# 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.

## 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.

### Performance of the foraging behaviour

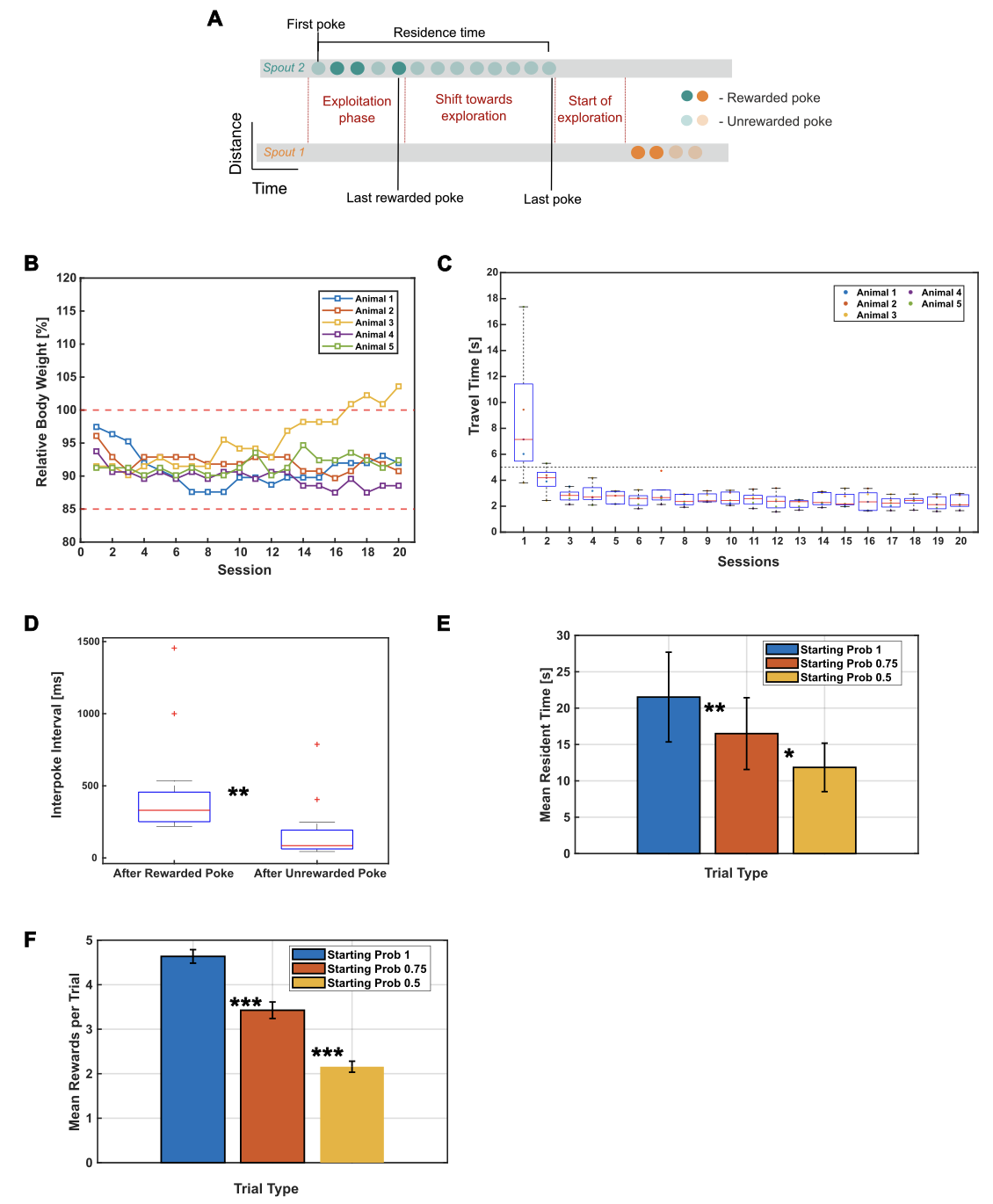
Over the course of 20 consecutive days, each of the five gerbils participated in daily foraging sessions, each comprising of N trials (Figure 1A). Each trial consists of a series of nose-pokes and the animals performed on an average 52 trials per session. From the 20 sessions, only 15 were deemed analytically 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 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 1B. 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 1C).

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 1D, 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.

