, 31(17):6339–6352.

Hanks, T. D. and Summerfield, C. (2017). Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron*, 93(1):15–31.

Hennig, J. A., Oby, E. R., Golub, M. D., Bahureksa, L. A., Sadtler, P. T., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., Batista, A. P., Chase, S. M., and Yu, B. M. (2021). Learning is shaped by abrupt changes in neural engagement. *Nature Neuroscience*, 24(5):727–736. Number: 5 Publisher: Nature Publishing Group.

Hermoso-Mendizabal, A., Hyafil, A., Rueda-Orozco, P. E., Jaramillo, S., Robbe, D., and de la Rocha, J. (2020). Response outcomes gate the impact of expectations on perceptual decisions. *Nature Communications*, 11(1):1057. Number: 1 Publisher: Nature Publishing Group.

Hirokawa, J., Vaughan, A., Masset, P., Ott, T., and Kepecs, A. (2019). Frontal cortex neuron types categorically encode single decision variables. *Nature*, 576(7787):446–451. Number: 7787 Publisher: Nature Publishing Group.

Horwitz, G. D. and Newsome, W. T. (1999). Separate Signals for Target Selection and Movement Specification in the Superior Colliculus. *Science*, 284(5417):1158–1161. Publisher: American Association for the Advancement of Science.

Hotelling, H. (1936). Relations Between Two Sets of Variates. *Biometrika*, 28(3/4):321–377. Publisher: [Oxford University Press, Biometrika Trust].

Huk, A. C., Katz, L. N., and Yates, J. L. (2017). The Role of the Lateral Intraparietal Area in (the Study of) Decision Making. *Annual Review of Neuroscience*, 40(1):349–372. _eprint: https://doi.org/10.1146/annurev-neuro-072116-031508.

Huk, A. C. and Shadlen, M. N. (2005). Neural Activity in Macaque Parietal Cortex Reflects Temporal Integration of Visual Motion Signals during Perceptual Decision Making. *Journal of Neuroscience*, 25(45):10420–10436. Publisher: Society for Neuroscience Section: Behavioral/Systems/Cognitive.

Hwang, E. J., Link, T. D., Hu, Y. Y., Lu, S., Wang, E. H.-J., Lilascharoen, V., Aronson, S., O'Neil, K., Lim, B. K., and Komiyama, T. (2019). Corticostriatal Flow of Action Selection Bias. *Neuron*, 104(6):1126–1140.e6.

Jazayeri, M. and Ostojic, S. (2021). Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. *arXiv:2107.04084 [q-bio]*. arXiv: 2107.04084.

Jeurissen, D., Shushruth, S., El-Shamayleh, Y., Horwitz, G. D., and Shadlen, M. N. (2022). Deficits in decision-making induced by parietal cortex inactivation are compensated at two timescales. *Neuron*, 110(12):1924–1931.e5. Publisher: Elsevier.

Jun, E. J., Bautista, A. R., Nunez, M. D., Allen, D. C., Tak, J. H., Alvarez, E., and Basso, M. A. (2021). Causal role for the primate superior colliculus in the computation of evidence for perceptual decisions. *Nature Neuroscience*, 24(8):1121–1131. Number: 8 Publisher: Nature Publishing Group.

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydın, ◆., Barbic, M., Blanche, T. J., Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Häusser, M., Karsh, B., Ledochowitsch, P., Lopez, C. M., Mitelut, C., Musa, S., Okun, M., Pachitariu, M., Putzeys, J., Rich, P. D., Rossant, C., Sun, W.-l., Svoboda, K., Carandini, M., Harris, K. D., Koch, C., O'Keefe, J., and Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551(7679):232–236. Number: 7679 Publisher: Nature Publishing Group.

Kang, B. and Druckmann, S. (2020). Approaches to inferring multi-regional interactions from simultaneous population recordings. *Current Opinion in Neurobiology*, 65:108–119.

Katahira, K. (2015). The relation between reinforcement learning parameters and the influence of reinforcement history on choice behavior. *Journal of Mathematical Psychology*, 66:59–69.

Kattner, F., Cochrane, A., and Green, C. S. (2017). Trial-dependent psychometric functions accounting for perceptual learning in 2-AFC discrimination tasks. *Journal of Vision*, 17(11):3–3. Publisher: The Association for Research in Vision and Ophthalmology.

Katz, L. N., Yates, J. L., Pillow, J. W., and Huk, A. C. (2016). Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature*, 535(7611):285–288. Number: 7611 Publisher: Nature Publishing Group.

