# CONTEXTUAL MODULATION OF PERCEPTUAL DECISIONS AND NEURONAL REPRESENTATIONS BY STIMULUS AND REWARD HISTORY

SHREYA KHANAL


Graduate School of
Systemic Neurosciences
LMU Munich

Dissertation der Graduate School of Systemic Neurosciences der Ludwig-Maximilians-Universität München

December, 2022

Shreya Khanal: *Contextual modulation of perceptual decisions and neuronal representations by stimulus and reward history.*

Supervisors:
Prof. Dr. Laura Busse
*Department of Biology II*
*Division of Neurobiology*
*Ludwig-Maximilians-Universität München*

Prof. Dr. Thomas Wachtler
*Department of Biology II*
*Computational Neuroscience*
*Ludwig-Maximilians-Universität München*

First reviewer: Prof. Dr. Laura Busse
Second reviewer: Prof. Dr. Thomas Wachtler
External reviewer: Prof. Dr. Ilka Diester

Date of Submission: 28$^{th}$ December, 2022
Date of defense: 21$^{st}$ June, 2023

*The illusion that we understand the past fosters overconfidence in our ability to predict the future.*

*- Daniel Kahneman*

# ABSTRACT

Utilizing regularities in this complex but stable environment is important for survival. Hence, our decisions do not only depend on current sensory input but are also influenced by previous experiences. This thesis investigates such effects on behavior during visual perceptual decision-making. I further implemented a long-term neuronal recording method and provide a proof of concept on how such a method can be utilized to reveal the neuronal encoding of previous experiences.

I designed behavioral setups and trained head-fixed mice to perform a lick-left / lick-right orientation discrimination task using a two-alternative forced choice paradigm. Using Logistic Regression, I estimated the contributions of task-relevant (current stimulus) and task-irrelevant (past stimuli, past response, past outcomes) factors on the animals' decisions. To investigate how mice trade-off stimulus-related and contextual information, I introduced blocks of trials with stimulus-specific imbalances of reward. I further designed a neuronal recording setup and implemented a long-term recording method using an immobile silicone probe.

Mice learned the task and reached stable performance in 6–8 weeks. For the current stimulus, model weights were negligible for trials with short reaction times, suggesting that on these trials, animals failed to use relevant sensory evidence to guide decisions. Irrespective of reaction time, significant weights were assigned to past stimulus, response, and outcome, indicating that history effects influence decisions across all trials. When contextual information (i.e., reward size) was manipulated, mice showed a consistent bias towards the response side associated with larger rewards and the model weight assigned to the current stimulus decreased. However, given the performance during balanced reward condition, mice performed at an optimal bias level to maximize reward. To understand how contextual signals influence neuronal responses and sensory representations in the primary visual

cortex, I performed long-term neuronal recording during learning. Stable recordings were obtained for 2 months and specific neurons were tracked for several days. Preliminary data suggest that there might be variation in responses of neurons in the primary visual cortex based on previous outcomes.

Taken together, even for simple visual stimuli well above threshold, perceptual decision-making in mice is influenced not only by the current visual stimulus but also by the history of past trials and the context of reward. In addition, it is important to account for trial-to-trial variability in the speed of the response to understand behavior. Finally, influences of prior outcomes might be reflected in the primary visual cortex.

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| **BDT** | **B**ayesian **D**ecision **T**heory |
| **CS** | **C**urrent **S**timulus |
| **CR** | **S**ide of **C**orrect **R**esponse |
| **DDM** | **D**rift **D**iffussion **M**odel |
| **DM** | **D**ecision **M**aking |
| **FOF** | **F**rontal **O**rienting **F**ield |
| **IncR** | **S**ide of **I**ncorrect **R**esponse |
| **LIP** | **L**ateral **I**ntra**P**arietal |
| **LR** | **L**ogistic **R**egression |
| **MT** | **M**edial **T**emporal |
| **OE** | **O**pen **E**phys |
| **PR** | **P**revious **R**esponse |
| **PS** | **P**revious **S**timulus |
| **RDK** | **R**andom **D**ot **K**inetogram |
| **RF** | **R**eceptive **F**ield |
| **RL** | **R**einforcement **L**earning |
| **RT** | **R**eaction **T**ime |
| **SAT** | **S**peed **A**ccuracy **T**rade-off |
| **SDT** | **S**ignal **D**etection **T**heory |

# 1

# INTRODUCTION

The ability to perceive our surroundings and make efficient decisions in response is essential for survival. Higher-level animals have evolutionary advantages because they are able to integrate information from various sources. This improved decision-making ability implies that they efficiently separate decision-relevant information from decision-irrelevant information (noise). To illustrate, picture yourself in the following situation:

> *You are enjoying holidays in a beautiful Mediterranean town. The sun is about to set and you are strolling down a lively street. At the side, relaxed people are sipping wine and indulging in culinary delicacies, while merchants are appraising their goods to passersby. Looking around, you suddenly spot a familiar face in the dim light. As your eyes meet, you wonder whether it might be a high-school friend, whom you had not met in years. Could it be her? Should you say hi? But wait, did you not recently see her posting on social media from elsewhere? As you contemplate how to react, she calls out your name. Hearing her voice immediately resolves all doubts, you respond, and are set for reminiscing joint memories.*

Information and expectation combine to form perception. Perceptual decision-making (PDM) refers to the above-illustrated act of choosing one decision (starting a conversation) from sets of alternatives (walking away, smiling without talking, etc.) based on perceived reality. While making such decisions, it is important to identify relevant sensory information and integrate it with previous experiences. For example, when you subconsciously recognized your friend's face, your attention is focused. You instantly cross-referenced her facial features with your dated memory. You also recalled contextual information, i.e., the

social media post. Given contradictory visual and contextual cues, you required additional auditory information (hearing her voice) to break the tie.

Understanding PDM is a primary goal of Neuroscience. Over the last few decades, improvements in behavioral paradigms as well as technological advancement in areas such as optogenetics, imaging, and electrophysiology created possibilities to investigate it. Researchers can isolate specific behavioral strategies and manipulate neural circuits to identify the causal mechanisms that underlie PDM. Particularly, they perform psychophysical studies to understand the relationship between subjective sensations and the physical stimuli that produce them. Early studies attempted to find a relationship between stimulus parameters (e.g., contrast and size) and perceived sensation. Over the last decade, the focus has shifted to identifying other factors that impact PDM. This is achieved by, for example, introducing uncertainties by manipulating reward amounts or stimulus probabilities. Here, the subject has to consider the task constraint and weigh alternative options to reach a decision (Akrami et al., 2018; Lak et al., 2020b; Busse et al., 2011). Hence, the focus of recent studies, including this thesis, has been on determining how various identified factors are internally weighed and how they affect perceptual judgment and neural responses.

In **section 1.1**, I discuss theoretical frameworks that have been employed to explain PDM. Subsequently, in **section 1.2**, I provide an overview of what is understood about the neurophysiological mechanisms that underlie PDM, with a focus on visual perception. Finally, in **section 1.3**, I review selected factors that explain variability in decision-making.

## 1.1 THEORETICAL FRAMEWORKS

Various theoretical frameworks can be employed to explain PDM processes. To accurately represent such processes, the influence of previous experience and sensory information needs to be modeled.

In this section, I introduce three theoretical frameworks that are frequently used to explain PDM processes: Bayesian Decision Theory (BDT), Signal Detection Theory (SDT) and the Drift Diffusion Model (DDM).

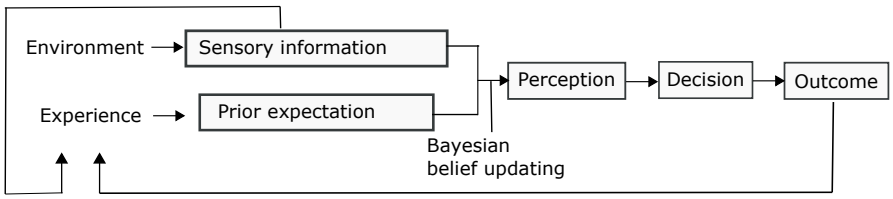


Figure 1.1: Illustrative description of Bayesian Decision Theory. Schematic depicting information integration and flow in Bayesian Decision Theory (BDT). Perception is based on integration of sensory information with experience-driven prior expectations. Experiences are constantly updated based on sensory information and decision outcomes.

BDT provides a general explanatory framework (Knill and Richards, 1996). As illustrated in **Figure 1.1**, BDT postulates that perception is based on both sensory information and prior expectation. Sensory information is extracted from the environment, while prior expectation arises from experience with sensation (e.g., what was previously seen) and outcomes (e.g., how were previous decisions rewarded). This perception guides decision-making (Wei and Stocker, 2015; Gold and Stocker, 2017; Wei and Stocker, 2016). Formally, perception can be described as a posterior probability distribution that is computed by the application of Bayes rule to the inputs.

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

Here, $P(H|E)$ is the conditional probability of a hypothesis ($H$) being true given evidence ($E$), $P(E|H)$ is the conditional probability of observing $E$ under $H$, $P(H)$ is the prior expectation of $H$, and $P(E)$ is the unconditional probability of observing $E$. While BDT is largely consistent with observed behavior, it does not provide specific theories of how perception translates into decisions.

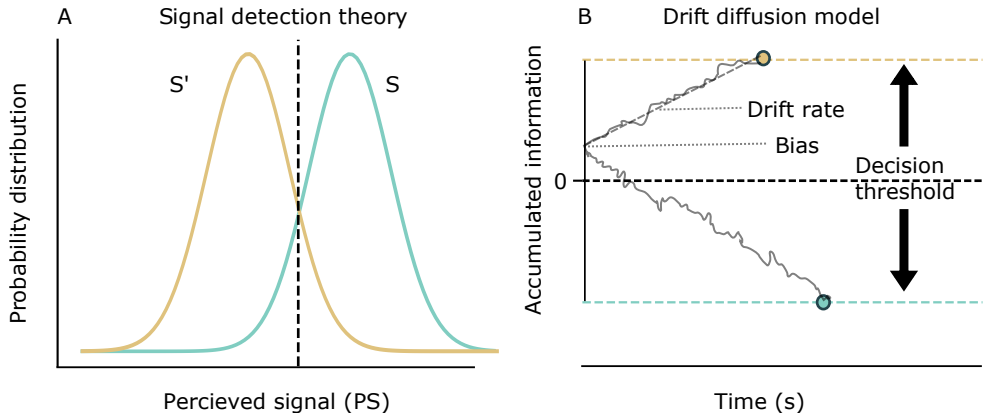


Figure 1.2: Illustrative description of Signal Detection Theory and the Drift Diffusion model.
(**A**) Schematic of Signal Detection Theory (SDT). The probability distribution held by the observer of Signal and internal noise (S) and external noise (S'). *Dotted line*: decision threshold, which indicates the minimum threshold to report a detection. (**B**) Schematic of Drift diffusion model (DDM). Example of two trials where noisy sensory information (*grey*) is accumulated over time with a specific drift rate. *Yellow / Turquoise; dotted line*: decision threshold for one and for the other stimulus, respectively. Offset from 0 indicates a bias towards a particular stimulus resulting, for example, from previous experiences.

SDT provides a simplistic model of discrimination of signal from noise (Green and Swets, 1966; Gold and Stocker, 2017). It relies on the concept that to make a decision, the decision-maker translates perceptual information into a decision variable. The decision variable classifies the decision-relevant evidence for one choice over the other. As illustrated in **Figure 1.2A**, the decision-maker is assumed to hold an opinion about the probability distribution of decision variables which is conditional on the observed stimuli and internal noise (S). The decision-maker compares this conditional probability density functions (PDF) to the one generated by external noise (S'). The decision criterion (*dashed black line*) informs the minimum criterion where the decision-maker detects a signal. The separation between the two PDF represents the decision maker's ability to detect the stimulus. SDT allows for varying prior expectations and flexible setting of the decision criterion.

However, it treats perception as static, thereby disregarding the process of information accumulation.

Sequential sampling models, such as the drift-diffusion model (DDM), build on SDT recognize that PDM is not an instantaneous process (Ratcliff, 1978b). DDMs rely on the concept of information accumulation, i.e., perceptual evidence is assumed to increase over time until a decision threshold is reached (Stone, 1960; Ratcliff and Smith, 2004; Gold and Stocker, 2017). **Figure 1.2B** illustrates this process. DDMs have three parameters (Gold and Stocker, 2017): First, the drift rate relates time with incremental evidence gain, i.e., the average slope of information accumulation. Second, the bias models predispositions towards a decision, i.e., the intercept of information accumulation. Third, the decision threshold indicates how much information is necessary to commit a decision. Jointly, these variables explain decisions and reaction time (RT).

In **section 1.2**, I provide an overview of what is understood about the neurophysiological mechanisms that underlie PDM, with a focus on visual perception.

## 1.2 NEURAL BASIS OF PERCEPTUAL DECISION-MAKING

The process of converting sensory input into decisions can be summarized along three stages: First, the sensory inputs are encoded. Second, these encoded sensory information are accumulated and integrated with previous experience. Third, once this integrated information reaches a threshold, a decision is made. In the following paragraphs, I discuss the state of research on the neural basis of each of these three stages, with an emphasis on visual perception.

Steinmetz et al. (2019) demonstrated that neurons in the mice visual cortex exhibit activity upon stimulus onset. Similarly, other studies in monkeys and rodents have also identified that neurons in the visual cortex respond to visual stimuli (Newsome et al., 1989; Salzman et al., 1990; Britten et al., 1992; Cai et al., 2018) and neurons in the

somatosensory cortex respond to tactile stimuli (Romo et al., 2002; Guo et al., 2014b). In early experiments, monkeys were trained to report the direction of randomly moving dots (RMD). During the RMD task, neural activity in the Medial Temporal (MT) area of the visual cortex is correlated with motion processing (Newsome et al., 1989). In fact, later a causal relationship was proven by micro-simulation and lesion studies (Salzman et al., 1990; Romo et al., 2000). A similar picture emerges when studying rodents using a visual discrimination task. Neural activity following sensory stimuli first appears in the primary visual cortex (V1), continues to the secondary visual cortex, and finally reaches the frontal cortex (Steinmetz et al., 2019).

To arrive to a decision efficiently, sensory information needs to be accumulated and integrated with previous experience. Brain areas involved in this process show increasing neural activity until a decision threshold is reached. In monkeys trained to make choices through saccadic eye movements, such neural activity occurs in the parietal cortex (Shadlen and Newsome, 1996), frontal eye field (FEF) (Ding and Gold, 2012a), prefrontal cortex (PFC) (Lin et al., 2020), and striatum (Ding and Gold, 2010). Studies indicate that these areas might be involved in the visual decision-making task, independent of specific movements (Gold and Shadlen, 2003; De Lafuente et al., 2015). In rodents, a similar picture unfolds. For example, Hanks et al. (2015) used auditory experiments to show that neural activity in the posterior parietal cortex (PPC) and the frontal orienting field (FOF) ramps up during decision formation. In another study, Yartsev et al. (2018) shows that optogenetic manipulation of striatal neurons in rats disrupts the evidence accumulation process (see also **subsection 1.3.1**). As mentioned in **section 1.1**, sensory information integrates with previous experiences to reach a decision. Recently, numerous studies have focused on understanding the underlying process of this history effect, which I discuss in **section 1.3**.

The neural signals for choice are identified both in lower and higher brain areas, suggesting that information flow in PDM is both feedforward and feedback. For example, in monkeys, the activity of sensory neurons in the visual cortex is correlated with decisions (Nienborg and Cumming, 2010). Similarly, choice signals are also identified in other areas of the brain such as the basal ganglia, the frontal cortex, the PPC,

and the motor cortex (Steinmetz et al., 2019; Buckley et al., 2009; Lee et al., 2007; Thura and Cisek, 2014; Pape and Siegel, 2016). For example, some studies show that prior to choice commitment, the neural activity in the motor cortex builds up reflecting competing choices (Cisek and Kalaska, 2010; Klaes et al., 2011) and upcoming response options (Donner et al., 2009; Lange et al., 2013). Other studies show that neural activities in the lateral PFC during a WM task correlate with optimal choice (Buckley et al., 2009; Lee et al., 2007). Choices in PDM are followed by outcomes. The impact of the outcome and its integration in the subsequent decision is discussed in **subsection 1.3.3**.

In summary, numerous studies identify neural correlates of PDM across multiple brain areas. However, research on understanding its causal mechanism is still in its early stages. Rigorous psychophysics studies combined with neural recording and circuit manipulation will give us a more detailed description of the underlying processes.

## 1.3 INTRINSIC VARIABILITY IN DECISION-MAKING

Decisions of both humans and animals show substantial variability, even when exposed to similar environments. While some variability remains unexplained, a substantial fraction can be explained. Explanatory factors include previous experience, expected reward size, motivation, and attention state, and the speed-accuracy trade-off (SAT). In **subsection 1.3.1** I discuss the SAT and its underlying neural mechanisms. Next, in **subsection 1.3.2**, I discuss the influence of history effects (previous stimulus, previous response, and outcome) and expected rewards on decision-making. Finally, in **subsection 1.3.3** I discuss three specific brain areas that play a role in integrating these effects with sensory information, i.e., the prefrontal cortex, the posterior parietal cortex, and the striatum.

### 1.3.1 *Speed-accuracy trade off*

There is a fundamental trade-off between the speed at which decisions are made and the probability that they are optimal (Wickelgren, 1977; Heitz, 2014). This speed-accuracy trade-off (SAT) exists across species (Franks et al., 2003; Marshall et al., 2009; Wenzlaff et al., 2011; Heitz and Schall, 2012) and modalities (Ings and Chittka, 2008; Palmer et al., 2005; Rinberg et al., 2006). It was first observed in 1911: When human subjects were shown two lines and asked to identify the longer one, their reaction time (RT) and accuracy were not independent (Henmon, 1911). In the following years, empirical understanding of the SAT progressed largely in parallel with the theoretical models of sequential decision-making (see section 1.1). Here, I discuss empirical evidence of the SAT as well as its neural representation.

The underlying framework of the SAT relies on the theory that decision thresholds can be adjusted to trade off speed and accuracy (Heitz, 2014). Decisions can be made faster by lowering the threshold, thus requiring less information. Empirical studies show that animals can be trained to adjust their decision threshold (Balci et al., 2011; Mendonça et al., 2020). For example, Heitz and Schall (2012) trained monkeys in a visual search task to analyze this phenomenon. Following a cue, when presented with different rewards and punishment contingencies, monkeys adjust their RT. In both humans and rodents, studies emphasize that RT increases with trial difficulty (Kurylo et al., 2015; Plainis and Murray, 2000), as animals spend more time accumulating evidence to reach the decision threshold. For example, using a visual task, Kurylo et al. (2015) shows that RT increases when contrast and/or stimulus duration decrease.

Neuroimaging studies demonstrated that neural activity in the prefrontal and subcortical areas varies during information accumulation. Specifically, such variation was identified in LIP (Roitman and Shadlen, 2002), FEF (Hanes and Schall, 1996; Boucher et al., 2007), superior colliculus (Ratcliff et al., 2003), and basal ganglia (Ding and Gold, 2010). For example, the firing rate of neurons in LIP increases during evidence accumulation and plateaus when the decision threshold is reached. The baseline activation of these neurons increases when prioritizing

speed, enabling them to reach the decision threshold faster (Heitz and Schall, 2012; Hanks et al., 2014). Studies with human subjects also identified a distinctive pattern of brain activity between those prioritizing speed and those prioritizing accuracy: The former shows higher baseline activation of the supplementary motor area, while the latter shows it in the prefrontal cortex (Perri et al., 2014).

### 1.3.2 *Influence of history effects and expected rewards on decision-making*

Decisions are influenced by previous experience. For example, behavior can often be influenced by both current and previous stimuli. The first evidence of such history effects was reported in the early $20^{th}$ century and is known as 'contraction bias' (Hollingworth, 1910). Hollingworth showed that the perception in humans of a stimulus is biased towards the average of previously shown stimuli. Contraction bias has since been demonstrated across species and modalities, ie., visual (Olkkonen et al., 2014), somatosensory (Fassihi et al., 2017), and auditory (Akrami et al., 2018; Raviv et al., 2014). Most of these studies applied working memory tasks in which the subject reported the perceived stimulus after a time delay (Akrami et al., 2018). Hence, they emphasize that contraction bias is integrated at a post-perceptual level. However, a few recent studies suggest that the influence of previous stimuli can already happen at the perceptual level (Fritsche et al., 2017; Patten et al., 2017). For example, using fMRI on human participants, Patten et al. (2017) shows that systematic bias in orientation perception correlates with neural activity in V1, suggesting previous stimuli influence neural processing from the earliest stage of cortical processing.

Choices in PDM tasks are often followed by external feedback: a reward for a correct decision and a punishment for an incorrect decision. Such feedback might influence subsequent choices, for example, through feedback-driven learning (Law and Gold, 2009; Seriès and Seitz, 2013). Feedback can influence expectations and thereby contribute to history effects. In fact, it is speculated that feedback changes a subject's perception of subsequent information and gives rise to decision-making strategies. A commonly observed strategy is "win-stay / lose-shift",

i.e., a bias to repeat a rewarded choice and avoid a punished choice (Fründ et al., 2014; Abrahamyan et al., 2016; Busse et al., 2011; Akrami et al., 2018). This strategy is frequently observed when the probability of making a correct response is low, for instance, in the presence of ambiguous stimuli such as low contrast (Akrami et al., 2018; Lak et al., 2020b). In fact, these observations are consistent with Bayesian Decision Theory (BDT; see also **section 1.1**).

Reward size also influences PDM. Under asymmetric reward conditions, where one alternative holds a larger reward than the other, subjects have a strong preference towards the larger reward alternative (Busse et al., 2011; Rorie et al., 2010). This preference is known to increase with trial difficulty and it influences decision-making strategies (Lak et al., 2020b). While the impact of reward expectation on behavior is unambiguous, the underlying neural mechanism remains unclear. Some studies suggest that reward information already affects sensory processing (Cicmil et al., 2015; Saproo and Serences, 2010; Weil et al., 2010), others provide evidence that later stages of decision-making are also affected (Rorie et al., 2010; Chen et al., 2015) (see **subsection 1.3.3**).

In summary, history effects and expected rewards influence decisions. In controlled randomized laboratory experiments, such influences lead to reduced performance. However, the real environment is temporally stable and predictable, and history effects are beneficial in such an environment with temporal structure. For example, such biases encourage heuristic decision-making and prioritize speed and resources to reach 'accurate enough' if not 'perfectly accurate' decisions. Hence, evolutionary pressure might have forced organisms to use necessarily limited neural resources efficiently by integrating previous information with current sensory information to reach a decision faster. In **subsection 1.3.3**, I discuss brain areas that are involved in integrating this information.

### 1.3.3 *Neural integration of prior information and reward expectation*

History effects and expected rewards influence PDM by modulating neural activity in brain areas such as PPC (Akrami et al., 2018; Hwang et al., 2017), PFC (Sul et al., 2010; Amiez et al., 2006; Lak et al., 2020b), and the striatum (Hwang et al., 2019; Balleine et al., 2007; Samejima et al., 2005; Lau and Glimcher, 2008; Lauwereyns et al., 2002; Schultz, 2022). In this section, I summarize the literature on the role of these areas in integrating history effects in PDM processes.

**PFC** receives information from various sensory areas such as the somatosensory and the parietal cortex and has extensive feedback and feedforward projections to mid-brain dopamine neurons (Blatt et al., 1990). PFC is known for its role in guiding complex cognitive behavior, retrieving both long- and short-term memory (Corcoran and Quirk, 2007; Narayanan et al., 2006; Seamans et al., 1995; Sul et al., 2010) and encoding future rewards from past outcomes (Sul et al., 2010; Amiez et al., 2006; Lak et al., 2020b). For example, a single-cell study in rats shows that the neural activity of the medial PFC and orbitofrontal cortex (OFC) correlates with the accumulated outcome and choice from past trials (Sul et al., 2010). Numerous studies suggest that neural activity in the medial PFC is associated with encoding reward prediction error and optimizing history-related strategy when task statistics abruptly change (Sul et al., 2010; Tervo et al., 2014). Contrarily, neural activity in the OFC is associated with expected outcomes, irrespective of it being rewarded or unrewarded (Sul et al., 2010; O'Reilly et al., 2013; Schoenbaum et al., 1998).

As discussed in **subsection 1.3.2**, **PPC** plays an important role in accumulating information during the decision-making process. PPC receives visual, somatosensory and auditory input and numerous studies have implicated its role in PDM (Bitzidou et al., 2018) (see **section 1.2**). Recent studies emphasize its role in integrating history effects (Akrami et al., 2018; Hwang et al., 2017). For example, Akrami et al. (2018) trained rats to report the relative loudness of two signals, following a delay period. The result showed that the inactivation of PPC, during the delay period, reduces contraction bias. Contradicting this observation, using mice Hwang et al. (2017) show that in an

orientation discrimination task, PPC inactivation during the inter-trial interval but not during the delay period reduces history effects. In both of these studies, PPC inactivation during stimulus presentation did not influence decisions. Taken together, the recent literature suggest a critical role of PPC in integrating history effects. However, it is still debated at what stage of information processing these effects are integrated in the decision-making processes.

