# 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. This task, designed to simulate a classic exploration-exploitation trade-off, featured a scenario where food rewards at a specific spout diminished over time, compelling the gerbils to decide whether to switch to an alternative spout with potentially greater food availability.

## Behavioural analysis

The behavioural experiment was designed to investigate attentional resource allocation in a probabilistic foraging paradigm. Mongolian gerbils had to decide when to switch from a depleting food spout to another that may offer more food. This task was designed to mimic the exploration-exploitation trade-off that is commonly observed in natural foraging behaviour. The task is particularly suited to study probabilistic decision-making as it requires subjects to make decisions based on uncertain outcomes, which is a fundamental aspect of many real-world decisions.

### Performance of the foraging behaviour

A total of 20 foraging sessions were conducted on 20 consecutive days. Every foraging session consists of N trials, with each trial comprising a sequence of nose pokes (Fig.1A). Each animal (n=5) performed on an average 52 trials per session. The average median inter-poke interval a after rewarded poke (M= 330.25 ms) was significantly higher than the one after an unrewarded poke (M= 85 ms, One-way ANOVA, p<0.01) (Fig.1B). Based on the animal’s performance, only 15 sessions were considered for the analysis resulting to a grand total of 3890 trials from all the animals. Animals were food deprived to instil a motivation to perform the task. Animal body weights were constantly monitored after every session to ensure that they didn’t drop below 15% of their original body weight measured before the start of the experiment (Fig.1C).

In order to see if the animal has learnt the foraging task and reduce the possibility of random exploration, we looked at their travel time. Here, travel time is defined as the time between the end of last poke in a trial and the start of the first poke in the succeeding trial (Fig1B, Eq.2). The consistent travel time after the first five sessions indicate that after each trial, the animals were directed towards the other spout and not randomly exploring the cage (Fig.1D). Therefore, only sessions from 6 to 20 were considered for further behavioural and electrophysiological analyses.

### 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 spoutes. Gerbils showed increased resident times with increasing spout quality (Fig.1E, One way ANOVA with Bonferroni correction, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001). The mean resident time for high quality spoutes (Mean = 21.52 ± 6.16 s) was significantly higher than medium (Mean = 16.49 ± 4.93 s, One-way ANOVA with Bonferroni correction, p<0.01) and low-quality spoutes (Mean = 11.84 ± 3.33 s, One-way ANOVA with Bonferroni correction, p<0.01).

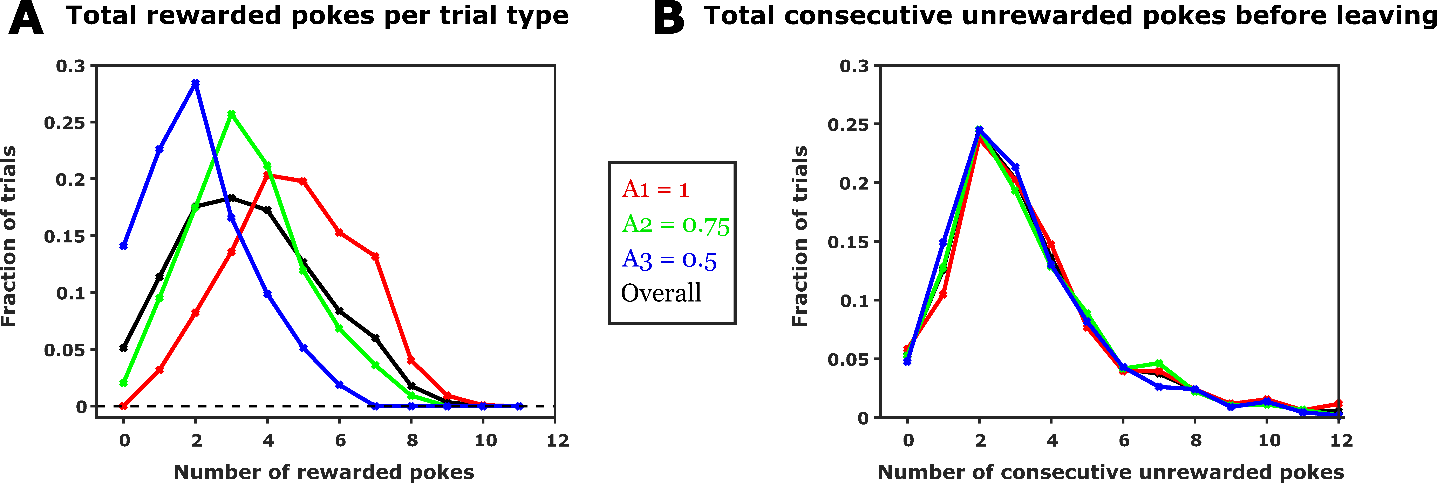
Consequently, they also on an average received a greater number of rewards as the quality of spoutes increased (Fig.1G, One way ANOVA with Bonferroni correction, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001). The number of rewards for high quality spoutes (Mean = 4.63 ± 0.15) was significantly higher than medium (Mean = 3.42 ± 0.18, One-way ANOVA with Bonferroni correction, p<0.001) and low-quality spoutes (Mean = 2.15 ± 0.12, One-way ANOVA with Bonferroni correction, p<0.001).