#### 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 1E). Specifically, the mean residence time at high-quality spouts (Mean = 21.52 ± 6.16 s) exceeded that at medium (Mean = 16.49 ± 4.93 s) and low-quality spouts (Mean = 11.84 ± 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 (Figure 1F). Rewards from high-quality spouts (Mean = 4.63 ± 0.15) were significantly more frequent than those from medium (Mean = 3.42 ± 0.18) and low-quality spouts (Mean = 2.15 ± 0.12) (One-way ANOVA with Bonferroni correction, \*\*\* p < 0.001).

****

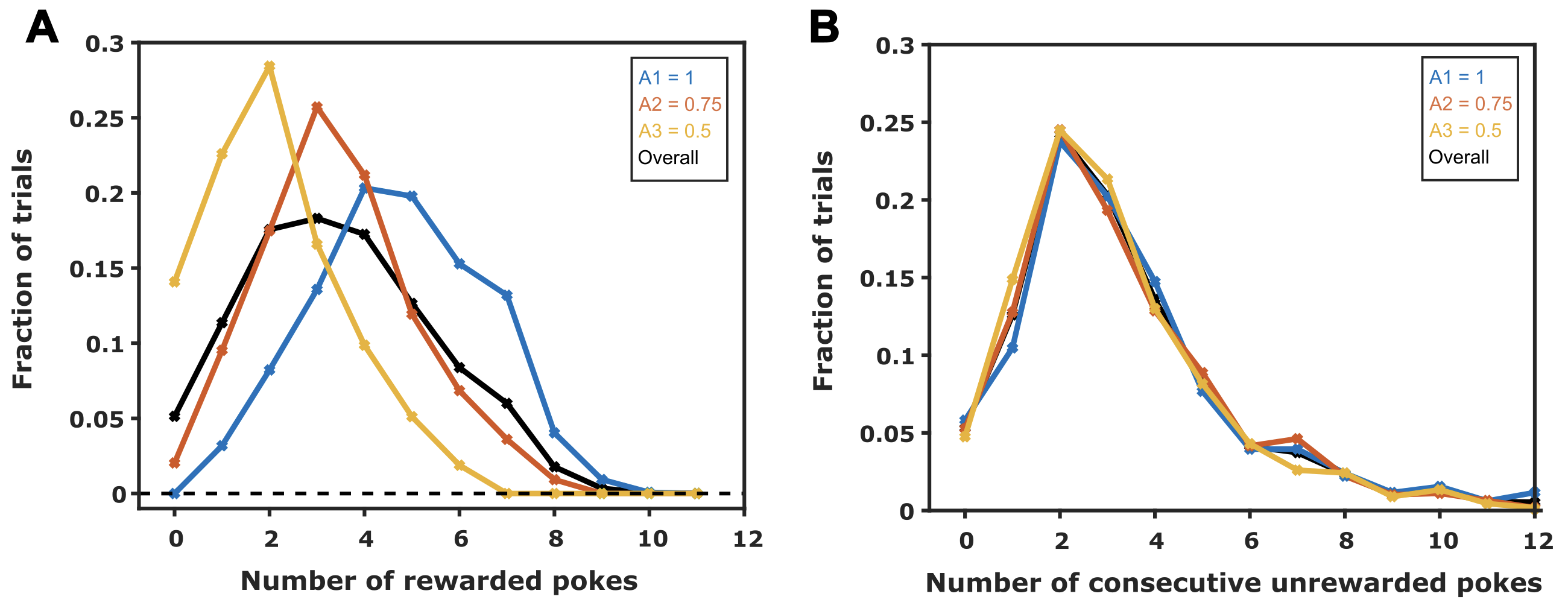
**Figure 1: Behavioural Metrics from the Probabilistic Foraging Task.** A) Schematic representation of the foraging task, illustrating the phases of exploitation and transition to exploration with the corresponding nose poke outcomes. B) Relative body weight percentage of individual animals over 20 sessions, 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. Data are represented as mean ± SEM, with statistical significance denoted by asterisks (\* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001).

#### Spout-leaving behaviour

Upon observing performance variations across trial types (different spout qualities), we investigated the decision-making behavior regarding the timing of spout-switching during the exploration-exploitation dilemma i.e., when is the right time to leave the current spout and explore the other option? This is important because each trial initiated with the exploitation of the current spout for rewards, transitioning to a critical decision on when to explore alternative options as the reward frequency diminished (Fig.1A).