Given that PPC neurons mediate history effects, the question arises, of where PPC sends this biased information to affect choices. One of the prime downstream candidates is the **Striatum** (Hwang et al., 2019; Balleine et al., 2007). The striatum receives input from both the prefrontal and the sensory cortex, and is an important part of the cortico-basal ganglia-thalamic circuit (Thibaut, 2016; Haber, 2022; Voorn et al., 2004). In the past decade, its role in encoding context-dependent learning and reward prediction error has been extensively emphasized (Samejima et al., 2005; Lau and Glimcher, 2008; Lauwereyns et al., 2002; Schultz, 2022). Particularly, researchers have shown that the striatum encodes the value of action in choice-task (Samejima et al., 2005; Lau and Glimcher, 2008). Supporting this claim, Tai et al. (2012) shows that optogenetic stimulation of neurons in the dorsal striatum bias choices towards the side with a larger reward value. Using micro-stimulation on monkeys performing the RMD task, studies suggest that striatal neurons incorporate this bias by shifting the starting point of the evidence accumulation (Ding and Gold, 2012b).

In summary, in this section, I give a brief overview of three brain areas associated with history-effect. However, one of the prime questions, that remains unanswered is the neural circuitry and mechanism by which these history effects are integrated into the decision-making process. A deeper understanding of this process requires identifying circuitry that governs flexible integration of such effects in the decision-making process. As discussed above, by combining carefully designed behavioral experiments with neural recordings, and optogentic or pharmacological manipulations, more recent research has focused on understanding the underlying causal mechanism of history effects.

# 2

# OBJECTIVES

The overall objective of this Ph.D. project is to understand how recent history of stimuli, choices, and outcomes influence subsequent behavior and neuronal activities. To this end, the specific objectives are:

1. **Design behavioral task, assemble setup and establish behavior**.

    a) Develop a new experimental computer-controlled setup to train head-fixed mice and to record behavioral activity

    b) Design, implement and execute behavioral training requiring minimal supervision

2. **Quantify history effects on decision-making during learning, task performance and reward manipulation.**

3. **Perform chronic recording across primary visual cortex**.

    a) Implement a computer-controlled setup to allow simultaneous recording of behavior and neuronal data

    b) Adapt chronic recording using an immobile-silicone probe for long-term neuronal recording

    c) Evaluate stability of neural responses over time

# 3

# MATERIALS AND METHODS

In this section, I discuss the methods used to acquire and analyze the data reported in this thesis. All procedures complied and were approved by the German animal welfare law (Tierschutzgesetz: ROB – 55.2 – 2532.Vet_02 – 16 – 33 and ROB – 55.2 – 2532.Vet_02 – 21 – 137).

In **section 3.1** I discuss the animal model used in the experiments. Subsequently, in **section 3.2**, I describe the behavioral setup and paradigms in which the animals were trained. In **section 3.3** I describe the method used for electrophysiology experiments. Finally, in **section 3.4** I give a overview of the data analysis method used to obtain the results presented in **chapter 4**.

## 3.1 ANIMAL MODEL

Both behavioral and electrophysiology experiments were performed on male C57BL/6 mice. The behavioral experiments were conducted on 5 animals, and electrophysiology experiments were performed on 3 other animals. The mice were either from the breeding colony of the Biocenter of the Ludwig Maximilian University of Munich or from the Charles River facility. Animals were housed in groups of 2 males in a transparent cage with nesting and environmental enrichment materials. When experiments were not being performed, the animals were kept in the animal facility rooms. The room was maintained at a temperature of 22.5°and 50 % humidity with 12 h light/dark circle. Three of the 5 behavioral animals were placed in inverse day/light circle. Hence, for these three animals, experiments were conducted in

the dark phase of their sleep cycle. Animals were at-least 8 weeks old when the head bar was implanted.

## 3.2 BEHAVIORAL EXPERIMENTS

In **subsection 3.2.1** I describe the behavioral setup that I designed and built. I further give an overview of various components of the setup and how they measure behavior. Subsequently, in **subsection 3.2.3** I describe the behavioral paradigm used in this thesis (adapted from Marques et al. (2018)). I also give a brief overview of different training stages the animals were trained in to successfully learn the task. Finally, in **subsection 3.2.4** I describe the surgical procedure for head bar implantation.

### 3.2.1 *Behavioral setup*

Mice were head-fixed and positioned in the center of a circular mounted plastic disk. The disk enabled them to freely sit or run. To measure running speed, a rotary encoder (MA3-A10-125-N Magnetic Encoder, Pewatron) sampling at 100 Hz was fixed at the bottom of the disk (see also **subsection 3.2.2**). Two lick spouts were positioned 4 mm apart in front of the animal's snout. Licking was measured by a custom-made lick sensor (see also **subsection 3.2.2**). A syringe pump (Aladdin AL-1000, WPI) delivered 4 µL water reward to the spout when the animal made a correct response. This was controlled by a computerized protocol written in MATLAB. Visual stimuli were presented using a custom written Matlab program using Psychotoolbox package (Wilson et al., 2011). The stimuli were presented on an LED monitor (FUJITSU B24W-7, refresh rate 60 Hz, mean luminance of 50 cd/m$^2$) positioned at 25 cm distance from the animals' eye. The stimulus presented was either a vertical (270°) or a horizontal (180°) drifting sinusoidal grating at a 50 % contrast, with a circular aperture and a diameter of 70°. The spatial frequency was 0.05 cyc/° and the temporal frequency

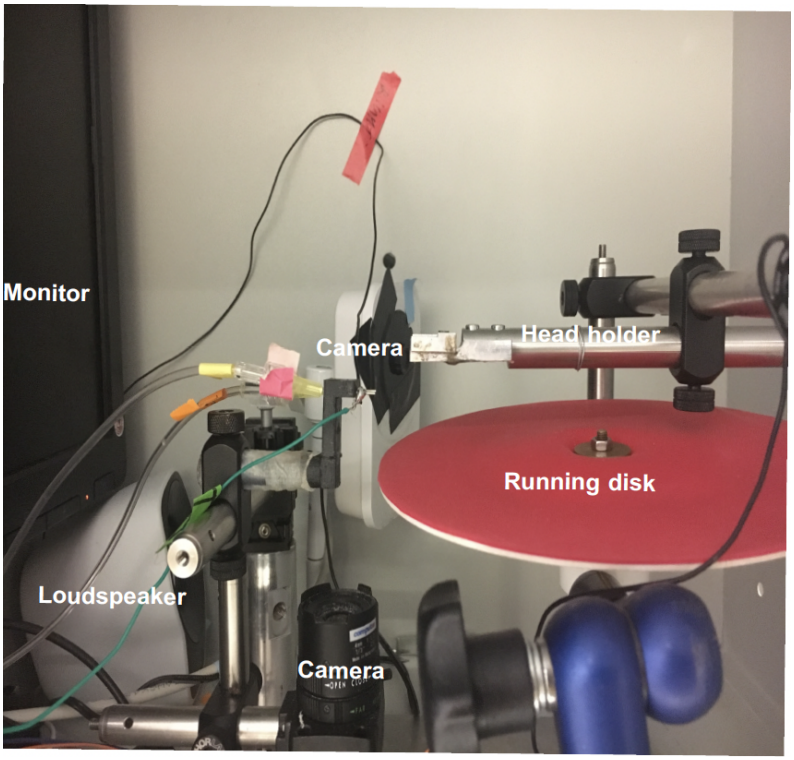


Figure 3.1: Behavioral Setup where the visual discrimination task was performed

was 1.5 Hz. The stimulus duration was a minimum of 0.2–0.3 s and maximal of 2–3 s.

### 3.2.2 *Behavioral measurements*

*Locomotion*

A miniature rotary absolute shaft encoder was placed on a rotating disk. The MA3 reports the shaft position over 360° and provides analogue voltage output in 10 bit resolution at a 2.6 kHz sampling rate, sampled at 100 Hz. Running speed was calculated using the reported voltage differences between two successive samples. In the task, animals were defined to be in stationary position when the animals' speed fell below 20°/s.

*Lick Sensor*

Based on Goltstein et al. (2018), Weijnen (1989) and Slotnick (2009), a custom-made lick detection circuit was used. Each lick spout was a part of a particular circuit, and the mouse short-circuited by licking the spout. This resulted in a voltage drop, which was registered by the digital port of the data acquisition system (Labjack U6). Refer to Goltstein et al. (2018) for detailed description of the sensor.

For experiments involving electrophysiological studies, due to electrical interference with the neuronal recording, a current-based sensor could not be used. Hence, a piezo-based sensor was used (Schwarz et al., 2010). A miniature piezo element, which detected vibration, was glued to the lick spout. The detected output voltage was amplified, high pass filtered, and a lick was detected when a certain threshold was crossed. Refer to Schwarz et al. (2010) for detailed description of the sensor.

*Behavior monitoring and eye tracking*

The setup were equipped with two cameras. One camera (IC-3140W, Edimax) was used to monitor the mouse behavior. The second camera (FMVU-03MTM-CS, FLIR Integrated Imaging Solutions and 4 mm 1:1.2 Camera Lens, Computar) was mounted below the lick spout and was used to monitor the licking behavior and for precise positioning of the lick spouts. For the electrophysiological experiments, eye movements were monitored under infrared illumination using a zoom lens (Navitar Zoom 6000) coupled to a camera (Guppy AVT, frame rate 50 Hz).

### 3.2.3 *Behavioral training*

One week after head-bar implantation, mice were habituated to head fixation on the disk. After at least 1 additional week, they were placed in a water restriction regime. Following the head-bar implantation, body weight, and health score were monitored daily, and mice with weight losses greater than 15% or any distress were removed from the experiment and placed on *ad libitium* water. After training started, mice received water for correctly performing the task. Training sessions

were typically performed 5 days a week. On days without training, mice were given at least 25 mL of water per kg.

*Behavioral Tasks*

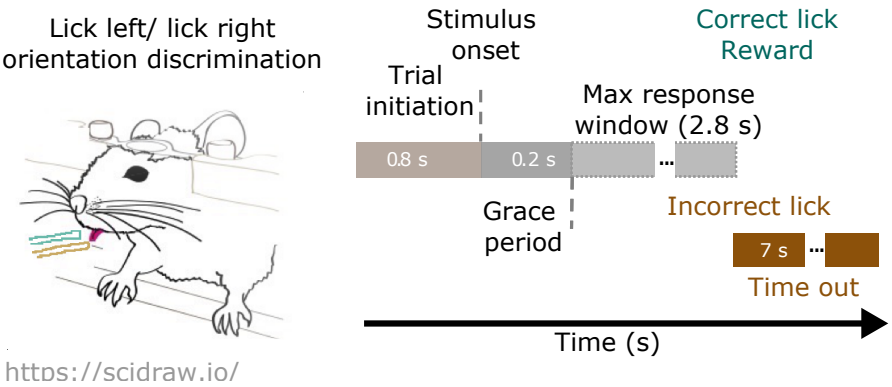


Figure 3.2: Schematic of the final task and trial structure.

During the initial stage of behavioral training, mice were given a water drop (~ 4 µL) as a reward for licking one of the two lick spouts. During this '0% contrast stimulus' training stage, no grating but only the mean luminance gray screen was presented. To discourage mice from developing any bias towards a particular side, the mice could only lick a particular lick spout a maximum of 5 consecutive time, after which they had to switch to the other spout to receive a reward. This training regime was followed until mice were able to fulfill the minimum daily water intake requirement (25 mL per kg) through the experiment alone (i.e., until there was no need to provide them with additional water outside the experiments). This typically took 1-2 training sessions. Following these initial training sessions, rewards were only provided when mice sat still (no running or licking) for 0.8 s. The animals typically moved to the 'switch task' training phase after 2-3 days.

During the 'switch task' phase of training, sitting still triggered the presentation of a grating stimulus (2–3 s duration). A water reward was delivered to the mice after the first correct lick, following a brief delay from the onset of the stimulus (0.2 to 0.3 s; grace period). During this grace period, the mice were still permitted to lick, but their licks

were not taken into account for decision-making purposes. Licks to the incorrect side were ignored and did not affect the possibility of rewarding the first correct lick later in the same trial. Trials without a response after 2–3 s of stimulus presentation and additional 1 s of gray screen were considered missed trials. After mice performed a few sessions in this training phase, incorrect trials were introduced. The decision in each trial was determined by the first lick after the grace period. A correct choice resulted in a water drop reward, while incorrect choices were penalized with a time-out period lasting 7 to 9 s. During this time-out, a flickering visual stimulus was displayed, and white noise was played. After a rewarded trial, the stimulus associated with the reward was shown for 1 to 2 seconds.

*Reward manipulation*

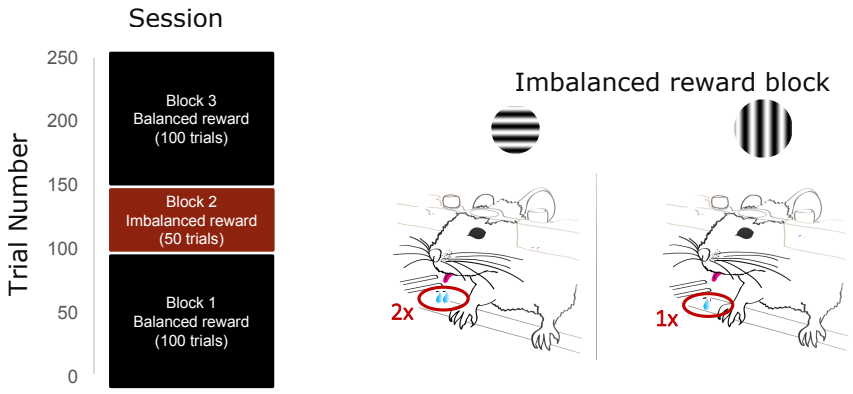


Figure 3.3: Session schematic for behavior paradigm with imbalanced reward condition

After animals reached stable performance and enough trials were accumulated for analysis, to understand how mice trade between stimulus-related and contextual information, a block of trials with stimulus-specific imbalances of reward was introduced. Specifically, in each session, mice experienced a middle block of 50 trials, in which they received double the reward for correctly choosing the lick spout associated with one of the two stimuli (e.g., horizontal). No other aspect of the task changed in these blocks, and hence the imbalanced reward blocks, in principle, did not introduce any new behavioral

requirements. To ensure that the animals did not develop any long term changes in strategy or biases, such imbalanced reward blocks were preceded and followed by blocks with balanced reward; in addition, the stimulus and hence lick side with double reward alternated between sessions.

### 3.2.4 *Surgery: head bar implantation*

Animals were subcutaneously administered Metamizole (200 mg/kg, sc, MSD Animal Health, Brussels, Belgium). 30 minutes later, they were transferred to an induction chamber and anesthetized using isoflurane (5% in oxygen, CP-Pharma, Burgdorf, Germany). Mice were then transferred to a heated-pad on the stereotactic apparatus (Drill & Microinjection Robot, Neurostar, Tuebingen, Germany). Anesthesia level was adjusted (0.5%–2% in oxygen) and consistently monitored by verifying the absence of the pedal reflex, breathing and heart rate. The body temperature was maintained at 37° C and monitored using a closed-loop temperature-control system (ATC 1000, WPI Germany, Berlin, Germany). During the entire surgery procedure following anesthetic induction the eyes were protected with an ointment (Bepanthen, Bayer, Leverkusen, Germany) and the head was stabilized using ear and bite bars.

Once the animal was in stable condition, Buprenorphine (Buprenorphine, 0.1 mg/kg, sc, Bayer, Leverkusen, Germany) was administered. Following which the animal's hair was thoroughly removed from the head using a depilatory ointment. The skin was then disinfected using iodine solution (Braun, Melsungen, Germany). A local anglesic Lidocaine hydrochloride (7 mg/kg, sc, bela-pharm, Vechta, Germany) was subcutaneously injected along the midline. Following this a small incision was made along the injection line using a scalpel. A small part of the skin where the head bar was to be placed was removed. The tissue residue around the incision site was removed and carefully cleaned using a drop of $H_2O_2$ (3%, AppliChem, Darmstadt, Germany).

Before head bar implantation, it was verified that the animal's head was in a skull-flat configuration. For this the four landmarks (lambda, bregma and two points 2 mm to the right and to the left, respectively) were utilized. For mice targeted for the electrophysiology study, V1 (AP: -2.8 mm, ML: -2.5 mm) was marked to locate the future craniotomy site (see **subsection 3.3.2**). For better adhesion of the head plate to the skull, OptiBond FL primer and adhesive Kerr dental (Kerr dental, Rastatt, Germany) was applied on the exposed skull. The V1 marked site and a position approximately 1.5 mm anterior and 1 mm to the right of bregma were left exposed, as these sites were for craniotomy and a miniature reference screw.

A lightweight stainless steel bar, with a round opening for the recording site, was used for the head plate. The head bar was positioned over the posterior part of the skull. For those mice used in electrophysiological recordings, the round opening was centered on the marked recording site. A thin layer of UV curing dental cement (Ivoclar Vivadent, Ellwangen, Germany) was applied to the primer/adhesive layer on the skull. The head-plate was placed on top of the cement and UV light was used to cure the cement. For the animals used only for behavioral studies the opening was also covered with dental cement, for the others it was filled with the silicone elastomer sealant Kwik-Cast 604 (WPI Germany, Berlin, Germany). For animals used in electrophysiology study, a miniature reference screw (00-96 X 1/16 stainless steel screw, Bilaney) was soldered to a miniature pin and was used as grounding. The screw was placed posterior to the head post and covered with dental cement.

At the end of the surgery the anesthesia was stopped, an antibiotic ointment (Imex, Merz Pharmaceuticals, Frankfurt, Germany) or iodine-based ointment (Braunodivon, 10%, B. Braun, Melsungen, Germany) was applied around the wound and long term analgesic (Meloxicam, 2 mg/kg, Bohringer) was subcutaneously administered. The animal was the transferred to an infrared light-heated cage with water and food for recovery. The cage was covered with surgical tissue instead of wood shaving. For 3 days post surgery (24 h interval), the animal was subcutaneously administered with long term analgesic (Meloxicam, 2 mg/kg, Bohringer). For at least 7 days post surgery, the animal's health status was monitored, and recorded in a scoresheet. After at

least 1 week of recovery, the animal was gradually habituated to handling and head fixation. The habituation phase consisted of at least 3 days of handling, followed by head fixation. During this recovery period, the animals were not trained and had food and water *ad libitum*.

## 3.3 ELECTROPHYSIOLOGICAL RECORDINGS

In order to achieve long-term neuronal recording, I chronically implanted an immobile silicon probe (see **subsection 3.3.2** for a detailed description). In **subsection 3.3.1** I describe the neuronal visual stimuli used in the electrophysiological studies, and in section **subsection 3.3.2** I describe the surgical procedure for craniotomy.

### 3.3.1 *Visual stimuli during neuronal recording*

On most of the recording days, the animal performed 4 passive viewing and one behavioral tasks The passive viewing experiments were used to monitor neuronal stability and responses over time. The first experiment was orientation tuning, followed by the behavioral task, sparse noise, checkerboard and finally an additional orientation tuning experiment. In the following section, I give a brief overview of the different stimuli presented during these experiments.

*Orientation tuning*

Drifting gratings with temporal and spatial frequencies of 0.05 Hz and 1.5 cyc/° respectively were used for orientation tuning experiments. The gratings were presented for 2 s with a 0.5 s inter-trial-interval and with 50% contrast. The grating direction was varied in step sizes of 30°. Each of the orientations were repeated 20 times. For all tuning experiments, the spontaneous firing rate was assessed by including trials in which only the mean luminance gray screen was presented. The stimulus parameters matched the parameters used for the gratings in the behavioral experiments.

*Sparse noise*

I used sparse noise stimulus to map the ON and OFF subfields, to measure receptive fields (RFs). The stimulus was presented for 100 ms on a gray background at a random location on a virtual 16 × 16 grid. The stimulus was 5°, full-contrast non-overlapping black and white squares. The RF centers for both ON and OFF sub-fields were determined by fitting the neural responses with a 2D Gaussian (Liu et al., 2010). I estimated the RF parameters online and relied on threshold crossings of spiking activity at each recording channel, to place the monitor in a position that covered as many RFs as possible.

*Checkerboard stimulus*

To determine the laminar location of the recording sites, mice were presented with full-screen reversing checkerboard stimuli at 100% contrast. The size of the stimulus was 5° and the temporal frequency was 0.5 cyc/s. The stimulus duration was 1 s without any inter-trial interval. Each experiment consisted of 150 contrast reversal.

*Behavior stimulus*

Visual stimuli matching the temporal and spatial frequency of orientation tuning experiments (see **subsection 3.3.1**) were used for behavior experiments. Stimuli were either vertical (270°) or horizontal (180°) drifting sinusoidal gratings.

### 3.3.2 *Craniotomy and Probe implantation*

Procedures were equivalent to that discussed in **subsection 3.2.4** with the following exceptions.

*Craniotomy*

At least 3 days before craniotomy, animals were provided with *ad libitum* food and water. 24 hours before the probe implantation, mice were fully anesthetized using the same protocol mentioned in **subsection 3.2.4**. The silicone Kwik-Cast was removed from the craniotomy site and a craniotomy of approximately 1.5 $mm^2$ was performed over

V1 using a drill size of 02. The craniotomy site was protected with Kwik-Cast. At the end of craniotomy, the analgesic Metacam was administered. Metacam was administered in subsequent days only if the animal showed any kind of discomfort.

*Probe implantation*

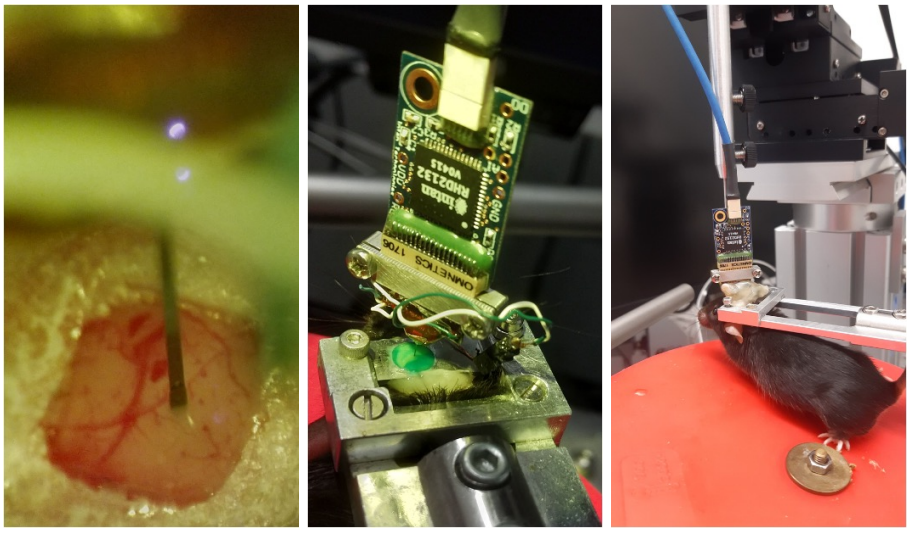

Figure 3.4: Implantation of chronic immobile silicone probe. (**A**) Craniotomy with inserted probe. (**B**) Brain covered with kwick-cast and inserted probe with head stage attached. (**C**) Animal with probe running on a disk.

The probes were implanted on three animals after training them on the behavioral task for approximately three weeks. Animals were first disk-trained, placed on a water-restriction regime, and trained to lick and receive rewards. After they learned to do so, they were again provided with *ad libitum* water for at least two days. The craniotomy (**subsection 3.3.2**) was performed and 24 hours later, the probe was implanted in a separate rig dedicated to electrophysiology.

For probe implantation, the animals were awake. As they were already accustomed to the disk and head fixation, this did not provide any additional discomfort. The animal was head-fixed and the probe was lowered using a micro-manipulator to the required depth. I aimed for

the primary visual cortex (V1), i.e., a depth of 1000-1200 µm below the cortical surface.

Throughout probe implantation, the probe was connected to the head stage and held by a neuronexus probe holder (IST-CM adapter). Neuronal recordings were performed using the OpenEphys system, and LFP and spiking activity were monitored. Neuronal recording during the implantation procedure allowed me to monitor any changes in neural activity during the implantation process and to position the probe at the desired depth and location. The location of the probe was first verified by using a mapping stimulus, which could be interactively controlled by the experimenter (hand-map). For example, I used hand-map to hear neural activity and to roughly check receptive field (RF) location. Before cementing the probe, the receptive fields were mapped and orientation tuning experiments (**subsection 3.3.1**) were also performed.

After verifying the probe position, I covered the craniotomy site with a layer of Kwik-Cast. Finally, I applied several layers of dental cement (Ivoclar Vivadent, Ellwangen, Germany), and exposed each layer with UV light before adding an addition layer. In the final stage, I disconnected the head stage, removed the holder, and applied the final layers of cement. Throughout the process, the eyes of the animals were protected from the UV light. The animals were returned to their home cage (along with their cage mates) and were allowed to recover for at least one week before performing additional experiments.