### Spout-leaving behaviour

Now that we have seen that the animal’s performance differs based on different trial types, it is important to investigate the critical decision-making behaviour during the exploration-exploitation dilemma situation i.e., when is the right time to leave the current spout and explore the other option? This is important because every trial begins with exploiting the current spout for rewards and slowly as the rewards gets exhausted, the important decision is to decide when to leave the current spout to explore the other.

From different possible decision rules such as fixed-time rule, fixed-number rule, incremental rule and giving up time rule (Introduction Figure), our previous collaborative cross-species study (Lasse, in prep) revealed that the Gerbils followed the *giving up time rule* (GUT rule). GUT rule states that a forager only tolerates a certain amount of time without a new reward since the last reward. Once this temporal threshold is exceeded, the forager leaves the spout. Each new reward, on the other hand, resets the tendency to stay. GUT rule does not require a prior judgment/knowledge about the spout-quality. In our study, GUT will be the time taken from the last rewarded poke to the end of the last poke in a trial. Although, every trial starts with hunger – the main motivation to perform the task – the motivation may reduce as the satiation increases. This may lead to task-unrelated behaviours such as grooming to which the GUT is sensitive. To overcome this, we chose the consecutive number of unrewarded pokes the animal makes before leaving the current spout as an alternative measure to study the spout leaving behaviour.

Fig.2B shows that the distinction between trials starting with different starting reward probabilities gets lost when it comes to the consecutive unrewarded pokes made before switching. In other words, it shows that the animals irrespective of the starting reward probability and the total number of rewards received in a trial maintains a consistent number of consecutive unrewarded pokes before leaving a particular spout. This behavioural consistency in spout-leaving decision may reflect the possibility that the animals follow the GUT rule or a similar rule which requires an inference about the hidden reward structure based on their learning of the task.



