



# **Navigating Uncertainty: Neural Mechanisms of Probabilistic Foraging in Mongolian Gerbils**

Thesis submitted to the Faculty of Natural Sciences of Otto-von-Guericke University Magdeburg for the degree of

**Master of Science (Integrative Neuroscience)**

on

06<sup>th</sup> February 2024

by

**Vishal Kannan**

Matriculation No. 240136

Thesis Supervisor:

**Prof. Dr. Max Happel**

# Acknowledgements

There are several individuals I need to acknowledge, whose support was pivotal in completing this thesis.

First and foremost, my deepest gratitude goes to my scientific advisor, Prof. Dr. Max Happel, for his invaluable guidance and advice on the subject matter. I will always cherish not only the honest scientific discussions that we have had that has enhanced my curiosity in neuroscience but also life advice, whenever I needed it. Words fall short of expressing my gratitude and having him as a mentor has been an inspirational journey through my academic career in Neuroscience.

Special thanks are also due to my colleague and mentor, Parthiban Saravanakumar, for enabling this thesis through the provision of high-quality data and training me with the behavioral paradigm. The extensive discussions and brainstorming sessions about data analysis were crucial in refining the data analysis pipeline. His mentorship, dedication, and effort are deeply appreciated. Additionally, my gratitude extends to the Technical Assistants of the Systems Physiology Department at LIN for their assistance with histology data, and to our secretary, Ms. Beate Traore, for her promptness in addressing queries and efficiently managing all administrative tasks, allowing me to focus more on my thesis.

The supportive environment within the lab played a significant role in the realization of this thesis. I am thankful to Ebru Acun and Katrina Deane for their support and encouragement during my initial days in the lab.

My thanks also go to Prof. Dr. Michael Brosch and Prof. Dr. Catherine Sweeney Reed for their technical and scientific guidance, enriching my understanding of neural data analysis beyond the scope of my master's thesis. Dr. Michael Katsillis deserves special mention for his inspirational teaching and in-depth conversations on statistics in neuroscience, which were immensely beneficial.

Finally, I owe a debt of gratitude to my mother and closest friends, especially Meera Ramadas, for their constant presence to share in my triumphs and support me through my challenges. Their support has been invaluable and has undeniably contributed to reaching this milestone.

# Index of Notation

ACC	– Anterior Cingulate Cortex
ANOVA	– Analysis of Variance
aPFC	– Anterior Prefrontal Cortex
AVREC	– Average Rectified Signal
CSD	– Current Source Density
DAC	– Digital to Analog Converter
dIPFC	– Dorsolateral Prefrontal Cortex
FPC	– Frontopolar Cortex
Fra	– Frontal Region A
GUT	– Giving Up Time
Hz	– Hertz
IFJ	– Inferior Frontal Junction
IR	– Infrared
kHz	– Kilo Hertz
LFP	– Local Field Potential
MATLAB	– Matrix Laboratory
mm	– Millimeter
mPFC	– Medial Prefrontal Cortex
ms	– Millisecond
MVT	– Marginal Value Theorem
OFC	– Orbitofrontal Cortex
PFC	– Prefrontal Cortex
RMS	– Root Mean Square
s	– Second
V	– Volt
VTA	– Ventral Tegmental Area
$\mu\text{m}$	– Micrometer

## Table of Contents

Acknowledgements .....	i
Index of Notation .....	ii
List of Figures .....	v
Abstract .....	ix
1 Introduction.....	2
1.1 Foraging behaviour and attentional resource allocation in the brain .....	2
1.1.1 Prefrontal cortex and exploratory decision-making: Insights from human and primate studies .....	3
1.2 Role of prefrontal cortex in exploratory behaviour in rodents .....	4
1.3 Mongolian Gerbils as a rodent model for exploratory behaviour .....	5
1.4 Probabilistic foraging as a tool to induce the exploration-exploitation dilemma .....	5
1.5 Layer-dependent processing in decision making: Insights from chronic laminar recordings and current source density analysis in rodents .....	6
2 Objectives and implications .....	9
3 Materials and methods .....	11
3.1 Neural recordings from frontal region A (FrA) .....	11
3.1.1 Experimental setup .....	12
3.2 Probabilistic foraging paradigm .....	13
3.3 Data analyses.....	15
3.3.1 Data storage and analysis pipeline .....	15
3.3.2 Analysis of behavioural data .....	16
3.3.3 Pre-processing of the neural data.....	17
3.3.4 Current source density (CSD) analysis.....	17
3.3.5 Feature extraction of CSD recordings.....	18
3.3.6 Statistics .....	20
4 Results .....	22
4.1 Behavioural analysis .....	22
4.1.1 Performance of the foraging behaviour.....	22

4.2	Distinct spatiotemporal activity patterns in the frontal field A.....	26
4.2.1	Overall frontal activity patterns in FrA .....	27
4.3	Shifts in frontal activity patterns: From exploitation to exploration .....	29
4.3.1	The evolution of frontal activity from exploitation to exploration.....	30
4.4	Layer specific spatiotemporal activity patterns in the frontal field A.....	31
1.1.1	Histological confirmation of laminar electrode positioning in frontal region A (FrA) .....	31
4.4.1	Distinct layer-specific shift in frontal activity patterns: Exploitation to exploration.....	32
5	Discussion .....	35
5.1	Inference-bound decision-making in gerbils .....	36
5.2	Neural encoding of reward expectation and evaluation in FrA .....	37
5.3	The dynamic transition of expectation is reflected in FrA activity. ....	38
5.4	Functional implications of layer-specific activity during foraging .....	39
5.5	Limitations and future considerations .....	41
6	Conclusion.....	43
7	Bibliography.....	44
8	Appendix .....	52
8.1	Individual CSD profiles and corresponding AVRECs .....	52

## List of Figures

**Figure 1: Schematic representation of the behavioural setup and behavioural paradigm.** **A** – The foraging box (37cm x 26cm x 48cm) containing two spouts on the right (orange) and left (green) separated by 36 cm. The animal is placed in the middle and the head connector is attached with the pre-amplifier of the data acquisition system. The animal freely moves within the box while the LFP signals are recorded simultaneously. **B** – Schematic representation of probabilistic foraging paradigm performed by the animal showing the inter-poke interval, trial duration and travel time. **C** – The exponential decay of reward probabilities for three different starting probabilities (Lottem et al., 2018). **D** – Timeline of the whole experiment from surgery to analysis..... 12

**Figure 2: Schematic representation of the data analysis pipeline.** The raw laminar local field potentials are pre-processed for artefact correction and channel rectification. The pre-processed LFP is then transformed into its respective current source profile by applying a second spatial derivative. Based on the activity profile and electrode depth information from histology, channel layer specifications are performed. Finally, the signals from current source profiles are rectified and averaged across the channels to obtain the overall cortical activity..... 15

**Figure 3: Chronic current source density (CSD) analysis in frontal field a (FrA).**  
**A)** The left panel displays in vivo multichannel local field potential (LFP) recordings obtained from a 32-channel silicon probe implanted perpendicularly in the FrA of awake, behaving gerbils. The probe captures activity across all cortical layers (I–VI), with t=0 corresponding to the end of a poke (black dashed line). The middle panel illustrates the CSD, showing task evoked CSD components appeared as current sink (in blue) and source (in red) activity. Channel-layer specification is informed by initial response to reward-related CSD components appearing 100 ms post-poke (highlighted in the red box), typically marking the infragranular layer V. The right panel presents a simplified schematic illustration of the cortical column in FrA, delineating the layered structure and the direction of reward/prediction error signals originating from the ventral tegmental area (VTA). This schematic aids in visualizing the depth and functional organization of the cortical layers as identified through the CSD analysis. **B)** A schematic representation of an electrode in a cortical column showing how the

current sink (blue)/ source (red) are interpreted with respect to the position of the synapses and cell bodies..... 19

**Figure 4: Behavioural performance of the probabilistic foraging task.** **A)** Schematic representation of the foraging task, illustrating the phases of exploitation and transition to exploration with the corresponding nose poke outcomes. **B)** Relative body weight percentage of individual animals over 20 sessions, with thresholds (red dashed lines) indicating the baseline (100%) and critical weight loss limits (85%). **C)** Boxplot of travel times for each animal across sessions, highlighting the learning curve and stabilization of task performance. **D)** Average median inter-poke intervals following rewarded and unrewarded pokes. **E)** Bar graph of mean residence times at spouts with different starting reward probabilities. **F)** Mean number of rewards obtained per trial type differentiated by different starting reward probabilities. **G)** Mean number of unrewarded pokes obtained per trial type differentiated by different starting reward probabilities. Data are represented as mean  $\pm$  SEM, with statistical significance denoted by asterisks (One way ANOVA with Bonferroni Correction, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ). ..... 24

**Figure 5: Behavioural patterns in spout-switching decision-making.** **A)** The distribution of trials relative to the number of rewarded pokes, delineated by different initial reward probabilities ( $A_1 = 1$ ,  $A_2 = 0.75$ ,  $A_3 = 0.5$ , and overall). **B)** The consistency of gerbil's spout-leaving behavior as shown by the fraction of trials against the number of consecutive unrewarded pokes prior to switching, irrespective of the initial reward probability. ..... 26

**Figure 6: Grand averaged current source density (CSD) profiles (n=5).** Distinct motor and reward related spatiotemporal neural activity in frontal field A. The selected epochs represent -1 to +2 seconds from the end of the poke (black dashed line,  $t=0$ ). The selected time interval was taken for four different events (pokes) and its corresponding consequence (reward): (top left) first poke with reward, (top right) first poke without reward, (bottom left) last rewarded poke, and (bottom left) last poke without reward. ..... 27

**Figure 7: Grand averaged average rectified signals (gaAVREC) (n=5).** AVREC displays the overall frontal cortical activity, revealing distinct motor and reward-related signals. The mean average rectified waveform (depicted in bold colours) together with its standard error (shown in lighter shades) is plotted for the selected time intervals

(epochs). These epochs span from one second before to two seconds after the end of the poke (t=0). AVRECs are presented for four distinct pokes: first poke with reward (top left), first poke without reward (top right), last rewarded poke (bottom left), and last poke without reward (bottom right). ..... 28

**Figure 8: Frontal activity change from exploitation to exploration. A)** An illustrative trial consisting of a sequence of pokes, with emphasis on the first unrewarded poke after the last rewarded poke (brown) and the last unrewarded poke (green). **B)** Distinct activation patterns in grand averaged AVREC for the first unrewarded poke after the last rewarded poke (brown) and the last unrewarded poke (green). From the grand AVREC data (B), two distinct time intervals (epochs) were identified for RMS computation: the early phase (0 – 100 ms, yellow), and the late phase (100 – 500 ms, light pink). **C)** The AVREC – RMS Z scores for unrewarded pokes between the last rewarded poke and the last poke (n) before the animal disengages from the current spout are shown (here, a scenario of 7 consecutive unrewarded pokes is depicted, where n-7 is the first unrewarded poke after the last rewarded poke (brown) and n represents the last unrewarded poke (green)). One-way ANOVA with Bonferroni correction was applied to detect differences between the pokes (\* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001). ..... 30

**Figure 9: Sample histology image of the frontal field A (FrA).** Sample histology image from Frontal field A (FrA), taken 4.85 mm anterior to the Bregma. The electrode track is evident at 1.5 mm lateral to the Bregma, reaching a cortical depth of approximately 1.29 mm. ..... 32

**Figure 10: Layer specific cortical activity in frontal region A (FrA). A)** Grand averaged rectified sink activity (n=5) is displayed for all identified cortical layers from the grand CSD profile (Figure 3). Epochs shown span from -1 to +2 seconds from the end of the poke (t = 0), focusing on the first unrewarded poke following the last rewarded poke (illustrated in brown) and the last unrewarded poke (illustrated in green). Two distinct time intervals (epochs) were selected for RMS computation based on the averaged rectified sinks: the early phase (0 – 100 ms, marked in yellow), and the late phase (100 – 500 ms, marked in light pink). **B)** Layer-wise Z scores for the RMS, derived from the average rectified sinks, quantify activity for unrewarded pokes ranging from the last rewarded poke to the last poke before the animal disengages from the current spout. As with Figure 5B, this illustration depicts a sequence of 7

consecutive unrewarded pokes, where  $n-7$  denotes the first unrewarded poke after the last rewarded poke, and  $n$  signifies the last unrewarded poke. ..... 33

## Abstract

This thesis investigates the neural underpinnings of decision-making in uncertain environments, using Mongolian gerbils as a rodent model. Specifically, it examines the role of the anterior frontal cortex—a key region in human decision-making—in managing attentional resources during an exploration-exploitation dilemma. The study employed a probabilistic foraging task alongside chronic laminar recordings in gerbil's frontal region A (FrA). This approach allowed an in-depth study of the FrA's mesoscale activity as the gerbils made decisions in environments simulating real-world uncertainties.

Our findings revealed that gerbils engage in sophisticated, inference-based decision-making strategies rather than following rigid foraging rules. Current Source Density (CSD) profiles within the FrA indicated a complex cognitive integration of past experiences with immediate action-outcome assessments, through encoding of reward anticipation and evaluation. Importantly, an increase in pre-decisional neural activity in the FrA, specifically observed just before the animals decide to explore, underscores its role in exploratory behaviour. Additionally, layer-specific analysis in the FrA showed enhanced engagement of supragranular layers during the crucial transition from exploitation to exploration, hinting at a layer-dependent mechanism within FrA that enables adapting adequate foraging strategies. Overall, these results contribute to our understanding of decision-making in uncertain environments, emphasizing the rodent anterior frontal cortex's role in attentional resource management for adaptive decision making.

Exploring these neural mechanisms within the exploration-exploitation dilemma—a critical aspect of cognitive functioning in uncertain environments— not only enriches our understanding of decision-making in rodents but also serves as a crucial model for investigating analogous processes in more complex beings, including humans. Such insights have the potential to inspire new approaches for treating decision-making disorders and enhancing cognitive adaptability across a spectrum of neurological and psychiatric conditions.

**Keywords:** Decision-making, exploration-exploitation dilemma, Mongolian gerbils, probabilistic foraging, frontal region A, chronic laminar recordings.

# **INTRODUCTION**

# 1 Introduction

"Should I take the familiar route to work or try a new path today? Should I choose for a new dish in the menu or go for my usual one?" Everyday dilemmas like this are an example of decision-making under uncertainty. In such scenarios, individuals must constantly weigh the known against the unknown, making choices that can range from trivial to life-altering. This process becomes more important when the environment is dynamic where one's adaptability and decision-making flexibility become key survival tools. One classic illustration of this is the exploration-exploitation dilemma, which poses a fundamental question: in an uncertain environment, how does one balance the choice between exploiting a known resource and exploring potential new ones? Exploration is a risk-taking behaviour that may come with a cost. However, without exploration, we will persevere with same strategies and miss the opportunity to find a better option. From the largest elephants roaming the African savannas to the smallest ants foraging in a rainforest, across a vast spectrum of species, the need to balance exploration and exploitation is a universal aspect of survival. While this exploratory/risk-taking behaviour have been extensively studied in various species including the humans, the neural mechanisms driving such intricate decision-making remain less explored. This thesis aims to address this gap by delving into the neural underpinnings of decision-making in uncertain situations, particularly focusing on the exploration-exploitation dynamics through the lens of animal behavior.

## 1.1 Foraging behaviour and attentional resource allocation in the brain

Foraging is a fundamental behavior seen in many species, where individuals search for food in their environment. It inherently involves a critical decision-making process where an individual must choose between staying with a known resource (exploitation) and venturing out in search of potentially better options (exploration). This balance, fundamental in foraging behavior, becomes a complex problem of resource allocation within the brain. When an individual is faced with the choice of staying with a known resource or exploring new options, the brain must allocate attention to process information about the current situation, assess potential risks and rewards, and decide based on this evaluation. This process involves dynamic shifts in attentional focus, from assessing the current resource's value to considering the unknown possibilities of unexplored options. The decision-making process thus becomes a cognitive

balancing act, where the allocation of attentional resources plays a crucial role in optimizing behavior to adapt to changing environmental conditions.

Many studies have focused on understanding the resource allocation in brain during various cognitive tasks. For instance, human studies on gambling (Beharelle et al., 2015; Daw et al., 2006), attentional shifts to environmental changes (Chetverikov et al., 2017; Muller et al., 1995), and visual search patterns (Kristjánsson et al., 2014; Wolfe et al., 2019) reveal the brain's efficiency in allocating attentional resources by adapting to subtle reward changes and environmental shifts, paralleling animal foraging behaviors.

### **1.1.1 Prefrontal cortex and exploratory decision-making: Insights from human and primate studies**

Numerous studies have established the crucial role of the prefrontal cortex in decision making across species (Bechara, 2005; Bechara et al., 2003). In humans, the anterior prefrontal cortex (aPFC), particularly the frontopolar cortex (FPC), plays a decisive role in decision-making, significantly influencing exploratory resource allocation. Notably, the FPC activates during exploratory decisions in contexts such as gambling, indicating its significant role in choosing to switch to alternative actions (Boorman et al., 2009; Daw et al., 2006). Additionally, anodal, and cathodal transcranial direct current stimulation experiments involving the FPC have demonstrated a direct and polarizing impact on exploratory behavior, highlighting its potential causative relationship (Beharelle et al., 2015).

In non-human primate studies, disruptions such as lesions in the aPFC have been linked to increased persistence in following task rules, suggesting its crucial role in shifting focus from current tasks to exploring alternative reward sources (Mansouri et al., 2015). This distinct causative role of the aPFC sets it apart from more posterior regions of the lateral prefrontal cortex, such as the inferior frontal junction area (IFJ), which primarily activates only when a task switch occurs. Furthermore, the coactivation of the frontopolar area (FP1) and IFJ observed during exploratory behaviour from functional imaging studies suggests a modulatory role of aPFC on IFJ to achieve exploratory attention shifts (Bludau et al., 2014). This functional evidence backed by the FP1's complex dendritic spine system (high spine density but low cell body density) that offers an excellent architecture for integrating different inputs (Jacobs et al., 2001; Ramnani & Owen, 2004) suggests an underlying neurocognitive circuit suitable for the

comparison of current task aspects with novel information required for exploratory decisions.

## **1.2 Role of prefrontal cortex in exploratory behaviour in rodents**

In rodents, although exploratory behaviours are observed (Lottem et al., 2018; Vertechi et al., 2020), the specific role of their prefrontal cortex (PFC), particularly in terms of attentional resource allocation, remains less clear. Studies have shown the medial prefrontal cortex (mPFC) to be central to decision-making processes, particularly in scenarios involving risk assessment and environmental adaptation (Orsini et al., 2020). For instance, sudden increase in mPFC neural activity was correlated to environmental changes during a tone-cued behavioural task in rats (Karlsson et al., 2012). Moreover, enhanced neuronal activity in the mPFC has been observed in mice during social approach behaviours, suggesting its active role in processing social cues, an essential requirement for exploratory behaviour (Lee et al., 2016). Additionally, accumulating evidence demonstrated that the mPFC plays an instrumental role in encoding various forms of decision-making including exploratory behaviour mainly due its strong projections from CA1, thalamus and amygdala that helps to integrate inputs from different regions and update its information about the surroundings (Bechara et al., 2003; Herry et al., 1999; Onge et al., 2012; Tang et al., 2021).

While the medial prefrontal cortex (mPFC) has been a primary focus in rodent decision-making research, controversies have emerged due to inconsistencies in nomenclature and anatomical definitions between rodent and primate studies of the prefrontal cortex. Laubach et al., 2018 emphasize that the varied use of terms, often borrowed or adapted from primate research, has led to a fragmented understanding of the PFC's role in rodents. Unlike in primates where the PFC is well-defined and its functions extensively studied, in rodents, the boundaries of this region are less clear, and its functions less distinct. The lack of a standardized approach in defining and researching the rodent PFC makes it challenging to draw direct comparisons with primate studies, thereby complicating the integration of rodent research into a broader neuroscientific context. This ambiguity in defining the boundaries of rodent PFC, encompassing areas like the medial prefrontal cortex (mPFC), orbitofrontal cortex (OFC), and anterior cingulate cortex (ACC), poses significant challenges in understanding its full functional spectrum, particularly in higher cognitive processes such as adaptive decision-making.