## 3.4 DATA ANALYSIS

Data analysis was performed using Python. Data were organized in a custom written schema using the relational database framework "Datajoint" (Yatsenko et al., 2018).

### 3.4.1 *Calculation of behavioral performance*

Behavioral performance was calculated as the ratio of the number of correct trials to the number of completed trials. Missed trials were removed from all analyses except when mentioned. The criterion for an animal to have reached stable performance was when the animal's performance was ($\geqslant$ 66% correct) for 3 consecutive days (Goltstein et al., 2018). Weekly performance was computed by pooling across all sessions within a week, with Sunday being the first day of the week.

### 3.4.2 *Logistic model analysis: integration of sensory evidence and recent history*

To quantify the behavior of the mice and to understand the task relevant and task irrelevant factors, I carried out analyses to weigh the contribution of the current sensory stimulus and various history-related features on decisions. To achieve this, I used a logistic regression (LR) Model. In the LR model, the log-odds of the dependent variable is modeled as a weighted linear combination of independent variable.

$$Y = A * B_0 + B * B_1 + C * B_2 + ..... + Bias \tag{3.1}$$

This linear combination is then mapped non-linearly using the logit function to infer a probability that the mouse selects the port associated with the vertical grating on each trial.

$$P(Vertical) = \frac{1}{1 + e^{Y}} \tag{3.2}$$

To understand the different strategies of the mice, I used various variants of the models. Weights were fitted to each individual mouse separately. In the following section, I give a brief overview of these variants.

*History Dependent Variables*

The sensory information comprised of the current (CS) and previous (PS) visual stimulus. The non-sensory information consisted of previous response (PR).

$$Y = CS * W_0 + PS * W_1 + PR * W_2 + Bias \tag{3.3}$$

- The CS measured the influence of the current visual stimulus on the decision, i.e. the strength of stimulus-choice associations. The model consisted of the CS which took the value 1 for 270° orientation and -1 for 180°.

- The PS measured the bias due to the presentation of the PS. Consistent with the coding for the CS, the PS took the value 1 for 270° orientation and -1 for 180°.

- The PR measured the influence of the animals' PR on the current decision. Here, PR towards the spout associated with the 270° orientation took the value 1, and those towards the spout associated with 180° orientation took the value -1.

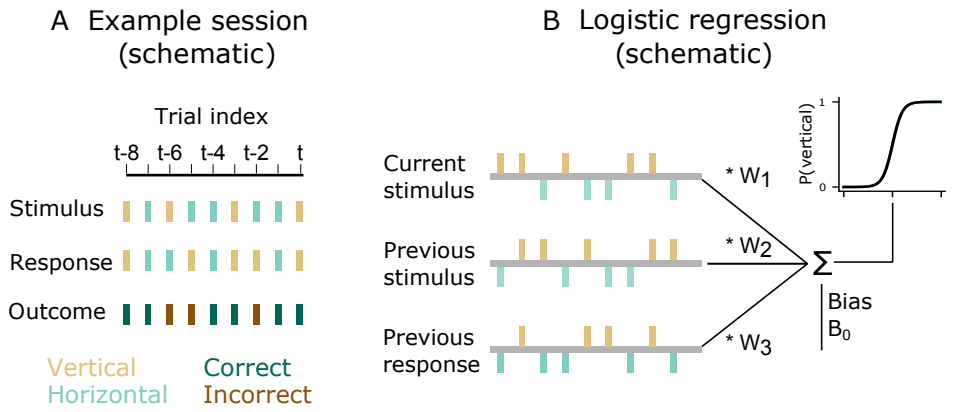


Figure 3.5: Schematics of LR model to capture the influence of CS, PS and PR on choice.
(**A**) Example stimulus sequence (*turquoise* for stimulus with horizontal grating and *yellow* for stimulus with vertical grating), response sequence (*turquoise* for lick on the spout associated with horizontal grating and *yellow* for lick on the spout associated with vertical grating), and associated outcomes (*green* for correct and *brown* for incorrect lick) (**B**) LR model predicting the probability to choose vertical based on a weighted combination of 'Current stimulus', 'Previous stimulus', and 'Previous response'.

Here, the probability to choose the lick port associated with the vertical stimulus was predicted by a weighted sum of influences related to the current visual stimulus (horizontal or vertical), the previous-trial stimulus (horizontal or vertical), and the previous-trial response. The model found optimal weights and a bias term, which were mapped non-linearly through a logistic function to yield the probability to lick the spout corresponding to the vertical stimulus.

*Outcome based model*

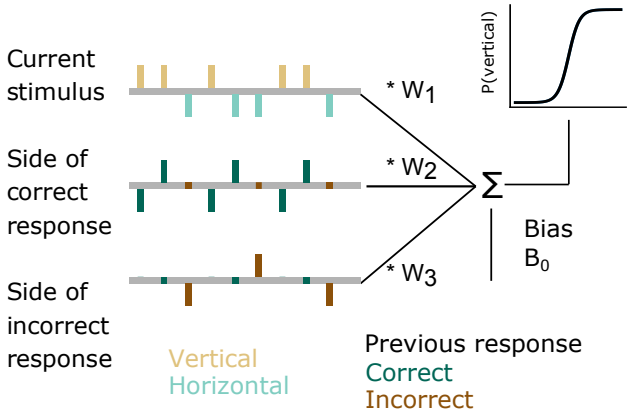


Figure 3.6: LR model that predicts the probability of choosing the spout associated with the vertical grating based on the previous outcome.

The second variation of the model captured the effect of bias resulting from previous outcome. Besides the current sensory stimulus, this variant of the model included side of correct and incorrect response, to explain the history dependent bias to repeat or alternate its previous response side (Hermoso-Mendizabal et al., 2020).

$$Y = CS * W_0 + CR * W_1 + IncR * W_2 + Bias \tag{3.4}$$

- The side of correct response (CR) is coded as -1 for correct response associated with the horizontal like spout, +1 for correct response associated with the vertical lick spout, and 0 for any incorrect response. Positive weight assigned to CR provide evidence for win-stay approach, i.e., bias towards the previously rewarded lick spout.

- The side of incorrect response (IncR) is coded as -1 for incorrect choices associated with the horizontal lick spout, 1 for incorrect choices associated with the vertical lick spout, and 0 for any correct response. Negative weight assigned to IncR captured a lose-shift approach, i.e., an avoidance bias towards the previously unrewarded lick spout.

In summary, in this variant, I modified the model and exchanged the predictors for PS and PR by two alternative predictors reflecting their

interaction. The interaction terms (side of correct response and side of incorrect response) reflect a "stay" (positive weight) or "shift" (negative weight) strategy towards a previously chosen side, depending on the previous outcome.

*Interaction with RT*

As discussed in **subsection 1.3.1**, RT influences decision-making. To understand this influence, I also included RT as a predictor. The predictors for CS, PS, and PR were divided into two bins each (S and L), with S referring to trials with short RT ($<$ 0.5 s), and L to trials with long RT ($>=$ 0.5 s) on the current trial.

During learning (**Figure 4.3** and **Figure 4.4**) only trials with long RT on the current trial were considered. In **Figure 4.5A**, **Figure 4.6A**, **Figure 4.10** , **Figure 4.11** and **Figure 4.12** the RT on the current trial was taken into account for all variables. In **Figure 4.5B**, **Figure 4.6B** the RT on the current trial was taken into account for task relevant variable and the RT on previous trial was taken into account for history related variables.

| Regressors | CS | | PS | | PR | |
|---|---|---|---|---|---|---|
| Orientation | Ver | Hori | Ver | Hori | Ver | Hori |
| Short RT | 0 | 0 | 0 | 0 | 0 | 0 |
| Long RT | 1 | -1 | 1 | -1 | 1 | -1 |

Table 3.1: Example variable coding for trials with long reaction time

*LR for manipulated trials*

For manipulated trials, I expressed choices with respect to the rewarded stimulus, hence the side with double reward was coded as +1 and that with single reward was coded as -1.

### 3.4.3 *Predictive accuracy test*

To measure the predictive accuracy reported in **Figure 4.3F** and **Figure 4.11C**, the classification table approach was used. In this table, the observed value for the dependent outcome and the predicted values are cross-classified. For the values reported in this thesis, a cut-off value of 0.5 was used: For example, all predicted values above 0.5 were classified as predicting that the animal chooses the lick spout associated with a vertical grating, and for the rest, the chosen side was predicted to be to the lick spout associated with the horizontal grating. The predicted decision was compared with the actual choice that the animal made, to find the accuracy of the model.

$$Accuracy = 100 * \frac{\text{n (Correctly predicted choices)}}{\text{n (All choices)}} \tag{3.5}$$

### 3.4.4 *Validation of logistic regression coefficient and model*

To calculate the significance of the model weight and accuracy, the bootstrapping method was used. The confidence interval for all coefficients, accuracy, and bias was computed by dropping 10 % of the data, refitting the model and extracting the parameters of interest. This was repeated 1000 times and 95 % confidence intervals were calculated for the parameters of interest.

### 3.4.5 *Optimal bias for manipulated trials*

When reward sizes vary across stimuli, it is possible to increase the expected reward by introducing a bias toward the stimulus that is associated with a higher reward. The expected reward value can be represented as

$$E[R] = P(Stim_H) * P(Choice_H \mid Stim_H, B_{ACC}, Bias) * Reward_H + P(Stim_V) * P(Choice_V \mid Stim_V, B_{ACC}, Bias) * Reward_V \tag{3.6}$$

Here, H and V refer to Horizontal and Vertical stimulus respectively, and $B_{ACC}$ represents baseline accuracy. $P(Stim_{H/V})$ and $Reward_{H/V}$ are fixed by the experimenter.

The conditional choice probability $P(Choice_i \mid Stim_i, B_{ACC}, Bias)$ can be represented as

$$P(Choice_i \mid Stim_i, B_{ACC}, Bias) = \frac{1}{1+e^{-(f(B_{Acc})+\delta bias)}} \tag{3.7}$$

where $f(B_{Acc})$ refers to the value that sets the probability of making a correct choice equal to baseline accuracy when there is no bias.

$$\begin{aligned} &P(Choice_i \mid Stim_i, B_{ACC}, Bias = 0) = B_{ACC} = \frac{1}{1+e^{-f(B_{Acc})}} \\ &\Rightarrow B_{ACC} = \frac{1}{1+e^{-f(B_{Acc})}} \\ &\Rightarrow f(B_{Acc}) = \ln(\frac{1-B_{ACC}}{B_{ACC}}) \end{aligned} \tag{3.8}$$

and delta controls whether to add or subtract the bias depending on the stimulus

$$\delta = \begin{cases} 1, & Stim_H \\ -1, & Stim_V \end{cases}$$

Conditioning **Equation 3.6** on bias yields:

$$E[R \mid Bias] = \frac{P(Stim_H) * Reward_H}{1 + e^{-(f(B_{Acc})+Bias)}} + \frac{P(Stim_V) * Reward_V}{1 + e^{-(f(B_{Acc})-Bias)}} \tag{3.9}$$

Maximizing **Equation 3.9** with respect to bias yields the reward maximizing bias.

### 3.4.6 *Statistical analyses*

Unless noted otherwise, the error-bars shown in all plots for behavioral analysis for individual mice represent 95 % confidence intervals, calculated via the bootstrapping method. Standard deviation across animals is reported when the mean value for the animals is reported. The P value is obtained for the mean value across animals from the Student's t-test.

### 3.4.7 *Electrophysiological analyses*

*Preprocessing and Spike Sorting*

The extracellular signal was continuously recorded using a 32 channel recording electrode (CM-32, NeuroNexus). The signals were first filtered between 1 and 7600 Hz and digitized using the Intan head stage (RHD 2132). I first used Kilosort (Pachitariu et al., 2016) to obtain single unit activity from extracellular recordings. This open source automated Matlab toolbox removed any saturations and clustered responses. Subsequently, I used Spyke (Spacek et al., 2009) to manually refine the clusters. This Python-based toolbox allowed me to select time ranges and channels around clusters for realignment of spikes and for representation of cluster in three dimensions. This representation was done via dimensionality reduction (multichannel PCA, ICA and/or spike time). The clusters were further split, using a gradient ascent based algorithm (GAC). Finally, following pairwise comparison, similar clusters were merged.

*Firing rates*

For each unit, the visual responses to a specific stimulus was computed as the mean firing rate over the time window defined by stimulus onset and offset. The error bar indicates the standard error of the mean.

For **Figure 4.16**, the difference in firing rate between two conditions is reported. The error bar indicates the standard error of the mean. The p value is calculated using a two sample t-test.

*Orientation tuning*

As mentioned in **subsection 3.3.1**, drifting sine wave gratings were used to test neuronal responses to different orientations. Orientation tuning curves in **Figure 4.15B, D** were fitted to the neurons' responses with a sum of two Gaussians, having peaks 180° apart. The amplitude of two Gaussians was allowed to vary, but baseline and width were restricted to be the same (Liu et al., 2009) .

$$R(\theta) = R_0 + R_p e^{\frac{-(\theta-\theta_p)^2}{2\sigma^2}} + R_n e^{\frac{-(\theta-\theta_p+180)^2}{2\sigma^2}} \tag{3.10}$$

Where,

- $R_0$: baseline response
- $R_p$: response at preferred direction
- $\theta$: stimulus orientation (0° – 360°)
- $\theta_p$: preferred orientation
- $\sigma$: tuning width
- $R_n$: neuron response at the null direction.

*Receptive field fitting*

Receptive field maps shown in **Figure 4.15A** were obtained in sparse noise experiments. The average activity during stimulus presentation to the sparse noise stimulus were fitted to the center of 2D Gaussian for both ON and OFF fields.

$$f(x,y) = \frac{A}{2\pi ab} e\left(\frac{x'^2}{2a^2} - \frac{y'^2}{2b^2}\right) + c \tag{3.11}$$

Where,

- $A$: the maximum amplitude
- $a$ and $b$: ellipse half-axes

- $x'$ and $y'$: the transformations of the stimulus coordinates x and y considering the angel ($\theta$) and the center of the ellipse ($x_c$, $y_c$)

- c: offset

*Euclidean distance*

To identify the same unit recorded across days in **Figure 4.14**, I calculated the Euclidean distances between each point of the spike wave forms between a pair of units.

$$d(p, q) = \sum_{i=1}^{n} (q_i - p_i)^2 \tag{3.12}$$

As single units recorded on the same day were obtained after exhaustive pair-wise comparison, I used the minimum Euclidean distance between these units, as the threshold for distinct units.

### 3.4.8 *Identification of recorded brain area*

To verify the site of probe implantation and to access damaged caused to the tissue, I performed histology.

Following the final recording session, mice were subcutaneously administered Metamizole (200 mg/kg, sc, MSD Animal Health, Brussels, Belgium). 30 minutes later, they were deeply anesthetized using a mixture of Medetomidin (Domitor, 0.5 mg/kg, Vetoquinol, Ismaning, Germany), Midazolam (Climasol, 5 mg/kg, Ratiopharm, Ulm, Germany) and Fentanyl (Fentadon, 0.05 mg/kg, Dechra Veterinary Products Deutschland, Aulendorf, Germany). Animals were then perfused first with Ringer's lactate solution and subsequently with 4 % paraformaldehyde (PFA) in 0.2 M sodium phosphate buffer (PBS). The brain and the probe were carefully removed. The brain was postfixated in PFA buffer for 24 hours and then rinsed with and stored in PBS at 4°C.

Before slicing, the brain was first washed in PBS. To preserve the brain shape during cutting, liquid agarose was poured and left to solidify. Afterwards, the brain was glued to the vibratome holder (Leica VT1200 S, Leica, Wetzlar, Germany). The primary visual cortex areas was cut in 40 µm thick slices and each slice were placed in microwells filled with PBS solution.

Some slices were stained with DAPI stain (DAPI, Thermo Fisher Scientific, Waltham, Massachusetts, USA; Vectashield H-1000, Vector Laboratories, Burlingame, USA) and mounted on glass slides, and cover-slipped. Other slices were Nissl stained. For the Nissl stained slices, a microwell was filled with a PBS solution with 1:300 green nissl concentration. The microplate was left for at least 12 hours in the shaker. Following this, in 20 min intervals between each wash, slices were washed 3X in PBS. The slices were finally placed on glass slides and cover-slipped. The stained brain slice was assessed for recording site and tissue damage.

# 4

# RESULTS

## 4.1 BEHAVIOR PERFORMANCE ANALYSIS

In the following section, behavioral data obtained from five mice are discussed. I first analyzed performance during the behavioral task. Next, I used LR to disentangle the impact of task relevant and irrelevant factors during learning. I further focused on behavior during stable performance to emphasize the impact of task-irrelevant factors on choice. Towards, the end of this section, I report on the use of reward expectation, to show that irrelevant factors further bias choices, even for trained animals.

### 4.1.1 *Mice learned the 2AFC orientation discrimination task*

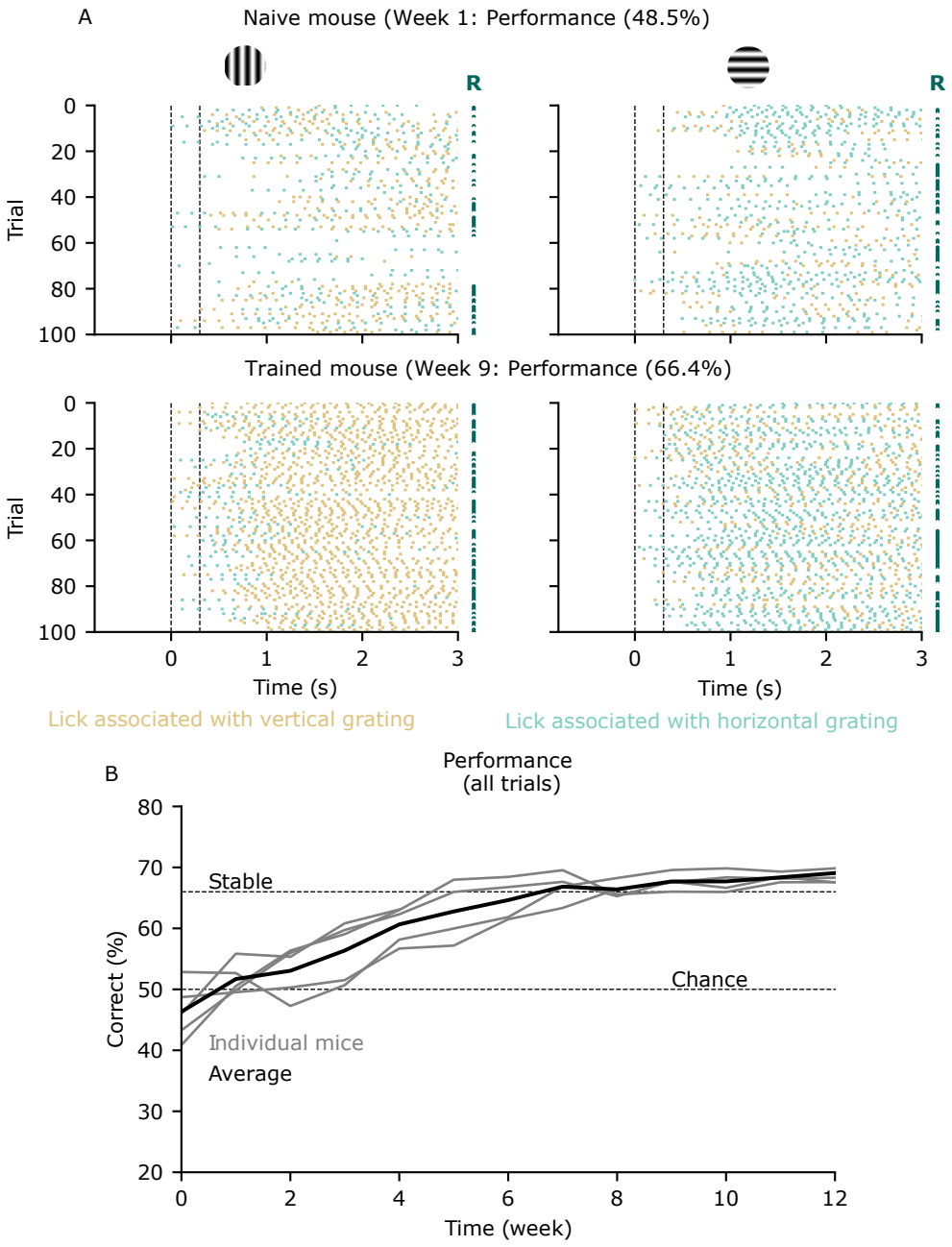


Figure 4.1: Mice learn to perform a lick-left / lick-right orientation discrimination task. *(Continued on next page)*

Figure 4.1: *(Continued)* (**A**) Two example sessions, showing licking behavior of an example mouse during the naive training stage and during stable performance. Each row corresponds to a trial, with trials sorted by orientation (*left*: vertical; *right*: horizontal). Each tick represents a lick (*turquoise*: lick on the spout associated with the horizontal grating; *yellow*: lick on the spout associated with the vertical grating). *Dark green dots*: correct trials. *Black dashed lines*: grace period. (**B**) Performance for all trials in the orientation discrimination task, computed by pooling across all sessions within a week.

To investigate how mice combined stimulus information and prior experiences during visual PDM, I trained head-fixed mice to perform a lick-left / lick-right orientation discrimination task. Refer to **subsection 3.2.3** for a detailed description of the behavioral task. I first studied how mice learned this task. As expected, early in training, the mice performed at chance level: for either the horizontal or the vertical orientation, they randomly licked on the right or left lick spout (**Figure 4.1A, top**). As training progressed, mice licked the rewarded side more consistently (**Figure 4.1A, bottom**). Mice learned the task in several weeks of training, reaching the criterion for stable performance ($\geqslant$ 66% correct) for 3 consecutive days in 6-8 weeks ($67.20 \pm 0.50$, $N = 5$ mice; **Figure 4.1B**).

### 4.1.2 *Mice performed better on trials with long RTs*

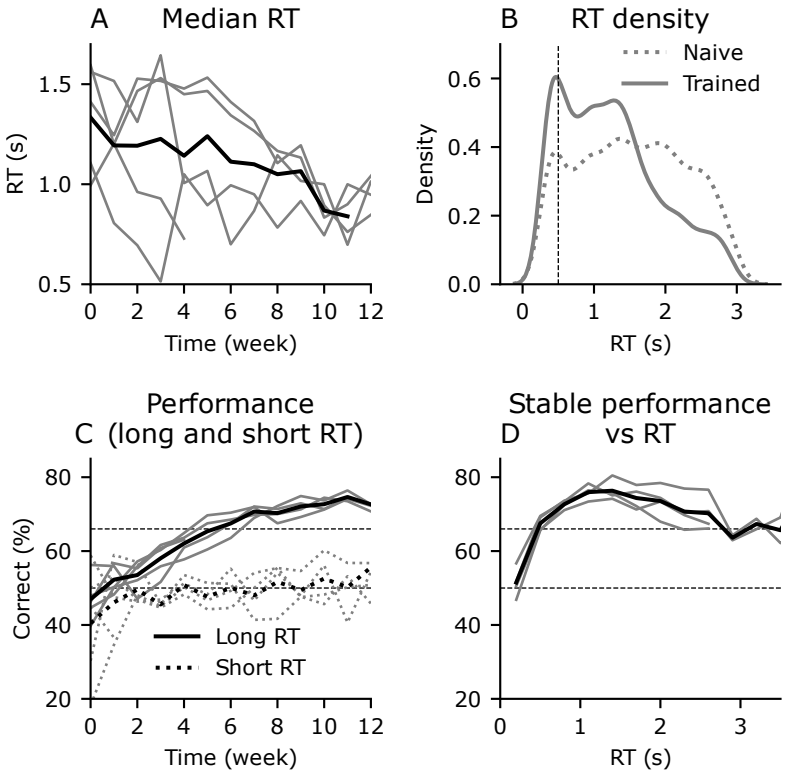


Figure 4.2: Development of RT and performance during learning.
(**A**) Median RT (with respect to stimulus start) in the orientation discrimination task, computed by pooling across all sessions within a week. (**B**) Cumulative RT distributions for an example mouse during the naive (*dotted line*) and trained stage (*solid line)*. Vertical dotted line indicates the separation boundary for short and long RT. (**C**) Performance for trials with short (*dotted lines*) and long RTs (*solid lines*). (**D**) Performance vs RT during stable performance. (A, C, D) *grey*: individual mice; *black*: mean across $n = 5$ mice.