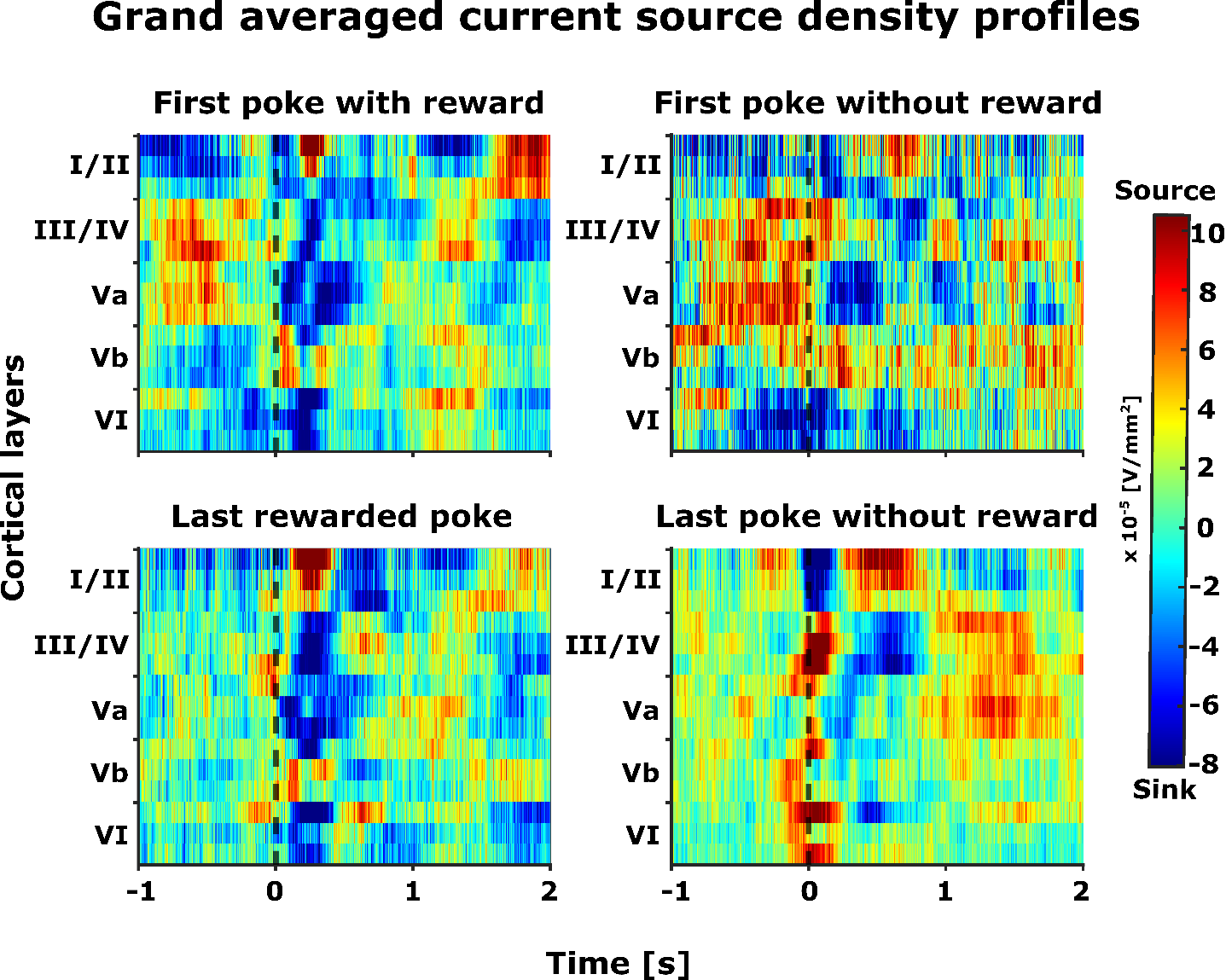
**Figure 2: Performance of the foraging behaviour. A** shows the median ± SD of travel time per session remains consistent and less than 5 seconds. Travel time is calculated as the time between the end of last poke in a trial to the start of the first poke in the succeeding trial. **B** shows the total number of rewarded pokes across trials for different starting probabilities (A = 1, 0.75 and 0.5). **C** shows the total consecutive unrewarded pokes before leaving a spout in each trial having different starting probabilities.

## Distinct spatiotemporal activity patterns in the frontal field A

To determine whether the FrA encodes distinct activity patterns reflective of the animal’s behavior, four specific time points were selected for analysis. 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. 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 current source density (CSD) profile corresponding to these time points reveals 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 (as depicted in Figure 3). 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 (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.