### **1.3 Mongolian Gerbils as a rodent model for exploratory behaviour**

In this study we used Mongolian gerbils as a rodent model to investigate the neural mechanisms behind exploratory resource allocation. The Mongolian gerbil is a popular animal model with a long history in multiple research topics such as animal's social cognition (Tchabovsky et al., 2019), neurological diseases – epilepsy (Cutler et al., n.d.), auditory processing (Happel et al., 2010, 2014) and hearing loss (Ohl et al., 1999; Otto & Jrge, 2012). Previous research has proved that gerbils are able to quickly learn experimental rules, adapt to environmental changes and learn complex auditory behavioural tasks such as reversal learning showcasing higher cognitive functions such as adaptive decision-making (Jarvers et al., 2016). Moreover, in the wild, the gerbils live in deserts making them natural foragers in one of the most dynamic environments in the world. Considering these adaptive behaviours of the gerbils, in this study, we decided to use them as a rodent model to investigate the neural underpinnings of decision making during an exploration-exploitation dilemma.

Furthermore, the gerbil atlas created by Radtke-Schuller et al., 2016 extensively maps their cortical boundaries and connections enabling researchers to develop techniques for chronic recordings and optogenetic manipulation on these animals (Brunk et al., 2019; Zempeltzi et al., 2020). Regarding prefrontal cortex, the atlas avoids the usage of common nomenclature such as mPFC or dorsolateral PFC. Instead, it just maps the anterior region of frontal cortex as frontal region A (FrA) that lies between the olfactory bulb and the secondary motor cortex. This clear anatomical distinction unlike in mice and rats forms a better analogue of human and monkey aPFC, whose role is significant in exploratory decision making.

### **1.4 Probabilistic foraging as a tool to induce the exploration-exploitation dilemma**

To investigate decision-making under uncertainty in Mongolian gerbils, we adapted the probabilistic foraging task from a study by Lottem et al., 2018, originally conducted with mice. This study presents a sophisticated setup to simulate an exploration-exploitation dilemma by incorporating probabilistic rules that govern the availability of rewards. In this task, mice were trained in a controlled environment featuring an elongated chamber with two water-reward ports at each end. In each trial, the mice performed a sequence of nose-pokes at one of the reward ports, with each nose-poke having a decreasing probability of yielding a water reward. In this task, mice had to make a

choice with each nose-poke: whether to continue exploiting the current water source, which offered diminishing rewards, or to venture out to explore another port that might provide greater rewards. This setup created a dynamic environment, requiring the animals to constantly evaluate their options. They needed to balance the benefits of a familiar but less rewarding resource against the potential gains from unexplored alternatives. The task was structured into three trial types, each defined by different starting reward probabilities. The introduction of varying initial reward probabilities further adds layers of uncertainty, compelling the animals to make decisions in an environment where the chances of reward are dynamic and unpredictable.

This paradigm effectively models real-life foraging scenarios and offers an excellent experimental framework to explore how rodents navigate the balance between exploitation and exploration under uncertainty. By adapting this task for Mongolian gerbils and introducing simultaneous chronic recordings, we closely investigated the neural mechanisms of attentional resource allocation in the frontal cortex during an exploration-exploitation dilemma.

### **1.5 Layer-dependent processing in decision making: Insights from chronic laminar recordings and current source density analysis in rodents**

The integration of chronic laminar electrophysiology recordings and current source density (CSD) analysis forms a cornerstone of our study in exploring decision-making processes in the frontal region A (FrA) of Mongolian gerbils. Chronic recordings, with their ability to capture long-term neural activity in awake, behaving animals, are instrumental in revealing the intricate dynamics of brain function. This approach, successfully employed in previous gerbil studies (Deane et al., 2020; Happel et al., 2014; Zempeltzi et al., 2020), allows for a mesoscale examination of population activity across cortical layers, providing a comprehensive view of neural circuit processes during perception, learning and cognitive tasks.

The CSD analysis is an approach used to approximate the location and magnitude of current sources and sinks within brain tissue, inferred from local field potential (LFP) recordings. CSD transformation of LFPs is reference free and thereby less affected from referencing artefacts and far-field potentials. CSD profile from laminar recordings is a refined measure that identifies regions of synaptic input (sinks) and output (sources), thus providing a detailed map of electrical current flow through the cortical layers, which is crucial for understanding neuronal circuitry at a mesoscopic scale.

They represent synaptic population activities with high spatial and temporal precision, uncovering the functional micro-circuitry of the cortex (Happel et al., 2010). It transcends the limitations of single-unit and calcium-imaging techniques by providing a mesoscopic understanding of brain function, essential for interpreting complex cognitive processes (Buzsáki et al., 2012; Douglas & Martin, 2007; Godlove et al., 2014).

Furthermore, CSD analysis is effective in revealing task-dependent differences in cortical activation patterns across layers from laminar recordings (Deane et al., 2020; Happel et al., 2014; Zempeltzi et al., 2020), highlighting the importance of layer-dependent processing in cognitive processes. Studies have shown that layers in the primary auditory cortex (A1) of mice exhibit task-related modulations (Francis et al., 2018), and similar differentiations in encoding task- and choice-related information are observed in the gerbil A1 (Zempeltzi et al., 2020). Accumulating evidence show that cortical layers encode relative value representations of stimulus, mirroring the principles observed in reinforcement learning models. Dopamine, a key neurotransmitter in reinforcement learning, modulates neural activity across cortical layers, influencing how information is integrated and processed for decision-making. For instance, optogenetic manipulation of VTA dopamine neurons showed reward-related modulation of auditory signal processing in the auditory cortex via a gain modulation of thalamic inputs in infragranular layers Vb/VI (Brunk et al., 2019).

As translaminar processing principles in the frontal cortex share certain common principles with the canonical cortical activation patterns (Douglas & Martin, 2007; Godlove et al., 2014), similar dopaminergic modulation of frontal layer-specific processing modes may set ground for the neural resource to code the salient representation of behaviourally relevant stimuli (Brunk et al., 2019; Happel, 2016). Therefore, in this study, we utilize the capabilities of CSD analysis from chronic laminar recordings to unravel the layer-specific activity patterns within the FrA underlying the decision-making during the probabilistic foraging task. This approach enables us to dissect the intricacies of neural processes as gerbils navigate the exploration-exploitation dilemma, providing insights into the fundamental neural mechanisms that underlie decision-making under uncertainty.

# **OBJECTIVE**

## **2 Objectives and implications**

This study delves into the complexities of decision-making under uncertainty in Mongolian gerbils. Particularly, it aims to uncover the neural underpinnings of attentional resource allocation during the exploration-exploitation dilemma. To this end, the research will utilize a probabilistic foraging task in conjunction with chronic laminar recordings from the frontal region A (FrA), an analogue of frontopolar cortex, recognized for its significance in human decision-making. This approach allows for an in-depth analysis of the gerbil's decision-making processes in an environment that simulates real-world uncertainties. The primary objectives of this study are:

1. To analyze the foraging behavior of gerbils during the probabilistic foraging task, particularly focusing on their spout-leaving behaviour in response to uncertain and changing reward probabilities.
2. To investigate the role of the FrA, examining how it influences the dynamics of decision boundaries in response to varying reward probabilities.
3. To examine the layer-dependent processing within the FrA, exploring how different cortical layers adaptively contribute to the selecting adequate search strategies under varying conditions of resource availability.

Achieving these objectives not only sheds light on the intricacies of rodent cognition but also holds significant implications for human decision-making research and clinical neuroscience. The capacity to navigate between exploring new possibilities and exploiting known resources is not only fundamental to animal survival, but also a key aspect in human cognitive processes across various contexts. Furthermore, impairments in decision-making and cognitive flexibility are hallmark features of numerous neurological and psychiatric disorders, such as schizophrenia, obsessive-compulsive disorder, and addiction. Insights gained from this research could therefore illuminate potential pathways for novel therapeutic approaches, enhancing our ability to address, manage, and potentially alleviate such conditions. Ultimately, this study bridges the gap between animals and human cognitive functions, promising to enhance the diagnosis and treatment of disorders characterized by decision-making impairments.

## **MATERIALS AND METHODS**

### **3 Materials and methods**

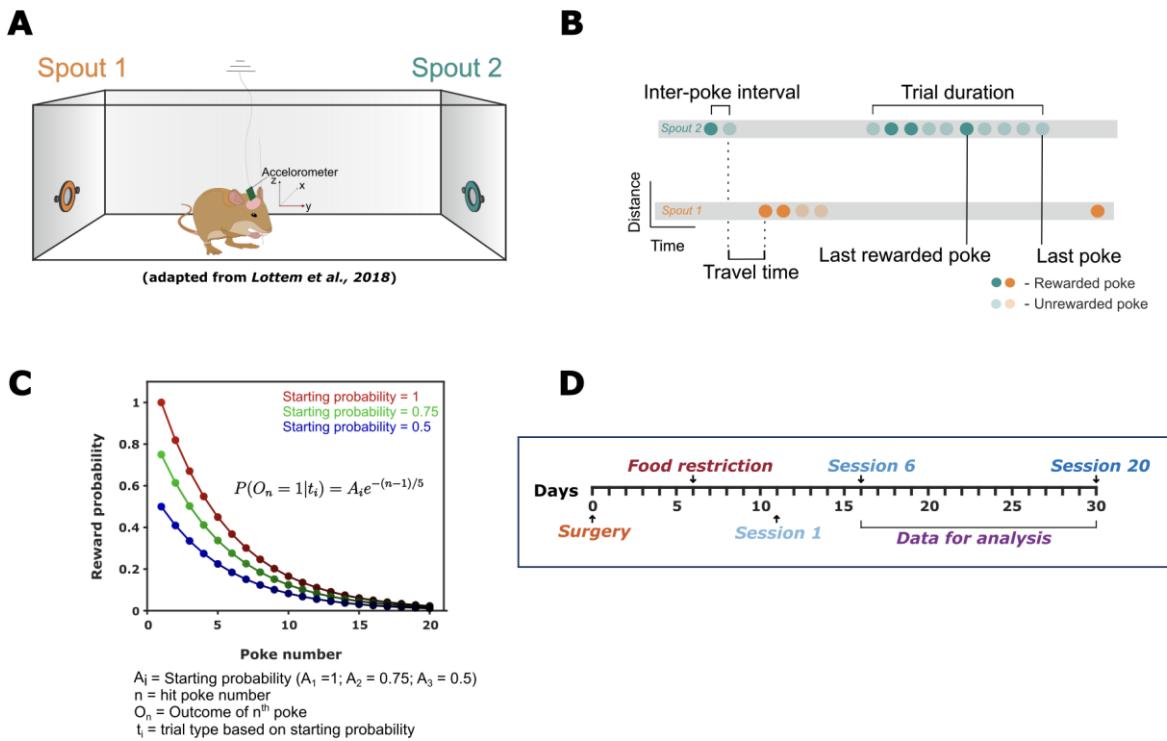
In this study, I performed analysis on current source density (CSD) profiles derived from chronic laminar local field potential (LFP) recordings that were previously collected from the frontal region A (FrA) of awake, behaving Mongolian gerbils (*Meriones unguiculatus*). The primary goal of this analysis was to elucidate the layer-specific, spatiotemporal population activity within the FrA on a mesoscopic scale. The animals (n=5) were made to perform a probabilistic foraging task designed to test their decision-making strategies while they exploit a known food source or to explore a new option. The analysis of the continuous foraging sessions, along with the electrophysiological recordings from the FrA, allowed us to explore the neural mechanisms underlying decision-making during the exploitation/exploration dilemma.

#### **3.1 Neural recordings from frontal region A (FrA)**

The chronic in vivo electrophysiological data used in this study were obtained by a fellow researcher using a 32-channel multilayer electrode (Neuronexus, A1x32-6mm-50-177\_H32\_21mm) while gerbils performed the probabilistic foraging task. The electrode was implanted into FrA, positioned 4.65 mm anterior to Bregma and 1.5 mm lateral to lambda. A comprehensive behavioural screening for epileptic seizures—a known genetic trait in gerbils—was conducted prior to the surgeries, utilizing a protocol developed by Gonzalo Arias Gil and Dr. Kentaroh Takagaki at the SPL Department – LIN, following the guidelines established by Seto-Ohshima et al., 1992. Only the animals that did not show epileptic seizure during the screening test were included for the study.

It is important to note that neither the surgical implantation of electrodes nor the recording the electrophysiological data form the central focus of the current thesis. The focus herein lies strictly on the post-hoc analysis of the neural recordings, and the surgical and recording procedures were completed independently by another researcher to ensure the quality and reliability of the data upon which this study is based.

### 3.1.1 Experimental setup



**Figure 1: Schematic representation of the behavioural setup and behavioural paradigm.** **A** – The foraging box (37cm x 26cm x 48cm) containing two spouts on the right (orange) and left (green) separated by 36 cm. The animal is placed in the middle and the head connector is attached with the pre-amplifier of the data acquisition system. The animal freely moves within the box while the LFP signals are recorded simultaneously. **B** – Schematic representation of probabilistic foraging paradigm performed by the animal showing the inter-poke interval, trial duration and travel time. **C** – The exponential decay of reward probabilities for three different starting probabilities (*Lottem et al., 2018*). **D** – Timeline of the whole experiment from surgery to analysis.

The foraging box (37cm x 26cm x 48cm) was placed in a chamber that is electrically and acoustically shielded. It contained two spouts on the right and left side separated by 36 cm and each spout had an infrared (IR) emitter/sensor pair on the sides to detect the nose poke (Fig.1). Each spout was attached to a food dispenser (Campden Instruments Ltd., USA) placed outside the foraging cage.

Once a poke has been detected by the IR sensor, the signal is communicated to an external Arduino device which converts it into a digital signal. This digital signal indicating a poke registration is communicated to the computer through a MATLAB (MathWorks, R2020b) interface. Consequently, the starting reward probability and the following probabilities of reward was generated in MATLAB according to Eq.1. The starting reward probability was randomly selected from the three different possibilities (1, 0.75, 0.5, Fig.1C). The generated digital reward outcome (1 – reward; 0 – no

reward) based on the reward probability was converted into an analog signal by a DAC (Arduino) and communicated to the commutator which provides the food pellet into the spout. The whole Arduino-MATLAB interface was performed using a custom Arduino and MATLAB script. There were two video cameras (Microsoft LifeCam HD-3000, top and side) to track the real-time behaviour of the animal inside the cage. The video recordings were captured using OBS 25.0.8 software.

Multichannel electrophysiology recordings were performed after connecting the head connector of the implanted electrode to the preamplifier (20-fold gain, HST/32V-G20, INTAN Technologies) which in turn is connected to a data acquisition system (INTAN Technologies). The electric cable was covered by a metal mesh for bite protection. Tension of the cable was relieved by a spring and a commutator that allows rotation and free movement of the animal inside the cage. Broadband LFPs were acquired using a hardware filter (0.1 Hz – 12 kHz), sampled at 30 kHz and digitally filtered with a maximum cut off frequency of 150 Hz. Proper grounding of the animal through its common ground was ensured to avoid ground loops between recording system, foraging cage and the animal.

### **3.2 Probabilistic foraging paradigm**

The probabilistic foraging task was adapted from (Lottem et al., 2018) in which the gerbils learn to do a nose poking behaviour to obtain rewards (Fig.1). Every foraging session consists of N trials, with each trial comprising a sequence of nose pokes (Fig.1B). A trial was defined as the period where the animal made a series of nose pokes on the same spout until it switched to the other side. Each individual nose poke has a probability of being rewarded with a 20 mg food pellet. After each reward, a dead time of 100 ms occurred. The reward probability for consequent pokes within a trial decreased exponentially forcing the gerbil to alternate between the spouts, thereby introducing the exploitation-exploration dilemma (Fig.1C). Only the pokes that lasted for at least 100 ms were assigned as hit pokes and followed the reward probability rule. The error pokes (poke duration < 100 ms) were unrewarded.

To instil and maintain the motivation of the animals to perform the task, they were food deprived. Animal body weights were constantly monitored after every session to ensure that they didn't drop beyond 15% of their original body weight measured before the start of the experiment (baseline). At times where the animal didn't perform well

(less than average number of trials) consecutively for two to three sessions, the amount of food restricted was increased to increase their motivation. However, it was ensured that the weight didn't drop beyond the critical limit (85% of the baseline weight).

Three different reward starting probabilities were used that followed the exponentially decreasing trend according to equation (1).

$$P(O_n = 1|t_i) = A_i e^{-(n-1)/5} \quad (1)$$

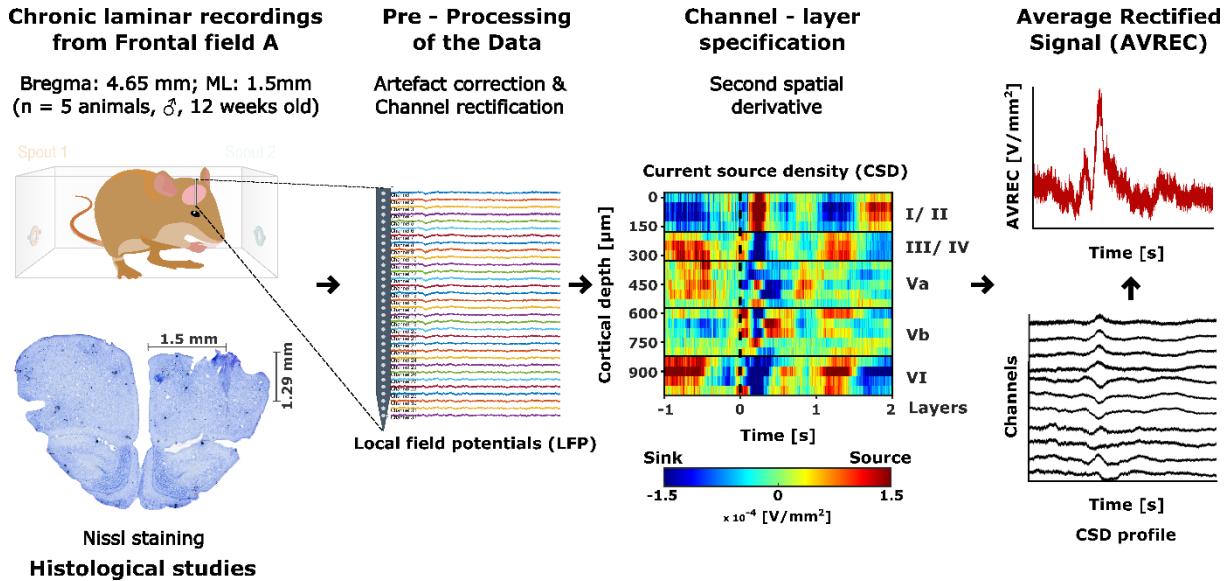
Where  $t_i$  is the  $i^{\text{th}}$  trial type ( $i = 1, 2, 3$ ) corresponding to different exponential scaling factors (starting probabilities) with  $A_1 = 1.0$ ,  $A_2 = 0.75$ ,  $A_3 = 0.5$ . 'n' denotes the hit poke number within a trial while  $O_n$  denotes the outcome of the  $n^{\text{th}}$  poke (1 – reward and 0 – no reward). Trial types ( $A_1$ ,  $A_2$ , and  $A_3$ ) were randomly ordered and the trial type was not cued to the animal. To obtain more trials within a session and to maintain the motivation to forage for longer period, the reward probability was forced to zero after the 20<sup>th</sup> poke in trial. A dead time of 100 ms was set to pause the session after every rewarded poke. Each session lasted a maximum of 30 minutes. There was a total of 20 sessions per animal that was performed continuously in 20 consecutive days without any break.

The implementation of three distinct starting probabilities aims to incentivize the goal-directed behavior in the animals. This approach ensures that they benefit directly from the rewards obtained in each trial, rather than following a strategy that does not consider reward outcomes. For instance, in a scenario with a uniform starting reward probability, the reduction in rewards over time would be consistent across all trials, potentially leading animals to expect a certain number of rewards in every trial. As a result, their goal-directed actions, such as nose-poking, might become less motivated by the rewards received and be driven more by an internal drive. The introduction of variable starting probabilities aims to prevent this by pushing the animals to make decisions in a setting where reward chances are dynamic and uncertain and therefore each additional reward in a trial is considered as incentive for their nose-poke behaviour. Additionally, varying the initial reward probabilities in a (pseudo)-random manner more accurately mirrors the unpredictable conditions found in natural environments. This offers a richer context for examining how animals manage the trade-off between exploiting known resources and exploring new possibilities when faced with uncertainty.

### 3.3 Data analyses

All the pre-processing and analyses were performed in MATLAB (MathWorks, R2022b) using our custom written script.