Kaufman, M. T., Churchland, M. M., Ryu, S. I., and Shenoy, K. V. (2014). Cortical activity in the null space: permitting preparation without movement. *Nature Neuroscience*, 17(3):440–448. Number: 3 Publisher: Nature Publishing Group.

Keeley, S. L., Zoltowski, D. M., Aoi, M. C., and Pillow, J. W. (2020). Modeling statistical dependencies in multi-region spike train data. *Current Opinion in Neurobiology*, 65:194–202.

Kiani, R., Cueva, C., Reppas, J., and Newsome, W. (2014). Dynamics of Neural Population Responses in Prefrontal Cortex Indicate Changes of Mind on Single Trials. *Current Biology*, 24(13):1542–1547.

Kim, C. K., Ye, L., Jennings, J. H., Pichamoorthy, N., Tang, D. D., Yoo, A.-C. W., Ramakrishnan, C., and Deisseroth, K. (2017). Molecular and Circuit-Dynamical Identification of Top-Down Neural Mechanisms for Restraint of Reward Seeking. *Cell*.

Kim, J.-N. and Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience*, 2(2):176–185. Number: 2 Publisher: Nature Publishing Group.

Kingma, D. P. and Ba, J. (2017). Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs].

Kleinman, M., Chandrasekaran, C., and Kao, J. C. (2019). Recurrent neural network models of multi-area computation underlying decision-making. Pages: 798553 Section: New Results.

Koay, S. A., Thiberge, S., Brody, C. D., and Tank, D. W. (2020). Amplitude modulations of cortical sensory responses in pulsatile evidence accumulation. *eLife*, 9:e60628. Publisher: eLife Sciences Publications, Ltd.

Kohn, A., Jasper, A. I., Semedo, J. D., Gokcen, E., Machens, C. K., and Yu, B. M. (2020). Principles of Corticocortical Communication: Proposed Schemes and Design Considerations. *Trends in Neurosciences*, page S016622362030165X.

Kopec, C., Erlich, J., Brunton, B., Deisseroth, K., and Brody, C. (2015). Cortical and Subcortical Contributions to Short-Term Memory for Orienting Movements. *Neuron*, 88(2):367–377.

Lak, A., Hueske, E., Hirokawa, J., Masset, P., Ott, T., Urai, A. E., Donner, T. H., Carandini, M., Tonegawa, S., Uchida, N., and Kepecs, A. (2020a). Reinforcement biases subsequent perceptual decisions when confidence is low, a widespread behavioral phenomenon. *eLife*, 9:e49834. Publisher: eLife Sciences Publications, Ltd.

Lak, A., Okun, M., Moss, M. M., Gurnani, H., Farrell, K., Wells, M. J., Reddy, C. B., Kepecs, A., Harris, K. D., and Carandini, M. (2020b). Dopaminergic and Prefrontal Basis of Learning from Sensory Confidence and Reward Value. *Neuron*, 105(4):700–711.e6.

Latimer, K. W. and Freedman, D. J. (2021). Low-dimensional encoding of decisions in parietal cortex reflects long-term training history. Pages: 2021.10.07.463576 Section: New Results.

Latimer, K. W., Yates, J. L., Meister, M. L. R., Huk, A. C., and Pillow, J. W. (2015). Single-trial Spike Trains in Parietal Cortex Reveal Discrete Steps During Decision-making. *Science (New York, N.Y.)*, 349(6244):184–187.

Law, C.-T. and Gold, J. I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature Neuroscience*, 12(5):655–663. Number: 5 Publisher: Nature Publishing Group.

Li, N., Chen, T. W., Guo, Z. V., Gerfen, C. R., and Svoboda, K. (2015). Amotor cortex circuit for motor planning and movement. *Nature*.

Li, N., Daie, K., Svoboda, K., and Druckmann, S. (2016). Robust neuronal dynamics in premotor cortex during motor planning. *Nature*, 532(7600):459–464.

Licata, A. M., Kaufman, M. T., Raposo, D., Ryan, M. B., Sheppard, J. P., and Churchland, A. K. (2017). Posterior Parietal Cortex Guides Visual Decisions in Rats. *Journal of Neuroscience*, 37(19):4954–4966. Publisher: Society for Neuroscience Section: Research Articles.

Linderman, S., Antin, B., Zoltowski, D., and Glaser, J. (2020). SSM: Bayesian Learning and Inference for State Space Models. original-date: 2018-06-07T15:03:39Z.