Drawing from our previous collaborative study (Lasse, in prep), 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 Giving Up Time (GUT) is operationalized as the duration from the last rewarded nose poke to the final poke in a given 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 2B illustrates that despite different initial reward probabilities, 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 maintains 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

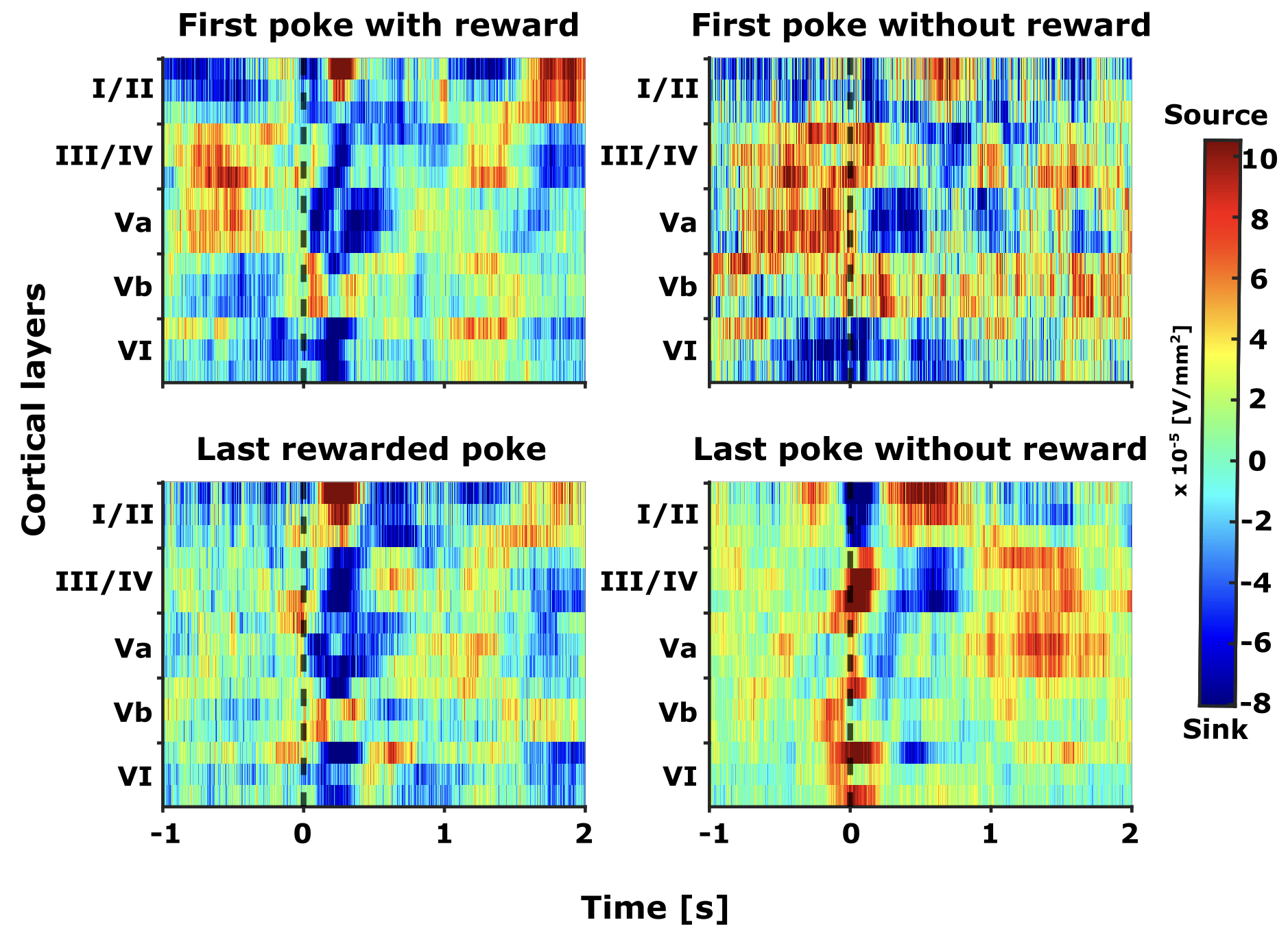
****

**Figure 2: 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 (A1 = 1, A2 = 0.75, A3 = 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.