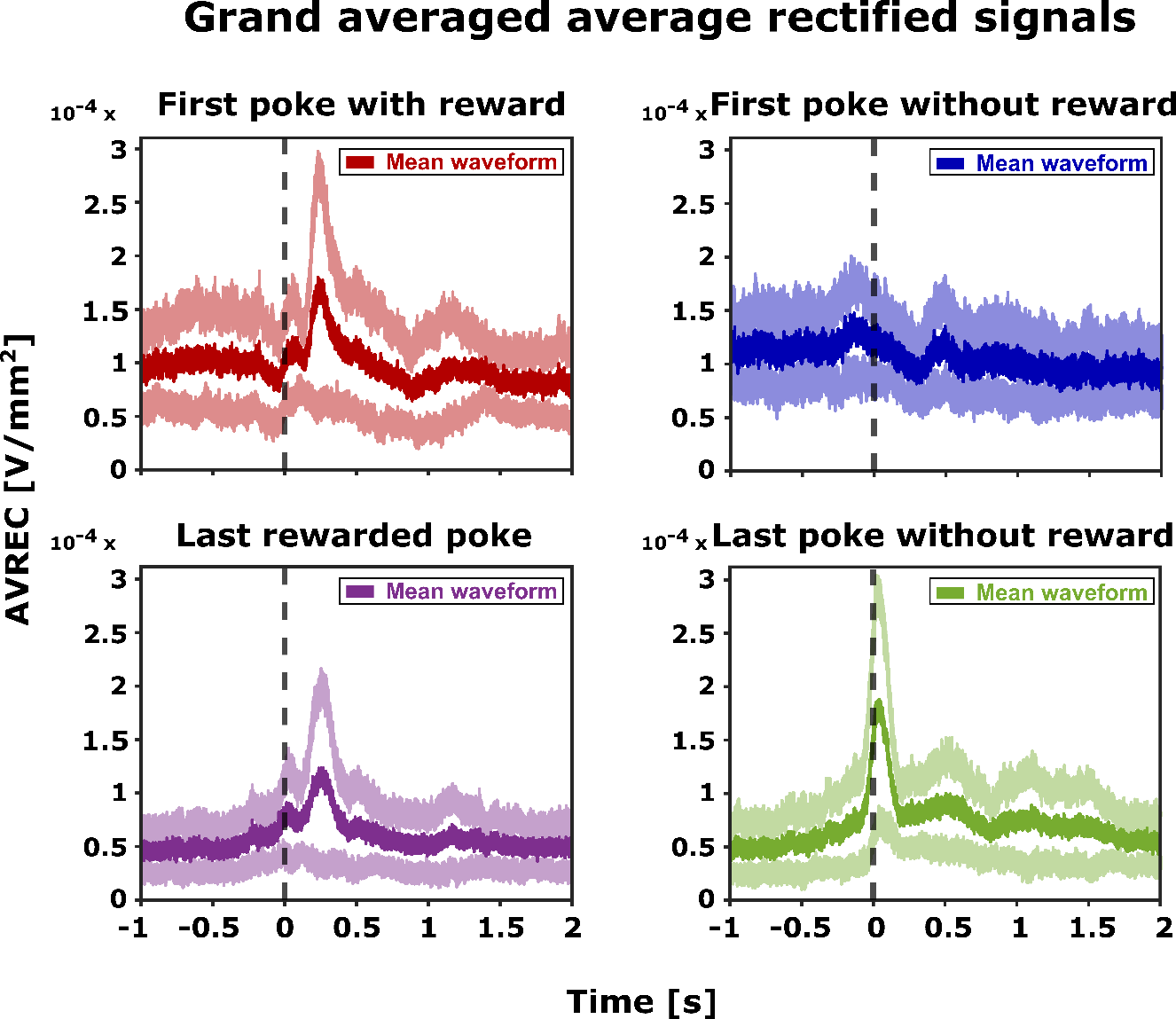
**Figure 3: Grand averaged current source density (CSD) profiles (n=5)** **– Distinct motor and reward related spatiotemporal neural activity in frontal field A**. The selected epochsrepresent -1 to +2 seconds from the end of the poke (black dashed line, t=0). The selected time interval was taken for 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.

### Overall frontal activity patterns in FrA

To assess the overall frontal cortical activity, the current source density (CSD) signals were rectified and averaged across all laminar electrodes, which intentionally obscured the spatial information. The grand averaged average rectified signals (AVREC) derived from the CSD profiles exhibit distinct overall activity patterns associated with different pokes, as illustrated in Figure 4. Notably, 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). 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 in amplitude within the first 100 ms, setting it apart from the other pokes.

The average rectified signals (AVREC) from the CSD profiles also show distinct motor (Fig.4, peak around end of the poke (t=0)) and reward related signals (Fig.4, peak around 250 ms from the end of the poke). From the perspective of expectation and evaluation of reward, it can be seen that in FrA, the encoding of the expected (towards the end of pokes) and received reward (subsequently) shows different activity patterns during the exploitation (first poke until last rewarded poke) and exploration (last poke) phases of the trial. Both the rewarded pokes (first rewarded and last rewarded pokes) show a similar waveform where there is an early peak in amplitude immediately after the end of the poke that may encode for the expectation of a reward followed by a slightly increased peak occurring in less than 250 ms from the end of the pokes that may encode for the reward evaluation. On the other hand, in the unrewarded pokes (first unrewarded and last poke), the reward evaluation shows a dip followed by a later peak in amplitude around 500 ms that may encode for a prediction error from the deeper layers. Compared to all other pokes, the last poke which is the starting point of exploration phase where the animal decides to switch to the other spout, shows a heightened frontal activity immediately after the end of the poke (<100 ms). These distinct activation patterns for different time points seen in Fig.3 and Fig. show that the frontal field A encodes not just the motor activity but also the reward related activity (expectation and prediction error).