Lo, C.-C. and Wang, X.-J. (2006). Cortico–basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nature Neuroscience*, 9(7):956–963. Number: 7 Publisher: Nature Publishing Group.

Luo, T. Z., Bondy, A. G., Gupta, D., Elliott, V. A., Kopec, C. D., and Brody, C. D. (2020). An approach for long-term, multi-probe Neuropixels recordings in unrestrained rats. Publisher: eLife Sciences Publications Limited.

Luo, T. Z., Hanks, T. D., Gupta, D., Bondy, A. G., and Brody, C. (2021). Dorsomedial frontal cortex participates in both evidence accumulation and history-based updating.

Lyamzin, D. R., Aoki, R., Abdolrahmani, M., and Benucci, A. (2021). Probabilistic discrimination of relative stimulus features in mice. *bioRxiv*, page 2020.12.20.423700. Publisher: Cold Spring Harbor Laboratory Section: New Results.

Macke, J. H. (2011). Empirical models of spiking in neural populations. page 9.

Maheswaranathan, N. and Sussillo, D. (2020). How recurrent networks implement contextual processing in sentiment analysis. *arXiv:2004.08013 [cs, stat]*. arXiv: 2004.08013.

Maheswaranathan, N., Williams, A., Golub, M., Ganguli, S., and Sussillo, D. (2019). Universality and individuality in neural dynamics across large populations of recurrent networks. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.

Mahn, M., Prigge, M., Ron, S., Levy, R., and Yizhar, O. (2016). Biophysical constraints of optogenetic inhibition at presynaptic terminals. *Nature Neuroscience*, 19(4):554–556.

Mamassian, P., Landy, M., and Maloney, L. T. (2002). Bayesian modelling of visual perception. In *Probabilistic models of the brain: Perception and neural function*, Neural information processing series, pages 13–36. The MIT Press, Cambridge, MA, US.

Mante, V., Sussillo, D., Shenoy, K. V., and Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474):78–84.

Masset, P., Ott, T., Lak, A., Hirokawa, J., and Kepecs, A. (2020). Behavior- and Modality-General Representation of Confidence in Orbitofrontal Cortex. *Cell*, 182(1):112–126.e18.

Mejías, J. F. and Wang, X.-J. (2022). Mechanisms of distributed working memory in a large-scale network of macaque neocortex. *eLife*, 11:e72136. Publisher: eLife Sciences Publications, Ltd.

Mendonça, A. G., Drugowitsch, J., Vicente, M. I., DeWitt, E. E. J., Pouget, A., and Mainen, Z. F. (2020). The impact of learning on perceptual decisions and its implication for speed-accuracy tradeoffs. *Nature Communications*, 11(1):2757. Number: 1 Publisher: Nature Publishing Group.

Miller, K. J., Botvinick, M. M., and Brody, C. D. (2019). From predictive models to cognitive models: An analysis of rat behavior in the two-armed bandit task. *bioRxiv*, page 461129. Publisher: Cold Spring Harbor Laboratory Section: New Results.

Mochol, G., Kiani, R., and Moreno-Bote, R. (2021). Prefrontal cortex represents heuristics that shape choice bias and its integration into future behavior. *Current Biology*, 31(6):1234–1244.e6.

Molano-Mazon, M., Duque, D., Yang, G. R., and Rocha, J. d. l. (2021). Pre-training RNNs on ecologically relevant tasks explains sub-optimal behavioral reset. Technical report. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results Type: article.

Morcos, A. S. and Harvey, C. D. (2016). History-dependent variability in population dynamics during evidence accumulation in cortex. *Nature Neuroscience*.

Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., and Forstmann, B. U. (2012). Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff. *Journal of Neuroscience*, 32(7):2335–2343. Publisher: Society for Neuroscience Section: Articles.

Murray, J. D., Jaramillo, J., and Wang, X.-J. (2017). Working Memory and Decision-Making in a Frontoparietal Circuit Model. *Journal of Neuroscience*, 37(50):12167–12186. Publisher: Society for Neuroscience Section: Research Articles.

Murray, J. M. (2019). Local online learning in recurrent networks with random feedback. *eLife*, 8:e43299. Publisher: eLife Sciences Publications, Ltd.