#### 3.3.1 Data storage and analysis pipeline



**Figure 2: Schematic representation of the data analysis pipeline.** The raw laminar local field potentials are pre-processed for artefact correction and channel rectification. The pre-processed LFP is then transformed into its respective current source profile by applying a second spatial derivative. Based on the activity profile and electrode depth information from histology, channel layer specifications are performed. Finally, the signals from current source profiles are rectified and averaged across the channels to obtain the overall cortical activity.

For each foraging session, the raw behaviour data was acquired as both “.csv” and “.mat” format while the raw electrophysiological data was acquired in “.dat” format. The size of raw electrophysiological data reached about (~ 423 GB). To reduce the complexity and combine the behaviour and electrophysiology data, a conversion routine was set up in MATLAB. The converted .mat file contained information about epochs of interest at the LFP level along with the important behavioural variables. The converted LFP data was stored as a three-dimensional matrix (channels x samples x trials) containing the spatial (channel) and temporal (samples) information for each trial.

A data analysis pipeline was created in MATLAB that converts the raw LFP data into epochs of interest followed by pre-processing and current source density analysis as shown in Fig.1. The analysis pipeline runs for all animals, all sessions and creates a data container for each session. Finally, all session data containers were used to create animal wise and grand averaged laminar current source density (CSD) profiles.

### 3.3.2 Analysis of behavioural data

The behavioural data output (.mat/.csv) files for each session included all the required information to investigate and reconstruct the behavioural of each animal during the probabilistic foraging paradigm. For example, time stamps of starting of a nose poke, ending of a nose poke, duration of a poke, trial number and type, each poke's reward probability and reward outcome (rewarded or not) are recorded. For this study, we mainly focused on the following behavioural features:

*Travel time.* Travel time is defined as the time taken for the animal to travel from one spout to another (Fig.1B, Eq.2). This crucial parameter was used to distinguish whether the animal was randomly exploring the cage or showing a more goal-directed exploratory behaviour. The idea here is when the animal learnt the task properly, the travel time should be less (< 5s, based on a pilot study) and consistent across different sessions.

$$\text{Travel time (s)} = (\text{Starting time of the first poke in the current trial}) - (\text{End time of the last poke in the preceding trial}) \quad (2)$$

*Relative body weight.* Relative body weight in percentage is defined as the body weight of the animal with respect to its baseline body weight (100 %) that was measured before the start of the experiment. As food restriction was performed to instil a motivation for the animal to perform the task, relative body weight was computed after each session to ensure that the body weight didn't drop beyond 85% of its initial weight.

*Inter-poke Interval.* Inter-poke interval is defined as the duration between two consecutive pokes in a trial. This was calculated separately for pokes after a rewarded poke and unrewarded poke.

*Resident time.* Resident time or trial duration is defined as the time spent at a particular spout in a trial. Mean resident time across different trial types was computed to capture the difference between different spout quality i.e., trials starting with different starting probability.

*Total rewarded and unrewarded pokes.* The total number of rewarded and unrewarded pokes for different trial types was obtained to measure the animal's performance in each session. Also, this measure was used to verify if the system correctly provided the rewards based on the probability rule.

*Total consecutive unrewarded pokes made before leaving a spout.* This is a critical measure to investigate the crucial decision-making behaviour of the animal i.e., the decision to explore the other spout. In each trial, this measure is completely decided by the animal and hence acts as a good proxy to measure how much time the animal required to make up the decision to explore.

### 3.3.3 Pre-processing of the neural data

Trial wise session data were analyzed to remove motion and chewing artefacts from the LFP using an amplitude cut-off factor (Threshold = mean  $\pm$  2 \*standard deviation of raw signal). Furthermore, LFPs and CSDs were visualized to identify broken or damaged channels. The damaged channels were corrected by a linear interpolation method based on the neighbouring unaffected channels at the LFP level. Trials with artefacts that couldn't be removed were discarded from further analysis (< 1% of total trials). Once the LFPs were channel rectified and artefact corrected, the current source density profile for each trial were re-obtained and averaged per session. The artefact corrected and channel interpolated data was then exported as a data container consisting of session averaged LFPs and CSDs for different epochs of interest.

### 3.3.4 Current source density (CSD) analysis

The current source density (CSD) analysis is an approach used to approximate the location and magnitude of current sources and sinks within brain tissue, inferred from local field potential (LFP) recordings. The CSD profile from laminar recordings is a refined measure that identifies regions of synaptic input (sinks) and output (sources), thus providing a detailed map of electrical current flow through the cortical layers, which is crucial for understanding neuronal circuitry at a mesoscopic scale.

The CSD profile was computed from the laminar LFPs by taking a second spatial derivative as shown in equation (3).

$$-\text{CSD} \approx \frac{\delta^2 \Phi(z)}{\delta z^2} = \frac{\Phi(z + n\Delta z) - 2\Phi(z) + \Phi(z - n\Delta z)}{(n\Delta z)^2} \quad (3)$$

$\Phi$  is the field potential,  $z$  is the spatial coordinate perpendicular to the cortical laminae,  $\Delta z$  is the spatial sampling interval, and  $n$  is the differential grid (Mitzdorf, 1985). LFP signals were smoothed using a Hamming window of 9 channels that corresponds to a spatial kernel filter of 400  $\mu\text{m}$  (Happel et al., 2010a).

CSD reflects the net amplitude of extracellular current flowing in (sinks) or out (source) of the neuronal tissue at a given point in time and space. Functionally, the current sinks represent the activation of excitatory synaptic populations while the source mainly represent the balancing return currents. This local functional spatiotemporal map of synaptic populations allows us to identify cortical layers by visualizing the spatiotemporal sequence of neuronal activation across the layers (Mitzdorf, 1985, Happel et al., 2010a).

Unlike the single- or multi-unit activity profile, the CSD profile provides a functional readout of cortical micro circuits in a wider mesoscopic scale. CSD transformation of LFPs is reference free and thereby less affected from referencing artefacts and far-field potentials. Furthermore, it improves the spatial resolution of the local synaptic current flow which is otherwise very poor in LFPs.

To get an overall columnar activity, the CSD profiles were transformed by averaging the rectified waveforms of each channel according to equation 4.

$$AVREC(t) = \frac{\sum_{i=1}^n |CSD_i|(t)}{n} \quad (4)$$

N is the number of recording channels and t is the time. The average rectified waveform (AVREC) represents the temporal profile of the whole columnar activity (Givre et al., 1994; Schroeder et al., 1998).

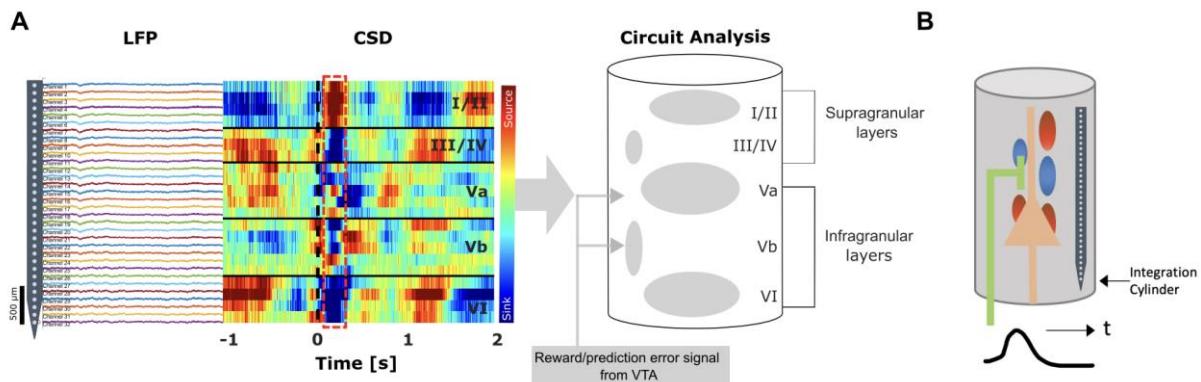
### 3.3.5 Feature extraction of CSD recordings

Early (0-100 ms) and late phases (100-500 ms) from the end of the poke were identified within the time locked (-1 to +2 seconds from the end of the poke) CSD and AVREC that could best separate the encoding of expectation and evaluation of reward respectively. Analyses were also carried out at individual poke level and the activity patterns were compared within and across these phases. CSDs were computed for individual pokes in each trial followed by grand AVREC per animal. The AVREC was quantified using the root mean square (RMS) to encapsulate the mean temporal activity for each poke during the specified early and late phases, enabling a comparative analysis across different pokes.

For layer specific analysis, cortical layers within the FrA were discerned for each animal from their session averaged CSD profiles, utilizing the functional spatiotemporal

sequence of activation. Unlike sensory cortices, where stimulus-driven activation of layer IV is more apparent, the FrA presents challenges in identifying such activations. Consequently, we leveraged the differential activation patterns elicited by rewarded versus unrewarded pokes to delineate cortical layers. This involved time-locking the CSD profile to the end of a poke (-1 to +2 seconds from the end of the poke) and contrasting activations under reward and no-reward conditions to pinpoint reward/prediction error responses in specific cortical layers. Typically, the initial response is indicative of infragranular layer V activity. Once infragranular layers were identified, the supragranular layers were demarcated and verified according to electrode depth data obtained from histological examination (see Figure 3). Three channels that best represents a layer activity was taken and the CSDs were averaged per layer across animals and arranged in a laminar fashion to obtain the grand averaged CSD.

Further, to compare the activity profiles at layer specific level, the source signals were removed (replaced with NaN) and only the sinks were considered for analysis to ensure that the signal is contributed only by layer specific local excitatory synaptic populations and eliminate the contribution of return currents from other cortical layers. For a more quantitative analysis, the RMS was computed for the average rectified sinks for each layer at individual poke level.



**Figure 3: Chronic current source density (CSD) analysis in frontal field a (FrA).** **A)** The left panel displays in vivo multichannel local field potential (LFP) recordings obtained from a 32-channel silicon probe implanted perpendicularly in the FrA of awake, behaving gerbils. The probe captures activity across all cortical layers (I–VI), with t=0 corresponding to the end of a poke (black dashed line). The middle panel illustrates the CSD, showing task evoked CSD components appeared as current sink (in blue) and source (in red) activity. Channel-layer specification is informed by initial response to reward-related CSD components appearing 100 ms post-poke (highlighted in the red box), typically marking the infragranular layer V. The right panel presents a simplified schematic illustration of the cortical column in FrA, delineating the layered structure and the direction of reward/prediction error signals originating from the ventral tegmental area (VTA). This schematic aids in visualizing the depth and functional organization of the cortical layers as identified through the CSD analysis. **B)** A schematic

representation of an electrode in a cortical column showing how the current sink (blue)/ source (red) are interpreted with respect to the position of the synapses and cell bodies.

### 3.3.6 Statistics

Both behavioural and electrophysiological results were tested for statistical significance by one-way ANOVA with Bonferroni correction (post-hoc test). For the electrophysiological analysis, each poke was considered a separate group, and the poke data was z-normalized within each animal. We used an overall significance level of 0.05 ( $\alpha = 0.05$ ).

# **RESULTS**

## 4 Results

In pursuit of understanding the neural mechanisms governing attentional resource allocation during probabilistic foraging in Mongolian gerbils, we conducted a comprehensive investigation into the frontal cortex's role. Our primary goal was to elucidate whether this brain region, recognized for its significance in human decision-making, also influences decision boundaries shaped by reward probabilities in rodents. To address this inquiry, we implemented multichannel electrode recordings in the gerbil frontal cortex, capturing neuronal responses during their participation in a probabilistic foraging task.

### 4.1 Behavioural analysis

In the probabilistic foraging task adapted from (Lottem et al., 2018), Mongolian gerbils were challenged with dynamic decision-making scenarios that required them to adaptively allocate attentional resources. This task was designed to mimic the exploration-exploitation trade-off that is commonly observed in natural foraging behaviour. The gerbil's task was to discern the optimal time to abandon a depleting food source in favour of exploring an alternative, potentially more rewarding option. Our foraging experimental setup provided a platform to study decision-making under uncertainty, as each nose poke's reward outcome was probabilistic, diminishing over consecutive pokes. Our analysis focused on quantifying how these gerbils balanced the trade-off between exploiting a known, but depleting resource and exploring new possibilities, reflecting a fundamental component of real-world decision-making.

#### 4.1.1 Performance of the foraging behaviour

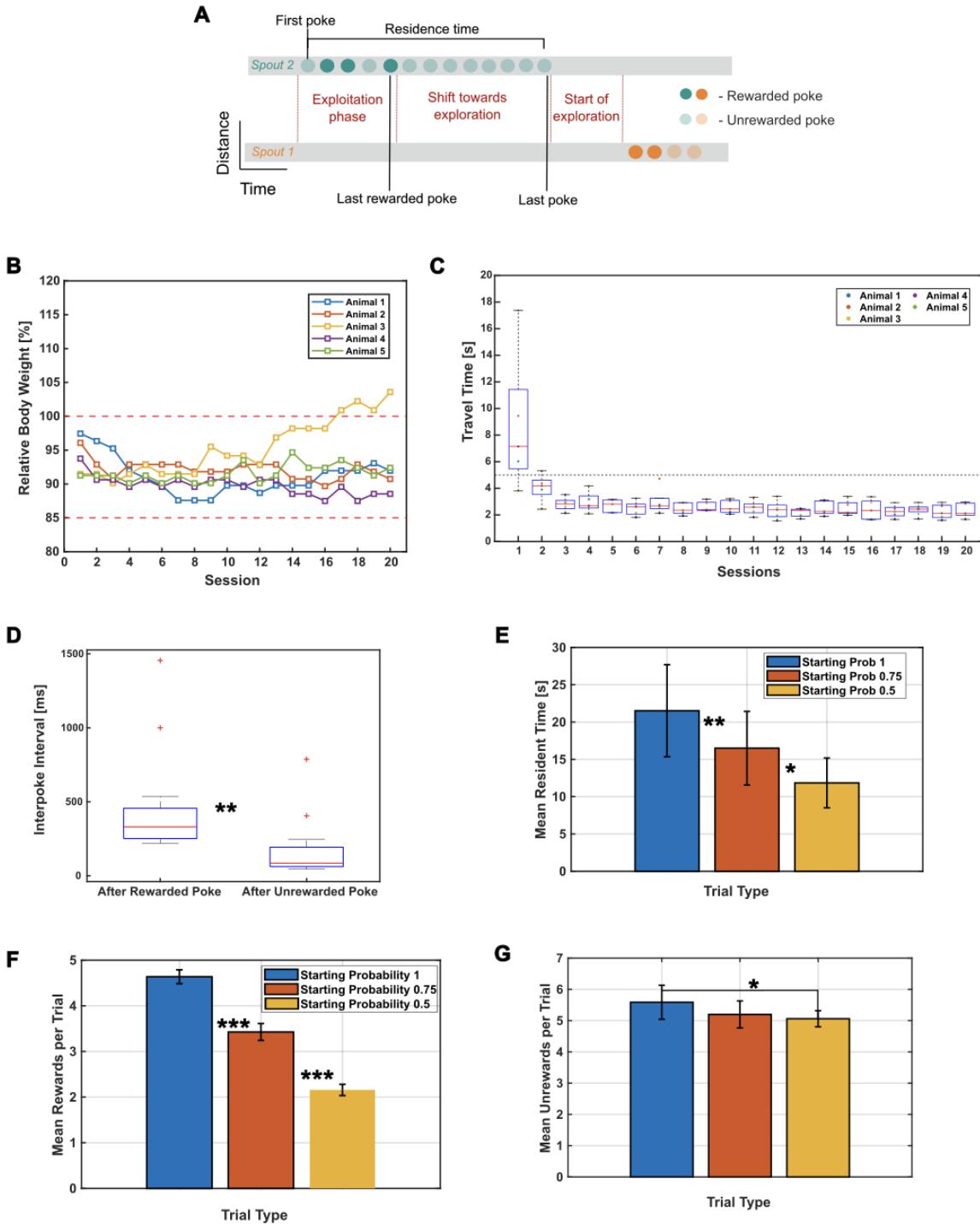
Over the course of 20 consecutive days, each of the five gerbils participated in daily foraging sessions, each comprising of N trials (Figure 4A). Each trial consists of a series of nose-pokes and the animals performed on an average 52 trials per session. From the 20 sessions, only 15 were deemed analytically relevant based on their performance, yielding in a dataset of 3,890 trials for further analysis. Behavioural analyses were performed to understand the animal's performance based on basic behavioural metrics such as inter-poke interval, resident time, travel time, number of rewards per trial etc. Analysis revealed that the median inter-poke interval was substantially longer after receiving a reward (330.25 ms on average) than after an unrewarded poke (85 ms; One-way ANOVA,  $p < 0.01$ ), as shown in Figure 4D. This

extended interval is attributed to the mandatory 100 ms dead time post-reward and the time taken by the animals to consume the food pellet. To ensure the animals remained motivated to participate, they were kept under a controlled food deprivation regime. Their body weight was regularly measured post-session to confirm that it did not fall below 85% of their initial weight, thereby preventing excessive weight loss (Figure 4B).

Travel time, defined as the duration from the end of the last poke of one trial to the commencement of the first poke in the subsequent trial, was analyzed to determine task learning and exploratory behavior (Figure 4C, Eq.2). After the initial five sessions, consistent travel times suggested goal-directed behavior rather than random exploration, leading to the exclusion of the first five sessions from further behavioural and electrophysiological analyses.

#### **4.1.1.1 Number of rewards and resident times increased with spout quality.**

Based on the starting reward probability, trials starting with a higher probability was defined as a high-quality spout (starting reward probability =1) followed by medium (starting reward probability =0.75) and low (starting reward probability =0.50) quality spouts. Gerbils exhibited longer residence times at spouts with higher reward probabilities (Figure 4E). Specifically, the mean residence time at high-quality spouts (Mean =  $21.52 \pm 6.16$  s) exceeded that at medium (Mean =  $16.49 \pm 4.93$  s) and low-quality spouts (Mean =  $11.84 \pm 3.33$  s) by 5.03 and 9.68 seconds, respectively (One-way ANOVA with Bonferroni correction, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001). Correspondingly, a greater number of rewards were obtained from spouts of higher quality (Fig.4F). Rewards from high-quality spouts (Mean =  $4.63 \pm 0.15$ ) were significantly more frequent than those from medium (Mean =  $3.42 \pm 0.18$ ) and low-quality spouts (Mean =  $2.15 \pm 0.12$ ) (One-way ANOVA with Bonferroni correction, \*\*\* p < 0.001).



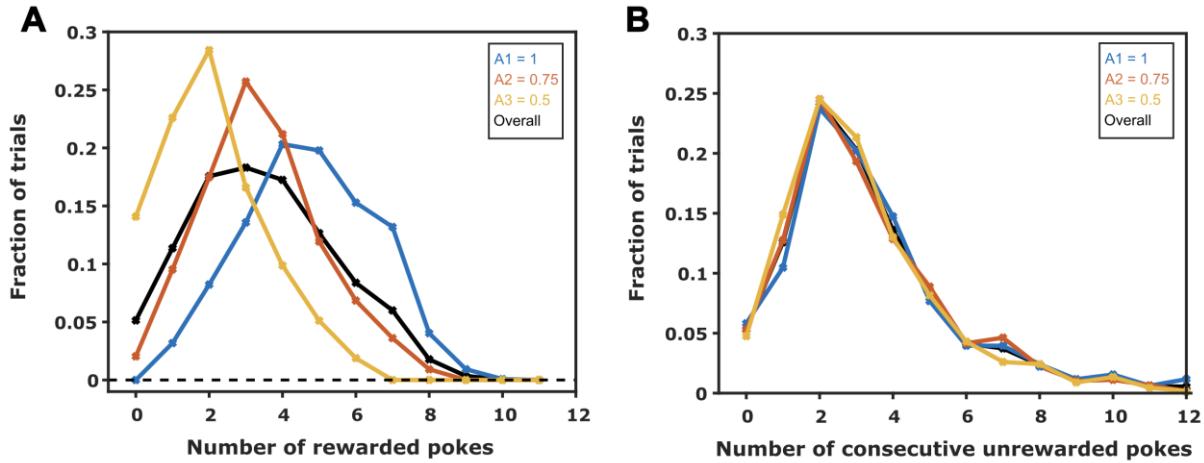
**Figure 4: Behavioural performance of the probabilistic foraging task.** **A)** Schematic representation of the foraging task, illustrating the phases of exploitation and transition to exploration with the corresponding nose poke outcomes. **B)** Relative body weight percentage of individual animals over 20 sessions, with thresholds (red dashed lines) indicating the baseline (100%) and critical weight loss limits (85%). **C)** Boxplot of travel times for each animal across sessions, highlighting the learning curve and stabilization of task performance. **D)** Average median inter-poke intervals following rewarded and unrewarded pokes. **E)** Bar graph of mean residence times at spouts with different starting reward probabilities. **F)** Mean number of rewards obtained per trial type differentiated by different starting reward probabilities. **G)** Mean number of unrewarded pokes obtained per trial type differentiated by different starting reward probabilities. Data are represented as mean  $\pm$  SEM, with statistical significance denoted by asterisks (One way ANOVA with Bonferroni Correction, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ).