During learning, not only did the overall performance improve, but also RT decreased systematically (**Figure 4.2A**), especially for those mice which had longer RTs in the early learning phase. In particular, the RT density plot revealed that there was a conspicuous increase in

short latency licks (RT $< 0.5$ s) after learning (**Figure 4.2B**). To understand whether short latency licks reflected better task performance or instead reflected responses driven by other factors, such as impulsive licking or timed responses without considerations of the visual stimulus, trials were separated into those with short ($< 0.5$ s) or long RTs ($\geqslant 0.5$ s) (**Figure 4.2C**). This analysis showed that for trials with short RT, the mice performed at chance level throughout and even after learning. In contrast, performance for long RT trials improved and plateaued around $72 \pm 1.5\%$. This phenomenon persisted even during stable performance. To illustrate this, I calculated performance for trained mice as a function of RT in $0.3$ s bins. For all mice, the performance gradually improved with increasing RT, plateauing at an optimal performance of $\approx 75.5\%$ at RTs of $1.1 - 1.7$s. Notably, the performance declined slightly following this plateau, however, the performance remained close to 66%. This further supports the hypothesis that short RT trials during and after learning correspond to trials with suboptimal performance and potentially overall lower task engagement.

### 4.1.3 *During learning, weight assigned to CS increased, while those assigned to PS and PR decreased.*

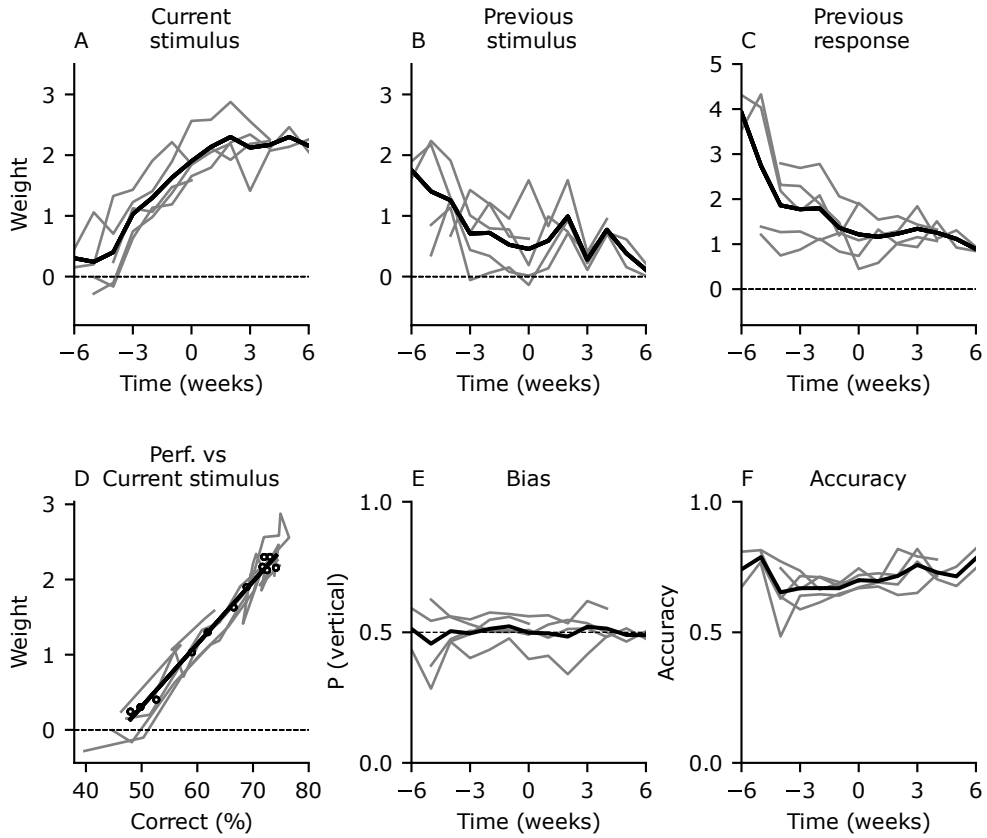


Figure 4.3: Adjustments of weights assigned to task-relevant and task-irrelevant influences during learning.
(**A**-**C**) Development of model weights across learning for 'Current stimulus' (A), 'Previous stimulus' (B), 'Previous response' (C). (**D**) Correlation between task performance and weight assigned to 'Current stimulus'. (**E**) Bias term $B_0$ of the model; dotted line i.e. 0.5 indicates no bias. (**F**) Model accuracy, computed as the proportion of correct prediction by the model. In (A, B, C), week 0 corresponds to the week in which individual animals reached stable performance (66% correct across three consecutive days). Trials with short RTs ($< 0.5$ s) were excluded from this analysis. *Grey*: individual mice; *black*: mean across $n = 5$ mice.

As discussed in **subsection 3.4.2**, in the first variant of the model I predicted the probability to choose the lick port associated with the vertical stimulus by a weighted sum of influences related to the current visual stimulus (CS) (horizontal or vertical), the previous-trial stimulus (PS) (horizontal or vertical), and the previous-trial response (PR). To quantify which factors might drive the change in mouse behavior during learning, I next disentangled the impact of several task-relevant and task-irrelevant influences on choices with a LR framework (**Figure 4.2C**). Since short RT trials were indicative of poor performance and potentially lack of task engagement, the trials with RTs $< 0.5$ s were excluded from this analysis.

To understand how CS and PS and the PR contributed to the improvement in task performance, I analyzed the development of model weights throughout the learning progress. The larger the magnitude of a particular weight, the more the animal's behavior relied on the corresponding factor. During learning, the weight assigned to the CS increased (**Figure 4.3B**). This observation is consistent with the improvement in behavioral performance over weeks (**Figure 4.2** C solid line). In contrast, the weights assigned to task-irrelevant influences, i.e., the PS and the PR, decreased (**Figure 4.3C–D**), suggesting that over time mice correctly identified the association between stimulus orientation and reward. In fact, week-by-week performance in the task was linearly related to the weight assigned to the CS ($R^2 = 0.98 \pm 0.0094$; **Figure 4.3D**).

Overall, the animal did not show a bias towards any particular side throughout learning and stable performance (**Figure 4.3E**). This was most likely influenced by the training phases introduced before the final task, where animals were discouraged to repeatedly choose the same side (see **subsection 3.2.3** for task description).

This simple LR model (**subsection 3.4.2**) also contained the majority of influences on choice, because it was able to correctly predict more than 70% of the variance on choice (**Figure 4.3F**). In addition, the overall accuracy of the model was consistent over the weeks during learning, suggesting that the major contributions to the animals' choices can be described by the same variables throughout learning and stable performance.

### 4.1.4 *During learning, the impact of the win-stay strategy decreased*

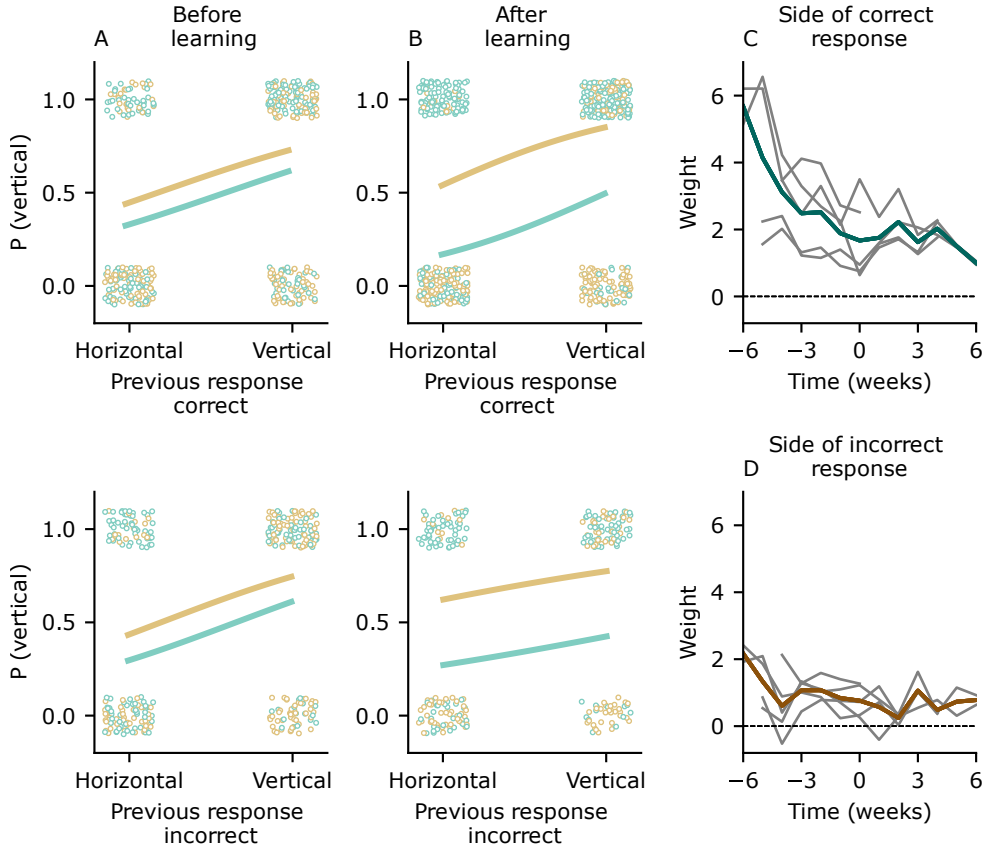


Figure 4.4: Impact of previous response and outcome on choice during and after learning.
(**A**, **B**) The probability to choose vertical on the current trial, separated based on previous correct (top) and incorrect (bottom) choices during (A) and after (B) learning for an example mouse. I and J are computed by pooling across all sessions in week 2 and week 6 for the example mouse. *Turquoise*: trials with horizontal stimulus in current trial; *Orange*: trials with vertical stimulus in current trial (**C**, **D**) Development of model weight across learning for 'Side of correct response' (C) and 'Side of incorrect response' (D). In (C, D), week 0 corresponds to the week in which individual animals reached stable performance; trials with short RTs (<0.5 s) were excluded from this analysis

Having found influences of both the previous response and the previous stimulus on the current decision, I next sought to differentiate their impact depending on the past trial's outcome. As discussed in detail in **subsection 3.4.2**, I modified the model and exchanged the predictors for the previous stimulus and the previous response by two alternative predictors reflecting their interaction: the side of correct and incorrect choice. The positive weight for the interaction terms (side of correct response and side of incorrect response) reflect a 'stay' (positive weight) or 'shift' (negative weight) strategy towards a previously chosen side, depending on the previous outcome.

To illustrate the impact of the previously chosen side on the previous outcome as closely as possible, **Figure 4.4A** and **Figure 4.4B** show data from an example mouse before (week 2; **Figure 4.4A**) and after (week 6) **Figure 4.4B** learning. In the early learning stage, irrespective of the stimulus, the overall tendency to repeat a previously chosen side was high following both a correct (0.65) and incorrect (0.61) response. However, the probability of correctly repeating a choice following both rewarded and unrewarded trial was closer to chance level (0.55 and 0.46 respectively). Interestingly, when the naive mouse made a switch, the probability for it to be incorrect was overall low (following rewarded: 0.28; unrewarded: 0.34). For the trained mouse (**Figure 4.4B**), the probability of correct stay increased compared to that of naive state (following rewarded: 0.65; unrewarded: 0.62). Notably, the probability of incorrect repeat was still high. However, the animal rarely made an incorrect switch (following rewarded: 0.17; unrewarded: 0.20).

Examining the progression of weights assigned to the side of correct response (**Figure 4.4C**) revealed that early in training, the tendency for mice to repeat their choice after a correct trial was high. In the course of learning, the impact of this 'win-stay strategy' declined, but was still present even during stable performance. Interestingly, as indicated by weights for this interaction term (**Figure 4.4D**) mice showed a general tendency to 'stay', but this tendency was higher for trials following a correct response. While this tendency to 'stay' with the previous response was not as strong after an incorrect trial as after a correct trial, it was present during learning and persisted even after learning. Taken together, progress in learning was paralleled by increases in weight assigned to the current stimulus, but behavioral choices were

also influenced by previous stimulus and response, with an overall tendency to repeat the previous response.

### 4.1.5 *History variables have stronger influence on choice when RT on the current trial was short*

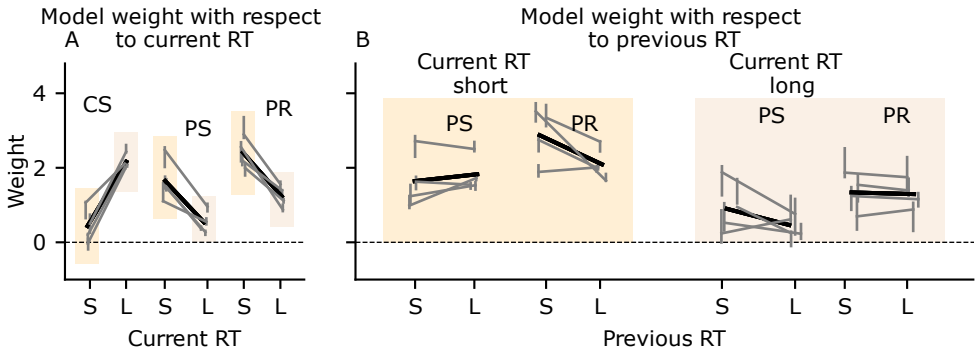


Figure 4.5: Impact of RT on the weights assigned to Current stimulus (CS) and previous stimulus (PS) and previous response (PR) during stable performance.
(**A**) Model weight for Current stimulus (CS), Previous stimulus (PS) and Previous response (PR) with respect to RT on the current trial. S: Short RT on current trial; L: Long RT on current trial. (**B**) Model weight for Previous stimulus (PS) and Previous response (PR) with respect to RT on previous trial. *Left*: RT on the current trial was short. *Right*: RT on the current trial was long. S: Short RT on previous trial; L: Long RT on previous trial. The shaded area represents the RT on the current trial. *Colours* indicate RT on the current trial.

Having studied how current and past information contribute to choices during learning, I next focused on the period of stable performance after learning, in which the larger number of trials allowed to test additional factors. Mice completed an average of 200 trials per session (range 168–230 trials), gathering an average of 5495 trials in total per animal (range 3629–8267 trials) during the stable performance.

Since behavioral results showed that short RTs were associated with poor performance, I first investigated, within the logistic modeling framework, the predictive power of RT on the differential weighting of task-relevant and irrelevant factors. I divided the predictors for the current and previous stimulus and previous responses into 2 bins each (S and L), with S referring to trials with short RT ($< 0.5$ s), and L to trials with long RT ($>= 0.5$ s) on the current trial (refer to section 3.4.2 for detailed description).

Consistent with the poor performance associated with short RTs, for trials with short RT, the weight assigned to the CS was negligible ($0.44 \pm 0.41, p = 0.16$ (**Figure 4.5A, (CS, S)**), suggesting that on these trials, animals failed to use relevant sensory evidence to guide decisions. In contrast, the weight assigned to the CS was significantly higher ($2.16 \pm 0.15, p < 0.001$) for trials with long RTs (**Figure 4.5C, (CS, L)**). Also, consistent with the earlier week-by-week analysis, substantial weights ($> 0$) were assigned to PS and PR irrespective of RT, indicating that PS and PR influence decisions across all trials.

The history effects were stronger for trials with short RTs (PS: $1.65 \pm 0.51$, PR: $2.41 \pm 0.34$), compared to those with long RTs (PS: $0.54 \pm 0.26$, PR: $1.24 \pm 0.26$) (**Figure 4.5B, (PS, PR)**, $p < 0.024$). This suggests that, in trials with short RT ($p < 0.001$), the choices are not random, but predominantly influenced by stimulus and response in the previous trial.

Next, I focused on the impact of the previous trial's RT on the weight assigned to task irrelevant variables. Interestingly, whether RT on the previous trial was short or long did not impact the weight assigned to these history-related variables (*Current RT short*: PS : $p = 0.69$, PR : $p = 0.13$, *Current RT long:* PS : $p = 0.32$, PR : $p = 0.89$) (**Figure 4.5B**). Notably, substantial weight was assigned to PR irrespective of the RT ($p < 0.012$). Regarding PS, the weight assigned was significant except on trials with long RT following a trial with long RT (PS: $0.89 \pm 0.62, p = 0.08$). Although the average weight is not significant at a 5% significance level, PS exhibited a positive weight for all 4 animals. To further clarify the significance of the observed effect, additional behavioral results from other animals need to be collected.

### 4.1.6 *The 'stay' tendency is higher on trials with short RT*

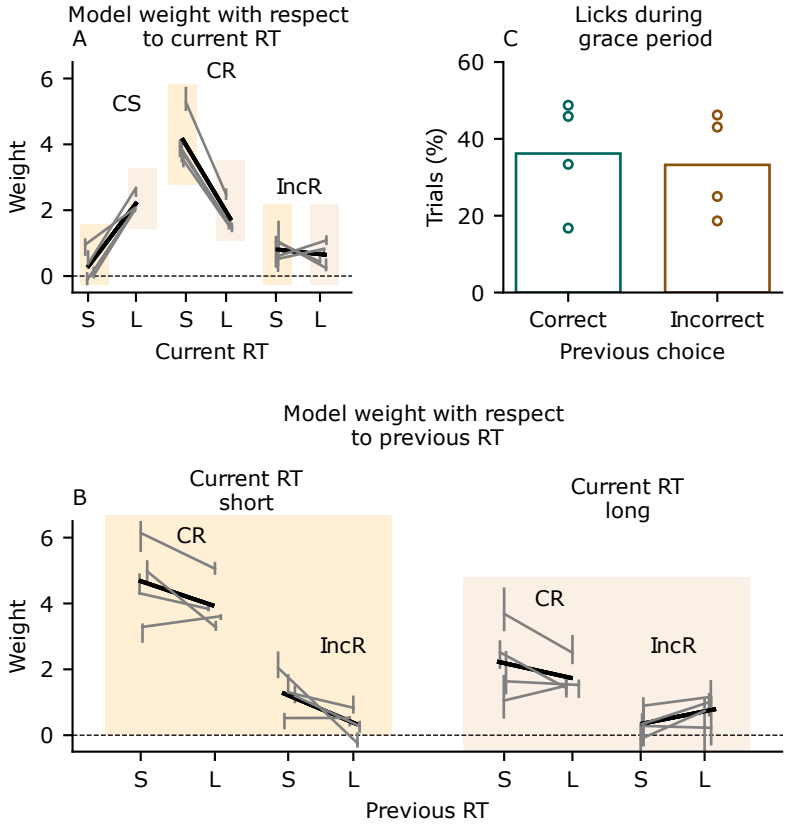


Figure 4.6: Impact of RT on the weights assigned to Current stimulus (CS) side of correct response (CR) and side of incorrect response (IncR) during stable performance.
(**A**) Model weight assigned to side of correct response (CR) and side of incorrect response (IncR) with respect to RT on the current trial. S: Short RT on current trial; L: Long RT on current trial. (**B**) Model weight for side of correct response (CR) and side of incorrect response (IncR) with respect to RT on previous trial. *Right*: RT on the current trial was short. *Left*: RT on the current trial was long. (**C**) Percentage of trials with short RT and licks during grace period separated according to previous outcome. *Colors in A-B* indicate stimulus on the current trial.

Similar to previous analyses during learning, I here re-expressed the regression model to include predictors capturing interactions between response and outcome (**Figure 4.6A**). In particular, for short-RT trials, a large positive weight ($p = 0.0027$) was assigned to the CR ($4.06 \pm 0.76$), which was even stronger than the weight for the CS ($2.16 \pm 0.15$) under the best possible behavioral conditions, i.e., in long RT trials ($p = 0.0056$). Substantial weights were also assigned to the CR on trials for long RT ($1.77 \pm 0.43, p = 0.0059$), however, the weight was not higher than that assigned to the CS. In contrast, the weight assigned to the IncR was overall small, irrespective of RT on the current trial (**Figure 4.6A, IncR**) (Long: $0.69 \pm 0.28, p = 0.024$, Short: $0.76 \pm 0.41, p = 0.047$).

The RT on the previous trial did not impact the weight assigned to the CR (*Current RT short*: $CR : p = 0.34$, *Current RT long:* $CR : p = 0.48$) (**Figure 4.6B**). With respect to InCR, the animals repeated previous choices more often when the RT on both current and previous trial was short (*Current RT short*: $IncR : p = 0.04$, *Current RT long:* $IncR : p = 0.23$) (**Figure 4.6B**). Importantly, it is unlikely that short RT trials were driven by lingering licks related to reward pickup in the previous trial, because the proportion of trials with licks during the grace period following rewarded trials ($36.18 \pm 12.61$) were not consistently higher than those following unrewarded trials ($33.22 \pm 11.67$; $p = 0.78$; **Figure 4.6C**).

### 4.1.7 *History influences are not limited to just a single trial into the past*

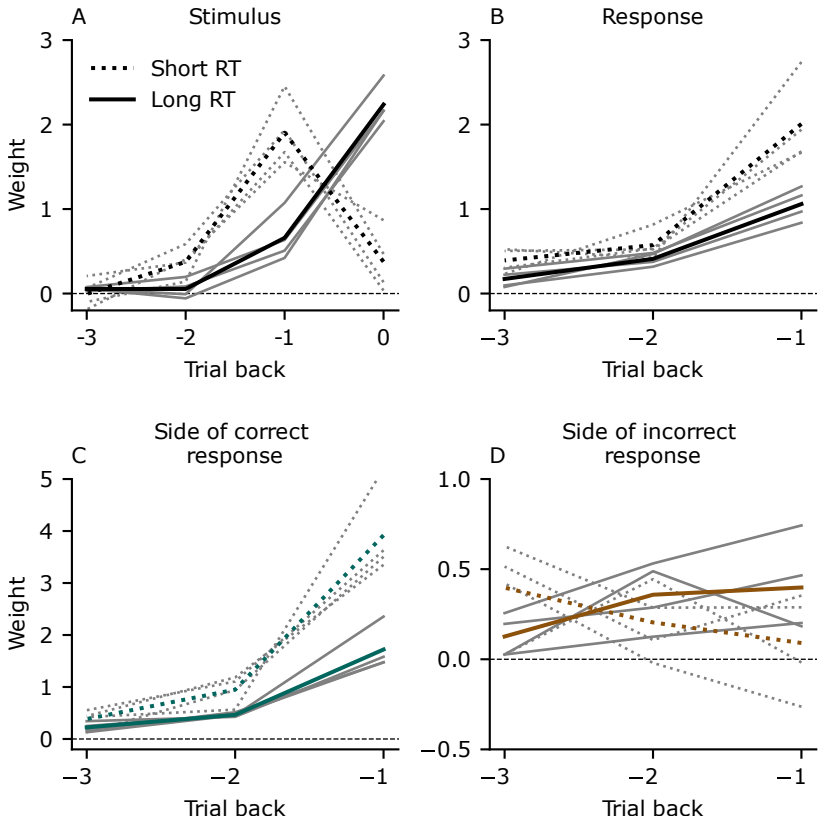


Figure 4.7: Model weight assigned to Stimulus, response, and outcome up to 4 trials back.
(**A**-**D**) Model weight assigned to **(A)** stimuli, **(B)**previous responses, **(C)** Side of correct responses and **(D)** Side of incorrect responses seen up to three trials back, separated according to RT on the current trial. *solid lines*: $\geqslant$ 0.5 s , *dotted lines*: $\leqslant$ 0.5 s

Since previous studies suggested that the history influence might not be limited to just a single trial into the past (Hermoso-Mendizabal et al., 2020; Akrami et al., 2018), I also investigated the degree to which the mice weighted stimuli choices and outcome that had occurred up to three trials back. I thus, built a LR model, containing weights for the current stimulus, and the past $n$ stimuli, choices and outcomes, where $n$ varied between -1 and -3. Given the pronounced impact of

RT that was observed in the earlier analyses, I also differentiated whether the current trial's RT was long or short. I first focused on the weights assigned to current and past stimuli (**Figure 4.7A**). I found that, for trials with long RTs, the influence of stimulus was high for the current trial, and dropped rapidly to 0 over the past 2 trials ($0.05 \pm 0.1$, $p = 0.37$). In contrast, for trials with short RT (*dotted lines*), the weight assigned to the stimulus was much lower for the current trial, was highest for the immediately preceding trial, and still had a non-zero influence at $n = -2$ trials back ($0.37 \pm 0.16$, $p = 0.006$). Turning to the impact of past choices (**Figure 4.7B**), a similar result emerged, where choices in the immediately preceding trial ($n = -1$) were assigned the highest weight, in particular for trials with short RTs (*dotted lines*). Next, I focused on the weight assigned to the side of correct (**Figure 4.7C**) and incorrect (**Figure 4.7D**) responses. The side of correct responses followed a similar trajectory as the weight assigned to past choices: highest for the immediately preceding trial, non-zero influence at $n = -2$ trials back (*Short RT:* $0.38 \pm 0.16$, $p = 0.006$, *Long RT:* $0.22 \pm 0.08$, $p = 0.003$ ). However, the weight assigned to the side of incorrect choices varied between animals.

### 4.1.8 *Implementation of trial-to-trial variation of model weight during learning and stable performance*

Recent studies have suggested that decision-making strategies might change on a trial-trial basis and may evolve rapidly during training, especially when animals are learning the task (Roy et al., 2021). To explore this impact on behavior, I used Psytrack package (Roy et al., 2021). Psytrack describes the decision-making behavior at the resolution of single trials, allowing us to estimate weight trajectories at shorter timescales.