Murugan, M., Jang, H. J., Park, M., Miller, E. M., Cox, J., Taliaferro, J. P., Parker, N. F., Bhave, V., Hur, H., Liang, Y., Nectow, A. R., Pillow, J. W., and Witten, I. B. (2017). Combined Social and Spatial Coding in a Descending Projection from the Prefrontal Cortex. *Cell*, 171(7):1663–1677.e16.

Musall, S., Sun, X. R., Mohan, H., An, X., Gluf, S., and Churchland, A. K. (2021). Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. Technical report. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results Type: article.

Musall, S., Urai, A. E., Sussillo, D., and Churchland, A. K. (2019). Harnessing behavioral diversity to understand neural computations for cognition. *Current Opinion in Neurobiology*, 58:229–238.

Márton, C. D., Schultz, S. R., and Averbeck, B. B. (2020). Learning to select actions shapes recurrent dynamics in the corticostriatal system. *Neural Networks*, 132:375–393.

Najafi, F. and Churchland, A. K. (2018). Perceptual Decision-Making: A Field in the Midst of a Transformation. *Neuron*, 100(2):453–462.

Nevin, J. A. (1969). Signal detection theory and operant behavior. *Journal of the Experimental Analysis of Behavior*, 12(3):475–480.

Nguyen, K. P., Josić, K., and Kilpatrick, Z. P. (2019). Optimizing sequential decisions in the drift–diffusion model. *Journal of Mathematical Psychology*, 88:32–47.

Nguyen, Q. N. and Reinagel, P. (2020). A qualitative difference in decision-making of rats vs. humans explained by quantitative differences in behavioral variability. preprint, Neuroscience.

Odoemene, O., Pisupati, S., Nguyen, H., and Churchland, A. K. (2018). Visual Evidence Accumulation Guides Decision-Making in Unrestrained Mice. *Journal of Neuroscience*, 38(47):10143–10155. Publisher: Society for Neuroscience Section: Research Articles.

Orsolic, I., Rio, M., Mrsic-Flogel, T. D., and Znamenskiy, P. (2021). Mesoscale cortical dynamics reflect the interaction of sensory evidence and temporal expectation during perceptual decision-making. *Neuron*, 109(11):1861–1875.e10. Publisher: Elsevier.

Otis, J. M., Namboodiri, V. M. K., Matan, A. M., Voets, E. S., Mohorn, E. P., Kosyk, O., McHenry, J. A., Robinson, J. E., Resendez, S. L., Rossi, M. A., and Stuber, G. D. (2017). Prefrontal cortex output circuits guide reward seeking through divergent cue encoding. *Nature*, 543(7643):103–107.

Palmer, J., Huk, A. C., and Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, 5(5):1–1. Publisher: The Association for Research in Vision and Ophthalmology.

Park, I. M., Meister, M. L. R., Huk, A. C., and Pillow, J. W. (2014). Encoding and decoding in parietal cortex during sensorimotor decision-making. *Nature Neuroscience*, 17(10):1395–1403. Number: 10 Publisher: Nature Publishing Group.

Park, J., Phillips, J. W., Guo, J.-Z., Martin, K. A., Hantman, A. W., and Dudman, J. T. (2022). Motor cortical output for skilled forelimb movement is selectively distributed across projection neuron classes. *Science Advances*, 8(10):eabj5167.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, ◆. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830.

Peixoto, D., Verhein, J. R., Kiani, R., Kao, J. C., Nuyujukian, P., Chandrasekaran, C., Brown, J., Fong, S., Ryu, S. I., Shenoy, K. V., and Newsome, W. T. (2021). Decoding and perturbing decision states in real time. *Nature*, 591(7851):604–609. Number: 7851 Publisher: Nature Publishing Group.

Perich, M. G., Arlt, C., Soares, S., Young, M. E., Mosher, C. P., Minxha, J., Carter, E., Rutishauser, U., Rudebeck, P. H., Harvey, C. D., and Rajan, K. (2021). Inferring brain-wide interactions using data-constrained recurrent neural network models. Technical report. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results Type: article.

Perich, M. G. and Rajan, K. (2020). Rethinking brain-wide interactions through multi-region 'network of networks' models. *Current Opinion in Neurobiology*, 65:146–151.

Peters, A. J., Fabre, J. M. J., Steinmetz, N. A., Harris, K. D., and Carandini, M. (2021). Striatal activity topographically reflects cortical activity. *Nature*, 591(7850):420–425. Number: 7850 Publisher: Nature Publishing Group.