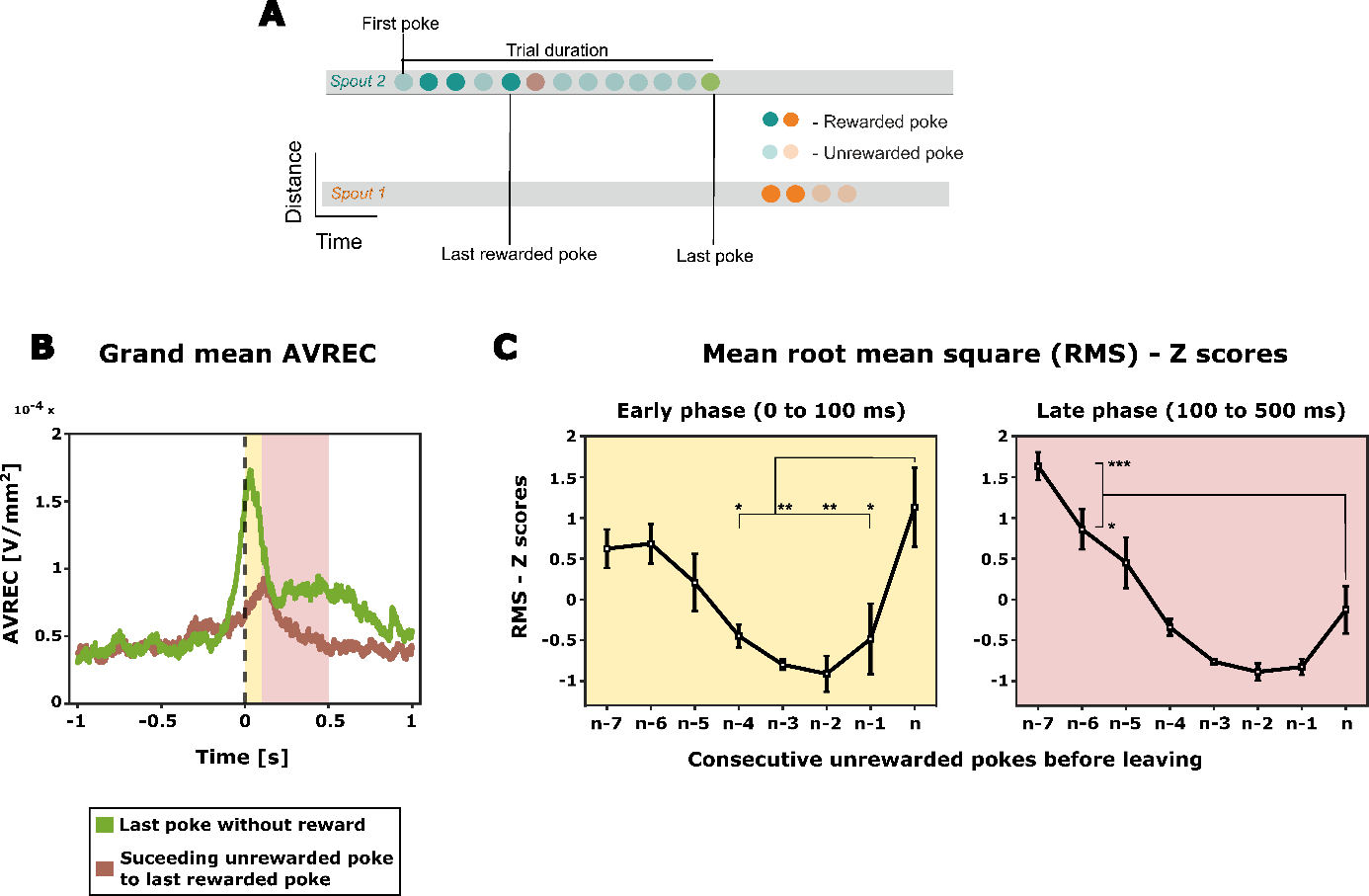
**Figure 4: Grand averaged AVREC (n=5)** displays the overall frontal cortical activity, revealing distinct motor and reward-related signals. The mean average rectified waveform (depicted in bold colours) together with its standard error (shown in lighter shades) is plotted for the selected time intervals (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).