## 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 (Fig.1A). 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.



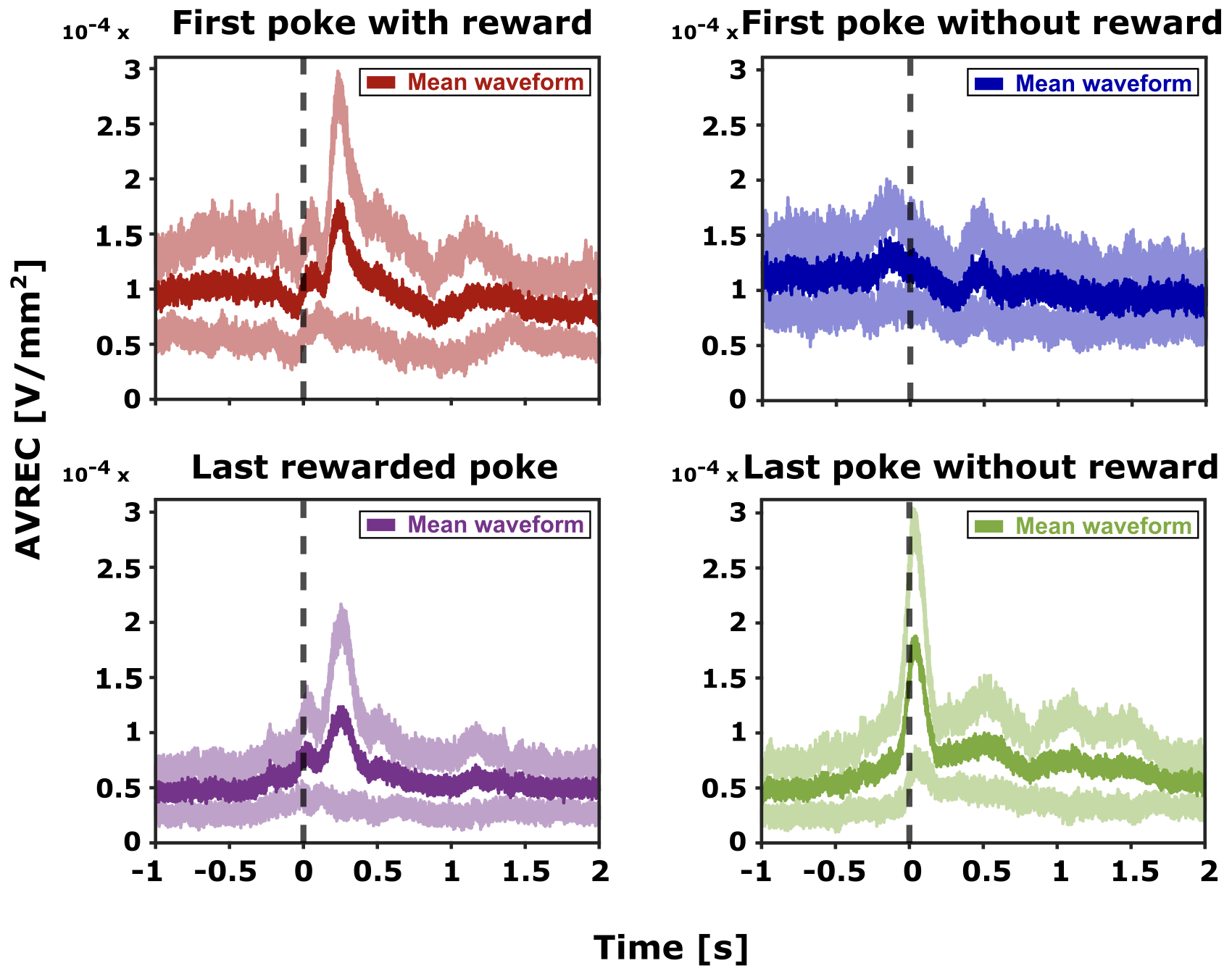
**Figure 3: Grand averaged current source density (CSD) profiles (n=5)** **– Distinct motor and reward related spatiotemporal neural activity in frontal field A**. The selected epochsrepresent -1 to +2 seconds from the end of the poke (black dashed line, t=0). The selected time interval was taken for 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 within the first 100 ms that is greater in amplitude, 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).