Piet, A. T., El Hady, A., and Brody, C. D. (2018). Rats adopt the optimal timescale for evidence integration in a dynamic environment. *Nature Communications*, 9(1):4265. Number: 1 Publisher: Nature Publishing Group.

Piet, A. T., Erlich, J. C., Kopec, C. D., and Brody, C. D. (2017). Rat Prefrontal Cortex Inactivations during Decision Making Are Explained by Bistable Attractor Dynamics. *Neural Computation*, 29(11):2861–2886.

Pinto, L., Koay, S. A., Engelhard, B., Yoon, A. M., Deverett, B., Thiberge, S. Y., Witten, I. B., Tank, D. W., and Brody, C. D. (2018). An Accumulation-of-Evidence Task Using Visual Pulses for Mice Navigating in Virtual Reality. *Frontiers in Behavioral Neuroscience*, 12. Publisher: Frontiers.

Pinto, L., Rajan, K., DePasquale, B., Thiberge, S. Y., Tank, D. W., and Brody, C. D. (2019). Task-Dependent Changes in the Large-Scale Dynamics and Necessity of Cortical Regions. *Neuron*, 104(4):810–824.e9.

Pisupati, S., Chartarifsky-Lynn, L., Khanal, A., and Churchland, A. K. (2021). Lapses in perceptual decisions reflect exploration. *eLife*, 10:e55490. Publisher: eLife Sciences Publications, Ltd.

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., and Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychological Review*, 117(4):1113–1143.

Putzeys, J., Raducanu, B. C., Carton, A., De Ceulaer, J., Karsh, B., Siegle, J. H., Van Helleputte, N., Harris, T. D., Dutta, B., Musa, S., and Mora Lopez, C. (2019). Neuropixels Data-Acquisition System: A Scalable Platform for Parallel Recording of 10 000+ Electrophysiological Signals. *IEEE transactions on biomedical circuits and systems*, 13(6):1635–1644.

Rahnev, D., Block, N., Jehee, J., and Denison, R. (2020). Is Perception Probabilistic?

Rajan, K., Harvey, C. D., and Tank, D. W. (2016). Recurrent Network Models of Sequence Generation and Memory. *Neuron*, 0(0):128–142.

Rao, R. P. N. (2010). Decision Making Under Uncertainty: A Neural Model Based on Partially Observable Markov Decision Processes. *Frontiers in Computational Neuroscience*, 4:146.

Raposo, D., Kaufman, M. T., and Churchland, A. K. (2014). A category-free neural population supports evolving demands during decision-making. *Nature Neuroscience*, 17(12):1784–1792. Number: 12 Publisher: Nature Publishing Group.

Ratcliff, R. and Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. *Psychological Science*, 9(5):347–356. Publisher: SAGE Publications Inc.

Remington, E. D., Egger, S. W., Narain, D., Wang, J., and Jazayeri, M. (2018). A Dynamical Systems Perspective on Flexible Motor Timing. *Trends in Cognitive Sciences*, 22(10):938–952.

Renart, A. and Machens, C. K. (2014). Variability in neural activity and behavior. *Current Opinion in Neurobiology*, 25:211–220.

Ridderinkhof, K. R., Nieuwenhuis, S., and Bashore, T. R. (2003). Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. *Neuroscience Letters*, 348(1):1–4.

Rigotti, M., Ben Dayan Rubin, D., Wang, X.-J., and Fusi, S. (2010). Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Frontiers in Computational Neuroscience*, 4:24.

Roitman, J. D. and Shadlen, M. N. (2002). Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task. *Journal of Neuroscience*, 22(21):9475–9489. Publisher: Society for Neuroscience Section: ARTICLE.

Rorie, A. E., Gao, J., McClelland, J. L., and Newsome, W. T. (2010). Integration of Sensory and Reward Information during Perceptual Decision-Making in Lateral Intraparietal Cortex (LIP) of the Macaque Monkey. *PLOS ONE*, 5(2):e9308. Publisher: Public Library of Science.

Rost, B. R., Wietek, J., Yizhar, O., and Schmitz, D. (2022). Optogenetics at the presynapse. *Nature Neuroscience*, pages 1–15. Publisher: Nature Publishing Group.

Roy, N. A., Bak, J. H., Akrami, A., Brody, C. D., and Pillow, J. W. (2021). Extracting the dynamics of behavior in sensory decision-making experiments. *Neuron*, 109(4):597–610.e6.