## Shifts in frontal activity patterns: From Exploitation to Exploration

We have established that Frontal Field A (FrA) encodes distinct neural activity patterns for various pokes, as demonstrated in Figures 3 and 4, which include the first poke (rewarded and unrewarded), the last rewarded poke, and the last poke. Notably, the last poke exhibits a neural activity pattern that is distinct from the others (Figure 4). 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 5A). This analysis is pivotal for understanding the decision-making process, specifically how an animal determines that a poke will be the last 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 5A) , 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 5B). 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 5B).

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, were differentiated (Figure 5B, with phases highlighted in yellow and light pink, respectively). The contrast is most pronounced during the transition from the exploitation phase, with reduced yet notable early activity beyond 100 ms, to the exploration phase, where the early phase activity is markedly elevated (Figure 5B). This shift in activity pattern 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.



**Figure 5: Shift 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** – Displays the 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).

### 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. This RMS was calculated for all unrewarded pokes occurring between the last rewarded poke and the last poke of the session (Figure 5B and 5C) and was z-normalized within each subject.

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). Conversely, the late phase displayed a consistent decrease in frontal activity, with a minor elevation observed just before the transition to exploration (nth poke).

The non-linear trend of activity change during the early phase indicates the involvement of a complex processing strategy underlying the inference-based decision-making behaviour seen in the animals (Fig.2C).

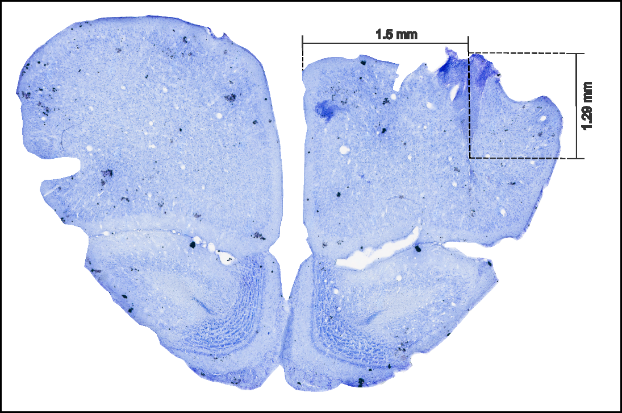
## Layer specific spatiotemporal activity patterns in the frontal field A

### Histological Confirmation of Laminar Electrode Positioning in Frontal Field A

For the accurate localization of neuronal activity across all cortical layers, it is imperative to position the laminar depth electrode perpendicularly 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, is capable of capturing 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.

Reference to the Gerbil brain atlas (Radtke-Schuller et al., 2016) situates the target region, Frontal field A (FrA), at coordinates 4.65 to 5 mm anterior and 1.5 mm lateral to the bregma. Figure 1 presents a histological slice from one specimen, located at 4.85 mm anterior and 1.5 mm lateral to the bregma, verifying 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 1: 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