#### 4.1.1.2 Spout-leaving behaviour

Following the establishment of the gerbil's capacity to perform the probabilistic foraging task, with discernible differences between various spout qualities (Fig.4E, F, G), our focus shifted to a crucial aspect of their decision-making behaviour during the exploration-exploitation dilemma: when is the right time to leave the current spout and explore the other option? This is significant as each trial, composed of a series of nose pokes at a specific spout, begins with the animal exploiting the current spout for rewards. The trial then progresses to a critical juncture, where the gerbil must decide when to leave the current spout and explore an alternative — a decision driven by the decreasing frequency of rewards in the current spout (Fig.4A).

Drawing from our previous collaborative study (Güldener et al.,2024; *in revision*), gerbils were found to adopt the Giving Up Time (GUT) rule—a forager tolerates a certain period without a reward following the last successful forage. Exceeding this threshold prompts the switch to another spout, with each reward resetting the GUT. The GUT rule does not rely on prior knowledge of spout quality; rather, it is a response to the temporal gap between rewards. In our investigation, the GUT is operationalized as the duration from the last rewarded nose poke to the final poke in each trial. Initially, hunger serves as the primary motivator for task engagement. However, as the gerbils become satiated, their motivation could diminish, potentially leading to an increase in non-task-related activities like grooming, which could affect GUT measurements. To address this issue and obtain a purer assessment of the decision-making behavior, we analyzed the number of consecutive unrewarded pokes made before a gerbil abandons the current spout. This metric serves as a robust alternative to GUT, mitigating the impact of satiation and ensuring a focus on the gerbil's spout-leaving strategy.

Figure 5B illustrates that despite different spout qualities, the gerbils maintained a consistent number of consecutive unrewarded pokes before leaving a spout. In other words, it shows that the animals irrespective of the starting reward probability and the total number of rewards received in a trial maintain a consistent number of consecutive unrewarded pokes before leaving a particular spout. This behavioral consistency suggests adherence to the GUT rule or a similar heuristic, which is likely shaped by their accumulated experience within the task environment.



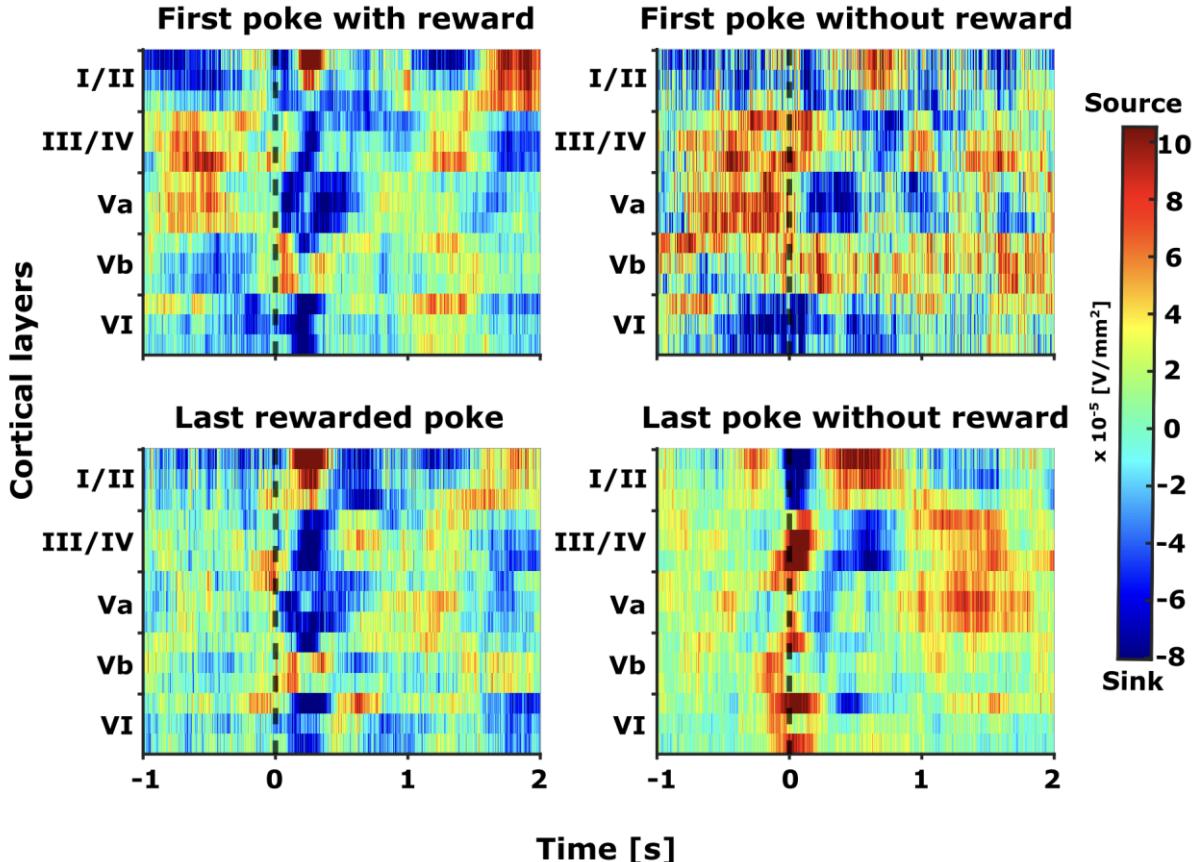
**Figure 5: Behavioural patterns in spout-switching decision-making.** **A)** The distribution of trials relative to the number of rewarded pokes, delineated by different initial reward probabilities ( $A_1 = 1$ ,  $A_2 = 0.75$ ,  $A_3 = 0.5$ , and overall). **B)** The consistency of gerbil's spout-leaving behavior as shown by the fraction of trials against the number of consecutive unrewarded pokes prior to switching, irrespective of the initial reward probability.

#### 4.2 Distinct spatiotemporal activity patterns in the frontal field A

Four specific time points in a trial were selected to determine whether the FrA encodes distinct activity patterns reflective of the animal's behavior. These time points correspond to unique and critical stages during the foraging session: the first poke (rewarded and unrewarded), the last rewarded poke, and the last poke (Fig.4A). The period from the first poke until the last rewarded poke is considered the exploitation phase, during which the animal, despite experiencing unrewarded pokes, continues to stay on the same side in anticipation of more rewards. The last rewarded poke marks a crucial transition point, indicating the end of the exploitation phase and the beginning of a shift toward exploration. This shift is most clearly represented by the last poke in a trial.

The CSD profile corresponding to these time points revealed distinct spatiotemporal neural activity within the FrA that may be associated with various features, such as motor activity related to the pokes and subsequent reward information (Fig.6). In anticipation of the decision-making process—whether to shift from the current spout or to continue exploiting it—the CSD signals were calculated for a time window spanning from one second before to two seconds after the end of the poke (Fig.6, indicated by a black dashed line at  $t=0$ ). This calculation allows for the comparison of spatiotemporal neural activity across different time points during the decision-making phase. Moreover, the differential neural activity patterns associated with rewarded and unrewarded pokes provided a basis for distinguishing between infragranular layers and

superficial layers in the laminar recordings, enabling a precise channel-layer specification.



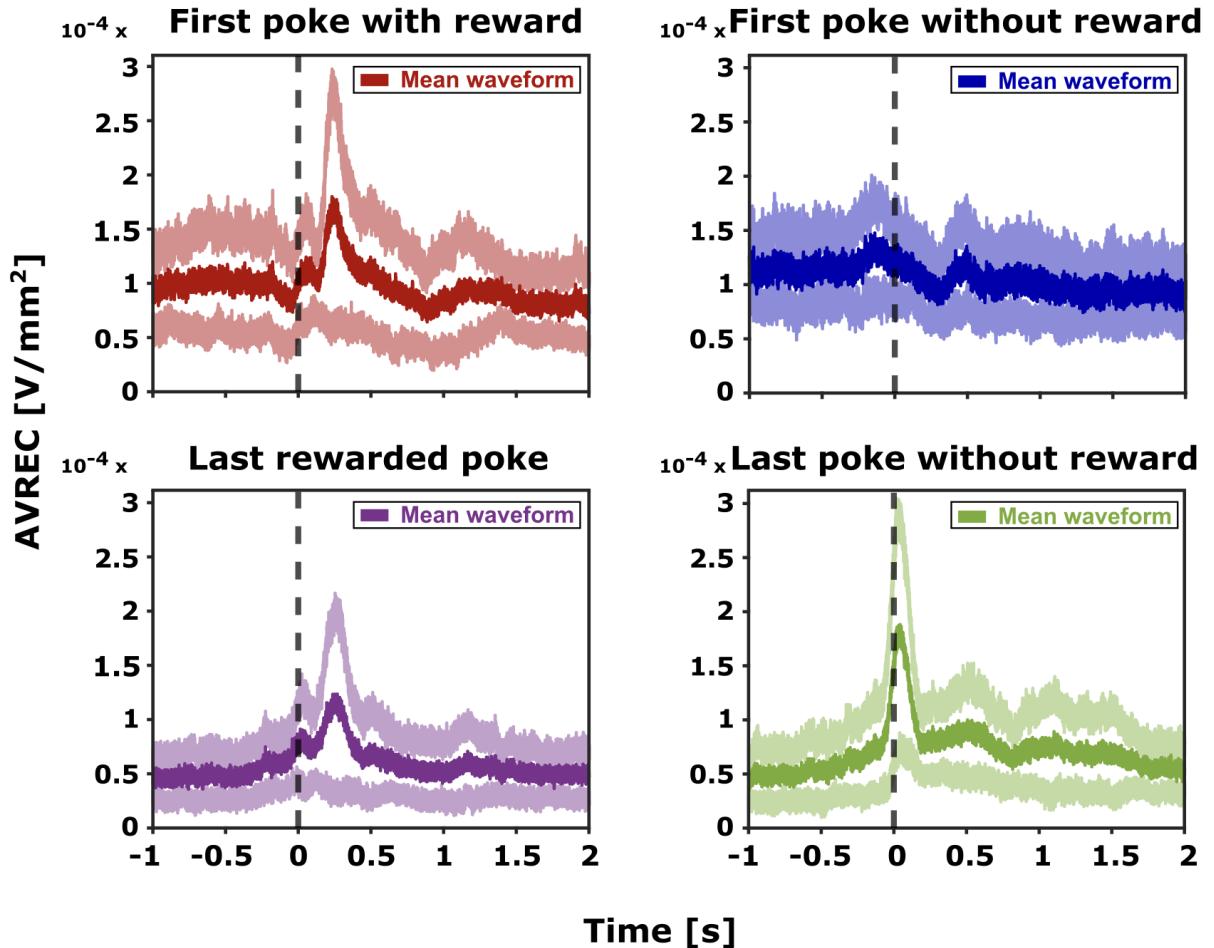
**Figure 6: Grand averaged current source density (CSD) profiles (n=5).** Distinct motor and reward related spatiotemporal neural activity in frontal field A. The selected epochs represent -1 to +2 seconds from the end of the poke (black dashed line, t=0). The selected time interval was taken for four different events (pokes) and its corresponding consequence (reward): (top left) first poke with reward, (top right) first poke without reward, (bottom left) last rewarded poke, and (bottom left) last poke without reward.

#### 4.2.1 Overall frontal activity patterns in FrA

To assess the overall frontal cortical activity, the CSD signals were rectified and averaged across all laminar electrodes, which intentionally obscured the spatial information. The grand averaged AVREC (gaAVREC) derived from the CSD profiles exhibit distinct overall activity patterns associated with different pokes, as illustrated in Figure 7. All pokes—except for the first unrewarded poke—demonstrate a bimodal waveform characterized by an initial peak shortly after the poke's end (0–100 ms) followed by a secondary peak beyond 100 ms.

In contrast, the first unrewarded poke displays a bimodal pattern with the early peak occurring prior to the poke's end (-100 – 0 ms) and the late peak occurring around 500 ms. Moreover, the overall activity patterns associated with rewarded pokes (first and

last rewarded pokes) exhibit a resemblance to one another, which stands in contrast to the patterns observed in unrewarded pokes (first unrewarded and last poke). Notably, the last poke, marking the transition to exploration, is distinguished by an early peak within the first 100 ms that is greater in amplitude, setting it apart from the other pokes.



**Figure 7: Grand averaged average rectified signals (gaAVREC) ( $n=5$ ).** AVREC displays the overall frontal cortical activity, revealing distinct motor and reward-related signals. The mean average rectified waveform (depicted in bold colours) together with its standard error (shown in lighter shades) is plotted for the selected time intervals (epochs). These epochs span from one second before to two seconds after the end of the poke ( $t=0$ ). AVRECs are presented for four distinct pokes: first poke with reward (top left), first poke without reward (top right), last rewarded poke (bottom left), and last poke without reward (bottom right).

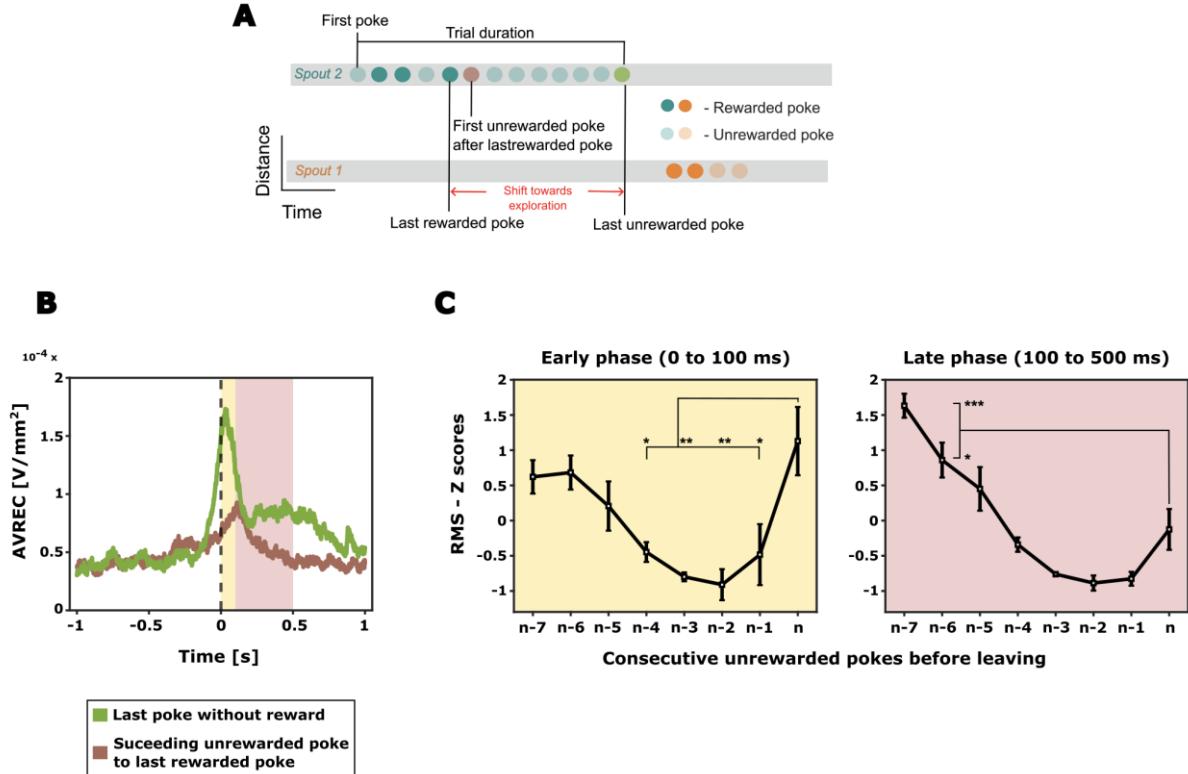
#### **4.3 Shifts in frontal activity patterns: From exploitation to exploration**

We have established that the FrA encodes distinct neural activity patterns for rewarded and unrewarded pokes, as demonstrated in Figures 6 and 7. Specifically, we observed that the last poke exhibits a neural activity pattern that is distinct from the others (Fig 7). To investigate whether this distinct pattern is exclusive to the last poke in a trial, and how it evolves, we examined the transition in frontal activity from the exploitation to exploration phases (Figure 8A). This analysis is pivotal for understanding the decision-making process, specifically how an animal determines that a poke will be the last one in a trial. This consideration is crucial because the animal often encounters multiple consecutive unrewarded pokes after the last rewarded poke in a trial (Figure 8A), necessitating a clear distinction between the last unrewarded poke and its predecessors.

In assessing the neural distinctions between the last unrewarded poke and the initial unrewarded poke following the last rewarded poke, a clear divergence in activity patterns was observed (Figure 8B). The last unrewarded poke exhibited a bimodal distribution, with a significant early peak of activity immediately after the poke (within 100 ms from its end), followed by a sustained activity phase (beginning after 250 ms from the end of the poke). On the other hand, the initial unrewarded poke following last rewarded poke displayed only a single, less pronounced peak after 100 ms from the end of the poke (Figure 8B).

Given that the two prominent peaks of the last poke fall within the initial 500 ms following the end of the poke, our analysis concentrated on this temporal window to ascertain if it represents the critical period during which the decision to switch to another spout is made. Our focused analysis within the first 500 ms post-poke revealed two distinct phases: the early phase (0 to 100 ms), denoted by an immediate sharp increase in activity, and the late phase (100 to 500 ms), characterized by sustained activity (Fig.8B, with phases highlighted in yellow and light pink, respectively). The contrast is most pronounced during the transition from the exploitation phase (Fig.8A, brown poke), with reduced yet notable early activity beyond 100 ms, to the exploration phase (Fig.8A, green poke), where the early phase activity is markedly elevated (Fig.8B). This shift in activity pattern from exploitation to exploration phase distinguishes the last poke and suggests that the early phase may be integral in encoding the decision to explore. To further probe the development of this distinct

activity pattern for the last poke, we analyzed individual pokes from the first unrewarded poke after the last rewarded one to the last poke of each trial (Fig.8A – shift towards exploration).



**Figure 8: Frontal activity change from exploitation to exploration.** **A)** An illustrative trial consisting of a sequence of pokes, with emphasis on the first unrewarded poke after the last rewarded poke (brown) and the last unrewarded poke (green). **B)** Distinct activation patterns in grand averaged AVREC for the first unrewarded poke after the last rewarded poke (brown) and the last unrewarded poke (green). From the grand AVREC data (B), two distinct time intervals (epochs) were identified for RMS computation: the early phase (0 – 100 ms, yellow), and the late phase (100 – 500 ms, light pink). **C)** The AVREC – RMS Z scores for unrewarded pokes between the last rewarded poke and the last poke (n) before the animal disengages from the current spout are shown (here, a scenario of 7 consecutive unrewarded pokes is depicted, where n-7 is the first unrewarded poke after the last rewarded poke (brown) and n represents the last unrewarded poke (green)). One-way ANOVA with Bonferroni correction was applied to detect differences between the pokes (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ).

#### 4.3.1 The evolution of frontal activity from exploitation to exploration

To elucidate the progression of frontal activity changes when shifting from exploitation to exploration, we examined the individual pokes more closely. The root mean square (RMS) of the average rectified signal (AVREC) was employed as a singular measure to capture the average temporal activity for each poke within the designated early and late phases, facilitating comparison across different pokes. The RMS was calculated for all unrewarded pokes occurring between the last rewarded poke and the last poke of the session (Figure 8B and 8C) and was z-normalized within each animal.

Throughout the transition from exploitation to exploration, a specific trend emerged during the early phase: the overall frontal activity initially showed a decrease from the first to the penultimate unrewarded poke (n-6 to n-2) and then exhibited an increase approaching the final decision to abandon the current spout (n-2 to nth poke). This trend was statistically significant (One way ANOVA with Bonferroni correction, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001). On the other hand, the late phase displayed a consistent decrease in frontal activity, with a minor elevation observed just before the transition to exploration (nth poke).