Similar to the LR model described in the prior section, PsyTrack uses Bernoulli GLM to model animals' choices. However, unlike the model used earlier, in this variant weight is assumed to evolve gradually from trial to trial over time. Similar to **subsection 3.4.2**, using Psytrack I predicted the probability to choose the lick port associated with the

vertical stimulus by a weighted sum of influences related to the current visual stimulus and the previous stimulus and response with respect to RT.

### 4.1.9 *The animals' decision-making strategies across trials and across weeks follow similar trajectories*

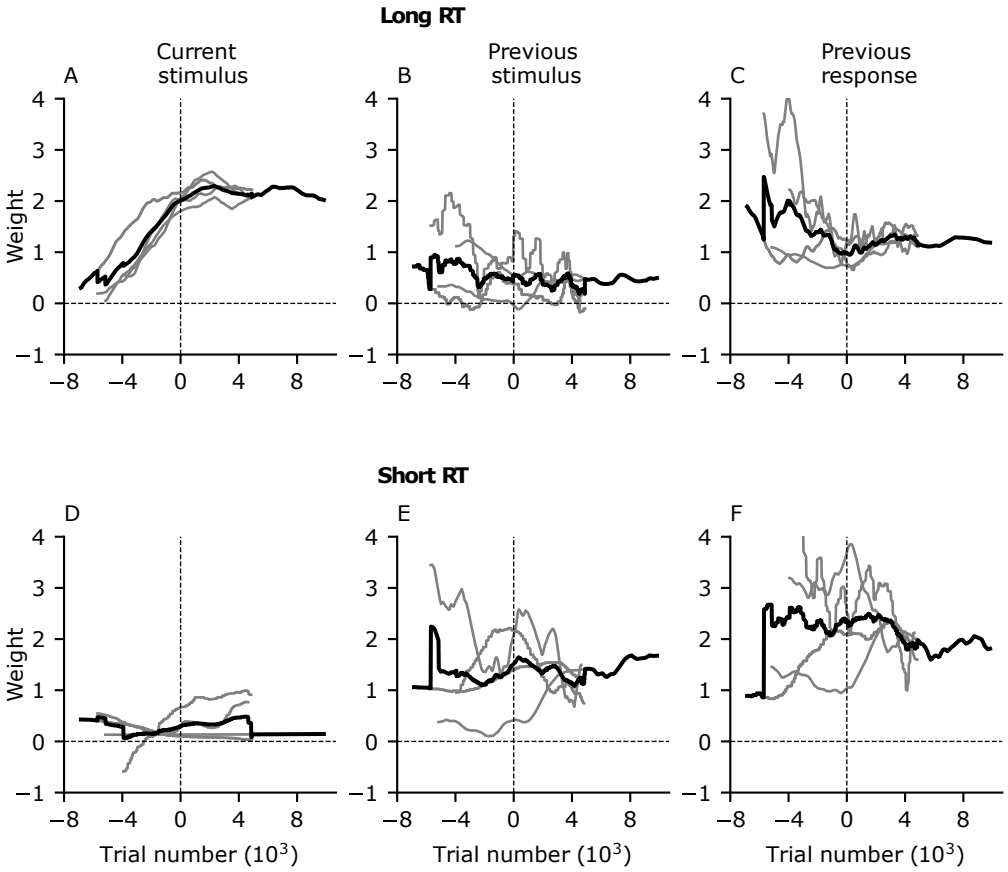


Figure 4.8: Trial-to-trial model weight during learning and stable performance obtained from Psytrack.
(**A**-**C**) Weight assigned to the current **(A)** and the previous stimulus **(B)** and the previous response **(C)** with respect to long **(top)** and short **(bottom)** RTs on current trial.

I next examined weight trajectories obtained from the PsyTrack model assigned to task-relevant and task-irrelevant factors. Consistent with the observation from the standard LR, weight assigned to the CS, particularly for trials with long RTs increased during learning and plateaued when the animals reached stable performance. For trials with short RT, as suggested by chance level behavioral performance and insignificant model weight during stable performance, the weight developments derived from the PsyTrack model were also insignificant. In fact, there were little to no deviations (from 0) in the weight trajectories through learning. Next, I focused on the weights assigned to the task irrelevant variables. Consistent with the earlier claims, the weight trajectories for PS and PR for trials with long RT decreased over learning. In particular, these decrements were only observed when task-irrelevant influences had a stronger impact on choice during the early training sessions. On the contrary, for the trials with short RT, the weight assigned to both PS and PR fluctuates, but is substantial. Notably, supporting our earlier observations, that the weight assigned to task-irrelevant factors remains higher when the RT on the current trial is short (Mean across trials; PS: $\Delta = 0.43 \pm 0.080$, PR: $\Delta = 0.36 \pm 0.14$. The result from the Psytrack model further confirms the claim that for trials with short RTs, animals fail to utilize task-relevant factors to guide choices and decisions are strongly influenced by previous experiences.

Taken together, similar to the standard LR model, the PsyTrack model can capture the development of the model weight during and after learning. An advantage of the Psytrack model is its ability to capture the decision-making trajectories on a trial-by-trial basis. As the weight trajectories obtained from the two models complement each other, I rely on the standard LR model for the analysis presented in this thesis.

### 4.1.10 *Animals' behavioral performance suggests a bias towards larger rewards*

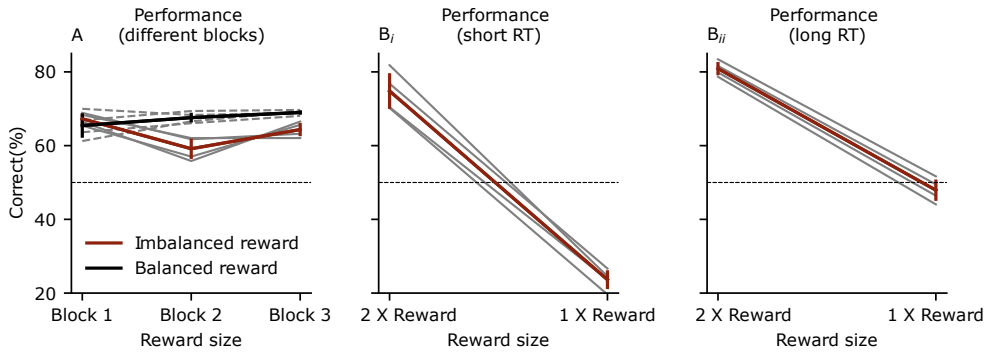


Figure 4.9: Impact of unequal reward on the performance of trained mice. (**A**) Overall performance for trials in the orientation discrimination task, computed by separating Balanced (Block 1) - Imbalanced (Block 2) - Balanced (Block 3) for manipulated reward condition (Red). *Black:* Overall performance on baseline condition, computed by trials separated in blocks of 100-50-100 trials, *grey:* Individual mice **($B_i$)** Proportion of correct trials in Imbalanced blocks for short RT, separated according to the reward size. **($B_{ii}$)** Proportion of correct trials in Imbalanced blocks for Long RT, separated according to the reward size.

Many studies have suggested that expectation of reward size biases decisions (Busse et al., 2011, Rorie et al., 2010, Lak et al., 2020b). To understand how mice trade between stimulus-related and contextual information, I introduced blocks of trials with stimulus-specific imbalances of reward (see **Figure 3.2.3**). To summarize, in each session, mice experienced a middle block of 50 trials, in which they received a double reward for correctly choosing the lick spout associated with one of the two stimuli (e.g., horizontal). Correctly choosing the spout associated with the other stimuli was normally rewarded, and incorrect choices for both stimuli were equally penalized.

To get the first insight into the animals' behavior during imbalanced reward regimes, I analyzed overall performance (**Figure 4.9A–B**). Despite performance above the criterion of 66% during Block 1 (67.21% ±

1.36%), performance of the mice declined during Block 2 (59.2% ± 2.83%) (**Figure 4.9A**; *red*). This decline was not mediated by general factors, such as time in the task or reduction of motivation due to satiation, because mice had the capacity to improve again in the subsequent block (Block3) with a balanced reward (64.57% ± 1.43%) (**Figure 4.9B**; *red*). In addition, performance decline in Block 2 was not evident in the sessions containing only the equal reward condition (**Figure 4.9A**; *black*). For these sessions, I split the trial in similar blocks and compared performance (block 1: 65.44% ± 3.30%, block 2: 67.59% ± 1.33%, block 3: 68.98% ± 0.60%). Splitting trials according to RT also in this version of the task (**Figure 4.9B**), revealed that in particular for trials with short RT(**Figure 4.9B$_i$**), mice rarely chose the stimulus associated with the single reward, leading to poor performance for this stimulus (22.64% ± 2.56% **Figure 4.9B**). Interestingly, mice performed at chance level for single reward stimuli for trials with long RTs (47.9% ± 2.90%, **Figure 4.9B**). This asymmetric performance for the two stimuli suggests that reward size seems to induce a bias for the double-rewarded stimulus during visually driven choices.

### 4.1.11 *Reward size biases the visually driven choice*

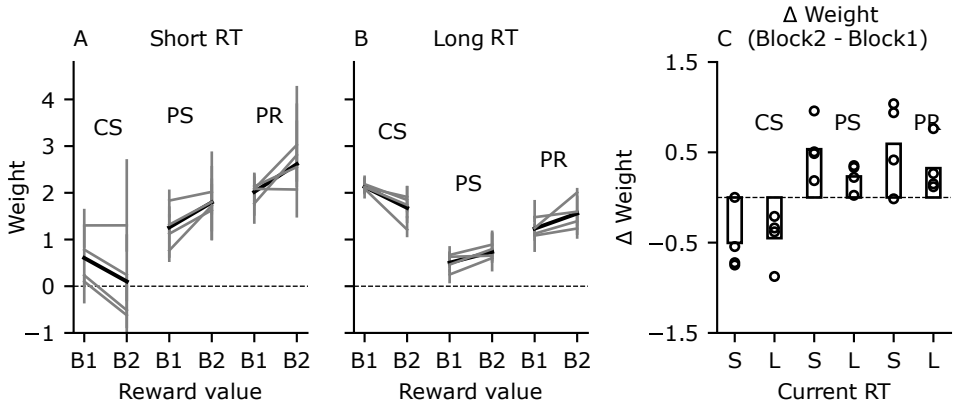


Figure 4.10: Impact of unequal reward on the weights assigned to Current stimulus (CS) and previous stimulus (PS) and previous response (PR) during stable performance.
(**A**, **B**) Model weight for Current stimulus (CS), Previous stimulus (PS) and Previous response (PR) for baseline and unequal rewarded condition with respect to short RT on current trial (**A**), trials with Long RT on current trial (**B**). B1: Equal reward block; B2: Unequal reward block. (**C**) Difference in model weight in block 2 (B2) vs block 1 (B1) separated according to RT on the current trial

To understand the dynamic adaptations of internal weights across blocks with changes in reward size, next, I applied the logistic modelling framework separately to behavior in the first (block 1; equal rewarded) and middle (block 2; unequal rewarded) blocks. Since inspection of performance suggested an improvement for the double-rewarded stimulus in the face of a decrement for the single-rewarded stimulus (**Figure 4.9B**), I expressed choices with respect to the rewarded stimulus and using LR predicted the probability that the animal would choose the side associated with the larger reward. For block 1 choices were still expressed with respect to the double rewarded side in block 2.

I first focused on the weight assigned to CS, PS, PR with respect to short (**Figure 4.10A-C**) and long RT (**Figure 4.10B, C**). The weight

assigned to the CS decreased (Long RT: $\Delta = -0.45 \pm 0.25$, $p = 0.031$; Short RT: $\Delta = -0.50 \pm 0.30$, $p = 0.37$), whereas those assigned to the PS (Long: $\Delta = 0.23 \pm 0.13$, $p = 0.091$; Short: $\Delta = 0.53 \pm 0.27$, $p = 0.065$) and PR (Long: $\Delta = 0.32 \pm 0.25$, $p = 0.13$; Short: $\Delta = 0.59 \pm 0.42$, $p = 0.036$) increased. Notably, the weight difference for CS and PS on trials with short RT and PS and PR on trials with long RT were not statistically significant. Hence more number of sessions and animals might be necessary to reach to the significant level.

### 4.1.12 *Animals' bias towards the larger reward size optimizes reward accumulation*

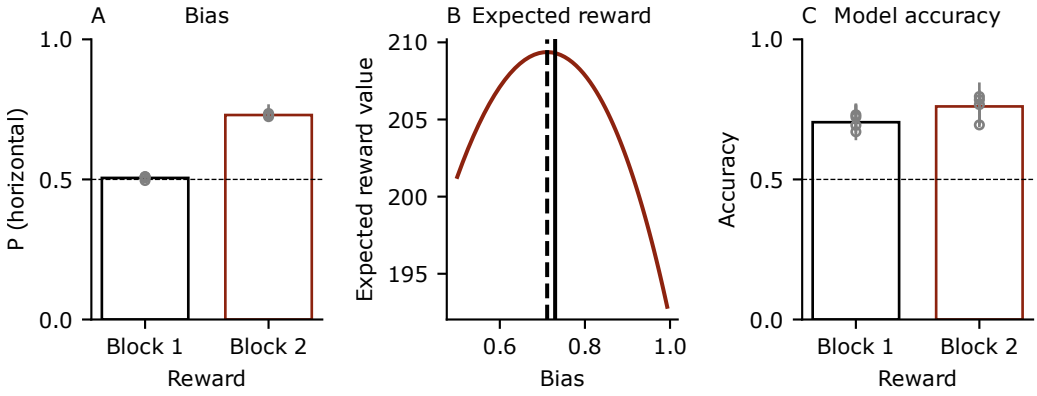


Figure 4.11: Comparison of Model bias and accuracy for equal and unequal rewarded blocks.
(**A**) The overall model bias, towards the side associated with double reward for imbalanced rewarded condition and towards the side associated with horizontal grating for equal rewarded condition. 0.5 indicates no bias. (**B**) Expected reward value for block 2 obtained from logistic regression simulation. *Dotted line*: the optimal bias to maximize reward, *solid line*: the average bias of the animals in block 2. (**C**)The overall model accuracy for balanced and imbalanced reward block

Next, I focused on model bias. The model bias here represents the overall probability to choose the side associated with a larger reward (**Figure 4.11B**). Consistent with the behavioral performance dur-

ing imbalanced reward condition, all mice had a bias $> 0.5$ (Block2 $0.73 \pm 0.004$; Block1 $0.50 \pm 0.005$), driving strongly the choice to double rewarded side. Although the variation in changes in weight assigned to CS, PS, and PR (**Figure 4.10C**) captured the impact of the reward size on animals' decisions, the strong bias towards the double rewarded side reiterates the extent to which the expectation of larger reward biases choice. Next, to evaluate if the animals' bias is optimal to maximize reward, I modelled the expected reward for different bias values (refer to **subsection 3.4.5** for detail description of the method). Given the performance in the equal reward condition, this analysis revealed that the ideal bias to maximize reward is 0.71. Despite the decline in overall performance, this shows that the animal adjusted its bias in an optimal manner to maximize its reward (**Figure 4.10B**).

Finally, the overall accuracy of the model was consistent during the balanced ($70.42\% \pm 2.41\%$) and the imbalanced block ($76.05\% \pm 3.94\%$) suggesting that the major contributions to the animals' choices can be described by the same variables in both balanced and imbalanced reward conditions (**Figure 4.10C**). In fact, these variables were necessary for the estimation of the choice. The reduced model with only CS with respect to RT, and a bias term as predictors had a worse model fit $\Delta = 9.0\% \pm 1.86\%$).

### 4.1.13 *The influence of side of correct response is higher during imbalanced reward conditions*

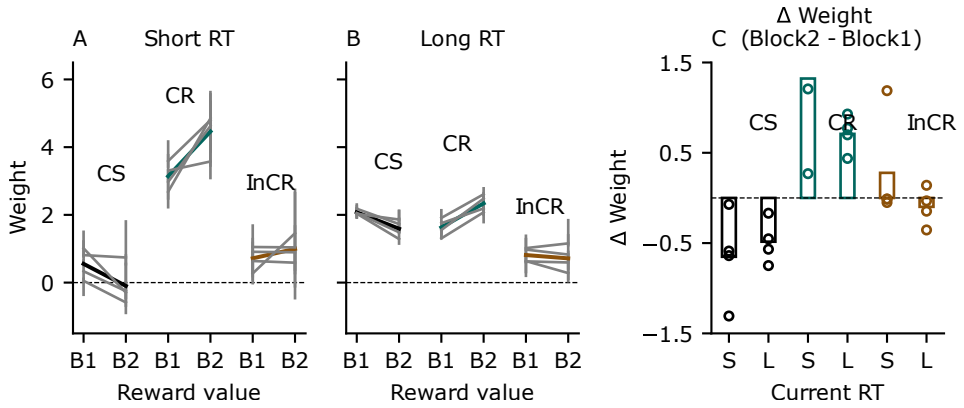


Figure 4.12: Impact of unequal reward on the weights assigned to current stimulus (CS) side of correct choice (CR) and side of incorrect choice (IncR) during stable performance.
Model weight assigned to current stimulus (CS) side of correct response (CR) and side of incorrect response (InCR) with respect to short RT (**A**) and long TR (**B**) a on the current trial. (**C**) Difference in model weight in balanced and Imbalanced reward block.

Next, similar to previous analyses during stable performance, I sought to determine the influence of previous response and outcome on choice for imbalanced reward condition. Consistent with previous findings, the reward size increased the dependency on history-related variables. In particular, the weight assigned to the side of the correct choice for both short (**Figure 4.12A, C**) ($\Delta = 1.12 \pm 0.55$, $p = 0.025$) and long RT (**Figure 4.12B, C**) ($\Delta = 0.45 \pm 0.25$, $p = 0.023$) is even higher for the imbalanced reward condition. In contrast, the differences in change in weight assigned to the side of incorrect choice between balanced and imbalanced reward were for most mice insignificant (Short: ($\Delta = -0.06 \pm 0.45$, $p = 0.82$); Long ($\Delta - 0.091 \pm 0.21$, $p = 0.73$).

Taken together, the results on imbalanced reward condition suggest that in conditions where external factors such as unequal reward biases choice, choices were strongly influenced by history related variables.

The majority of the sensory decision-making literature is concerned with characterizing neuronal responses and decision-making behavior of fully trained animals. Over the last decades, this approach has given us insight both in decision-making strategies and in identifying neuronal population involved in decision-making process (**section 1.2**). Recently, Roy et al. (2021) studies have focused on understanding the impact of dynamic decision-making behavior, especially during the learning period. With this aim in mind, the final part of the thesis was dedicated to adapting a recording set-up and establishing in the lab a method for long-term neuronal recording.

Traditionally, in chronic neuronal recordings, the probe is attached to a microdrive, allowing to advance to the interest region, with small increment. This mobile configuration, allows one to move the probe further if the recording quality detoriates. However, the implantation process is tedious, and the configuration is heavier. Here, I adapted a method based on Okun et al. (2016) where silicone probes were rigidly affixed to the skull. In addition to involving a lighter implant and a less tedious implantation process, Okun et al. (2016) suggests, this method may allow for more stable and longer term recordings. The authors recorded neuronal activity for over 3 months and reported stable waveforms in at least 70 % of neurons recorded in consecutive days (Okun et al., 2016). Refer to **subsection 3.3.2** for a detailed description of the method.

In the following section, I first show that neuronal recording can be obtained over several weeks and assess the stability of the recordings across days. Next, I evaluate the presence of visually responsive neurons in the recorded neuronal population. Finally, I discuss the preliminary results obtained from neuronal recordings during task performance.

### 4.2.1 *Recording with the chronically implanted probe*

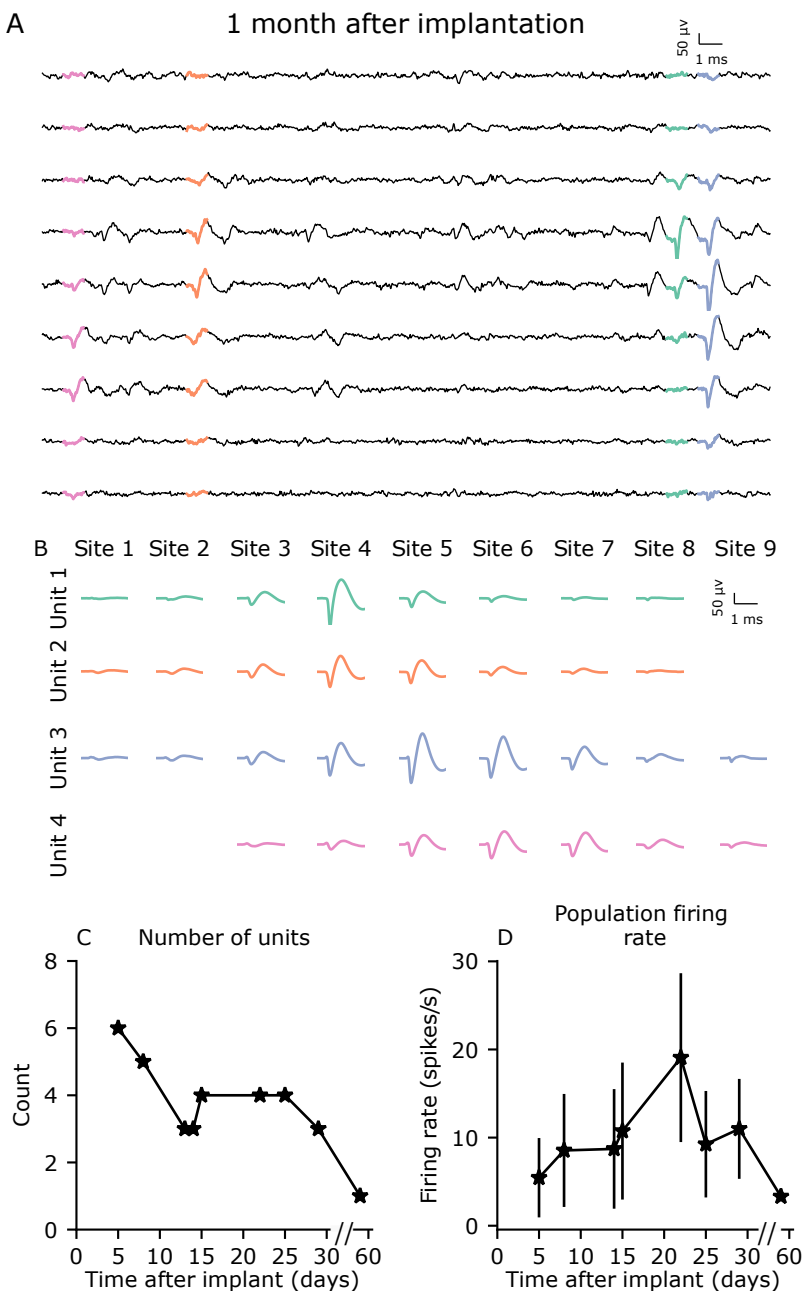


Figure 4.13: Recording with chronically implanted, immobile silicon probes.
*(Continued on next page)*

Figure 4.13: *(Continued)* (**A**) Raw traces after high pass filtering across five sites, recorded 20 days after implantation of the silicon probe. *Colour*: waveforms of four single units. (**B**) Average spike waveforms of the four single units shown in (A) at nine different recording sites, also shown in (A). (**C**) Number of well-isolated units recorded across days. (**D**) Mean population firing rate of all isolated units across recording days during periods of sitting. Error bars indicate the standard deviation across units.

To investigate whether long-term neuronal recordings could be obtained from an immobile silicone probe, I examined the recordings acquired on the fourth week following implantation (**Figure 4.13**). Inspection of the voltage traces across nine recording sites after high-pass filtering revealed the presence of neuronal spikes (**Figure 4.13A**). This is further supported by the average waveform plot of these units across the same nine recording sites (**Figure 4.13A, B**). Next, I compared the number of single units obtained post spike sorting across days (**Figure 4.13C**). Six single units were obtained on the first recording day (after a week of implantation), however this number declined over the weeks. After approximately two months, only one single unit was isolated. Although I recorded activity from several other units, the auto-correlogram did not exhibit a refractory period gap. This observation suggests the presence of multi-unit activity and a relatively low signal-to-noise ratio. Finally, I assessed the stability of the firing rate (**Figure 4.13D**) by comparing the population firing rate for spontaneous condition when the animal was sitting. This analysis revealed that the population firing rate varied between days ($\mathrm{std} = 4.35$). In conclusion, these data demonstrate that single units can be acquired after several weeks of implantation of the immobile silicone probe, however the number of units acquired decreased over time.