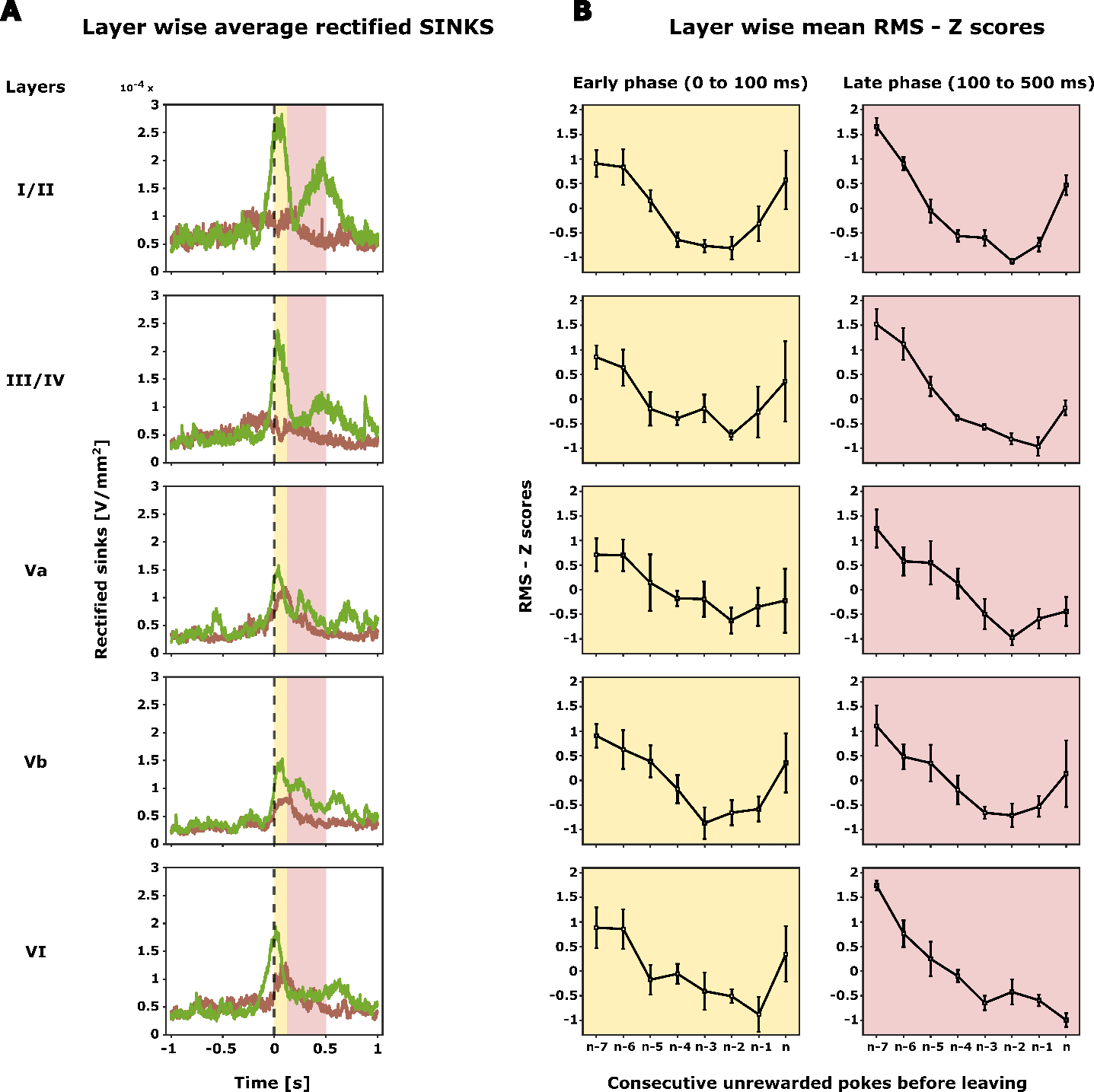
### Distinct layer wise shift in frontal activity patterns: Exploitation to exploration

Five distinct cortical layers were identified from the current source profiles of the laminar recordings (Fig.3). The source signal was removed and only the sinks were considered to ensure that the signal is contributed only by layer specific local excitatory synaptic populations (See methods section). Comparison of layer-wise averaged sink activity between the last poke and first unrewarded poke succeeding the last rewarded poke indicates that the large difference in activity patterns observed in overall frontal activity (Fig,5A) is mainly contributed by the large difference seen in the superficial layers (layer I/II and III/IV) (Fig.6A). Further, shortly before the decision of the animal to explore the other spout, particularly layers I/II and III/IV showed two prominent activity peaks in amplitude directly at the retraction of the spout (early phase, 0-100 ms) and afterwards (late phase, 100-500 ms).

In the individual poke RMS analysis, similar to overall frontal activity, there is an initial decrease in all cortical layers (Fig.6B). However, at the beginning of the exploration phase (n-2 to nth poke), the activity of upper layers I/II and III/IV, as well as Vb, increases. The increase is mainly observed in the early phase following the nose poke, and then with regard to the evaluation of the absence of rewards (Fig.6B). In the late phase, with the exception of layer I/II and layer III/IV which showed a steep increase in activity just before the last poke, all the cortical layers showed a decrease in activity across all the pokes similar to that of overall frontal activity.

Through the analysis of current source density profiles from laminar recordings, we identified five distinct cortical layers (Figure 3). 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 5A), are primarily driven by variations in the superficial layers (layers I/II and III/IV) (Figure 6A). 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 6B). 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. Similar to 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 6: Layer specific cortical activity in frontal region A**. **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.