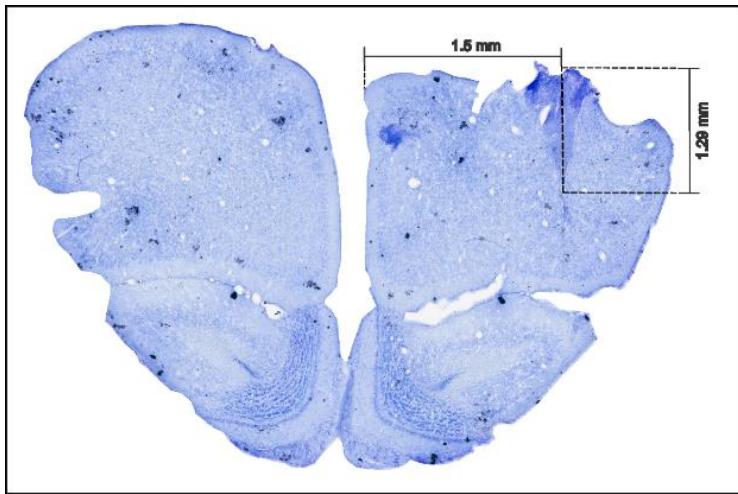
#### **4.4 Layer specific spatiotemporal activity patterns in the frontal field A**

##### **1.1.1 Histological confirmation of laminar electrode positioning in frontal region A (FrA)**

For the accurate localization of neuronal activity across all cortical layers, it is imperative to position the laminar depth electrode perpendicular to the cortical surface. Histological verification forms a critical step in this process, confirming that the electrode indeed traverses the full cortical thickness and, therefore, can capture layer-specific activity within our behavioural paradigm.

The histological examination involved tissue processing post-experiment, followed by sectioning and Nissl staining, which allowed for visualization of the electrode track. This examination confirmed the electrode's orientation in relation to the cortical layers. Correlating the electrode's location with the established cytoarchitecture of the frontal cortex, we could deduce that our recordings represent the integrative neuronal dynamics from the entire cortical depth, which is postulated to play a vital role in attentional resource allocation during decision-making tasks.

The gerbil brain atlas (Radtke-Schuller et al., 2016) situates the target region, frontal region A (FrA), at coordinates 4.65 to 5 mm anterior and 1.5 mm lateral to the bregma. Figure 9 presents a histological slice from one specimen, located at 4.85 mm anterior and 1.5 mm lateral to the bregma, confirming that the electrode placement is within the FrA. The trace of the electrode is distinctly visible, penetrating to a maximal cortical depth of approximately 1.29mm. This image serves as a qualitative representation within a larger series of analyses being conducted on additional specimens, which, while not the central focus of this thesis, underpins the integrity of the electrophysiological data presented.



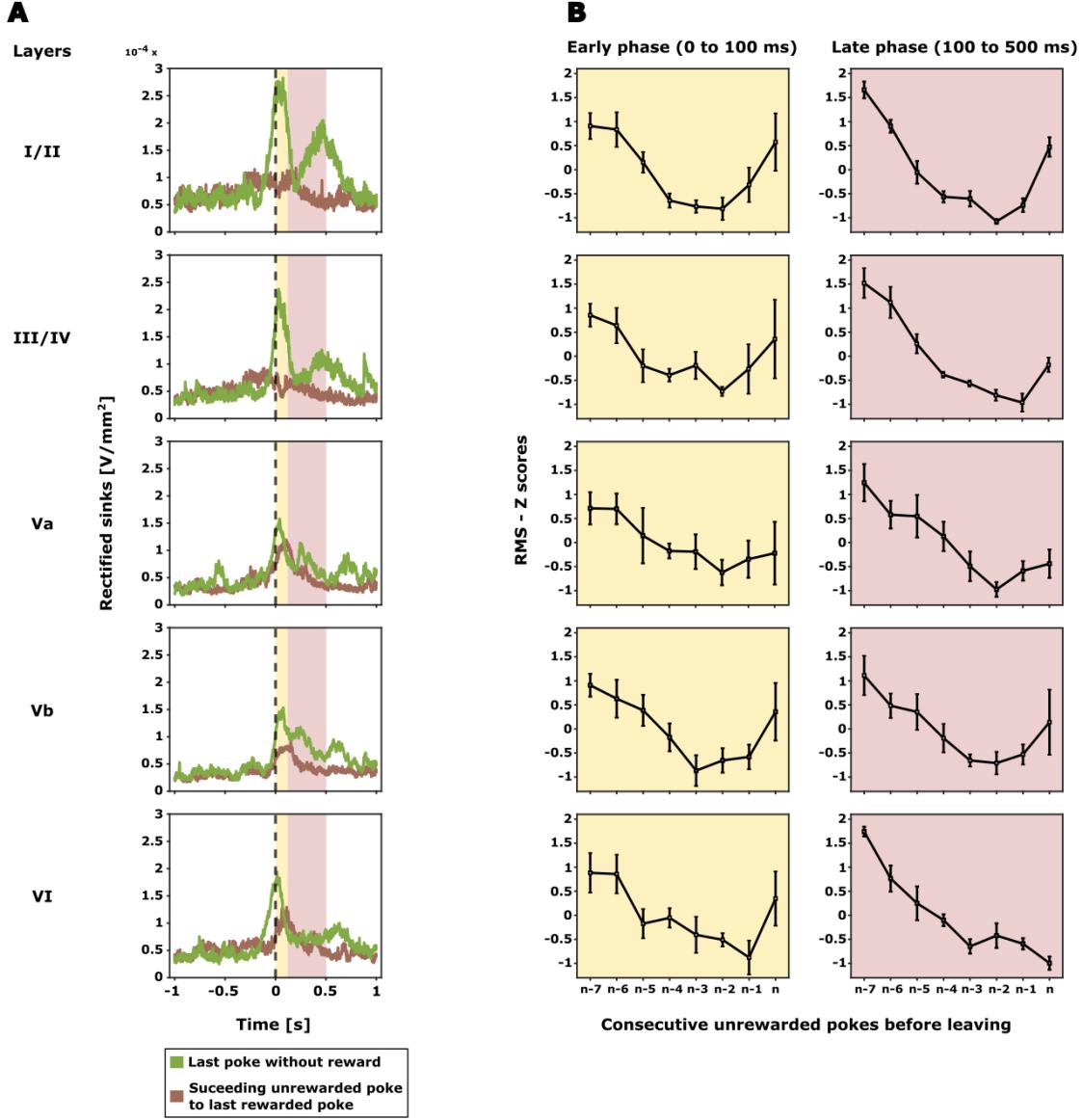
**Figure 9: Sample histology image of the frontal field A (FrA).** Sample histology image from Frontal field A (FrA), taken 4.85 mm anterior to the Bregma. The electrode track is evident at 1.5 mm lateral to the Bregma, reaching a cortical depth of approximately 1.29 mm.

#### 4.4.1 Distinct layer-specific shift in frontal activity patterns: Exploitation to exploration

Through the analysis of current source density profiles from laminar recordings, we identified five distinct cortical layers (Figure 6). To focus on the local excitatory synaptic populations, we isolated and analyzed only the sink signals from each layer (as detailed in the methods section). The comparison of layer-wise averaged sink activity between the last poke and the first unrewarded poke following the last rewarded poke revealed that the differences in overall frontal activity, previously noted (Figure 8A), are primarily driven by variations in the superficial layers (layers I/II and III/IV) (Figure 10A). Particularly, for the last poke, these layers exhibited two pronounced peaks in amplitude: one immediately at the retraction from the spout (early phase, 0-100 ms) and another following this period (late phase, 100-500 ms), coinciding with the time just before the animal decides to explore an alternative spout.

When evaluating the RMS of individual pokes, an initial decline in activity was noted across all cortical layers, reflecting the trend observed in the overall frontal activity (Figure 10B). However, a pivotal shift occurs as the animal nears the exploration phase (from the penultimate unrewarded poke, n-2, to the last poke, n). During this transition, a marked escalation in activity was observed in the superficial layers I/II and III/IV, as well as in layer Vb. This increased activity was especially pronounced in the early phase immediately following the nose poke. Like the early phase, during the late phase, a general decrease in activity was noted across all pokes. However, in contrast to early phase where layers I/II, III/IV, and Vb showed a steep increase in activity, here,

these layers only exhibited a subtle uptick in activity during the moments leading up to the decision to explore (from n-2 to nth poke), providing a nuanced view of the layer-specific activity dynamics associated with the shift from exploitation to exploration.



**Figure 10: Layer specific cortical activity in frontal region A (FrA).** **A)** Grand averaged rectified sink activity ( $n=5$ ) is displayed for all identified cortical layers from the grand CSD profile (Figure 3). Epochs shown span from -1 to +2 seconds from the end of the poke ( $t = 0$ ), focusing on the first unrewarded poke following the last rewarded poke (illustrated in brown) and the last unrewarded poke (illustrated in green). Two distinct time intervals (epochs) were selected for RMS computation based on the averaged rectified sinks: the early phase (0 – 100 ms, marked in yellow), and the late phase (100 – 500 ms, marked in light pink). **B)** Layer-wise Z scores for the RMS, derived from the average rectified sinks, quantify activity for unrewarded pokes ranging from the last rewarded poke to the last poke before the animal disengages from the current spout. As with Figure 5B, this illustration depicts a sequence of 7 consecutive unrewarded pokes, where  $n-7$  denotes the first unrewarded poke after the last rewarded poke, and  $n$  signifies the last unrewarded poke.

## **DISCUSSION**

## 5 Discussion

In our exploration of the neural underpinnings of decision-making during uncertain foraging situations, we examined the role of the frontal cortex in Mongolian gerbils through multichannel electrode recordings in the anterior frontal field A (FrA). While studies in humans and non-human primates have highlighted the importance of the anterior prefrontal cortex, particularly the frontopolar cortex, in exploratory decisions and attention shifts (Boorman et al., 2009; Daw et al., 2006), our research provides novel insights into the layer-dependent neural activity within the rodent frontal cortex during exploration-exploitation dilemmas.

On the behavioural front, our study illuminated the gerbil's reliance on a more sophisticated decision-making based on evidence accumulation rather than rigid foraging rules. At the neuronal level, the observed increase in pre-decisional neural activity in the FrA of Mongolian gerbils, especially in the supragranular layers, aligns with the modulatory role of the human aPFC on exploratory behavior. However, our study diverges from previous rodent studies that have primarily focused on the medial prefrontal cortex (mPFC) by demonstrating the nuanced role of the FrA in attentional resource allocation. This discrepancy may be attributed to the unique anatomical and functional characteristics of the gerbil FrA, which henceforth may provide a clearer analogue to the human aPFC than the mPFC in mice and rats. Furthermore, our use of chronic laminar recordings and CSD analysis has uncovered complex cognitive integration processes that are not evident in single-unit or calcium-imaging techniques, thereby contributing a mesoscopic perspective to the understanding of decision-making processes (Deane et al., 2020; Happel et al., 2014; Zempeltzi et al., 2020). The described distinct neural patterns within the FrA suggest an intricate encoding of motor actions, as well as reward expectation and evaluation, hinting at a complex cognitive process that integrates past experiences with immediate action-outcome assessments. Additionally, our findings on the layer-specific activity within the FrA suggests a nuanced, layer-dependent processing mechanism, potentially underpinning the adaptive decision-making observed in these animals.

These findings not only confirm the pivotal role of the frontal cortex in decision-making but also reveal the significance of layer-specific activity, which may have implications for future studies on the neural circuitry underlying cognitive flexibility and decision-making disorders.

## 5.1 Inference-bound decision-making in gerbils

In the world of foraging, optimal decision-making is pivotal. It is not merely a question of when an animal feeds, but rather how it determines the opportune moment to abandon a known resource in search of potentially greater yields elsewhere. One of the most prominent models of optimal foraging is the Marginal Value Theorem (MVT). Proposed by Charnov, 1976, the theory suggests that an animal should leave a food source when the immediate rate of reward collection falls to the environmental average. This model presupposes a static environment with predictable reward probabilities, allowing for the calculation of current versus average reward rates.

Our foraging paradigm presents a more complex scenario. Gerbils must navigate an environment with hidden and shifting probabilities, making it implausible for them to accurately calculate reward rates. So, what guides their decision in this uncertainty? Do they operate under any specific rules, or they act in a seemingly random manner? Our examination of spout-leaving behavior aimed to shed light on this dynamic decision-making process.

The gerbil's varied resident times and rewards across different spout qualities (Fig.4) indicate that they did not conform to fixed-time or fixed-reward rules, which prescribe consistent duration or reward counts before leaving a spout. These strategies would only be beneficial in a stable environment with constant probabilities. For instance, under a fixed-reward rule, a gerbil might overstay in a lower-quality spout, expending more effort for fewer rewards, thereby missing optimal opportunities to switch.

In an unpredictable setting such as our foraging paradigm, the likelihood of reliably estimating a spout's quality upon entry is low. Instead, each successful forage might provide incremental information about the spout's value, potentially influencing the decision to stay. While our findings initially suggest an alignment with the incremental rule i.e., the more the number of rewards the animal receives, more time it takes to leave the current spout, the pattern of consecutive unrewarded pokes before switching suggests that the behaviour is not entirely driven by this rule. Regardless of the spout's starting probability or the number of rewards accumulated, gerbils exhibited a consistent behavior of leaving after a consistent number of unrewarded pokes (Fig.5). This consistency suggests a reliance on the Giving up Time (GUT) rule, where a forager tolerates only a specific duration without reward following the last successful forage (last reward). Exceeding this temporal threshold prompts a switch to another

spout while each new reward resets this temporal tolerance. In other words, a single reward appeared to reset their attempt to explore, hinting that they valued this positive outcome more than the absence of it—a hallmark of inference over mere stimulus-response.

This propensity of the gerbils to form an inference about the action-outcome in the foraging task and alter their action based on accumulated evidence rather than immediate stimuli aligns with observed behaviours in other rodents, primates, and humans, where adaptive strategies are formed through experience (Brody & Hanks, 2016; Kira et al., 2015; Twomey et al., 2016). The significance of such a process lies in its reliance on evidence accumulation — a cognitive approach that is fundamentally more complex and versatile than simple stimulus-response conditioning. This finding suggests the presence of an intricate neural circuitry in the gerbils, facilitating a decision-making strategy that involves valence coding of outcomes that is both adaptive and informed by experience.

## 5.2 Neural encoding of reward expectation and evaluation in FrA

The spatiotemporal dynamics within the frontal region A (FrA), as elucidated by current source density (CSD) profiles, provide a window into the neural encoding of reward expectations and outcomes during the gerbil's engagement in a probabilistic foraging task (Fig.6). The average rectified signals (AVREC) captured from the CSD profiles reveal an intricate interplay between action and anticipation. This suggests that the process of expecting a reward and the subsequent evaluation of the outcome is integral to the gerbil's subsequent motor actions.

Diving into the specifics, the patterns of neural activity during the trial's exploitation phase—evidenced in both the first and last rewarded pokes—display a consistent waveform (Fig.7). Immediately following the poke's end, an initial peak in amplitude is observed, potentially encoding the anticipation of a reward. This is swiftly followed by a secondary peak, occurring less than 250 ms, which aligns with the actual receipt of the food pellet and thus likely signifies the evaluation of the reward (Fig.7).

A departure from this pattern is starkly evident in the first unrewarded poke, where the early peak arises before the poke's conclusion, accompanied by a subsequent dip instead of a reward peak as observed in rewarded pokes. This divergence indicates a neural encoding of prediction error rather than reward receipt (Fig.7). If the FrA activity

merely represented motor actions, we would expect the first peak to consistently occur at the end of all pokes, reflecting the common motor activity of nose retraction. However, the nuanced depiction of expected and unanticipated outcomes in the FrA underscores its role in outcome evaluation relative to expectations.

### **5.3 The dynamic transition of expectation is reflected in FrA activity.**

Building on the understanding that the frontal region A (FrA) encodes both reward expectation and evaluation, we delved into studying the neural activity transitions as gerbils move from exploitation to exploration. This transition is marked by a remarkable change in neural dynamics, particularly evident towards the last unrewarded poke, which signals the commencement of exploratory behavior.

Our initial comparative analysis revealed a qualitative shift in FrA activity patterns. During the exploitation phase's tail end, marked by the first unrewarded poke post-reward (brown poke in Fig.8B), there is a dip in early activity but a sustained pattern, thereafter, signifying a persistent focus on reward acquisition. In contrast, the transition to the exploration phase (last unrewarded poke, green poke in Fig.8B) is characterized by a sharp increase in activity within 100 ms post-poke. This heightened early-phase activity may imply the integration of expectation of outcome with motor planning required for the switch to a new foraging location. This distinct neural coding of the last unrewarded poke, which signals the shift to exploration, makes it unique from previous pokes. Besides, it hints that the early phase could be the crucial where the primary signal to explore may get encoded. Furthermore, this observed qualitative shift in FrA activity patterns prompted a more quantitative analysis at individual poke level to understand the dynamics of activity change from exploitation to exploration.

To this end, we took an example where the animal made seven unrewarded pokes before leaving to study the evolution of the FrA activity towards exploration behaviour. The RMS analysis of individual pokes illustrated a non-linear evolution of FrA activity, forming a U-shaped curve as the trial progresses from exploitation to exploration (Fig.8C). This U-shaped pattern in neural activity—rising, dipping, then rising again—may reflect a dynamic alteration in reward expectation. Initially, the animal experiences a prediction error, expecting a reward that does not materialize (n-7<sup>th</sup> poke). This is characterized by a dip in activity 100 ms post-poke where the reward receipt is usually encoded (brown waveform, Fig.8B). As unrewarded pokes accumulate (Fig.8C, n-6 to n-2<sup>nd</sup> poke), the animal enters a state of uncertainty (not knowing what to expect),

diminishing overall frontal activity. This prompts for a shift in attention from the anticipated reward to the act of nose poking as the action is now more informative for updating the expectations. This follows up with a slow rise in FrA activity in the early phase (0-100 ms) that is closer to the nose-poke action. Eventually, a new expectation sets in—the anticipation of no reward—which, upon confirmation through continued unrewarded pokes ( $n-2$  to  $n^{\text{th}}$  poke), leads to an overall increase in activity in the early phase followed by the subsequent spout switch. This dynamic transition of FrA activity from exploitation to exploration highlights a decision-making process within the FrA that goes beyond simple stimulus-response patterns, involving an alteration of expectations shaped by experience—an essential element underlying the inference-driven behavior.

The increasing activity observed in the FrA during the transition from exploitation to exploration in gerbils resonates with patterns observed in other species. Specifically, this activity mirrors the pre-decisional neural activity buildup seen in frontal-parietal circuits in humans and rats, which is indicative of evidence accumulation during decision-making tasks (Brosnan et al., 2020; Scott et al., 2017). This parallel suggests a fundamental neural process shared across species, highlighting the FrA's role in integrating evolving information before reaching a decision.

Additionally, the nuanced modulation of expectations within the FrA points to a sophisticated neural mechanism, where layer-specific processing could be critical. The progression of expectation from an initial certainty about receiving a reward, through a phase of uncertainty, to a new certainty of no reward, for the same action hints at the differential involvement of cortical layers. This may likely reflect the layered architecture of the FrA's role in integrating past experiences with current decision-making processes, underscoring the importance of understanding layer-specific dynamics in this complex decision-making behaviour.

#### **5.4 Functional implications of layer-specific activity during foraging**

In the pursuit of understanding foraging strategies, the laminar current source density recordings from the FrA have provided a detailed map of the local layer-specific micro-circuitry activity. The gerbil's foraging behavior has allowed us to discern how the FrA influences the differential allocation of attentional resources during critical decision-making moments—specifically, when shifting focus from exploitation at a current resource to exploration of a new one. This differentiation in strategic behavior appears

to be influenced within the FrA in a layer-dependent manner, as revealed in our results (Fig.10).

The transition from exploitation to exploration is fundamentally an attentional reallocation challenge in the brain. According to established reinforcement learning theories, goal-directed behavior is underpinned by the brain's capacity to predict the expected reward from specific actions or stimuli. The predictive coding of this expectation relies on a neural computation of prediction error—the discrepancy between expected and received rewards (Schultz, 2015). Brain structures such as the ventral tegmental area (VTA) play a crucial role in this process by evaluating reinforcement and conveying salience and valence information to various brain regions, including the frontal cortex (Bromberg-Martin et al., 2010).