Ryali, C., Reddy, G., and Yu, A. J. (2018). Demystifying excessively volatile human learning: A Bayesian persistent prior and a neural approximation. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., editors, *Advances in Neural Information Processing Systems 31*, pages 2781–2790. Curran Associates, Inc.

Schall, J. D. (2001). Neural basis of deciding, choosing and acting. *Nature Reviews Neuroscience*, 2(1):33–42. Number: 1 Publisher: Nature Publishing Group.

Schütt, H. H., Harmeling, S., Macke, J. H., and Wichmann, F. A. (2016). Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research*, 122:105–123.

Scott, B. B., Constantinople, C. M., Akrami, A., Hanks, T. D., Brody, C. D., and Tank, D. W. (2017). Fronto-parietal Cortical Circuits Encode Accumulated Evidence with a Diversity of Timescales. *Neuron*, 95(2):385–398.e5.

Scott, B. B., Constantinople, C. M., Erlich, J. C., Tank, D. W., and Brody, C. D. (2015). Sources of noise during accumulation of evidence in unrestrained and voluntarily head-restrained rats. *eLife*, 4:e11308. Publisher: eLife Sciences Publications, Ltd.

Semedo, J. D., Gokcen, E., Machens, C. K., Kohn, A., and Yu, B. M. (2020). Statistical methods for dissecting interactions between brain areas. *Current Opinion in Neurobiology*, 65:59–69.

Semedo, J. D., Jasper, A. I., Zandvakili, A., Krishna, A., Aschner, A., Machens, C. K., Kohn, A., and Yu, B. M. (2022). Feedforward and feedback interactions between visual cortical areas use different population activity patterns. *Nature Communications*, 13(1):1099. Number: 1 Publisher: Nature Publishing Group.

Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M., and Kohn, A. (2019). Cortical Areas Interact through a Communication Subspace. *Neuron*, 102(1):249–259.e4.

Serre, T., Oliva, A., and Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences*, 104(15):6424–6429. Publisher: Proceedings of the National Academy of Sciences.

Seung, H. S. (1996). How the brain keeps the eyes still. *Proceedings of the National Academy of Sciences of the United States of America*, 93(23):13339–13344.

Shadlen, M. N. and Kiani, R. (2013). Decision Making as a Window on Cognition. *Neuron*, 80(3):791–806. Publisher: Elsevier.

Shadlen, M. N. and Newsome, W. T. (1996). Motion perception: seeing and deciding. *Proceedings of the National Academy of Sciences*, 93(2):628–633. Publisher: Proceedings of the National Academy of Sciences.

Shadlen, M. N. and Newsome, W. T. (2001). Neural Basis of a Perceptual Decision in the Parietal Cortex (Area LIP) of the Rhesus Monkey. *Journal of Neurophysiology*, 86(4):1916–1936. Publisher: American Physiological Society.

Sharma, A., Johnson, R., Engert, F., and Linderman, S. (2018). Point process latent variable models of larval zebrafish behavior. *Advances in Neural Information Processing Systems*, 31:10919–10930.

Shen, S. and Ma, W. J. (2019). Variable precision in visual perception. *Psychological Review*, 126(1):89–132.

Shevinsky, C. A. and Reinagel, P. (2019). The Interaction Between Elapsed Time and Decision Accuracy Differs Between Humans and Rats. *Frontiers in Neuroscience*, 13:1211.

Shushruth, S. and Shadlen, M. N. (2021). Sequential sampling from memory underlies action selection during abstract decision making. Technical report. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results Type: article.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., and Cohen, J. D. (2009). Reward rate optimization in two-alternative decision making: Empirical tests of theoretical predictions. *Journal of Experimental Psychology-Human Perception and Performance*, 35(6):1865–1897.

Sindreu, C., Marín, R., Fuentealba, Y., Duque, D., Jimenez, L., and de la Rocha, J. (2021). The causal role of the striatum in the encoding of task-adaptive expectation-based choice biases. In *Computational and Systems Neuroscience (Cosyne) 2021*, page 117.

Siniscalchi, M. J., Wang, H., and Kwan, A. C. (2019). Enhanced Population Coding for Rewarded Choices in the Medial Frontal Cortex of the Mouse. *Cerebral Cortex*, 29(10):4090–4106.

Smith, A. C., Frank, L. M., Wirth, S., Yanike, M., Hu, D., Kubota, Y., Graybiel, A. M., Suzuki, W. A., and Brown, E. N. (2004). Dynamic Analysis of Learning in Behavioral Experiments. *Journal of Neuroscience*, 24(2):447–461. Publisher: Society for Neuroscience Section: Behavioral/Systems/Cognitive.