### 4.2.2 *Stability of neural recordings across days*

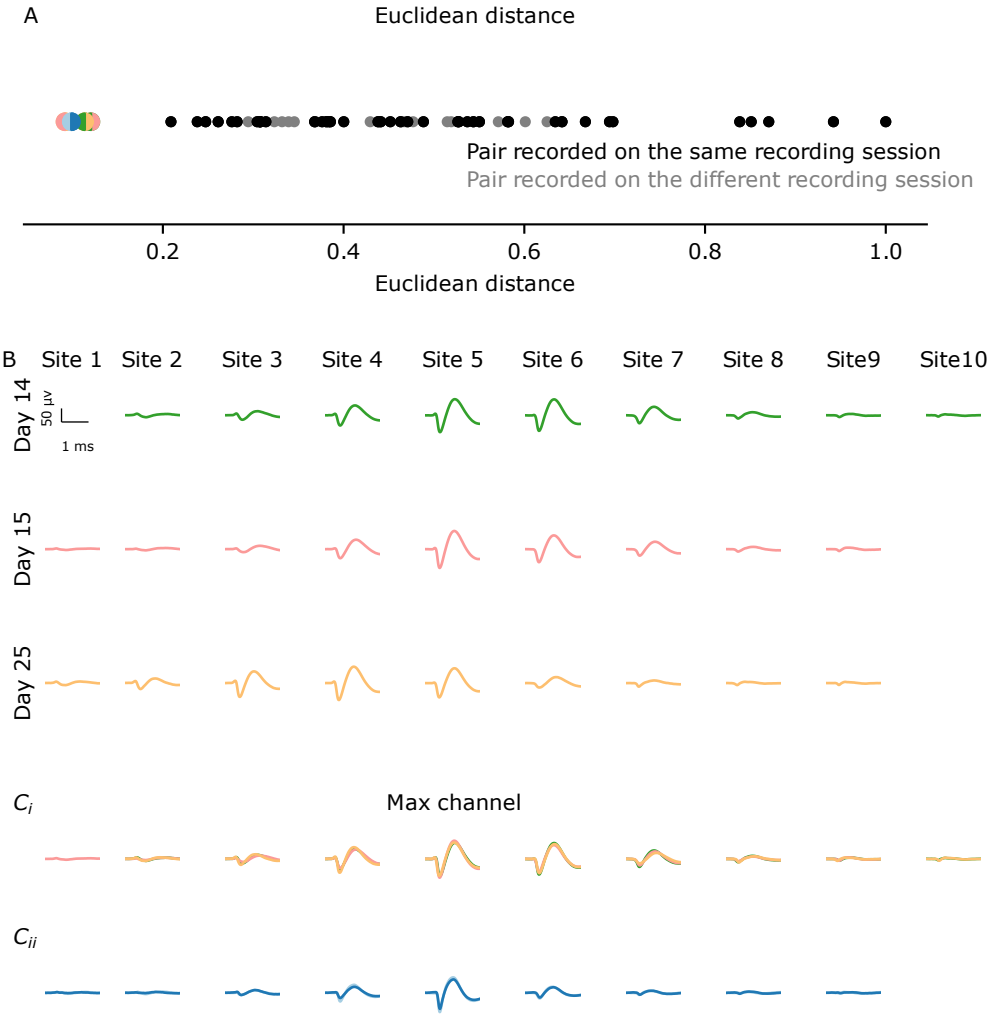


Figure 4.14: Example of a stable unit recorded across days.
(**A**) Euclidean distance measured between each pair of waves. The distance was measured between voltage at each time point. *Colored circles*: color reflects the day units were recorded on (*light blue*: day 8, *dark blue*: day 10, *green*: day 14, *pink*: day 15, *orange*: day 25); same units as shown in $B_i$ and $B_{ii}$; *black*: non-matched units recorded on same day; *grey*: other, non-matched units recorded on different days. (**B**) Example wave form across 10 recording sites of three units across three different recording days. *Color*: unit recorded on a specific day. **(**$C_i$**)** Overlaid waveforms from B, aligned with respect to the maximum amplitude channel. **(**$C_{ii}$**)** Overlaid waveforms for another example neuron, with respect to the maximum amplitude channel.

To assess the stability of neural recording, I compared the extracellular wave shapes of the recorded units across several days. As neural responses acquired on different sessions/days were sorted separately, I first identified units which had similar wave shape but were recorded on different days. To quantify the similarities in units, I calculated the Euclidean distance between the voltage at each time point between pairs of waves (refer to **Equation 3.4.7** for detail description of the method). The single units, identified on a particular day, were distinct from each other as they were obtained after exhaustive pair-wise comparison (Euclidean distance:$0.49 \pm 0.20$). I used the minimum Euclidean distance between them as a threshold for unmatched units (0.21). Next, to illustrate the similarity in the wave shape, **Figure 4.14B** shows examples of three units acquired on three different days across 10 recording sites but having similar wave shape properties (Euclidean distance: $0.11 \pm 0.01$). Here, the presence of drift is evident by the displacement of the wave towards the top of the probe (towards the left). To better match the units across days, I shifted the waveform with respect to the maximum channel and overlaid the waveform on different channels over days. **Figure 4.14C** shows two such example neurons $C_i$ (the same neuron as shown in B) and $C_{ii}$ (Euclidean distance: 0.099). Taken together, the result here suggests that the method discussed in **subsection 3.3.2** may allow recording of the same unit across several recording days.

### 4.2.3 *Presence of visually responsive neuron in the recorded neuronal population*

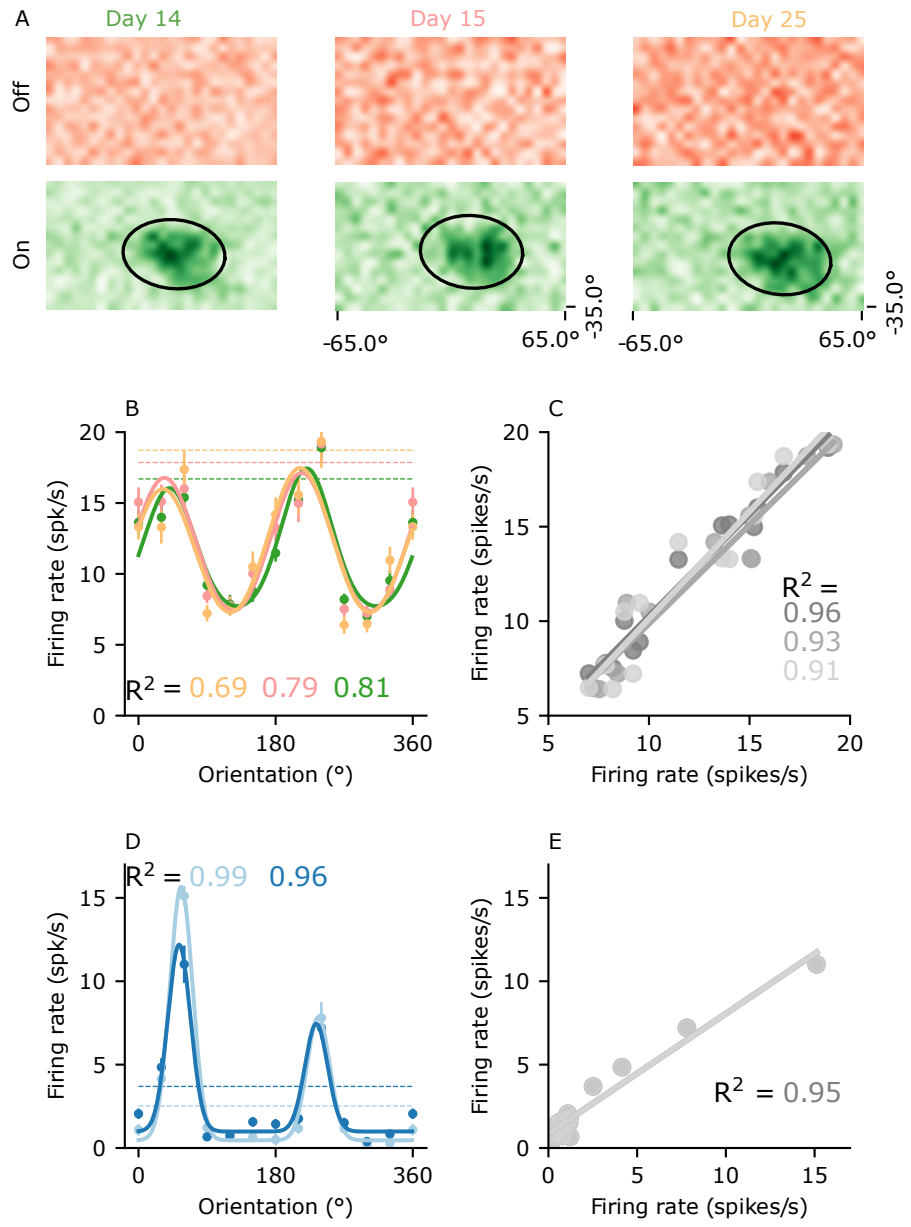


Figure 4.15: Example of visually responsive units.
(**A**) Spatial RF map obtained for three particular units across three recording days. *Red:* on response. *Green:* off response. (**B**, **D**) Tuning curve generated from drifting grating (12 directions) for two (B, D) example units across different days. *Colours:* units recorded on different days. *Markers:* average firing rate for a specific orientation on a specific day. *Dotted lines:* spontaneous firing rate. (**C**, **E**) Comparison of firing rate for different orientations for units shown in B and D respectively. The X-axis and Y-axis represent firing rate across different pairs of units. In C, grey level indicates $R^2$ across different pairs.

To understand the stability of visually responsive neurons, I compared the spatial RF maps obtained in sparse noise experiments of three example units putatively representing the same neuron across days. On all three days, the units responded only to the 'on' stimulus (**Figure 4.14A**). The 'off' stimulus did not elicit any responses. The fitted 2D Gaussian (Mean fit across days: $R^2 = 0.65 \pm 0.037$) was also similar across all three days for the units. The area of the fit was consistent over days ($1949.45 \pm 40.13\ deg^2$) and the majority of the field area overlapped ($82\% \pm 0.5\%$ ).

Next, I focused on understanding the similarities between orientation tuning curves generated from the drifting gratings (**Figure 4.15B-E**). Consider two example units recorded across 3 weeks (**Figure 4.15B, C**) and 3 days apart (**Figure 4.15D, E**), respectively. The orientation tuning were stable over time. For the first putative example neuron (**Figure 4.15E**), the responses to the gratings were consistently suppressed below the spontaneous firing rate and the preferred orientations were 30° and 210°. In contrast, for the second putative neuron (**Figure 4.15D**), the preferred orientation across consecutive days peaked at 60°. To further assess the stability of these neurons, I computed the correlation between the firing rate in response to various orientations across days and found them to be strongly correlated (**Figure 4.15C, E**).

Taken together, both the neuronal wave shape **Figure 4.14** and the visual response properties **Figure 4.15** suggest that these units recorded on different days might actually capture responses from the same neuron.

### 4.2.4 *Preliminary neuronal responses recorded during the behavioral task*

Studies have characterized that prior stimulus history and reward expectation can influence early stages of cortical sensory processing, and thus profoundly change neuronal response properties (Aton, 2013; Schoups et al., 2001; Yang and Maunsell, 2004). However, it is still unclear how a particular sensory event alters the distribution of repres-

entation of information in the visual system. As a next step, I aimed to quantify history effects on neuronal activity in the visual cortex during task learning. Below, I analyzed the firing rate of 3 neurons during task performance based on the outcome. Unfortunately, the animal did not learn the task, and therefore the animals' performance in the sessions discussed here was on average 53% $\pm$ 0.65 (n = 2 sessions). Unit 2 and unit 3 are the same neurons, shown in **Figure 4.15B** (pink and orange respectively).

### 4.2.5 *Firing rate is higher on trials following a correct outcome*

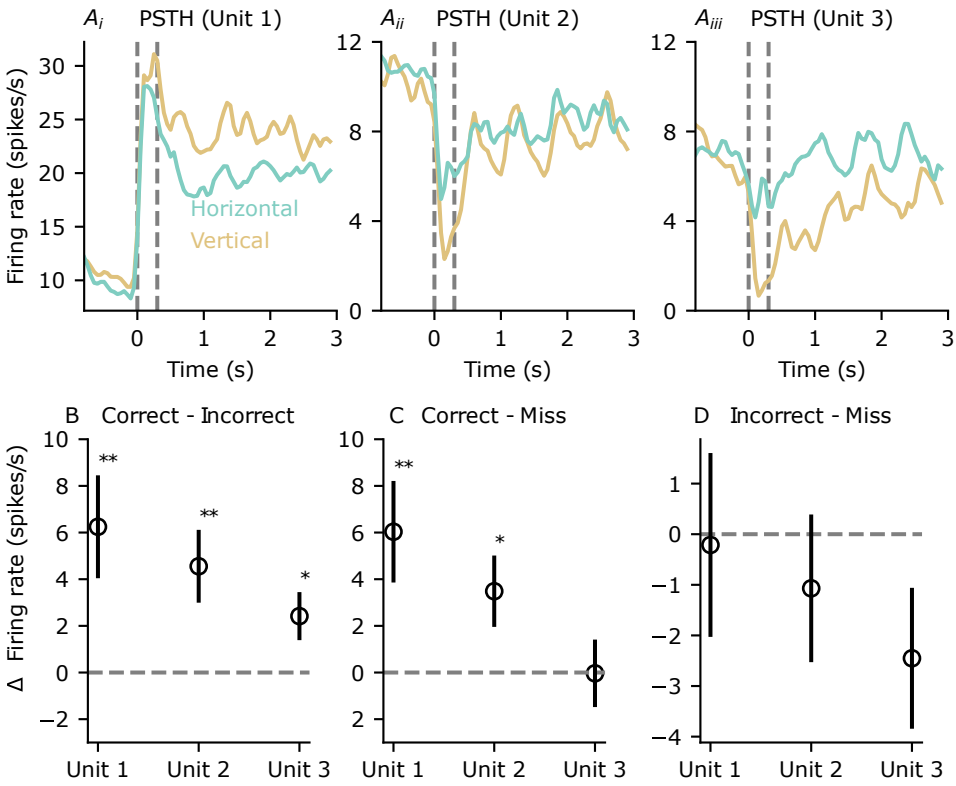


Figure 4.16: Example neuronal responses during the behavioral task. (**A**) PSTHs for 3 example units for the 2 grating stimuli presented during the task. *Dotted line* at 0 s indicates stimulus onset; *dotted line* at 0.3 s indicates end of grace period. *Turquoise*: horizontal orientation; *yellow*: vertical orientation. Unit 2 and unit 3 are the same units shown in **Figure 4.15B**. (**B**-**C**) The difference in mean firing rate during trial initiation following a previous correct vs incorrect (**B**), correct vs missed (**C**), and incorrect vs missed (**D**) trials. In B-D The error bars represent the standard error of the mean difference.

To illustrate neuronal responses during task performance, **Figure 4.16A** shows the PSTHs of three example neurons in response to two orientations used in the task. To understand, the impact of visual stimulation on neuronal firing rates, I compared the firing rate before visual stim-

ulation with the firing rate during the grace period (time between dotted lines). Visual stimulation increased the firing rate for unit 1 ($\Delta$ =270°: $18.80 \pm 2.31$ spikes/s and 180°: $18.5 \pm 2.19$ spikes/s). Unit 2 ($\Delta$ = 270°: $-7.24 \pm 1.30$ spikes/s and 180°: $-4.76 \pm 1.33$ spikes/s) and unit 3 ($\Delta$ = 270°: $-5.68 \pm 0.84$ spikes/s and 180°: $-2.44 \pm 1.14$ spikes/s) were suppressed by the visual stimuli. Observations for unit 2 and unit 3 are supported by the tuning curve illustrated in, **Figure 4.15B** as both of these orientations suppressed the firing rate.

To understand the relationship between firing rate and trial outcome, I first focused on the time period between the end of the grace period and the animal's decision. This analysis revealed that at the 5% significance level, there was no difference in firing rate across units.

Having shown that previous outcomes have an impact on decision-making, I next sought to determine their impact on the firing rate. For this analysis, I focused on the trial initiation phase following the decision. As mentioned in **subsection 3.2.3**, during this phase the behavior was stable as the animal was neither licking nor running. The analysis shows that the firing rate was higher for all units following a correct trial than those following an incorrect trial (unit 1 $6.24 \pm 2.20$ spikes/s, $p = 0.01$; unit 2 $4.55 \pm 1.56$ spikes/s, $p = 0.005$; unit 3 $2.47 \pm 1.02$ spikes/s, $p = 0.02$; **Figure 4.16B**). Similarly, the firing rate was significantly higher for unit 1 and unit 2 following correct vs missed trials (unit 1 $6.04 \pm 2.17$ spikes/s, $p = 0.006$ ; unit 2 $3.48 \pm 1.52$ spikes/s, $p = 0.02$; unit 3 $-0.03 \pm 1.44$ spikes/s, $p = 0.98$; **Figure 4.16C**). Finally, a comparison of firing rates between trials with previous incorrect vs missed outcomes revealed that there were no differences in firing rates (**Figure 4.16D**).

Taken together, the results in **Figure 4.16** suggest that there might be variations in V1 firing rates based on the previous outcome. However, a full data set is required to reach valid conclusions. Notably, the animal's internal state following a reward consumption might have also been reflected in the firing rate difference following a correct vs incorrect trial.

# 5

# DISCUSSION

In this chapter, I discuss the results presented in **chapter 4**. In **section 5.1**, I discuss the behavioral results and its implications and in **section 5.2**, I reflect on the neuronal recording approach and discuss possible ways to improve such recordings.

## 5.1 BEHAVIORAL RESULTS AND ITS IMPLICATIONS

I implemented a lick-left/lick-right orientation discrimination task. The mice learned the task and performed well above chance. However, even after rigorous training, performance was suboptimal. To understand why, I analyzed RT. I show that on short RT trials, animals perform on chance level throughout both learning and stable performance. Next, using the standard logistic modeling approach, I capture week-by-week learning progress and show how the perceptual choices of mice are affected by the history of past stimuli, responses, and outcomes. I further characterize the decision-making process during stable performance and show that history influences are stronger on trials with short RT. Finally, I manipulate reward size and show how visually driven choices are biased by reward expectations.

In **subsection 5.1.1**, I discuss the influence of RT on task performance and decision-making. Subsequently, in **subsection 5.1.2**, I discuss the impact of previous outcomes on decision-making. Next in **subsection 5.1.3**, I discuss behavior adaptability during learning and under asymmetric reward conditions. In **subsection 5.1.4**, I discuss why biases in decision-making exist. Finally, in **subsection 5.1.5**, I discuss

behavior experiments that would further deepen our understanding of PDM.

### 5.1.1 *Influence of reaction time on decision-making*

The relationship between RT and performance is widely discussed in the literature. In this section, I first discuss the empirical evidence for SAT and relate it to motivation and impulsive behavior. Next, I discuss the suboptimal performance during long RT trials and relate it to lapses in decision-making. Finally, I describe the relationship between RT and bias and discuss its implications.

*Baseline performance on short RT trials*

As outlined in **subsection 1.3.1**, numerous empirical studies provide evidence in support of the theory that performance decreases as RT decreases. In **Figure 4.2C**, I provide evidence for SAT: mice performed better on long RT trials, and at chance level on short RT trials. However, prior studies also suggest that both short and (very) long RTs are indicative of low task engagement (Shevinsky and Reinagel, 2019; Robbins, 2002; Lyamzin et al., 2021; Ratcliff, 1978a; Yerkes and Dodson, 1908). The former has been linked to impulsive behavior and hyper-aroused states, and the latter to low-aroused states. The result in **Figure 4.2D** did not provide evidence for low task engagement on trials with (very) long RT: the performance in the longest RT bin is above the threshold. However, as the animal had limited time to respond in the task (3 – 4 s), the response time might have not been sufficient to capture such an effect.

Poorly performed short RT trials were also prevalent in trained animals. Studies have reported that even for highly trained animals, internal factors such as motivation and attention vary within sessions (Andermann et al., 2010; Carandini and Churchland, 2013; McGinley et al., 2015; Berditchevskaia et al., 2016; Groblewski et al., 2020). Especially in studies using water restriction similar to that used in this thesis, motivation changes as the sessions progress. For example, Berditchevskaia et al. (2016) report a hyper-aroused state and poor performance in the

beginning of a session, likely resulting from thirst. Although water restriction might be effective in the early learning stages, later it could be important to maintain animals at a satisfied state to understand behavior precisely. In fact, researchers have successfully used alternative restriction methods, such as free home-cage access to citric-acid water. These studies report a slight decline in the number of completed trials, however, they show that animals readily performed and learned the behavioral task (Urai et al., 2021; Reinagel, 2018). Taken together, to better understand animals' behavior, in the absence of unaccounted internal factors, alternative restrictions should be considered.

In the following section, I relate the suboptimal performance during long RT to lapses in decision-making.

*Suboptimal performance on long RT trials*

In **Figure 4.2**, I show that while performance on trials with long RT is higher than on trials with short RT it is still suboptimal. Studies on both humans and animals have shown that, independent of stimulus strength, subjects often have judgements errors (Carandini and Churchland, 2013; Busse et al., 2011; Gold and Ding, 2013). Recent literature relate such 'errors' in behavior to 'lapses' in decision-making (Pinto et al., 2018; The International Brain Laboratory et al., 2021; Ashwood et al., 2022; Pisupati et al., 2021). Although the frequency of lapses decreases with learning, they still exist in trained animals (Law and Gold, 2009; Cloherty et al., 2020). Especially, in rodent sensory decision-making, it is speculated that lapses underlie a significant number (10–20%) of decisions (Ashwood et al., 2022). While in earlier studies, a lapse was associated with attention deficiency and motor error, a recent study relates it to strategic exploratory behavior i.e., trade-off between exploring uncertain rewarded action vs repeating a known rewarded action (Pisupati et al., 2021).

In the following section, I discuss the influence of previous experiences with respect to short and long RT.

*Reaction time and biases in decision-making*

Even in highly trained animals, previous experiences play a vital role in decision-making, especially when conditions are uncertain (Lak et al.,

2020a; Akrami et al., 2018; Busse et al., 2011). Here, I show that in trials with short RTs, animals fail to perform above chance (**subsection 4.1.5**, **subsection 4.1.6**) even when stimulus information remains unaltered. A question to consider is whether poor performance on short RT trials is associated with the inability to accumulate enough information and/or due to a stronger bias from previous experiences.

The results of my experiments suggest that on these trials, decisions are mostly biased by previous experiences. As discussed in **subsection 1.3.1**, on trials with long RT, animals can accumulate additional evidence, and reach greater decision confidence. However, short and long RT trials influence the following decisions equally (**Figure 4.5B**). Therefore, animals might collect enough evidence on short RT trials to make a correct choice. However, the influence of previous experiences on these trials might be comparatively stronger, biasing decisions and leading to poor performance. In fact, studies show that biases can be reduced by a long stimulus exposure (Wolfe, 1984; Kaneko et al., 2017; Dekel and Sagi, 2020). For example, Dekel and Sagi (2020) hypothesize that over time bias from previous experiences and noisy evidence accumulate. With long stimulus exposure time, the influence of previous experiences decreases due to noise accumulation, hence reducing the bias in decision-making.

How do animals transition between short and long RT states? Some studies associate the transition between short and long RT as a transition between separate systems: impulsive bias-susceptible system to slow bias-free system (Tversky and Kahneman, 1974; Evans and Stanovich, 2013). Similarly, the decreases in (history) biases on long RT trials can also be explained by dual processing (Zoest and Hunt, 2011; Wolfe, 1984). Others argue that a simple evidence accumulation model (i.e., DDM, also see **section 1.1**) can describe this behavior (Dekel and Sagi, 2020). Importantly, although this simple approach can describe the basic mechanistic phenomenon, there are consecutive processes in the brain. For example, depending on the behavioral task, the starting threshold (i.e., bias in **Figure 1.2B**) might be determined and integrated by higher-level brain areas, while sensory encoding is possibly performed by lower-level brain areas.

In the following section, I discuss the influence of previous outcomes on decision-making.

### 5.1.2 *Strategies based on previous outcome*

One of the extensively discussed strategies is the win-stay / lose-shift strategy (Fründ et al., 2014; Akrami et al., 2018; Braun et al., 2018; Busse et al., 2011; Hermoso-Mendizabal et al., 2020). The data presented in **Figure 4.6** support the former and provide evidence that the win-stay strategy is strongest for trials with short RT. Previous studies suggested that short RTs are indicative of low task engagement and linked them to impulsive behavior and hyper-aroused states (Shevinsky and Reinagel, 2019; Robbins, 2002; Lyamzin et al., 2021; Ratcliff, 1978a; Yerkes and Dodson, 1908). Following a reward pickup, the animal might be in a hyper-aroused state and hence makes impulsive decisions.

Contrary to what is mostly reported in the literature (lose-shift), **subsection 4.1.4** and **subsection 4.1.6** suggest negligible effects of the 'Side of incorrect response'. A correct response in PDM task often only requires correct stimulus processing, however, an incorrect choice might be influenced by various internal factors, such as, attention, motivation and impulsive behavior. Supporting this, a recent study provides evidence of diverse behavioral influences within a session that follows an incorrect trial (Lak et al., 2020a). My result here shows the overall low impact of incorrect outcomes on animals' decisions.

In the following section, I discuss contextual behavior adaptivity.