Drawing parallels from studies on other neocortical regions (Atencio & Schreiner, 2010; Krienen et al., 2016; Lin et al., 2015; Wester & Contreras, 2012), it has been noted that supragranular layers possess anatomical connections conducive to cross-columnar activations and long-range inter-cortical communications. These connections are ideally positioned to facilitate the attentional resource reallocation required for exploratory behavior. Conversely, the infragranular layers are known for their corticoefferent feedback loops (Avery & Krichmar, 2015; Happel, 2016), which are implicated in updating working memory content through reward-prediction error signaling play an integral role in maintaining exploitation strategies.

Our hypothesis posited that exploratory behavior, such as a change in foraging site, would correlate with heightened activity in the upper layers, whereas deeper layer activity would align with the continuation of exploitation at the current site. Layer-specific analysis substantiated this; upper layers (Layer I/II and III/IV) exhibited a pronounced increase in neural activity during exploratory phases relative to deeper layers (Fig.10). Moreover, the activity within these upper layers paralleled the U-shaped non-linear trend observed in the overall FrA activity, particularly escalating during the early phase preceding the animal's decision to switch foraging sites.

These findings endorse our hypothesis, suggesting that the FrA employs a layer-dependent processing mechanism to allocate attentional resources effectively, thereby facilitating the gerbil's dynamic shift in search strategies during foraging behavior.

## 5.5 Limitations and future considerations

Our investigation into the neural mechanisms of probabilistic foraging in Mongolian gerbils has shed light on the complexities of decision-making under uncertainty. While providing valuable initial insights on how the FrA tactfully allocates attentional resources, it also opens opportunities to delve deeper and broaden our understanding in several key areas.

While our experimental setup offers valuable insights by simulating a naturalistic environment, incorporating elements such as competitive foraging in a larger, more complex space could further enhance the ecological validity of our findings. Introducing another gerbil into the paradigm could introduce real-world selection pressures and more intricate decision-making scenarios, thus providing a richer context for behavioral analysis. On the behavioral front, the optimality of gerbil behavior within our current setup warrants deeper exploration. A behavior modelling approach, such as reinforcement learning-based models, could be instrumental in dissecting trial-by-trial parameters. This approach could yield profound insights into the emergence of optimal behaviors and the decision-making parameters driving these behaviors. Additionally, strengthening the evidence for inference-based behavior in gerbils requires an in-depth investigation into the correlation between consecutive unrewarded pokes before spout switching and the total number of rewards received in a trial. Demonstrating a minimal to zero correlation would further substantiate the hypothesis of inference-driven decision-making in these animals.

From a neural perspective, integrating time-frequency analysis to examine frontal theta activity changes preceding exploratory decisions could illuminate the neural mechanisms overcoming Pavlovian biases (Cavanagh et al., 2013). Such an analysis would complement our findings by providing insights into the temporal dynamics of the shift from exploitation to exploration in the FrA. Furthermore, building upon recent advancements in primate research, future studies could explore the causal role of the anterior frontal cortex and dopaminergic pathways in balancing risk-return decisions (Sasaki et al., 2024). Optogenetic or pharmacologic manipulation of these pathways could provide pivotal insights into the role of meso-frontal circuits in rodents in balancing the exploitation-exploration trade off, significantly enhancing our understanding of neural decision-making processes in uncertain environments.

# **CONCLUSION**

## 6 Conclusion

This thesis has delved into the neural underpinnings of decision-making under uncertainty, with a focus on the dynamic role of the frontal region A (FrA) in gerbils. In our probabilistic foraging paradigm, the gerbils demonstrated not a reliance on fixed foraging patterns, but rather an advanced, inference-based decision-making process informed by their experiences. The complex neural activity within the FrA, depicting reward anticipation and evaluation, highlights a cognitive process where past learning is intricately merged with immediate situational assessments. Furthermore, our investigation into layer-specific neural activity in the FrA revealed a sophisticated, layer-dependent processing mechanism. This is crucial for the adaptation of search strategies during foraging and reflects the nuanced way the brain can dynamically allocate attentional resources.

The implications of this research extend far beyond gerbil behavior, offering profound insights for the broader field of neuroscience, particularly in understanding human cognitive functions. By demonstrating the adaptive decision-making capabilities in gerbils, this study serves as a valuable model for dissecting similar neural mechanisms in humans. The ability to balance exploration and exploitation reflects a fundamental cognitive process crucial across species, including humans. Besides, impairments in decision-making and cognitive flexibility are hallmark features of numerous neurological and psychiatric disorders, such as schizophrenia, obsessive-compulsive disorder, and addiction. Therefore, the insights into decision-making and cognitive flexibility gained here can have potential therapeutic implications for treating neurological and psychiatric disorders marked by decision-making impairments. Ultimately, by bridging animal behavior with human cognitive science, this study lays the groundwork for future research into the complex neural underpinnings of decision-making in uncertain environments. It underscores the beginning of a deeper inquiry into decision neuroscience, inviting further exploration into the vast, uncharted territories of how brains navigate decisions in the real world.

## 7 Bibliography

- Atencio, C. A., & Schreiner, C. E. (2010). Columnar connectivity and laminar processing in cat primary auditory cortex. *PLoS ONE*, 5(3). <https://doi.org/10.1371/journal.pone.0009521>
- Avery, M. C., & Krichmar, J. L. (2015). Improper activation of D1 and D2 receptors leads to excess noise in prefrontal cortex. *Frontiers in Computational Neuroscience*, 9(MAR). <https://doi.org/10.3389/fncom.2015.00031>
- Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. In *Nature Neuroscience* (Vol. 8, Issue 11, pp. 1458–1463). <https://doi.org/10.1038/nn1584>
- Bechara, A., Damasio, H., & Damasio, A. R. (2003). Role of the Amygdala in Decision-Making. *Annals of the New York Academy of Sciences*, 985(1), 356–369. <https://doi.org/10.1111/J.1749-6632.2003.TB07094.X>
- Beharelle, A. R., Polanía, R., Hare, T. A., & Ruff, C. C. (2015). Transcranial stimulation over frontopolar cortex elucidates the choice attributes and neural mechanisms used to resolve exploration–exploitation trade-offs. *Journal of Neuroscience*, 35(43), 14544–14556. <https://doi.org/10.1523/JNEUROSCI.2322-15.2015>
- Bludau, S., Eickhoff, S. B., Mohlberg, H., Caspers, S., Laird, A. R., Fox, P. T., Schleicher, A., Zilles, K., & Amunts, K. (2014). Cytoarchitecture, probability maps and functions of the human frontal pole. *NeuroImage*, 93, 260–275. <https://doi.org/10.1016/j.neuroimage.2013.05.052>
- Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., & Rushworth, M. F. S. (2009). How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. *Neuron*, 62(5), 733–743. <https://doi.org/10.1016/j.neuron.2009.05.014>
- Brody, C. D., & Hanks, T. D. (2016). Neural underpinnings of the evidence accumulator. In *Current Opinion in Neurobiology* (Vol. 37, pp. 149–157). Elsevier Ltd. <https://doi.org/10.1016/j.conb.2016.01.003>
- Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. In *Neuron* (Vol. 68, Issue 5, pp. 815–834). <https://doi.org/10.1016/j.neuron.2010.11.022>

- Brosnan, M. B., Sabaroedin, K., Silk, T., Genc, S., Newman, D. P., Loughnane, G. M., Fornito, A., O'Connell, R. G., & Bellgrove, M. A. (2020). Evidence accumulation during perceptual decisions in humans varies as a function of dorsal frontoparietal organization. *Nature Human Behaviour*, 4(8), 844–855. <https://doi.org/10.1038/s41562-020-0863-4>
- Brunk, M. G. K., Deane, K. E., Kisse, M., Deliano, M., Vieweg, S., Ohl, F. W., Lippert, M. T., & Happel, M. F. K. (2019). Optogenetic stimulation of the VTA modulates a frequency-specific gain of thalamocortical inputs in infragranular layers of the auditory cortex. *Scientific Reports* 2019 9:1, 9(1), 1–15. <https://doi.org/10.1038/s41598-019-56926-6>
- Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents-EEG, ECoG, LFP and spikes. In *Nature Reviews Neuroscience* (Vol. 13, Issue 6, pp. 407–420). <https://doi.org/10.1038/nrn3241>
- Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal theta overrides Pavlovian learning biases. *Journal of Neuroscience*, 33(19), 8541–8548. <https://doi.org/10.1523/JNEUROSCI.5754-12.2013>
- Charnov, E. L. (1976). Optimal Foraging, the Marginal Value Theorem. In *POPULATION BIOLOGY* (Vol. 9).
- Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017). Learning features in a complex and changing environment: A distribution-based framework for visual attention and vision in general. In *Progress in Brain Research* (Vol. 236, pp. 97–120). Elsevier B.V. <https://doi.org/10.1016/bs.pbr.2017.07.001>
- Cutler, M. G., Mackintosh, J. H., & And, J. H. M. (n.d.). Epilepsy and Behaviour of the Mongolian Gerbil: An Ethological Study. In *Physiology & Behavior* (Vol. 46). ¢ Pergamon Press plc.
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095), 876–879. <https://doi.org/10.1038/nature04766>
- Deane, K. E., Brunk, M. G. K., Curran, A. W., Zempeltzi, M. M., Ma, J., Lin, X., Abela, F., Aksit, S., Deliano, M., Ohl, F. W., & Happel, M. F. K. (2020). Ketamine anaesthesia induces gain enhancement via recurrent excitation in granular input

layers of the auditory cortex. *The Journal of Physiology*, 598(13), 2741–2755.  
<https://doi.org/10.1113/JP279705>

Douglas, R. J., & Martin, K. A. C. (2007). Recurrent neuronal circuits in the neocortex. In *Current Biology* (Vol. 17, Issue 13). Cell Press.  
<https://doi.org/10.1016/j.cub.2007.04.024>

Francis, N. A., Winkowski, D. E., Sheikhattar, A., Armengol, K., Babadi, B., & Kanold, P. O. (2018). Small Networks Encode Decision-Making in Primary Auditory Cortex. *Neuron*, 97(4), 885-897.e6. <https://doi.org/10.1016/j.neuron.2018.01.019>

Givre, S. J., Schroeder, C. E., & Arezzo, J. C. (1994). Contribution of extrastriate area V4 to the surface-recorded flash VEP in the awake macaque. *Vision Research*, 34(4), 415–428. [https://doi.org/10.1016/0042-6989\(94\)90156-2](https://doi.org/10.1016/0042-6989(94)90156-2)

Godlove, D. C., Maier, A., Woodman, G. F., & Schall, J. D. (2014). Microcircuitry of agranular frontal cortex: Testing the generality of the canonical cortical microcircuit. *Journal of Neuroscience*, 34(15), 5355–5369.  
<https://doi.org/10.1523/JNEUROSCI.5127-13.2014>

Happel, M. F. K. (2016). Dopaminergic impact on local and global cortical circuit processing during learning. In *Behavioural Brain Research* (Vol. 299, pp. 32–41). Elsevier. <https://doi.org/10.1016/j.bbr.2015.11.016>

Happel, M. F. K., Deliano, M., Handschuh, J., & Ohl, F. W. (2014). Dopamine-modulated recurrent corticoefferent feedback in primary sensory cortex promotes detection of behaviorally relevant stimuli. *Journal of Neuroscience*, 34(4), 1234–1247. <https://doi.org/10.1523/JNEUROSCI.1990-13.2014>

Happel, M. F. K., Jeschke, M., & Ohl, F. W. (2010). Spectral integration in primary auditory cortex attributable to temporally precise convergence of thalamocortical and intracortical input. *Journal of Neuroscience*, 30(33), 11114–11127.  
<https://doi.org/10.1523/JNEUROSCI.0689-10.2010>

Herry, C., Vouimba, R. M., & Garcia, R. (1999). Plasticity in the mediodorsal thalamo-prefrontal cortical transmission in behaving mice. *Journal of Neurophysiology*, 82(5), 2827–2832.  
<https://doi.org/10.1152/JN.1999.82.5.2827/ASSET/IMAGES/LARGE/9K1190588005.jpeg>

- Jacobs, B., Schall, M., Prather, M., Kapler, E., Driscoll, L., Baca, S., Jacobs, J., Ford, K., Wainwright, M., & Treml, M. (2001). Regional dendritic and spine variation in human cerebral cortex: a quantitative golgi study. *Cerebral Cortex (New York, N.Y. : 1991)*, 11(6), 558–571. <https://doi.org/10.1093/CERCOR/11.6.558>
- Jarvers, C., Brosch, T., Brechmann, A., Woldeit, M. L., Schulz, A. L., Ohl, F. W., Lommerzheim, M., & Neumann, H. (2016). Reversal learning in humans and gerbils: Dynamic control network facilitates learning. *Frontiers in Neuroscience*, 10(NOV). <https://doi.org/10.3389/FNINS.2016.00535/ABSTRACT>
- Karlsson, M. P., Tervo, D. G. R., & Karpova, A. Y. (2012). Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science (New York, N.Y.)*, 338(6103), 135–139. <https://doi.org/10.1126/SCIENCE.1226518>
- Kira, S., Yang, T., & Shadlen, M. N. (2015). A Neural Implementation of Wald's Sequential Probability Ratio Test. *Neuron*, 85(4), 861–873. <https://doi.org/10.1016/j.neuron.2015.01.007>
- Krienen, F. M., Yeo, B. T. T., Ge, T., Buckner, R. L., & Sherwood, C. C. (2016). Transcriptional profiles of supragranular-enriched genes associate with corticocortical network architecture in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 113(4), E469–E478. <https://doi.org/10.1073/pnas.1510903113>
- Kristjánsson, Á., Jóhannesson, Ó. I., & Thornton, I. M. (2014). Common attentional constraints in visual foraging. *PLoS ONE*, 9(6). <https://doi.org/10.1371/JOURNAL.PONE.0100752>
- Laubach, M., Amarante, L. M., Swanson, K., & White, S. R. (2018). What, if anything, is rodent prefrontal cortex? In *eNeuro* (Vol. 5, Issue 5). Society for Neuroscience. <https://doi.org/10.1523/ENEURO.0315-18.2018>
- Lee, E., Rhim, I., Lee, J. W., Ghim, J. W., Lee, S., Kim, E., & Jung, M. W. (2016). Enhanced neuronal activity in the medial prefrontal cortex during social approach behavior. *Journal of Neuroscience*, 36(26), 6926–6936. <https://doi.org/10.1523/JNEUROSCI.0307-16.2016>
- Lin, I. C., Okun, M., Carandini, M., & Harris, K. D. (2015). The Nature of Shared Cortical Variability. *Neuron*, 87(3), 644–656. <https://doi.org/10.1016/j.neuron.2015.06.035>

- Lottem, E., Banerjee, D., Vertechi, P., Sarra, D., Lohuis, M. O., & Mainen, Z. F. (2018). Activation of serotonin neurons promotes active persistence in a probabilistic foraging task. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-018-03438-y>
- Mansouri, F. A., Buckley, M. J., Mahboubi, M., & Tanaka, K. (2015). Behavioral consequences of selective damage to frontal pole and posterior cingulate cortices. *Proceedings of the National Academy of Sciences of the United States of America*, 112(29), E3940–E3949. <https://doi.org/10.1073/pnas.1422629112>
- Mitzdorf, U. (1985). Current Source-Density Method and Application in Cat Cerebral Cortex: Investigation of Evoked Potentials and EEG Phenomena. In *PHYSIOLOGICAL REVIEWS* (Vol. 65, Issue 1). [www.physiology.org/journal/physrev](http://www.physiology.org/journal/physrev)
- Muller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. In *Perception & Psychophysics* (Vol. 57, Issue I).
- Ohl, F. W., Wetzel, W., Wagner, T., Rech, A., & Scheich, H. (1999). *Bilateral Ablation of Auditory Cortex in Mongolian Gerbil Affects Discrimination of Frequency Modulated Tones but not of Pure Tones*.
- Onge, J. R. S., Stopper, C. M., Zahm, D. S., & Floresco, S. B. (2012). Separate Prefrontal-Subcortical Circuits Mediate Different Components of Risk-Based Decision Making. *Journal of Neuroscience*, 32(8), 2886–2899. <https://doi.org/10.1523/JNEUROSCI.5625-11.2012>
- Orsini, C. A., Blaes, S. L., Dragone, R. J., Betzhold, S. M., Finner, A. M., Bizon, J. L., & Setlow, B. (2020). Distinct relationships between risky decision making and cocaine self-administration under short- and long-access conditions. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 98. <https://doi.org/10.1016/j.pnpbp.2019.109791>
- Otto, G., & Jrge, S. (2012). The Mongolian Gerbil as a Model for the Analysis of Peripheral and Central Age-Dependent Hearing Loss. In *Hearing Loss*. InTech. <https://doi.org/10.5772/33569>

- Radtke-Schuller, S., Schuller, G., Angenstein, F., Grosser, O. S., Goldschmidt, J., & Budinger, E. (2016). Brain atlas of the Mongolian gerbil (*Meriones unguiculatus*) in CT/MRI-aided stereotaxic coordinates. *Brain Structure & Function*, 221 Suppl 1(Suppl 1), 1–272. <https://doi.org/10.1007/S00429-016-1259-0>
- Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. In *Nature Reviews Neuroscience* (Vol. 5, Issue 3, pp. 184–194). Nature Publishing Group. <https://doi.org/10.1038/nrn1343>
- Sasaki, R., Ohta, Y., Onoe, H., Yamaguchi, R., Miyamoto, T., Tokuda, T., Tamaki, Y., Isa, K., Takahashi, J., Kobayashi, K., Ohta, J., & Isa, T. (2024). Balancing risk-return decisions by manipulating the mesofrontal circuits in primates. *Science*, 383(6678), 55–61. <https://doi.org/10.1126/SCIENCE.ADJ6645>
- Schroeder, C. E., Mehta, A. D., & Givre, S. J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cerebral Cortex (New York, N.Y.: 1991)*, 8(7), 575–592. <https://doi.org/10.1093/CERCOR/8.7.575>
- Schultz, W. (2015). Neuronal Reward and Decision Signals: From Theories to Data. *Physiol Rev*, 95, 853–951. <https://doi.org/10.1152/physrev.00023.2014.-Re>
- Scott, B. B., Constantinople, C. M., Akrami, A., Hanks, T. D., Brody, C. D., & Tank, D. W. (2017). Fronto-parietal Cortical Circuits Encode Accumulated Evidence with a Diversity of Timescales. *Neuron*, 95(2), 385-398.e5. <https://doi.org/10.1016/j.neuron.2017.06.013>
- Seto-Ohshima, A., Ito, M., Kudo, T., & Mizutani, A. (1992). Intrinsic and drug-induced seizures of adult and developing gerbils. *Acta Neurologica Scandinavica*, 85(5), 311–317. <https://doi.org/10.1111/J.1600-0404.1992.TB04049.X>
- Tang, W., Shin, J. D., & Jadhav, S. P. (2021). Multiple time-scales of decision-making in the hippocampus and prefrontal cortex. *eLife*, 10. <https://doi.org/10.7554/eLife.66227>
- Tchabovsky, A. V., Savinetskaya, L. E., Ovchinnikova, N. L., Safonova, A., Ilchenko, O. N., Sapozhnikova, S. R., & Vasilieva, N. A. (2019). Sociability and pair-bonding in gerbils: A comparative experimental study. *Current Zoology*, 65(4), 363–373. <https://doi.org/10.1093/cz/zoy078>

Twomey, D. M., Kelly, S. P., & O'Connell, R. G. (2016). Abstract and effector-selective decision signals exhibit qualitatively distinct dynamics before delayed perceptual reports. *Journal of Neuroscience*, 36(28), 7346–7352. <https://doi.org/10.1523/JNEUROSCI.4162-15.2016>

Vertechi, P., Lottem, E., Sarra, D., Godinho, B., Treves, I., Quendera, T., Oude Lohuis, M. N., & Mainen, Z. F. (2020). Inference-Based Decisions in a Hidden State Foraging Task: Differential Contributions of Prefrontal Cortical Areas. *Neuron*, 106(1), 166-176.e6. <https://doi.org/10.1016/j.neuron.2020.01.017>

Wester, J. C., & Contreras, D. (2012). Columnar interactions determine horizontal propagation of recurrent network activity in neocortex. *Journal of Neuroscience*, 32(16), 5454–5471. <https://doi.org/10.1523/JNEUROSCI.5006-11.2012>