Song, H. F., Yang, G. R., and Wang, X.-J. (2016). Training Excitatory-Inhibitory Recurrent Neural Networks for Cognitive Tasks: A Simple and Flexible Framework. *PLOS Computational Biology*, 12(2):e1004792. Publisher: Public Library of Science.

Steinemann, N. A., Stine, G. M., Trautmann, E. M., Zylberberg, A., Wolpert, D. M., and Shadlen, M. N. (2022). Direct observation of the neural computations underlying a single decision. preprint, Neuroscience.

Steinmetz, N. A., Zatka-Haas, P., Carandini, M., and Harris, K. D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature*, 576(7786):266–273. Number: 7786 Publisher: Nature Publishing Group.

Stine, G. M., Trautmann, E. M., Jeurissen, D., and Shadlen, M. N. (2022). A neural mechanism for terminating decisions. Pages: 2022.05.02.490327 Section: New Results.

Stuber, G. D., Sparta, D. R., Stamatakis, A. M., van Leeuwen, W. A., Hardjoprajitno, J. E., Cho, S., Tye, K. M., Kempadoo, K. A., Zhang, F., Deisseroth, K., and Bonci, A. (2011). Excitatory transmission from the amygdala to nucleus accumbens facilitates reward seeking. *Nature*, 475(7356):377–380.

Sugawara, M. and Katahira, K. (2021). Dissociation between asymmetric value updating and perseverance in human reinforcement learning. *Scientific Reports*, 11(1):3574. Number: 1 Publisher: Nature Publishing Group.

Summerfield, C. and Koechlin, E. (2010). Economic Value Biases Uncertain Perceptual Choices in the Parietal and Prefrontal Cortices. *Frontiers in Human Neuroscience*, 4. Publisher: Frontiers.

Sussillo, D. and Barak, O. (2013). Opening the Black Box: Low-Dimensional Dynamics in High-Dimensional Recurrent Neural Networks. *Neural Computation*, 25(3):626–649.

Tervo, D. G. R., Kuleshova, E., Manakov, M., Proskurin, M., Karlsson, M., Lustig, A., Behnam, R., and Karpova, A. Y. (2021). The anterior cingulate cortex directs exploration of alternative strategies. *Neuron*, 109(11):1876–1887.e6.

The International Brain Laboratory, Aguillon-Rodriguez, V., Angelaki, D., Bayer, H., Bonacchi, N., Carandini, M., Cazettes, F., Chapuis, G., Churchland, A. K., Dan, Y., Dewitt, E., Faulkner, M., Forrest, H., Haetzel, L., Häusser, M., Hofer, S. B., Hu, F., Khanal, A., Krasniak, C., Laranjeira, I., Mainen, Z. F., Meijer, G., Miska, N. J., Mrsic-Flogel, T. D., Murakami, M., Noel, J.-P., Pan-Vazquez, A., Rossant, C., Sanders, J., Socha, K., Terry, R., Urai, A. E., Vergara, H., Wells, M., Wilson, C. J., Witten, I. B., Wool, L. E., and Zador, A. M. (2021). Standardized and reproducible measurement of decision-making in mice. *eLife*, 10:e63711. Publisher: eLife Sciences Publications, Ltd.

Tye, K. M., Prakash, R., Kim, S.-Y., Fenno, L. E., Grosenick, L., Zarabi, H., Thompson, K. R., Gradinaru, V., Ramakrishnan, C., and Deisseroth, K. (2011). Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature*, 471(7338):358–362.

Urai, A. E., de Gee, J. W., Tsetsos, K., and Donner, T. H. (2019). Choice history biases subsequent evidence accumulation. *eLife*, 8:e46331.

Vallat, R. (2018). Pingouin: statistics in Python. *Journal of Open Source Software*, 3(31):1026.

Veuthey, T. L., Derosier, K., Kondapavulur, S., and Ganguly, K. (2020). Single-trial cross-area neural population dynamics during long-term skill learning. *Nature Communications*, 11(1):4057. Number: 1 Publisher: Nature Publishing Group.

Wang, H.-T., Smallwood, J., Mourao-Miranda, J., Xia, C. H., Satterthwaite, T. D., Bassett, D. S., and Bzdok, D. (2018). Finding the needle in high-dimensional haystack: A tutorial on canonical correlation analysis. *arXiv:1812.02598 [cs, stat]*. arXiv: 1812.02598.