### 5.1.3 *Behavioral adaptability*

Earlier studies extensively focused on understanding context-based behavioral adaptation of fully trained animals (Romo and Salinas, 2003; Gold, Shadlen et al., 2007; Nienborg and Cumming, 2009). Recent studies extended this focus to learning (Roy et al., 2021; Ashwood et al.,

2022). These studies suggest a Bayesian approach to decision-making. Animals update their beliefs about reward probability and adjust their strategy. Using both the standard LR approach and a dynamic generalized linear model with time-varying weights (Roy et al., 2021), I first show that animals' adapt their decision-making strategies during learning. In particular, as the animal learns to associate a stimulus with choice and reward, the influence of task-relevant sensory stimulus increases. Perhaps animals learn that the stimulus sequence is random and that the stimulus repetition and alternation probabilities both are equal. Therefore, as learning progresses, previous experiences are considered less while relevant sensory information are considered more.

Contextual behavioral adaptivity is not limited to sensory processes. It involves various behavior biasing factors, for example, asymmetry in the reward or stimulus statistic (Busse et al., 2011; Akrami et al., 2018; Waiblinger et al., 2022; Stüttgen et al., 2013; Lak et al., 2020b; Gao et al., 2011; Rorie et al., 2010; Teichert and Ferrera, 2010). Also in fully trained animals, changes in behavioral strategy with respect to variation in reward or stimulus statistics are linked to behavioral adaptability (Abrahamyan et al., 2016; Treviño et al., 2020; Trevino, 2014). I show that asymmetry of rewards can also induce a bias to choose the side associated with the higher expected reward (**Figure 4.10**, **Figure 4.12**, **Figure 4.11**). But is the animal maximizing the expected reward? Earlier studies investigated the optimality in bias across different species. For example in humans, the bias is lower than optimal (Maddox, 2002), and in monkeys, it is higher than optimal (Feng et al., 2009; Teichert and Ferrera, 2010). In rodents, studies report a lower than optimal bias (Funamizu, 2021). In contrast, **Figure 4.11** suggests that during the asymmetric reward condition, mice had reward maximizing bias levels. Notably, Funamizu (2021) used an auditory task, with a reward value of 3.8 µL and 1.0 µL (i.e., 3.8:1 ratio) while I used a visual task with a reward value of 8 µL and 4 µL (i.e., 2:1 ratio). Hence, it is important to conduct further experiments with varying reward sizes and ratios to understand the underlying behavior phenomenon more precisely.

The experimental paradigm discussed in **Figure 3.2.3** involved reward contingencies that were alternated across the training block, giving me further opportunities to probe behavioral adaptation. I show that

animals successfully suppress previously rewarded responses and disengage from ongoing behavior when context (i.e., experimental blocks), changed (**Figure 4.9**). This provides further evidence of context-dependent behavioral adaptability.

In the following section, I discuss why biases exist during decision-making.

### 5.1.4 *Why do such biases in decision-making occur?*

Similar to most previous PDM experiments, in my experiments the sequence of sensory stimuli was random. Under such controlled conditions, history-dependent biases result in suboptimal performance (Akrami et al., 2018; Trevino, 2014; Treviño et al., 2020; Roy et al., 2021; Lak et al., 2020b). If so, why do such biases in decision-making occur? Some studies suggest that they could be influenced both by uncertainty in the decision and by strategies learned for maximizing reward accumulation (Abrahamyan et al., 2016; Killeen et al., 2018; Chen et al., 2021). Although the latter is true for the manipulated reward paradigm used in this thesis, for the other experiments, there should be no trial-by-trial variability in decision confidence, as the discrimination task involved only two orientations. Considering there are long-term regularities in the real-life environment (Akrami et al., 2018), history-related biases could also reflect innate strategies (Zador, 2019; Treviño et al., 2020). In fact, many behavioral patterns that are necessary for survival involve multiple repetitive or alternative actions (Langen et al., 2011). Therefore, strategies such as win-stay/lose shift might represent innate strategical preferences such as avoiding or exploring behavior that could be encoded by evolution: avoid predators and/or explore food.

Biases could also reflect reinforced habits. These learned habits become persistent and difficult to abolish in a particular context (Langen et al., 2011). However, they are adaptable if the context changes (Abrahamyan et al., 2016; Treviño et al., 2020; Trevino, 2014). The results presented in **Figure 4.9** suggest that some choice biases are adaptable. These types

of bias are often based on purely motivational processes. For example, while the animals had a strong bias towards the double-rewarded side, they were able to quickly adapt when the reward size was balanced again.

### 5.1.5 *Future research*

Based on the behavioral results presented here, there are several experiments that would further deepen our understanding. In the described orientation discrimination task, mice discriminate between two distinct stimuli. It would be interesting to introduce more stimuli, by varying either contrast or orientation. For example, the grating direction could be varied in step sizes of 30° from 0° to 360°, and the mouse could be trained to associate orientations between 0° to 180° to the right and from 180° to 360° to the left lick spout. I expect history biases to be stronger when the difference in orientation between prior and current stimulus is smaller. In addition, I hypothesize that the dependencies on history would increase when discrimination would get more difficult (close to 180°).

Past studies show how asymmetry of reward biases decision-making, especially when trial difficulty increases. Here, I show that during the asymmetric condition used in this thesis, mice's bias is optimal with regard to reward maximization **Figure 4.11**. However, it remains unclear whether this optimality in bias is dependent on magnitude of reward size differences. Further experiments with varying reward sizes would help to better understand this phenomenon. Perhaps, with a small reward size difference, the animal might not adapt its strategy. With increasing differences in reward sizes, the incentive to have a bias towards the greater reward size increases. In addition to reward statistics, the probability that a particular side choice is rewarded could be varied (Sugawara and Katahira, 2021). Furthermore, both reward statistic and stimulus probability could be varied to understand whether mice integrate both of these variables while making a decision or evaluate one before the other (Nachev et al., 2021).

Finally, the lick-left, lick-right orientation discrimination task that I used in the behavioral paradigm has also been extensively used in the literature (Burgess et al., 2017; Akrami et al., 2018; Guo et al., 2014a). Two types of motor movements to consider in such tasks are the movement made to make a decision and the one made to pick up a reward. It might be beneficial to disentangle the two motor movements in the design of the task. Specifically, this would allow the researcher to isolate impulsive reward pickup movement from the animal's decision. An alternative way to measure behavioral responses such as 'lever push' or a 'wheel movement' should be considered (The International Brain Laboratory et al., 2021; Rossi and Yin, 2012).

## 5.2 Neuronal recording approach and improvement opportunities

The results presented in this thesis show that neuronal activity can be acquired from the mouse visual cortex through neuronal recording for approximately 2 months. Furthermore, the acquired neuronal units were stable for several days. However, as shown in **Figure 4.13**, only 6 single units could be isolated on the first day of recording, and this number declined over time. In order to capture both learning and stable performance phases, the method needs to be adapted to acquire recordings from more units and track them for longer durations.

In **subsection 5.2.1**, I discuss specific ways to improve the probe implantation method, to enable longer-term neuronal recording. Subsequently, in **subsection 5.2.2**, I discuss the ideal time to implant the probe to capture neuronal responses during different stages of learning. In **subsection 5.2.3**, I discuss the advantages and disadvantages of immobile and microdrive probe configurations. Finally, in **subsection 5.2.4**, I discuss electrophysiology experiments that would further deepen our understanding of the neuronal mechanisms of PDM.

### 5.2.1 *Method improvement of probe implantation*

In this section, I discuss four ways to improve the method of probe implantation. First, an ideal craniotomy size should be carefully considered. For the animals used in this thesis, a $1.5 \times 1.5$ mm craniotomy was made. Given the thickness of the probe (15 µm), the size of the craniotomy could be reduced. A smaller craniotomy, would expose a smaller section of the brain, hence reducing the risk of infection and tissue degeneration. Second, unlike the self-curing acrylic dental cement used in Okun et al. (2016), I used UV-curing dental cement, because it is stable and easy to handle. Although the exposed brain was covered with Kwik cast, the UV light may still have reached and deteriorated the tissue, hence reducing recording quality. Third, the hardening of cement around the probe causes ~ 10–20 µm displacement after implantation (Okun et al., 2016; Lee et al., 2014). While the acrylic cement takes 10 mins to self-cure, UV curing dental cement cures within seconds of UV exposure. A disadvantage of such instantaneous UV-curing might be sudden probe displacement, which would result in more neuronal damage than slow curing with acrylic cement. Lastly, after implantation, the probe was not enclosed with an protective enclosure. Building such a protective enclosure might help reduce trauma caused by external events.

In the next section, I discuss the ideal time to implant the probe to capture neuronal responses during different stages of learning.

### 5.2.2 *Probe implantation timing*

Timing of probe implantation needs to be carefully considered. To capture the learning curve, I implanted the probe 2–3 weeks after starting to train the animals on the final task. During those prior weeks, I aimed to acclimatize the animals to the behavioral set-up and to teach them to associate the correct response with rewards. Unfortunately, even after 8 weeks of training, the animals did not learn. To researchers attempting similar experiments, I suggest implanting the probe early

in training, i.e. before the beginning of the final task. Due to protocol and animal welfare constraints, the animals were required to be taken-off training for an extended period of time. First, several days before performing the craniotomy, it is necessary to provide the "ad libitum" water. Second, after performing the implantation, they need a recovery time of at least one week. This disruption of training likely interfered with learning. As neuronal signals were present for ~ 2 months after implantation, probe implantation before the final task still allows capturing both the naive and trained stages.

PDM strategies evolve and fluctuate even during the stable performance. Hence, it is important to understand the sensory decision-making process of fully trained animals. Especially on tasks with varying stimulus and reward statistics, animals are continually required to change strategies to maximize reward. It would be insightful to track neuronal responses under such conditions. For this, the probe implantation should be performed after the animal reached stable performance. First, in fully trained animals, training disruption did not disrupt performance for long. For example, for fully trained animals used only on behavioral studies in this thesis, after five days of training, following a week of disruption, animals' performance was back to the pre-disruption level. Second, given that I was able to record neuronal responses for ~ 2 months, implanting the probe before the final task might not give sufficient time to capture the majority of stable performance and the reward manipulation phase. In summary, it is important to carefully consider the study objectives, to identify the timing of probe implantation.

### 5.2.3 *Comparison of microdrive and immobile probe*

Compared to the standard microdrive method, the configuration of the immobile silicone probe is lighter and the implantation procedure is both easier and faster. This is especially useful when performing neuronal recordings on small animals such as mice (Okun et al., 2016). Some studies also suggest that this method might provide neuronal recording stability over longer time periods (Juavinett et al., 2019;

Okun et al., 2016; Steinmetz et al., 2021), however, this remains to be demonstrated. Furthermore, since this method does not require any movable parts, it can be completely sealed off after implantation, minimizing the risk of infection after surgery. However, this also comes with a disadvantage: the microdrive allows the probe to move further into the brain when the recording depth needs to be adjusted, e.g., due to deteriorating recording quality. Here, I was unable to acquire any signal after approximately 2 months. As the probe configuration that I used was immobile, I was unable to change its position after the signal decayed and, hence, had to terminate the experiment after 2 months. Therefore, while immobile chronic implantation is advantageous under certain circumstances, experimental requirements should be considered when evaluating recording approaches.

### 5.2.4 *Future research*

Based on the results presented in this thesis, there are experiments that would further deepen our understanding of PDM. In this section, I give a few examples of experiments that could deepen our understanding of the neuronal mechanisms that underlie PDM.

The behavioral paradigm and long-term neuronal recording technique that I used provide an opportunity to capture the dynamic learning progress at a single cell level. First, tracking neuronal responses from the same neuron and/or neuronal population over learning might enable us to identify how neuronal responses and preferences vary as animals learn the behavioral relevance of sensory information (Jurjut et al., 2017; Jehee et al., 2012; Hua et al., 2010; Li et al., 2004). For example, precisely tracking orientation preferences during task learning and during passive orientation tuning experiments would help us understand whether learning shifts the preferences of individual V1 neurons. If it does, a sharper discriminability during task performance, but not during passive tuning measurements, is to be expected (Jurjut et al., 2017). Second, the neuronal recording method described here would allow one to understand at what stage of task training does neuronal responses indicate learning. For example, a study using acute

recording has reported that neural responses in V1 indicate learning before the animal behaviorally demonstrated it (Jurjut et al., 2017). The method described in **subsection 3.3.2** may allow us to observe changes in neuronal encoding throughout different learning stages and hence get a deeper understanding of this mechanism.

Successfully establishing the neuronal correlates of bias provide opportunities to investigate the role of brain areas in integrating history-related biases (Akrami et al., 2018; Hwang et al., 2017; Tai et al., 2012). Intervention via, for example, optogenetic manipulation would give us further insights into the neural mechanisms. For example, one could investigate, whether history dependent V1 response modulation was implemented via a top-down feedback circuit. If so, silencing the long-range feedback circuit, terminating in V1 for example the PFC, PPC would reduce the impact of history-dependent biases and thereby improve the animal's performance.

The task is well suited to isolate when and where sensory information and history-related information are integrated in the decision-making process. While some studies identified that integration happens at the perceptual level (Fritsche et al., 2017; Patten et al., 2017), others suggested that it happens at a post-perceptual level (Akrami et al., 2018). Utilizing the task structure described in **subsection 3.2.3**, one could optogenetically manipulate areas of interest at different stages of the task (trial initiation, grace period, stimulus presentation, reward pickup / timeout), thereby providing further evidence on when and where biases are integrated (Akrami et al., 2018; Hwang et al., 2017). Finally, a similar approach could be used to identify the role of the visual cortex. For example, optogenetically silencing V1 during various stages of the trial would provide us with additional information on the importance of V1 in visual PDM.

# REFERENCES


Abrahamyan, Arman, Laura Luz Silva, Steven C Dakin, Matteo Carandini and Justin L Gardner (2016). 'Adaptable history biases in human perceptual decisions'. *Proceedings of the National Academy of Sciences* 113.25, E3548–E3557.

Akrami, Athena, Charles D. Kopec, Mathew E. Diamond and Carlos D. Brody (2018). 'Posterior Parietal Cortex Represents Sensory History and Mediates Its Effects on Behaviour'. *Nature* 554.7692, pp. 368–372.

Amiez, Céline, Jean-Paul Joseph and Emmanuel Procyk (2006). 'Reward encoding in the monkey anterior cingulate cortex'. *Cerebral Cortex* 16.7, pp. 1040–1055.

Andermann, Mark L, Aaron M Kerlin and Clay Reid (2010). 'Chronic cellular imaging of mouse visual cortex during operant behavior and passive viewing'. *Frontiers in Cellular Neuroscience* 4, p. 3.

Ashwood, Zoe C, Nicholas A Roy, Iris R Stone, Anne E Urai, Anne K Churchland, Alexandre Pouget and Jonathan W Pillow (2022). 'Mice alternate between discrete strategies during perceptual decision-making'. *Nature Neuroscience* 25.2, pp. 201–212.

Aton, Sara J (2013). 'Set and setting: how behavioral state regulates sensory function and plasticity'. *Neurobiology of Learning and Memory* 106, pp. 1–10.

Balci, Fuat, Patrick Simen, Ritwik Niyogi, Andrew Saxe, Jessica A Hughes, Philip Holmes and Jonathan D Cohen (2011). 'Acquisition of decision making criteria: reward rate ultimately beats accuracy'. *Attention, Perception, & Psychophysics* 73.2, pp. 640–657.

Balleine, Bernard W, Mauricio R Delgado and Okihide Hikosaka (2007). 'The role of the dorsal striatum in reward and decision-making'. *Journal of Neuroscience* 27.31, pp. 8161–8165.

Berditchevskaia, Aleksandra, RD Cazé and Simon R Schultz (2016). 'Performance in a GO/NOGO perceptual task reflects a balance between impulsive and instrumental components of behaviour'. *Scientific Reports* 6.1, pp. 1–15.

Bitzidou, Malamati, Michael R Bale and Miguel Maravall (2018). 'Cortical lifelogging: the posterior parietal cortex as sensory history buffer'. *Neuron* 98.2, pp. 249–252.

Blatt, Gene J, Richard A Andersen and Gene R Stoner (1990). 'Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque'. *Journal of Comparative Neurology* 299.4, pp. 421–445.

Boucher, Leanne, Thomas J Palmeri, Gordon D Logan and Jeffrey D Schall (2007). 'Inhibitory control in mind and brain: an interactive race model of countermanding saccades.' *Psychological Review* 114.2, p. 376.

Braun, Anke, Anne E Urai and Tobias H Donner (2018). 'Adaptive history biases result from confidence-weighted accumulation of past choices'. *Journal of Neuroscience* 38.10, pp. 2418–2429.

Britten, Kenneth H, Michael N Shadlen, William T Newsome and J Anthony Movshon (1992). 'The analysis of visual motion: a comparison of neuronal and psychophysical performance'. *Journal of Neuroscience* 12.12, pp. 4745–4765.

Buckley, Mark J, Farshad A Mansouri, Hassan Hoda, Majid Mahboubi, Philip GF Browning, Sze C Kwok, Adam Phillips and Keiji Tanaka (2009). 'Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions'. *Science* 325.5936, pp. 52–58.

Burgess, Christopher P, Armin Lak, Nicholas A Steinmetz, Peter Zatka-Haas, Charu Bai Reddy, Elina AK Jacobs, Jennifer F Linden, Joseph J Paton, Adam Ranson, Sylvia Schröder et al. (2017). 'High-yield methods for accurate two-alternative visual psychophysics in head-fixed mice'. *Cell reports* 20.10, pp. 2513–2524.

Busse, Laura, Asli Ayaz, Neel T. Dhruv, Steffen Katzner, Aman B. Saleem, Marieke L. Schölvinck, Andrew D. Zaharia and Matteo

Carandini (2011). 'The Detection of Visual Contrast in the Behaving Mouse'. *The Journal of Neuroscience* 31.31, pp. 11351–11361.

Cai, Lei, Bian Wu and Shuiwang Ji (2018). 'Neuronal activities in the mouse visual cortex predict patterns of sensory stimuli'. *Neuroinformatics* 16, pp. 473–488.

Carandini, Matteo and Anne K Churchland (2013). 'Probing perceptual decisions in rodents'. *Nature Neuroscience* 16.7, pp. 824–831.

Chen, Cathy S, R Becket Ebitz, Sylvia R Bindas, A David Redish, Benjamin Y Hayden and Nicola M Grissom (2021). 'Divergent strategies for learning in males and females'. *Current Biology* 31.1, pp. 39–50.

Chen, Mei-Yen, Koji Jimura, Corey N White, W Todd Maddox and Russell A Poldrack (2015). 'Multiple brain networks contribute to the acquisition of bias in perceptual decision-making'. *Frontiers in Neuroscience* 9, p. 63.

Cicmil, Nela, Bruce G Cumming, Andrew J Parker and Kristine Krug (2015). 'Reward modulates the effect of visual cortical microstimulation on perceptual decisions'. *Elife* 4, e07832.

Cisek, Paul and John F Kalaska (2010). 'Neural mechanisms for interacting with a world full of action choices'. *Annual Review of Neuroscience* 33, pp. 269–298.

Cloherty, Shaun L, Jacob L Yates, Dina Graf, Gregory C DeAngelis and Jude F Mitchell (2020). 'Motion perception in the common marmoset'. *Cerebral Cortex* 30.4, pp. 2659–2673.

Corcoran, Kevin A and Gregory J Quirk (2007). 'Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears'. *Journal of Neuroscience* 27.4, pp. 840–844.

De Lafuente, Victor, Mehrdad Jazayeri and Michael N Shadlen (2015). 'Representation of accumulating evidence for a decision in two parietal areas'. *Journal of Neuroscience* 35.10, pp. 4306–4318.

Dekel, Ron and Dov Sagi (2020). 'Perceptual bias is reduced with longer reaction times during visual discrimination'. *Communications biology* 3.1, pp. 1–12.

Ding, Long and Joshua I Gold (2010). 'Caudate encodes multiple computations for perceptual decisions'. *Journal of Neuroscience* 30.47, pp. 15747–15759.

Ding, Long and Joshua I Gold (2012a). 'Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field'. *Cerebral Cortex* 22.5, pp. 1052–1067.

Ding, Long and Joshua I Gold (2012b). 'Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task'. *Neuron* 75.5, pp. 865–874.

Donner, Tobias H, Markus Siegel, Pascal Fries and Andreas K Engel (2009). 'Buildup of choice-predictive activity in human motor cortex during perceptual decision making'. *Current Biology* 19.18, pp. 1581–1585.

Evans, Jonathan St BT and Keith E Stanovich (2013). 'Dual-process theories of higher cognition: Advancing the debate'. *Perspectives on Psychological Sfroncience* 8.3, pp. 223–241.

Fassihi, Arash, Athena Akrami, Francesca Pulecchi, Vinzenz Schönfelder and Mathew E. Diamond (2017). 'Transformation of Perception from Sensory to Motor Cortex'. *Current Biology* 27.11, 1585–1596.e6.

Feng, Samuel, Philip Holmes, Alan Rorie and William T Newsome (2009). 'Can monkeys choose optimally when faced with noisy stimuli and unequal rewards?' *PLoS Computational Biology* 5.2, e1000284.

Franks, Nigel R, Anna Dornhaus, Jon P Fitzsimmons and Martin Stevens (2003). 'Speed versus accuracy in collective decision making'. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270.1532, pp. 2457–2463.

Fritsche, Matthias, Pim Mostert and Floris P de Lange (2017). 'Opposite effects of recent history on perception and decision'. *Current Biology* 27.4, pp. 590–595.

Fründ, Ingo, Felix A. Wichmann and Jakob H. Macke (2014). 'Quantifying the Effect of Intertrial Dependence on Perceptual Decisions'. *Journal of Vision* 14.7, pp. 9–9.

Funamizu, Akihiro (2021). 'Integration of sensory evidence and reward expectation in mouse perceptual decision-making task with various sensory uncertainties'. *Iscience* 24.8, p. 102826.

Gao, Juan, Rebecca Tortell and James L McClelland (2011). 'Dynamic integration of reward and stimulus information in perceptual decision-making'. *PloS one* 6.3, e16749.

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

Gold, Joshua I and Michael N Shadlen (2003). 'The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands'. *Journal of Neuroscience* 23.2, pp. 632–651.

Gold, Joshua I, Michael N Shadlen et al. (2007). 'The neural basis of decision making'. *Annual review of Neuroscience* 30.1, pp. 535–574.

Gold, Joshua I and Alan A Stocker (2017). 'Visual decision-making in an uncertain and dynamic world'. *Annual Review of Vision Science* 3.1, pp. 227–250.

Goltstein, Pieter M., Sandra Reinert, Annet Glas, Tobias Bonhoeffer and Mark Hübener (2018). 'Food and water restriction lead to differential learning behaviors in a head-fixed two-choice visual discrimination task for mice'. *PLoS ONE* 13.9, e0204066.

Green, David M. and John A. Swets (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

Groblewski, Peter A, Douglas R Ollerenshaw, Justin T Kiggins, Marina E Garrett, Chris Mochizuki, Linzy Casal, Sissy Cross, Kyla Mace, Jackie Swapp, Sahar Manavi et al. (2020). 'Characterization of learning, motivation, and visual perception in five transgenic mouse lines expressing GCaMP in distinct cell populations'. *Frontiers in Behavioral Neuroscience* 14, p. 104.

Guo, Zengcai V, S Andrew Hires, Nuo Li, Daniel H O'Connor, Takaki Komiyama, Eran Ophir, Daniel Huber, Claudia Bonardi, Karin Morandell, Diego Gutnisky et al. (2014a). 'Procedures for behavioral experiments in head-fixed mice'. *PloS one* 9.2, e88678.

Guo, Zengcai V, Nuo Li, Daniel Huber, Eran Ophir, Diego Gutnisky, Jonathan T Ting, Guoping Feng and Karel Svoboda (2014b). 'Flow of cortical activity underlying a tactile decision in mice'. *Neuron* 81.1, pp. 179–194.

Haber, Suzanne N (2022). 'Corticostriatal circuitry'. *Dialogues in Clinical Neuroscience*.

Hanes, Doug P and Jeffrey D Schall (1996). 'Neural control of voluntary movement initiation'. *Science* 274.5286, pp. 427–430.

Hanks, Timothy, Roozbeh Kiani and Michael N Shadlen (2014). 'A neural mechanism of speed-accuracy tradeoff in macaque area LIP'. *eLife* 3.

Hanks, Timothy D, Charles D Kopec, Bingni W Brunton, Chunyu A Duan, Jeffrey C Erlich and Carlos D Brody (2015). 'Distinct relationships of parietal and prefrontal cortices to evidence accumulation'. *Nature* 520.7546, pp. 220–223.

Heitz, Richard P (2014). 'The speed-accuracy tradeoff: history, physiology, methodology, and behavior'. *Frontiers in Neuroscience* 8, p. 150.

Heitz, Richard P and Jeffrey D Schall (2012). 'Neural mechanisms of speed-accuracy tradeoff'. *Neuron* 76.3, pp. 616–628.

Henmon, Vivian Allen Charles (1911). 'The relation of the time of a judgment to its accuracy.' *Psychological Review* 18.3, p. 186.

Hermoso-Mendizabal, Ainhoa, Alexandre Hyafil, Pavel E. Rueda-Orozco, Santiago Jaramillo, David Robbe and Jaime de la Rocha (2020). 'Response Outcomes Gate the Impact of Expectations on Perceptual Decisions'. *Nature Communications* 11.1, p. 1057.