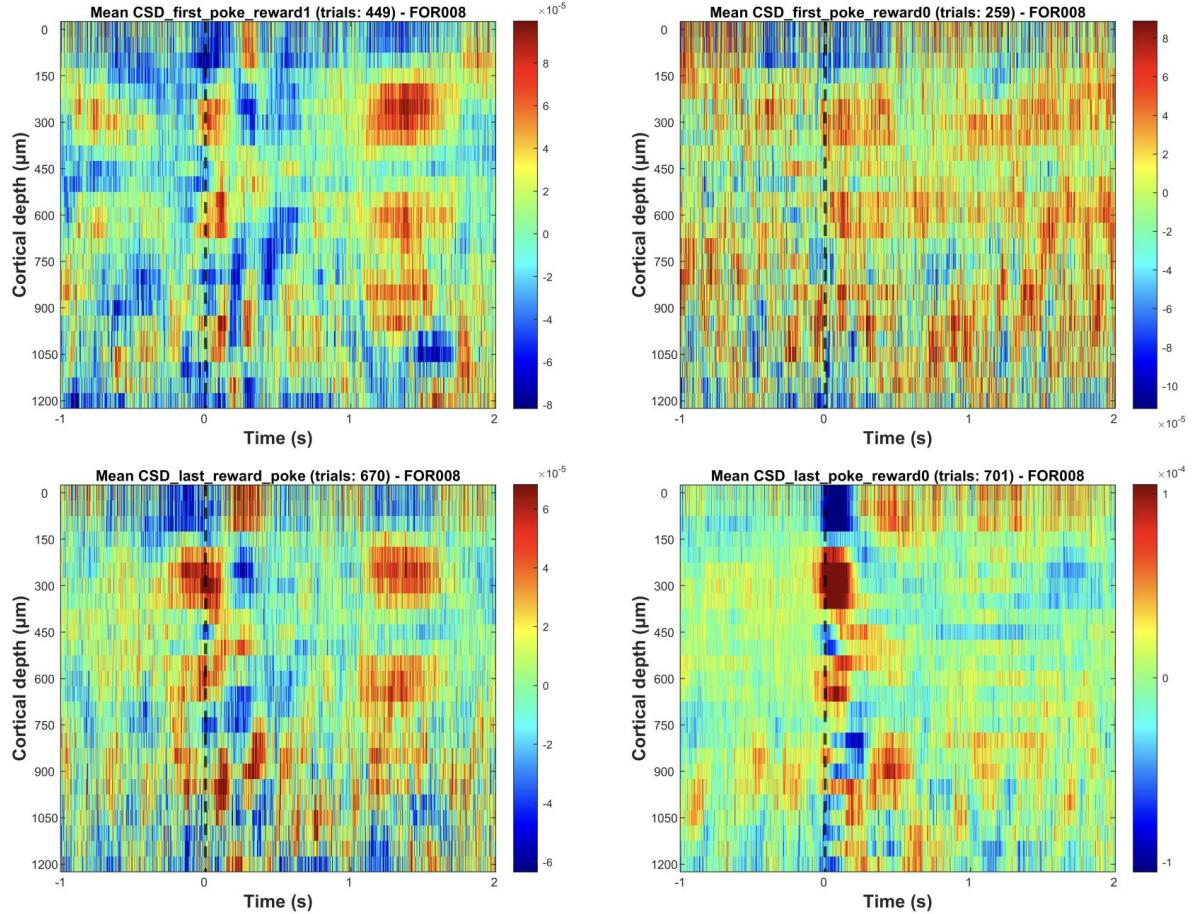
Wolfe, J. M., Cain, M. S., & Aizenman, A. M. (2019). Guidance and selection history in hybrid foraging visual search. *Attention, Perception, and Psychophysics*, 81(3), 637–653. <https://doi.org/10.3758/s13414-018-01649-5>

Zempeltzi, M. M., Kisse, M., Brunk, M. G. K., Glemser, C., Aksit, S., Deane, K. E., Maurya, S., Schneider, L., Ohl, F. W., Deliano, M., & Happel, M. F. K. (2020). Task rule and choice are reflected by layer-specific processing in rodent auditory cortical microcircuits. *Communications Biology*, 3(1). <https://doi.org/10.1038/s42003-020-1073-3>

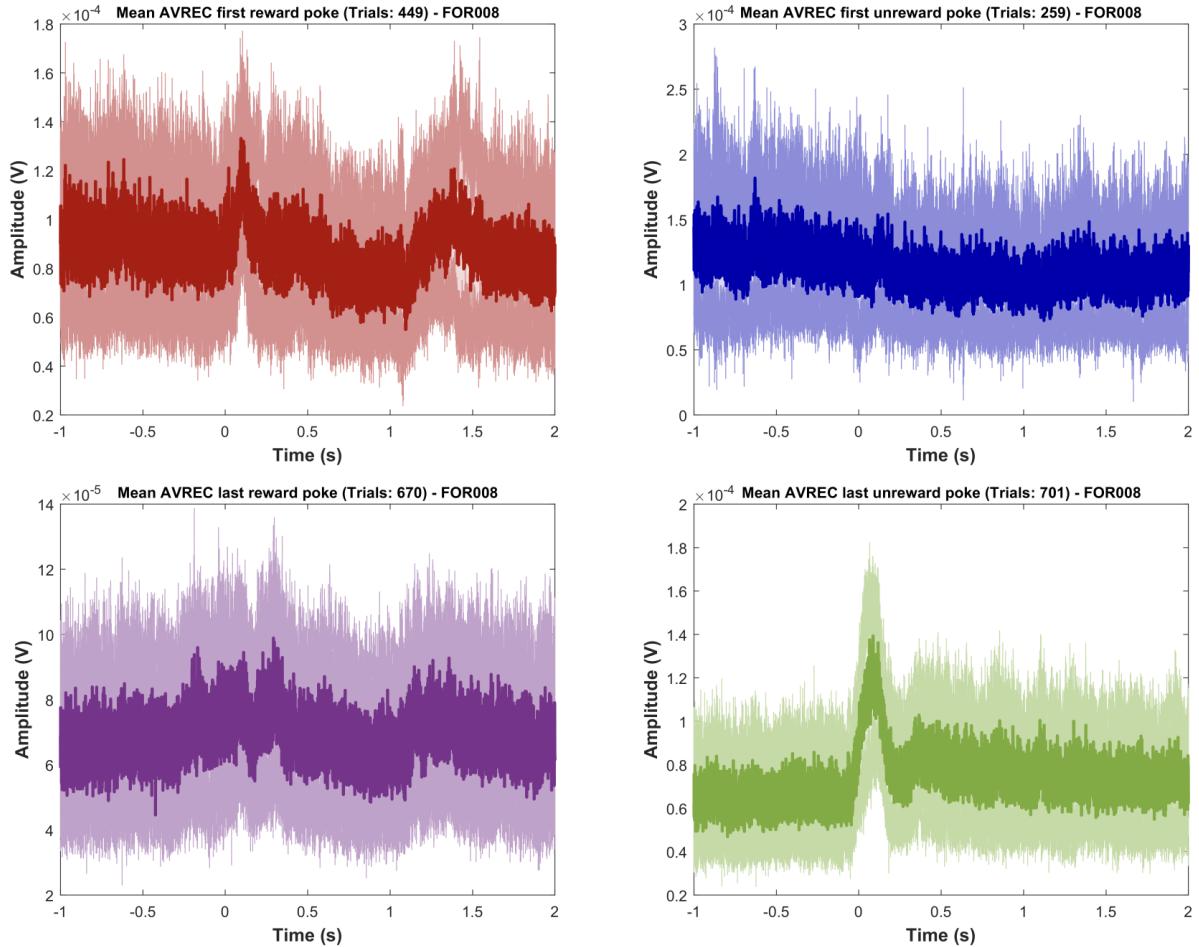
# **APPENDIX**

## 8 Appendix

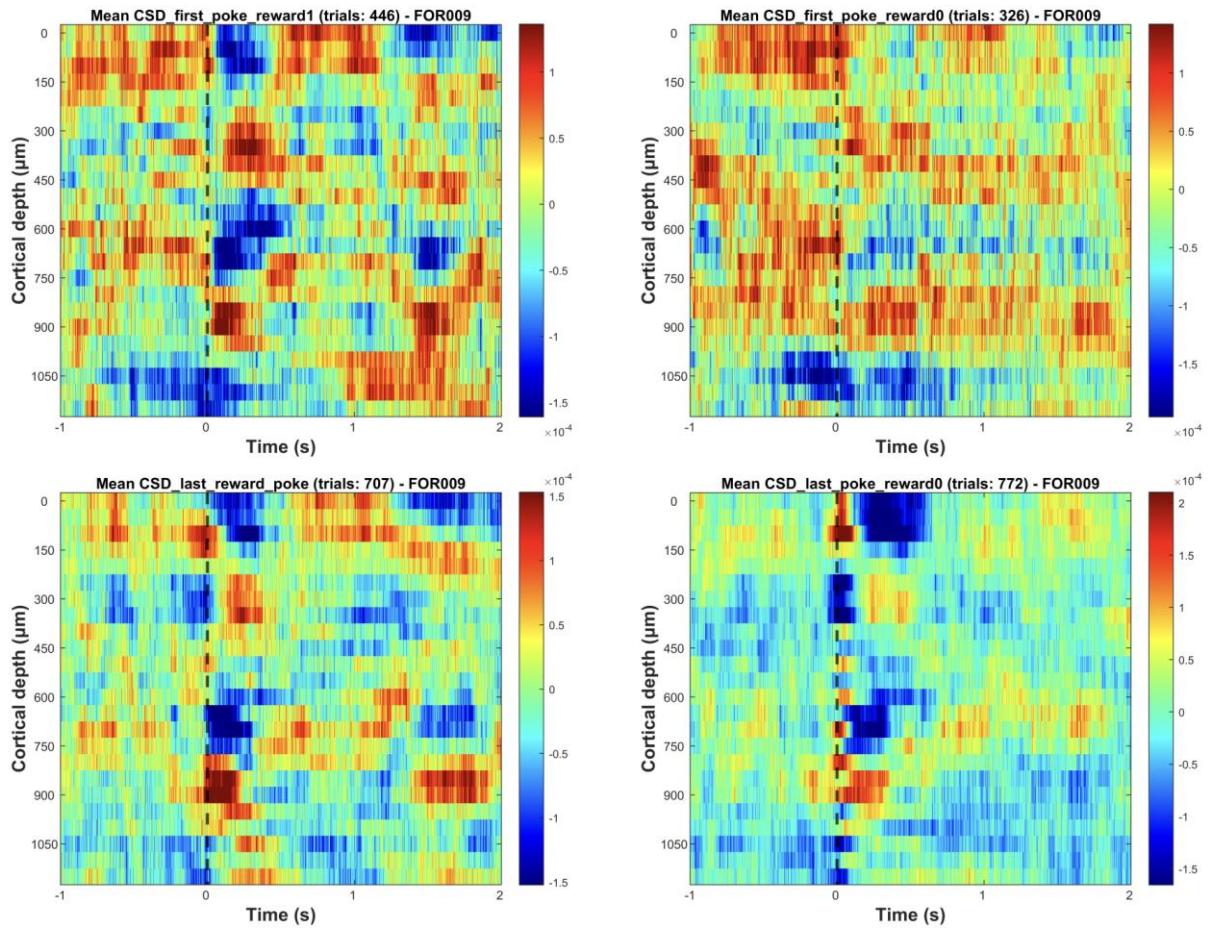
### 8.1 Individual CSD profiles and corresponding AVRECs



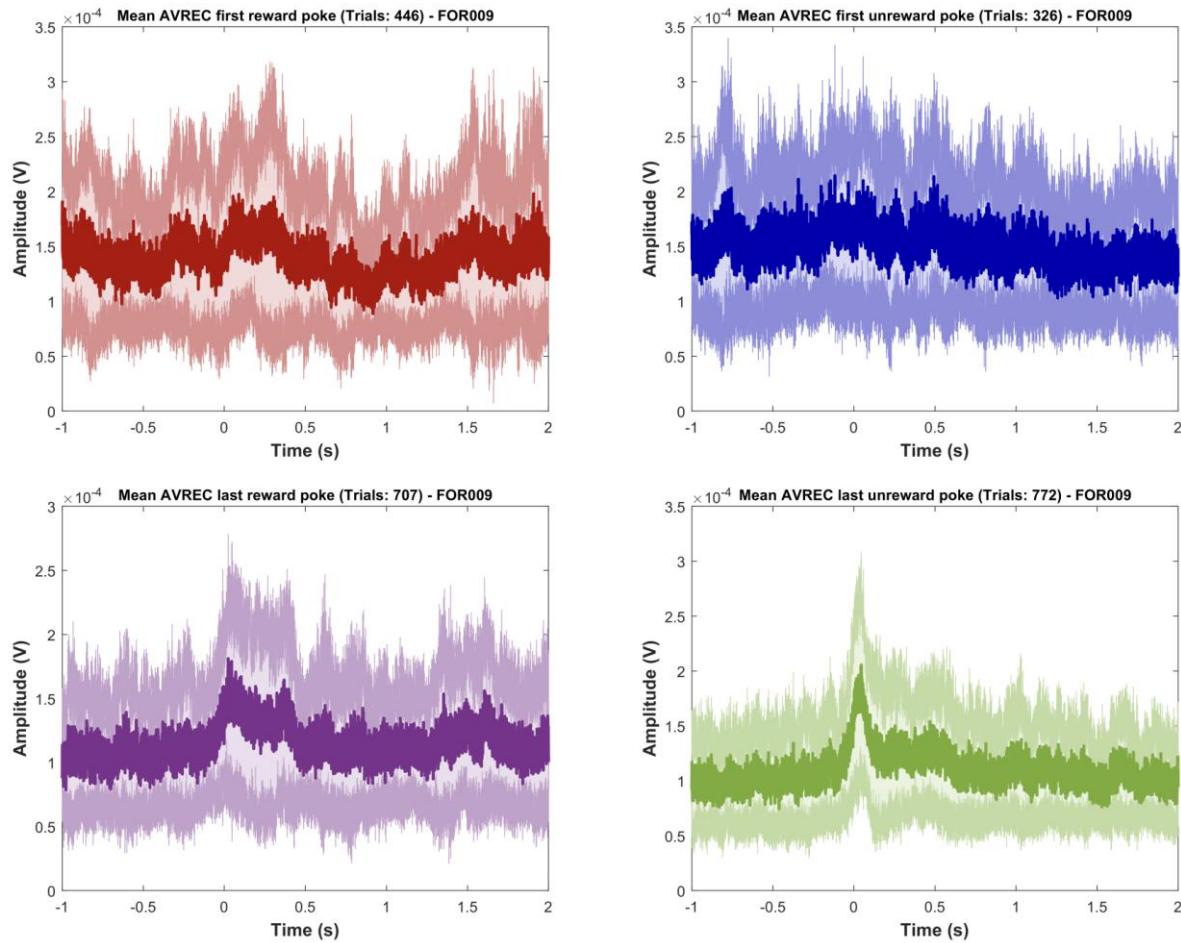
**Figure\_S1: Individual current source density (CSD) profile (Animal-1).** Distinct motor and reward related spatiotemporal neural activity in frontal field A. The selected epochs represent -1 to +2 seconds from the end of the poke (black dashed line,  $t=0$ ). The selected time interval was taken for four different events (pokes) and its corresponding consequence (reward): (top left) first poke with reward, (top right) first poke without reward, (bottom left) last rewarded poke, and (bottom right) last poke without reward. Each figure is scaled independently to better visualize the layer-specific activation profiles.



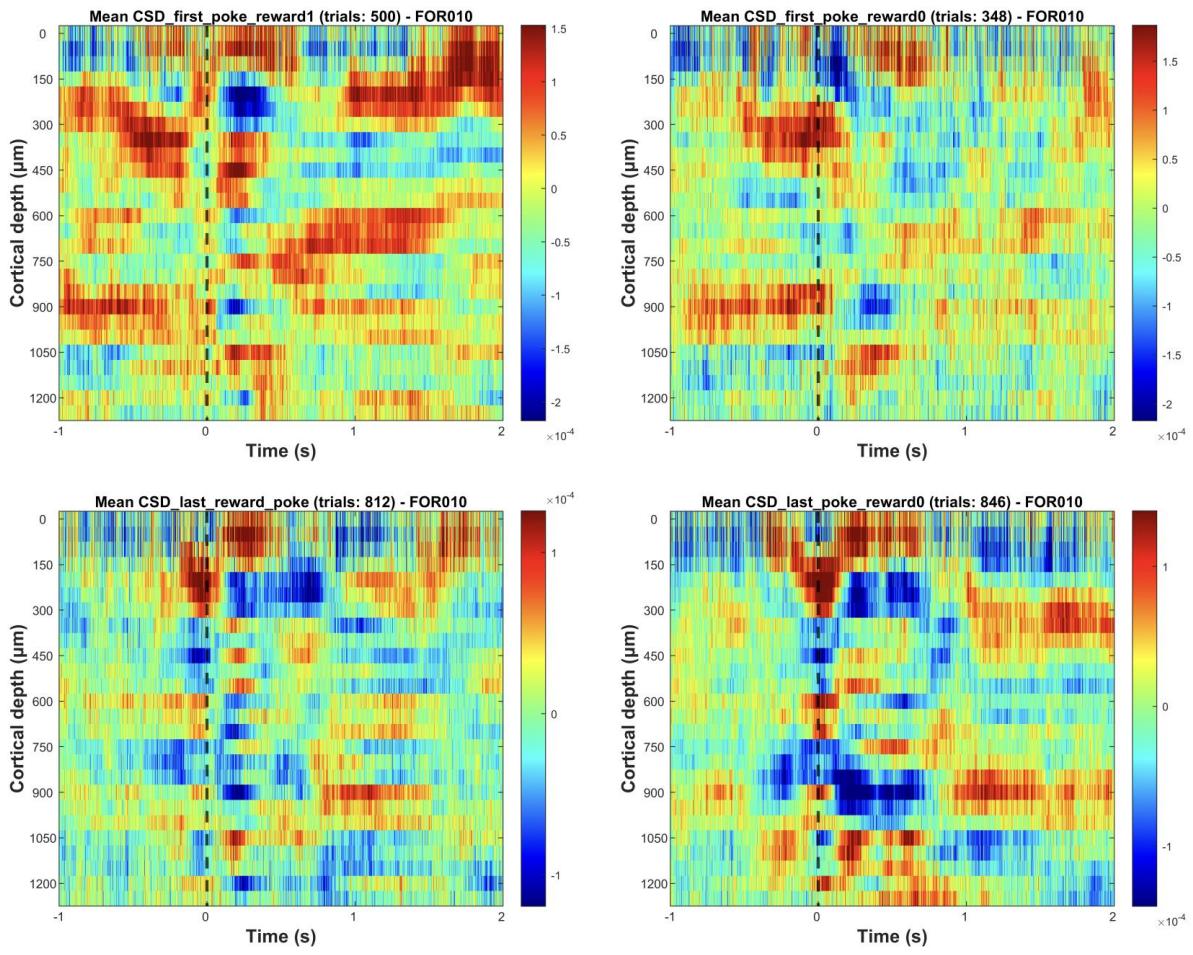
**Figure\_S2: Individual average rectified signals (AVREC) (Animal-1).** AVREC displays the overall frontal cortical activity, revealing distinct motor and reward-related signals. The mean average rectified waveform (depicted in bold colours) together with its standard error (shown in lighter shades) is plotted for the selected time intervals (epochs). These epochs span from one second before to two seconds after the end of the poke ( $t=0$ ). AVRECs are presented for four distinct pokes: first poke with reward (top left), first poke without reward (top right), last rewarded poke (bottom left), and last poke without reward (bottom right).