Wang, X.-J. (2008). Decision Making in Recurrent Neuronal Circuits. *Neuron*, 60(2):215–234.

Warden, M. R., Selimbeyoglu, A., Mirzabekov, J. J., Lo, M., Thompson, K. R., Kim, S.-Y., Adhikari, A., Tye, K. M., Frank, L. M., and Deisseroth, K. (2012). A prefrontal cortex-brainstem neuronal projection that controls response to behavioural challenge. *Nature*, 492(7429):428–432.

Waskom, M. L., Okazawa, G., and Kiani, R. (2019). Designing and Interpreting Psychophysical Investigations of Cognition. *Neuron*, 104(1):100–112.

Wichmann, F. A. and Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, 63(8):1293–1313.

Wiegert, J. S., Mahn, M., Prigge, M., Printz, Y., and Yizhar, O. (2017). Silencing Neurons: Tools, Applications, and Experimental Constraints. *Neuron*, 95(3):504–529.

Wiltschko, A., Johnson, M., Iurilli, G., Peterson, R., Katon, J., Pashkovski, S., Abraira, V., Adams, R., and Datta, S. (2015). Mapping Sub-Second Structure in Mouse Behavior. *Neuron*, 88(6):1121–1135.

Wolff, S. B. E., Ko, R., and Ölveczky, B. P. (2022). Distinct roles for motor cortical and thalamic inputs to striatum during motor skill learning and execution. *Science Advances*, 8(8):eabk0231. Publisher: American Association for the Advancement of Science.

Yamins, D. L., Hong, H., Cadieu, C., and DiCarlo, J. J. (2013). Hierarchical Modular Optimization of Convolutional Networks Achieves Representations Similar to Macaque IT and Human Ventral Stream. In *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc.

Yang, G. R. and Molano-Mazón, M. (2021). Towards the next generation of recurrent network models for cognitive neuroscience. *Current Opinion in Neurobiology*, 70:182–192.

Yartsev, M. M., Hanks, T. D., Yoon, A. M., and Brody, C. D. (2018). Causal contribution and dynamical encoding in the striatum during evidence accumulation. *eLife*, 7:e34929. Publisher: eLife Sciences Publications, Ltd.

Yates, J. L., Park, I. M., Katz, L. N., Pillow, J. W., and Huk, A. C. (2017). Functional dissection of signal and noise in MT and LIP during decision-making. *Nature Neuroscience*, 20(9):1285–1292. Number: 9 Publisher: Nature Publishing Group.

Yu, A. J. and Cohen, J. D. (2009). Sequential effects: Superstition or rational behavior? In Koller, D., Schuurmans, D., Bengio, Y., and Bottou, L., editors, *Advances in Neural Information Processing Systems 21*, pages 1873–1880. Curran Associates, Inc.

Yu, A. J., Dayan, P., and Cohen, J. D. (2009). Dynamics of Attentional Selection Under Conflict: Toward a Rational Bayesian Account. *Journal of Experimental Psychology-Human Perception and Performance*, 35(3):700–717.

Zhang, S., Huang, H. C., and Yu, A. J. (2014). Sequential effects: A Bayesian analysis of prior bias on reaction time and behavioral choice. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36(36).

Zhong, L., Zhang, Y., Duan, C. A., Deng, J., Pan, J., and Xu, N.-l. (2019). Causal contributions of parietal cortex to perceptual decision-making during stimulus categorization. *Nature Neuroscience*, 22(6):963–973. Number: 6 Publisher: Nature Publishing Group.

Zhou, Y. and Freedman, D. J. (2019). Posterior Parietal Cortex Plays a Causal Role in Perceptual and Categorical Decisions. *Science (New York, N.Y.)*, 365(6449):180–185.

Znamenskiy, P. and Zador, A. M. (2013). Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature*, 497(7450):482–485. Number: 7450 Publisher: Nature Publishing Group.

Zoltowski, D. M., Latimer, K. W., Yates, J. L., Huk, A. C., and Pillow, J. W. (2019). Discrete Stepping and Nonlinear Ramping Dynamics Underlie Spiking Responses of LIP Neurons during Decision-Making. *Neuron*, 102(6):1249–1258.e10.

Zoltowski, D. M., Pillow, J. W., and Linderman, S. W. (2020). Unifying and generalizing models of neural dynamics during decision-making. *arXiv:2001.04571 [q-bio, stat]*. arXiv: 2001.04571.