Hollingworth, Harry Levi (1910). 'The central tendency of judgment'. *The Journal of Philosophy, Psychology and Scientific Methods* 7.17, pp. 461–469.

Hua, Tianmiao, Pinglei Bao, Chang-Bing Huang, Zhenhua Wang, Jinwang Xu, Yifeng Zhou and Zhong-Lin Lu (2010). 'Perceptual learning improves contrast sensitivity of V1 neurons in cats'. *Current biology* 20.10, pp. 887–894.

Hwang, Eun Jung, Jeffrey E Dahlen, Madan Mukundan and Takaki Komiyama (2017). 'History-based action selection bias in posterior parietal cortex'. *Nature Communications* 8.1, pp. 1–14.

Hwang, Eun Jung, Trevor D Link, Yvonne Yuling Hu, Shan Lu, Eric Hou-Jen Wang, Varoth Lilascharoen, Sage Aronson, Keelin O'Neil, Byung Kook Lim and Takaki Komiyama (2019). 'Corticostriatal flow of action selection bias'. *Neuron* 104.6, pp. 1126–1140.

Ings, Thomas C and Lars Chittka (2008). 'Speed-accuracy tradeoffs and false alarms in bee responses to cryptic predators'. *Current Biology* 18.19, pp. 1520–1524.

Jehee, Janneke FM, Sam Ling, Jascha D Swisher, Ruben S van Bergen and Frank Tong (2012). 'Perceptual learning selectively refines orientation representations in early visual cortex'. *Journal of Neuroscience* 32.47, pp. 16747–16753.

Juavinett, Ashley L, George Bekheet and Anne K Churchland (2019). 'Chronically implanted Neuropixels probes enable high-yield recordings in freely moving mice'. *Elife* 8, e47188.

Jurjut, Ovidiu, Petya Georgieva, Laura Busse and Steffen Katzner (2017). 'Learning enhances sensory processing in mouse V1 before improving behavior'. *Journal of Neuroscience* 37.27, pp. 6460–6474.

Kaneko, Sae, Stuart Anstis and Ichiro Kuriki (2017). 'Brief presentation enhances various simultaneous contrast effects'. *Journal of vision* 17.4, pp. 7–7.

Killeen, Peter R, Thomas J Taylor and Mario Treviño (2018). 'Subjects adjust criterion on errors in perceptual decision tasks.' *Psychological Review* 125.1, p. 117.

Klaes, Christian, Stephanie Westendorff, Shubhodeep Chakrabarti and Alexander Gail (2011). 'Choosing goals, not rules: deciding among rule-based action plans'. *Neuron* 70.3, pp. 536–548.

Knill, David C and Whitman Richards (1996). *Perception as Bayesian inference*. Cambridge University Press.

Kurylo, Daniel D, Caroline Chung, Sowmya Yeturo, Joseph Lanza, Arina Gorskaya and Farhan Bukhari (2015). 'Effects of contrast,

spatial frequency, and stimulus duration on reaction time in rats'. *Vision Research* 106, pp. 20–26.

Lak, Armin, Emily Hueske, Junya Hirokawa, Paul Masset, Torben Ott, Anne E Urai, Tobias H Donner, Matteo Carandini, Susumu Tonegawa, Naoshige Uchida et al. (2020a). 'Reinforcement biases subsequent perceptual decisions when confidence is low, a widespread behavioral phenomenon'. *ELife* 9, e49834.

Lak, Armin, Michael Okun, Morgane M. Moss, Harsha Gurnani, Karolina Farrell, Miles J. Wells, Charu Bai Reddy, Adam Kepecs, Kenneth D. Harris and Matteo Carandini (2020b). 'Dopaminergic and Prefrontal Basis of Learning from Sensory Confidence and Reward Value'. *Neuron* 105.4, 700–711.e6.

Lange, Floris P de, Dobromir A Rahnev, Tobias H Donner and Hakwan Lau (2013). 'Prestimulus oscillatory activity over motor cortex reflects perceptual expectations'. *Journal of Neuroscience* 33.4, pp. 1400–1410.

Langen, Marieke, Martien JH Kas, Wouter G Staal, Herman van Engeland and Sarah Durston (2011). 'The neurobiology of repetitive behavior: of mice…' *Neuroscience & Biobehavioral Reviews* 35.3, pp. 345–355.

Lau, Brian and Paul W Glimcher (2008). 'Value representations in the primate striatum during matching behavior'. *Neuron* 58.3, pp. 451–463.

Lauwereyns, Johan, Katsumi Watanabe, Brian Coe and Okihide Hikosaka (2002). 'A neural correlate of response bias in monkey caudate nucleus'. *Nature* 418.6896, pp. 413–417.

Law, Chi-Tat and Joshua I Gold (2009). 'Reinforcement learning can account for associative and perceptual learning on a visual-decision task'. *Nature neuroscience* 12.5, pp. 655–663.

Lee, Daeyeol, Matthew FS Rushworth, Mark E Walton, Masataka Watanabe and Masamichi Sakagami (2007). 'Functional specialization of the primate frontal cortex during decision making'. *Journal of Neuroscience* 27.31, pp. 8170–8173.

Lee, Doyun, Gleb Shtengel, Jason E Osborne and Albert K Lee (2014). 'Anesthetized-and awake-patched whole-cell recordings in freely moving rats using UV-cured collar-based electrode stabilization'. *nature protocols* 9.12, pp. 2784–2795.

Li, Wu, Valentin Piëch and Charles D Gilbert (2004). 'Perceptual learning and top-down influences in primary visual cortex'. *Nature neuroscience* 7.6, pp. 651–657.

Lin, Zhongqiao, Chechang Nie, Yuanfeng Zhang, Yang Chen and Tianming Yang (2020). 'Evidence accumulation for value computation in the prefrontal cortex during decision making'. *Proceedings of the National Academy of Sciences* 117.48, pp. 30728–30737.

Liu, Bao-hua, Pingyang Li, Ya-tang Li, Yujiao J Sun, Yuchio Yanagawa, Kunihiko Obata, Li I Zhang and Huizhong W Tao (2009). 'Visual receptive field structure of cortical inhibitory neurons revealed by two-photon imaging guided recording'. *Journal of Neuroscience* 29.34, pp. 10520–10532.

Liu, Bao-hua, Pingyang Li, Yujiao J Sun, Ya-tang Li, Li I Zhang and Huizhong Whit Tao (2010). 'Intervening inhibition underlies simple-cell receptive field structure in visual cortex'. *Nature Neuroscience* 13.1, pp. 89–96.

Lyamzin, Dmitry R., Ryo Aoki, Mohammad Abdolrahmani and Andrea Benucci (2021). 'Probabilistic discrimination of relative stimulus features in mice'. *Proceedings of the National Academy of Sciences* 118.30.

Maddox, W Todd (2002). 'Toward a unified theory of decision criterion learning in perceptual categorization'. *Journal of the experimental analysis of behavior* 78.3, pp. 567–595.

Marques, Tiago, Mathew T. Summers, Gabriela Fioreze, Marina Fridman, Rodrigo F. Dias, Marla B. Feller and Leopoldo Petreanu (2018). 'A Role for Mouse Primary Visual Cortex in Motion Perception'. English. *Current Biology* 28.11, 1703–1713.e6.

Marshall, James AR, Rafal Bogacz, Anna Dornhaus, Robert Planqué, Tim Kovacs and Nigel R Franks (2009). 'On optimal decision-making in brains and social insect colonies'. *Journal of the Royal Society Interface* 6.40, pp. 1065–1074.

McGinley, Matthew J, Martin Vinck, Jacob Reimer, Renata Batista-Brito, Edward Zagha, Cathryn R Cadwell, Andreas S Tolias, Jessica A Cardin and David A McCormick (2015). 'Waking state: rapid variations modulate neural and behavioral responses'. *Neuron* 87.6, pp. 1143–1161.

Mendonça, André G, Jan Drugowitsch, M Inês Vicente, Eric EJ DeWitt, Alexandre Pouget and Zachary F Mainen (2020). 'The impact of learning on perceptual decisions and its implication for speed-accuracy tradeoffs'. *Nature Communications* 11.1, pp. 1–15.

Nachev, Vladislav, Marion Rivalan and York Winter (2021). 'Two-dimensional reward evaluation in mice'. *Animal cognition* 24.5, pp. 981–998.

Narayanan, NS, NK Horst and Mark Laubach (2006). 'Reversible inactivations of rat medial prefrontal cortex impair the ability to wait for a stimulus'. *Neuroscience* 139.3, pp. 865–876.

Newsome, William T, Kenneth H Britten and J Anthony Movshon (1989). 'Neuronal correlates of a perceptual decision'. *Nature* 341.6237, pp. 52–54.

Nienborg, Hendrikje and Bruce Cumming (2010). 'Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality?' *Current opinion in neurobiology* 20.3, pp. 376–381.

Nienborg, Hendrikje and Bruce G Cumming (2009). 'Decision-related activity in sensory neurons reflects more than a neuron's causal effect'. *Nature* 459.7243, pp. 89–92.

O'Reilly, Jill X, Urs Schüffelgen, Steven F Cuell, Timothy EJ Behrens, Rogier B Mars and Matthew FS Rushworth (2013). 'Dissociable effects of surprise and model update in parietal and anterior cingulate cortex'. *Proceedings of the National Academy of Sciences* 110.38, E3660–E3669.

Okun, Michael, Armin Lak, Matteo Carandini and Kenneth D. Harris (2016). 'Long Term Recordings with Immobile Silicon Probes in the Mouse Cortex'. *PLoS ONE* 11.3, e0151180.

Olkkonen, Maria, Patrice F McCarthy and Sarah R Allred (2014). 'The central tendency bias in color perception: Effects of internal and external noise'. *Journal of Vision* 14.11, pp. 5–5.

Pachitariu, Marius, Nicholas Steinmetz, Shabnam Kadir, Matteo Carandini et al. (2016). 'Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels'. *BioRxiv*, p. 061481.

Palmer, John, Alexander C Huk and Michael N Shadlen (2005). 'The effect of stimulus strength on the speed and accuracy of a perceptual decision'. *Journal of Vision* 5.5, pp. 1–1.

Pape, Anna-Antonia and Markus Siegel (2016). 'Motor cortex activity predicts response alternation during sensorimotor decisions'. *Nature Communications* 7.1, pp. 1–10.

Patten, Matthew L, Damien J Mannion and Colin WG Clifford (2017). 'Correlates of perceptual orientation biases in human primary visual cortex'. *Journal of Neuroscience* 37.18, pp. 4744–4750.

Perri, Rinaldo Livio, Marika Berchicci, Donatella Spinelli and Francesco Di Russo (2014). 'Individual differences in response speed and accuracy are associated to specific brain activities of two interacting systems'. *Frontiers in Behavioral Neuroscience* 8, p. 251.

Pinto, Lucas, Sue A. Koay, Ben Engelhard, Alice M. Yoon, Ben Deverett, Stephan Y. Thiberge, Ilana B. Witten, David W. Tank and Carlos D. Brody (2018). 'An Accumulation-of-Evidence Task Using Visual Pulses for Mice Navigating in Virtual Reality'. *Frontiers in Behavioral Neuroscience* 12.

Pisupati, Sashank, Lital Chartarifsky-Lynn, Anup Khanal and Anne K Churchland (2021). 'Lapses in perceptual decisions reflect exploration'. *Elife* 10, e55490.

Plainis, S and IJ Murray (2000). 'Neurophysiological interpretation of human visual reaction times: effect of contrast, spatial frequency and luminance'. *Neuropsychologia* 38.12, pp. 1555–1564.

Ratcliff, Roger (1978a). 'A theory of memory retrieval'. *Psychological Review* 85.2, pp. 59–108.

Ratcliff, Roger (1978b). 'A theory of memory retrieval.' *Psychological review* 85.2, p. 59.

Ratcliff, Roger, Anil Cherian and Mark Segraves (2003). 'A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions'. *Journal of Neurophysiology* 90.3, pp. 1392–1407.

Ratcliff, Roger and Philip L Smith (2004). 'A comparison of sequential sampling models for two-choice reaction time.' *Psychological review* 111.2, p. 333.

Raviv, Ofri, Itay Lieder, Yonatan Loewenstein and Merav Ahissar (2014). 'Contradictory behavioral biases result from the influence of past stimuli on perception'. *PLoS Computational Biology* 10.12, e1003948.

Reinagel, Pamela (2018). 'Training rats using water rewards without water restriction'. *Frontiers in Behavioral Neuroscience* 12, p. 84.

Rinberg, Dmitry, Alexei Koulakov and Alan Gelperin (2006). 'Speed-accuracy tradeoff in olfaction'. *Neuron* 51.3, pp. 351–358.

Robbins, T. (2002). 'The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry'. *Psychopharmacology* 163.3-4, pp. 362–380.

Roitman, Jamie D and Michael N Shadlen (2002). 'Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task'. *Journal of Neuroscience* 22.21, pp. 9475–9489.

Romo, Ranulfo, Adrián Hernández, Antonio Zainos, Carlos D Brody and Luis Lemus (2000). 'Sensing without touching: psychophysical performance based on cortical microstimulation'. *Neuron* 26.1, pp. 273–278.

Romo, Ranulfo, Adrián Hernández, Antonio Zainos, Luis Lemus and Carlos D Brody (2002). 'Neuronal correlates of decision-making in secondary somatosensory cortex'. *Nature neuroscience* 5.11, pp. 1217–1225.

Romo, Ranulfo and Emilio Salinas (2003). 'Flutter discrimination: neural codes, perception, memory and decision making'. *Nature Reviews Neuroscience* 4.3, pp. 203–218.

Rorie, Alan E, Juan Gao, James L McClelland and William T Newsome (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.

Rossi, Mark A and Henry H Yin (2012). 'Methods for studying habitual behavior in mice'. *Current protocols in neuroscience* 60.1, pp. 8–29.

Roy, Nicholas A., Ji Hyun Bak, Athena Akrami, Carlos D. Brody and Jonathan W. Pillow (2021). 'Extracting the dynamics of behavior in sensory decision-making experiments'. *Neuron* 109.4, 597–610.e6.

Salzman, C Daniel, Kenneth H Britten and William T Newsome (1990). 'Cortical microstimulation influences perceptual judgements of motion direction'. *Nature* 346.6280, pp. 174–177.

Samejima, Kazuyuki, Yasumasa Ueda, Kenji Doya and Minoru Kimura (2005). 'Representation of action-specific reward values in the striatum'. *Science* 310.5752, pp. 1337–1340.

Saproo, Sameer and John T Serences (2010). 'Spatial attention improves the quality of population codes in human visual cortex'. *Journal of neurophysiology* 104.2, pp. 885–895.

Schoenbaum, Geoffrey, Andrea A Chiba and Michela Gallagher (1998). 'Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning'. *Nature Neuroscience* 1.2, pp. 155–159.

Schoups, Aniek, Rufin Vogels, Ning Qian and Guy Orban (2001). 'Practising orientation identification improves orientation coding in V1 neurons'. *Nature* 412.6846, pp. 549–553.

Schultz, Wolfram (2022). 'Dopamine reward prediction error coding'. *Dialogues in clinical neuroscience*.

Schwarz, Cornelius, Harald Hentschke, Sergejus Butovas, Florent Haiss, Maik Stüttgen, Todor Gerdjikov, Caroline Bergner and Christian Waiblinger (2010). 'The Head-fixed Behaving Rat—Procedures and Pitfalls'. *Somatosensory and Motor Research* 27, pp. 131–48.

Seamans, Jeremy K, Stanley B Floresco and Anthony G Phillips (1995). 'Functional differences between the prelimbic and anterior cingulate regions of the rat prefrontal cortex.' *Behavioral Neuroscience* 109.6, p. 1063.

Seriès, Peggy and Aaron R Seitz (2013). 'Learning what to expect (in visual perception)'. *Frontiers in Human Neuroscience* 7, p. 668.

Shadlen, Michael N and William T Newsome (1996). 'Motion perception: seeing and deciding.' *Proceedings of the National Academy of Sciences* 93.2, pp. 628–633.

Shevinsky, Carly A. and Pamela Reinagel (2019). 'The Interaction Between Elapsed Time and Decision Accuracy Differs Between Humans and Rats'. *Frontiers in Neuroscience* 13.

Slotnick, Burton (2009). 'A SIMPLE 2-TRANSISTOR TOUCH OR LICK DETECTOR CIRCUIT'. *Journal of the Experimental Analysis of Behavior* 91.2, pp. 253–255.

Spacek, Martin A, Tim Blanche and Nick Swindale (2009). 'Python for large-scale electrophysiology'. *Frontiers in Neuroinformatics*, p. 9.

Steinmetz, Nicholas A, Cagatay Aydin, Anna Lebedeva, Michael Okun, Marius Pachitariu, Marius Bauza, Maxime Beau, Jai Bhagat, Claudia Böhm, Martijn Broux et al. (2021). 'Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings'. *Science* 372.6539, eabf4588.

Steinmetz, Nicholas A, Peter Zatka-Haas, Matteo Carandini and Kenneth D Harris (2019). 'Distributed coding of choice, action and engagement across the mouse brain'. *Nature* 576.7786, pp. 266–273.

Stone, Mervyn (1960). 'Models for choice-reaction time'. *Psychometrika* 25.3, pp. 251–260.

Stüttgen, Maik C, Nils Kasties, Daniel Lengersdorf, Sarah Starosta, Onur Güntürkün and Frank Jäkel (2013). 'Suboptimal criterion setting in a perceptual choice task with asymmetric reinforcement'. *Behavioural processes* 96, pp. 59–70.

Sugawara, Michiyo and Kentaro Katahira (2021). 'Dissociation between asymmetric value updating and perseverance in human reinforcement learning'. *Scientific reports* 11.1, pp. 1–13.

Sul, Jung Hoon, Hoseok Kim, Namjung Huh, Daeyeol Lee and Min Whan Jung (2010). 'Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making'. *Neuron* 66.3, pp. 449–460.

Tai, Lung-Hao, A Moses Lee, Nora Benavidez, Antonello Bonci and Linda Wilbrecht (2012). 'Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value'. *Nature Neuroscience* 15.9, pp. 1281–1289.

Teichert, Tobias and Vincent P Ferrera (2010). 'Suboptimal integration of reward magnitude and prior reward likelihood in categorical decisions by monkeys'. *Frontiers in Neuroscience* 4, p. 186.

Tervo, Dougal GR, Mikhail Proskurin, Maxim Manakov, Mayank Kabra, Alison Vollmer, Kristin Branson and Alla Y Karpova (2014). 'Behavioral variability through stochastic choice and its gating by anterior cingulate cortex'. *Cell* 159.1, pp. 21–32.

The International Brain Laboratory Aguillon-Rodriguez, Valeria et al. (2021). 'Standardized and reproducible measurement of decision-making in mice'. *eLife* 10. Ed. by Naoshige Uchida and Michael J Frank, e63711.

Thibaut, Florence (2016). 'Basal ganglia play a crucial role in decision making'. *Dialogues in Clinical Neuroscience* 18.1, pp. 3–3.

Thura, David and Paul Cisek (2014). 'Deliberation and commitment in the premotor and primary motor cortex during dynamic decision making'. *Neuron* 81.6, pp. 1401–1416.

Trevino, Mario (2014). 'Stimulus similarity determines the prevalence of behavioral laterality in a visual discrimination task for mice'. *Scientific Reports* 4.1, pp. 1–12.

Treviño, Mario, Ricardo Medina-Coss y León and Belén Haro (2020). 'Adaptive choice biases in mice and humans'. *Frontiers in Behavioral Neuroscience* 14, p. 99.

Tversky, Amos and Daniel Kahneman (1974). 'Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty.' *Science* 185.4157, pp. 1124–1131.

Urai, Anne E., Valeria Aguillon-Rodriguez, Inês C. Laranjeira, Fanny Cazettes, Zachary F. Mainen and Anne K. Churchland and (2021). 'Citric Acid Water as an Alternative to Water Restriction for High-Yield Mouse Behavior'. *eneuro* 8.1, ENEURO.0230–20.2020.

Voorn, Pieter, Louk JMJ Vanderschuren, Henk J Groenewegen, Trevor W Robbins and Cyriel MA Pennartz (2004). 'Putting a spin on the dorsal–ventral divide of the striatum'. *Trends in Neurosciences* 27.8, pp. 468–474.

Waiblinger, Christian, Megan E McDonnell, April R Reedy, Peter Y Borden and Garrett B Stanley (2022). 'Emerging experience-dependent dynamics in primary somatosensory cortex reflect behavioral adaptation'. *Nature communications* 13.1, pp. 1–15.

Wei, Xue-Xin and Alan Stocker (2016). 'A new law defining the relationship between perceptual bias and discrimination threshold'. *Journal of Vision* 16.12, pp. 1112–1112.

Wei, Xue-Xin and Alan A Stocker (2015). 'A Bayesian observer model constrained by efficient coding can explain'anti-Bayesian'percepts'. *Nature neuroscience* 18.10, pp. 1509–1517.

Weijnen, J.A.W.M. (1989). 'Lick sensors as tools in behavioral and neuroscience research'. *Physiology and Behavior* 46.6, pp. 923–928.

Weil, Rimona Sharon, Nicholas Furl, Christian C Ruff, Mkael Symmonds, Guillaume Flandin, Raymond J Dolan, Jon Driver and Geraint Rees (2010). 'Rewarding feedback after correct visual discriminations has both general and specific influences on visual cortex'. *Journal of neurophysiology* 104.3, pp. 1746–1757.

Wenzlaff, Hermine, Markus Bauer, Burkhard Maess and Hauke R Heekeren (2011). 'Neural characterization of the speed–accuracy tradeoff in a perceptual decision-making task'. *Journal of Neuroscience* 31.4, pp. 1254–1266.

Wickelgren, Wayne A (1977). 'Speed-accuracy tradeoff and information processing dynamics'. *Acta psychologica* 41.1, pp. 67–85.

Wilson, Andrew D, James Tresilian and Friederike Schlaghecken (2011). 'The masked priming toolbox: an open-source MATLAB toolbox for masked priming researchers'. *Behavior Research Methods* 43.1, pp. 210–214.

Wolfe, Jeremy M (1984). 'Short test flashes produce large tilt aftereffects'. *Vision research* 24.12, pp. 1959–1964.

Yang, Tianming and John HR Maunsell (2004). 'The effect of perceptual learning on neuronal responses in monkey visual area V4'. *Journal of Neuroscience* 24.7, pp. 1617–1626.

Yartsev, Michael M, Timothy D Hanks, Alice Misun Yoon and Carlos D Brody (2018). 'Causal contribution and dynamical encoding in the striatum during evidence accumulation'. *Elife* 7, e34929.

Yatsenko, Dimitri, Edgar Y Walker and Andreas S Tolias (2018). 'DataJoint: a simpler relational data model'. *arXiv preprint arXiv:1807.11104*.

Yerkes, Robert M. and John D. Dodson (1908). 'The relation of strength of stimulus to rapidity of habit-formation'. *Journal of Comparative Neurology and Psychology* 18.5, pp. 459–482.

Zador, Anthony M (2019). 'A critique of pure learning and what artificial neural networks can learn from animal brains'. *Nature Communications* 10.1, pp. 1–7.

Zoest, Wieske van and Amelia R Hunt (2011). 'Saccadic eye movements and perceptual judgments reveal a shared visual representation that is increasingly accurate over time'. *Vision Research* 51.1, pp. 111–119.

# ACKNOWLEDGEMENTS

I am grateful to my supervisors, Prof. Dr. Laura Busse and Prof. Dr. Thomas Wachtler. Thank you, Laura and Thomas, for having me in the lab and for your guidance, constructive feedback, support, and motivation throughout the past years.

I would also like to thank Dr. Steffen Katzner. Thank you, Steffen, for always being eager to problem-solve, whether it was setup-related or behavior-related.

Thank you to PD Dr. Kay Thurley and Prof. Dr. Zhuanghua Shi for your insightful suggestions during TAC meetings and for your help in shaping this project.

Thank you to each and every member of the Busse lab for making this such a great learning experience and for your friendship. Thank you for all the feedback, suggestions, and moral support. In particular, thank you, Magda, whether it was anything work-related or personal, thanks for always being there. Thank you, Melanie and Ann for all the coffee breaks, dinner evenings and your constant support.

Thank you, Diana, for your friendship, encouragement and your humor and thank you, Dieter, Steffi and Babsi for making our office full of laughter.

Furthermore, I would like to convey my gratitude to the Graduate School of Systemic Neurosciences for providing me with resources. I was financially supported by RTG 2175 Perception in Context and its Neural Basis as well as by LMU's Graduate School of Systemic Neurosciences.

I would not be here without the constant love and support from my family. Thank you, Bua and Thulobua. I am sure you would have been

proud of me. Thank you, Ama, Baba, and Sarda, for always being my safe place. Thank you, Amsul, I wonder if I would have been here if it weren't for the siblings' ego.

Finally, a big thank you to Mogli. Thank you for being my greatest critic and an even greater cheerleader.