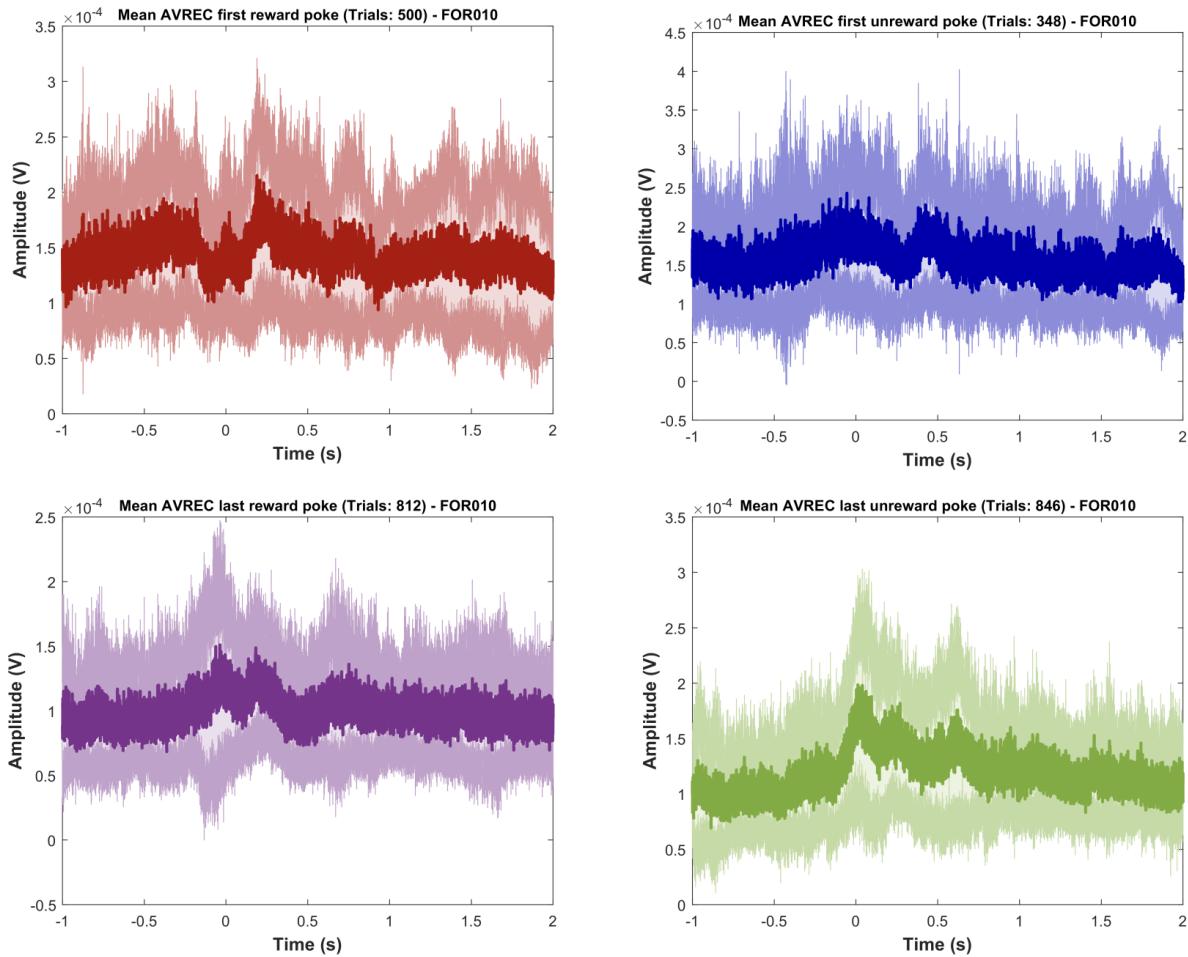
**Figure\_S3: Individual current source density (CSD) profile (Animal-2).** Distinct motor and reward related spatiotemporal neural activity in frontal field A. The selected epochs represent -1 to +2 seconds from the end of the poke (black dashed line,  $t=0$ ). The selected time interval was taken for four different events (pokes) and its corresponding consequence (reward): (top left) first poke with reward, (top right) first poke without reward, (bottom left) last rewarded poke, and (bottom right) last poke without reward. Each figure is scaled independently to better visualize the layer-specific activation profiles.



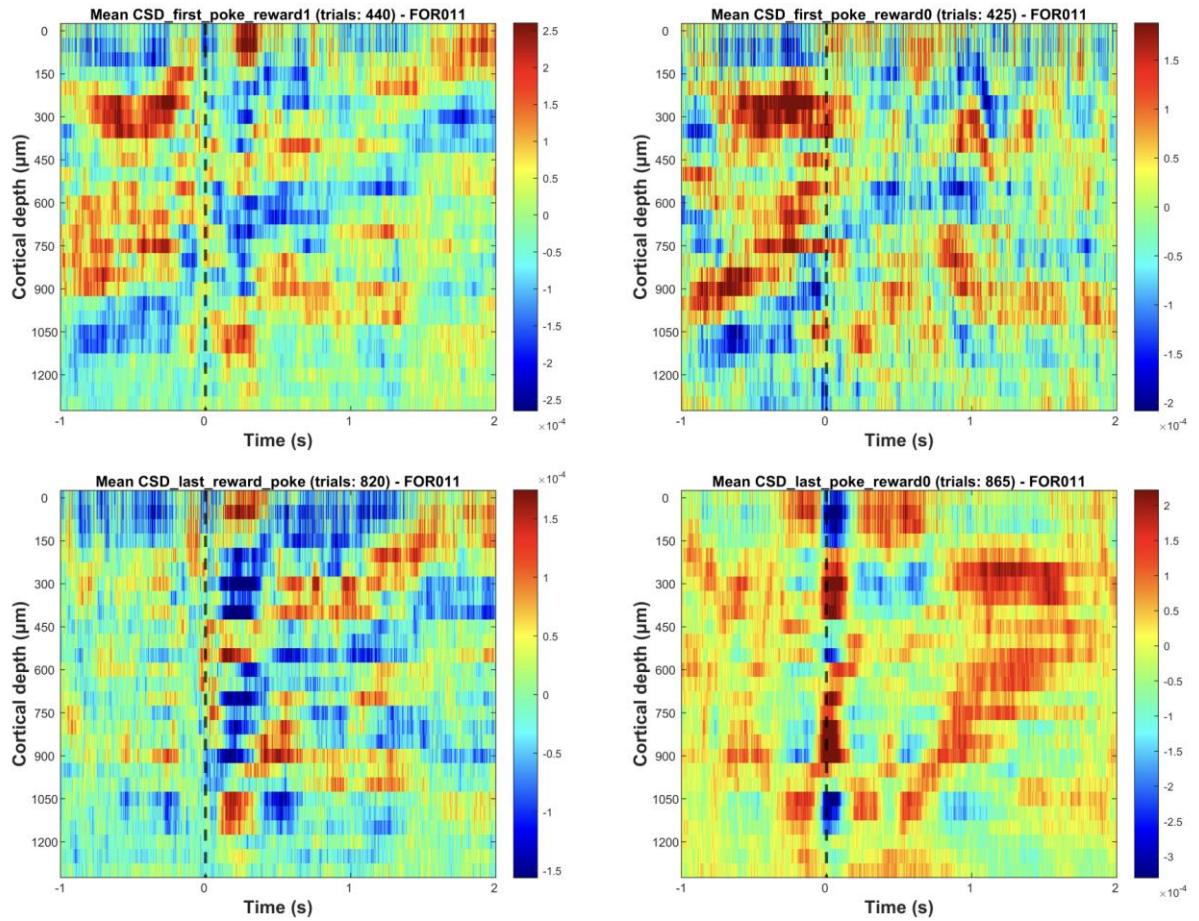
**Figure\_S4: Individual average rectified signals (AVREC) (Animal-2).** AVREC displays the overall frontal cortical activity, revealing distinct motor and reward-related signals. The mean average rectified waveform (depicted in bold colours) together with its standard error (shown in lighter shades) is plotted for the selected time intervals (epochs). These epochs span from one second before to two seconds after the end of the poke ( $t=0$ ). AVRECs are presented for four distinct pokes: first poke with reward (top left), first poke without reward (top right), last rewarded poke (bottom left), and last poke without reward (bottom right).



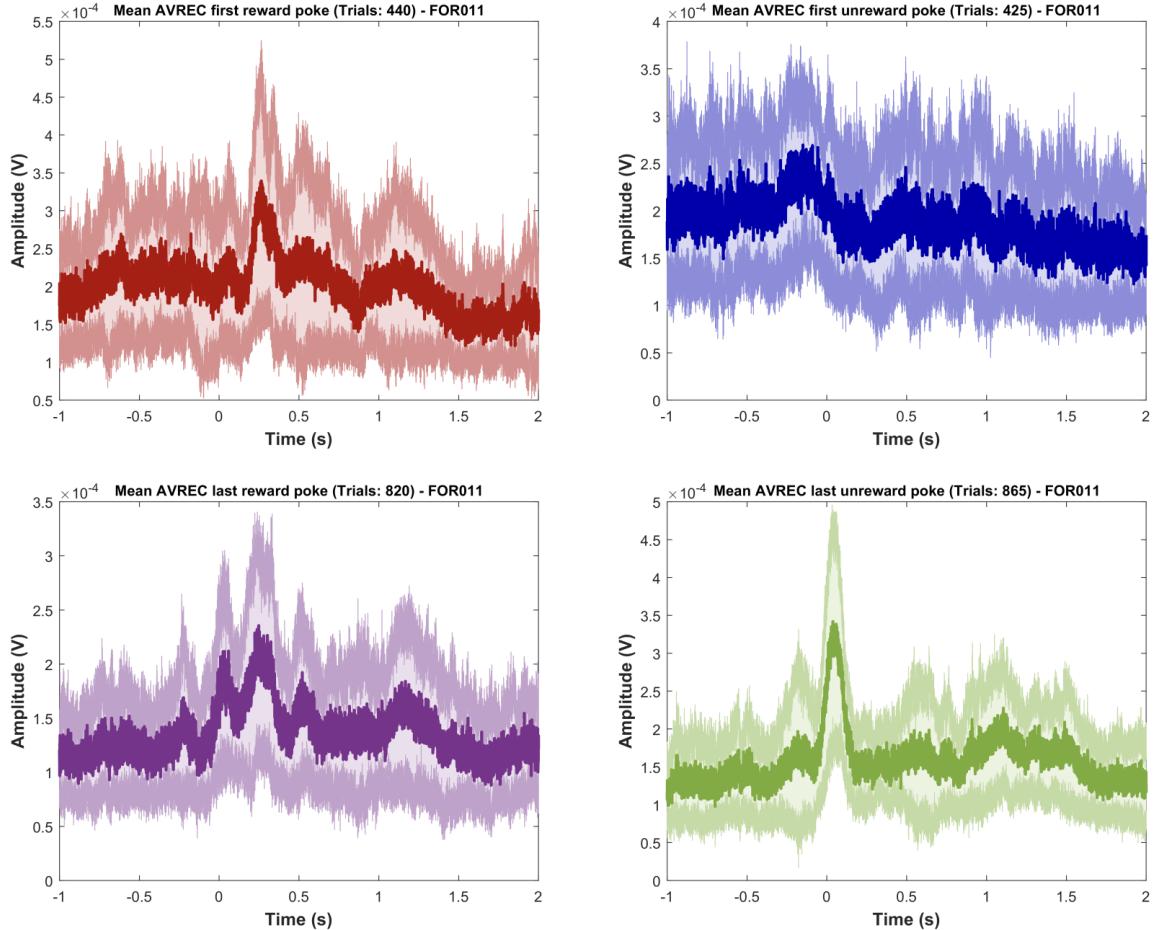
**Figure\_S5: Individual current source density (CSD) profile (Animal-3).** Distinct motor and reward related spatiotemporal neural activity in frontal field A. The selected epochs represent -1 to +2 seconds from the end of the poke (black dashed line,  $t=0$ ). The selected time interval was taken for four different events (pokes) and its corresponding consequence (reward): (top left) first poke with reward, (top right) first poke without reward, (bottom left) last rewarded poke, and (bottom right) last poke without reward. Each figure is scaled independently to better visualize the layer-specific activation profiles.



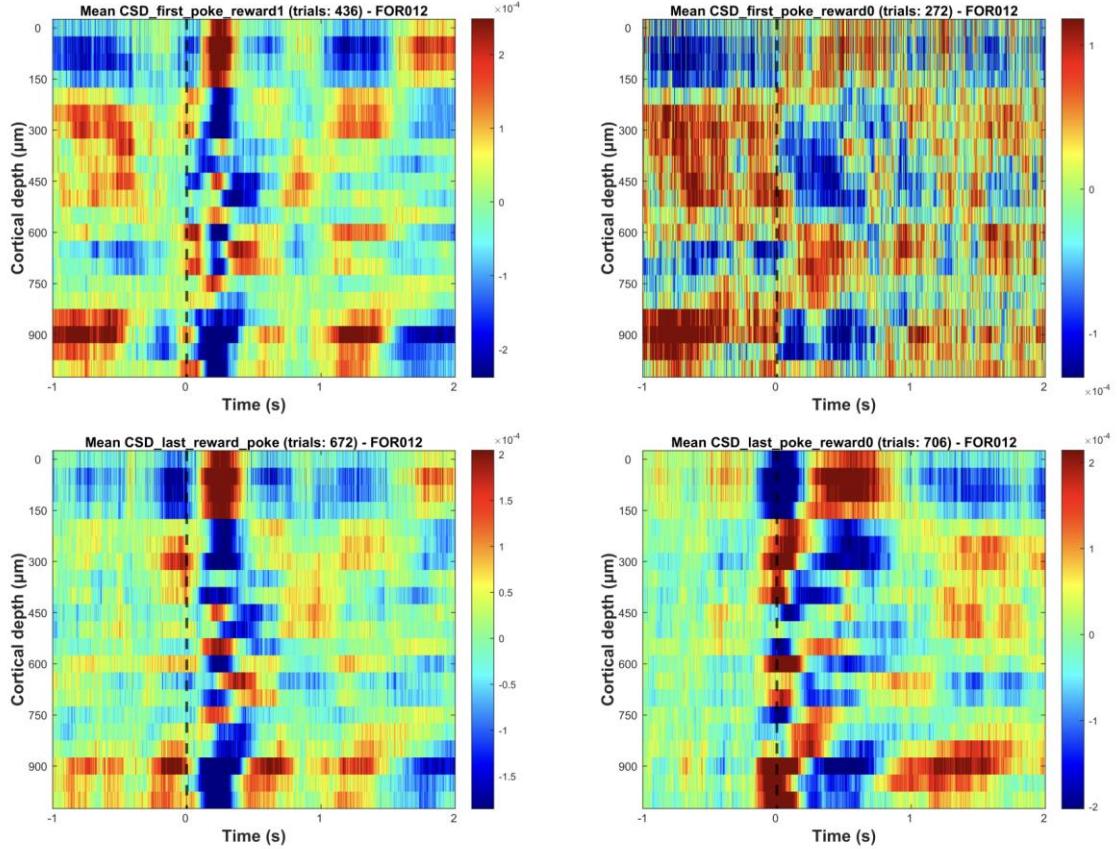
**Figure\_S6: Individual average rectified signals (AVREC) (Animal-3).** AVREC displays the overall frontal cortical activity, revealing distinct motor and reward-related signals. The mean average rectified waveform (depicted in bold colours) together with its standard error (shown in lighter shades) is plotted for the selected time intervals (epochs). These epochs span from one second before to two seconds after the end of the poke ( $t=0$ ). AVRECs are presented for four distinct pokes: first poke with reward (top left), first poke without reward (top right), last rewarded poke (bottom left), and last unrewarded poke (bottom right).



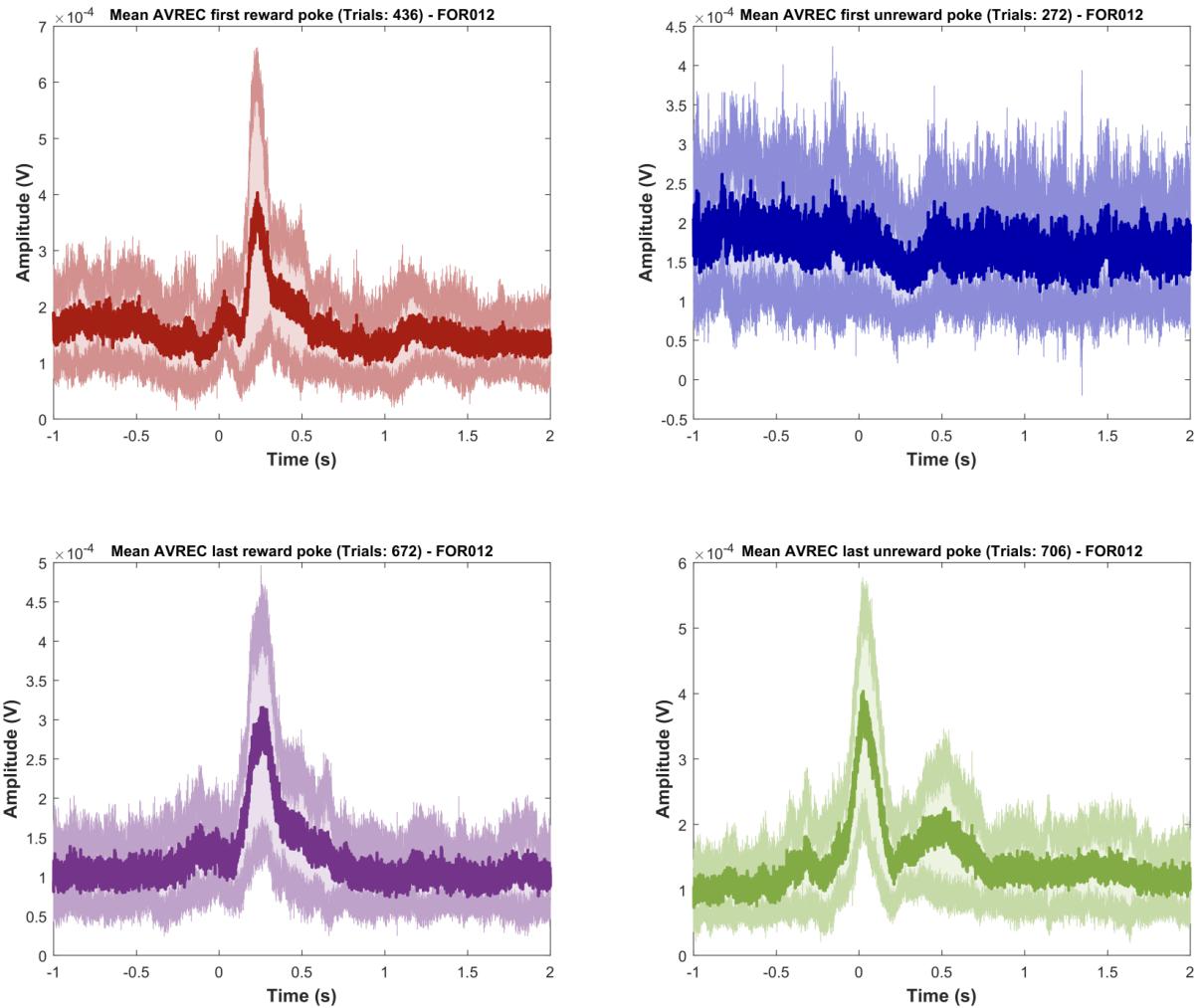
**Figure\_S7: Individual current source density (CSD) profile (Animal-4).** Distinct motor and reward related spatiotemporal neural activity in frontal field A. The selected epochs represent -1 to +2 seconds from the end of the poke (black dashed line,  $t=0$ ). The selected time interval was taken for four different events (pokes) and its corresponding consequence (reward): (top left) first poke with reward, (top right) first poke without reward, (bottom left) last rewarded poke, and (bottom right) last poke without reward. Each figure is scaled independently to better visualize the layer-specific activation profiles.



**Figure\_S8: Individual average rectified signals (AVREC) (Animal-4).** AVREC displays the overall frontal cortical activity, revealing distinct motor and reward-related signals. The mean average rectified waveform (depicted in bold colours) together with its standard error (shown in lighter shades) is plotted for the selected time intervals (epochs). These epochs span from one second before to two seconds after the end of the poke ( $t=0$ ). AVRECs are presented for four distinct pokes: first poke with reward (top left), first poke without reward (top right), last rewarded poke (bottom left), and last poke without reward (bottom right).



**Figure\_S9: Individual current source density (CSD) profile (Animal-5).** Distinct motor and reward related spatiotemporal neural activity in frontal field A. The selected epochs represent -1 to +2 seconds from the end of the poke (black dashed line,  $t=0$ ). The selected time interval was taken for four different events (pokes) and its corresponding consequence (reward): (top left) first poke with reward, (top right) first poke without reward, (bottom left) last rewarded poke, and (bottom right) last poke without reward. Each figure is scaled independently to better visualize the layer-specific activation profiles.



**Figure\_S10: Individual average rectified signals (AVREC) (Animal-5).** AVREC displays the overall frontal cortical activity, revealing distinct motor and reward-related signals. The mean average rectified waveform (depicted in bold colours) together with its standard error (shown in lighter shades) is plotted for the selected time intervals (epochs). These epochs span from one second before to two seconds after the end of the poke ( $t=0$ ). AVRECs are presented for four distinct pokes: first poke with reward (top left), first poke without reward (top right), last rewarded poke (bottom left), and last poke without reward (bottom right).