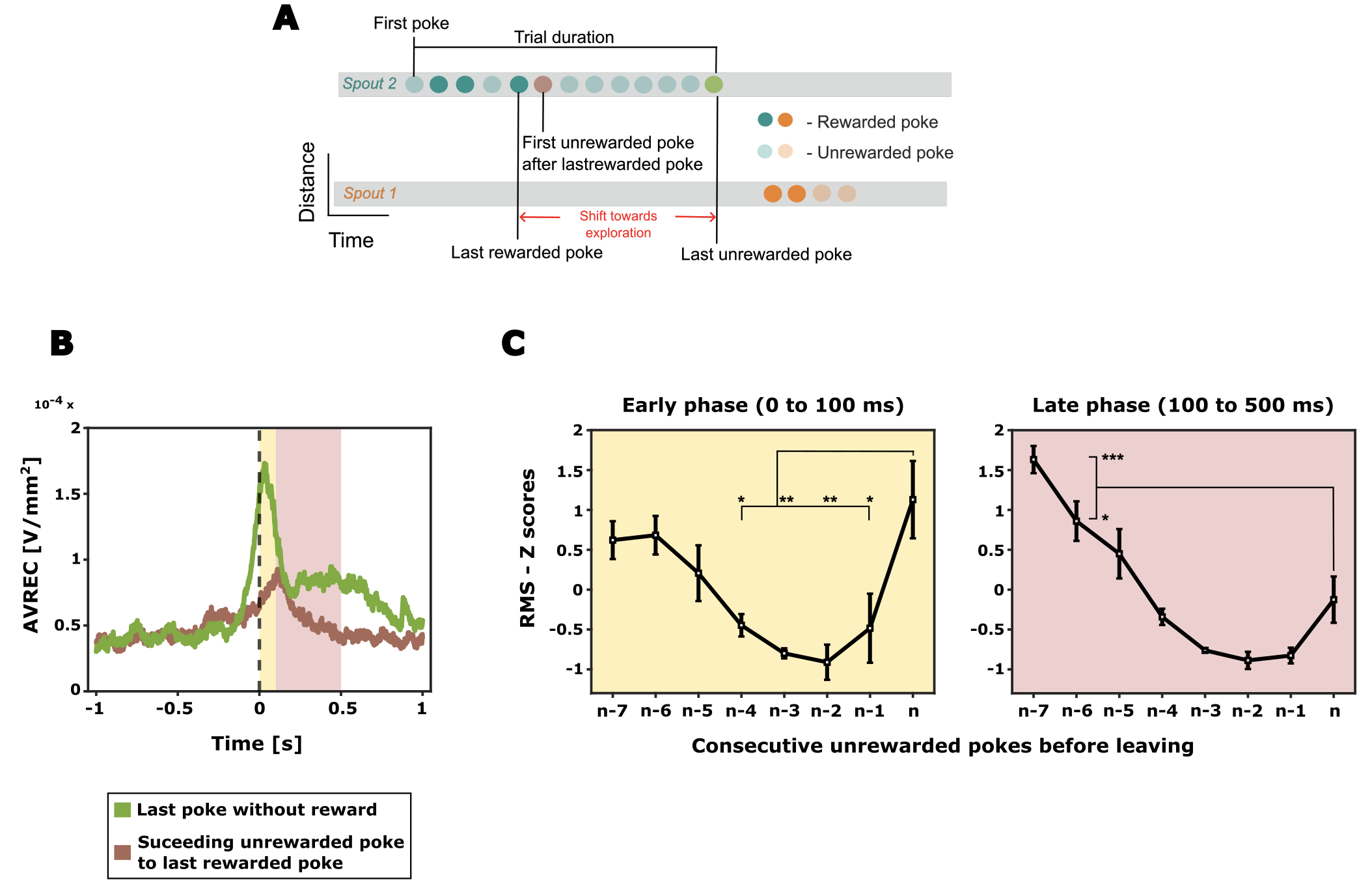
**Figure 4: Grand averaged AVREC (n=5)** displays the overall frontal cortical activity, revealing distinct motor and reward-related signals. The mean average rectified waveform (depicted in bold colours) together with its standard error (shown in lighter shades) is plotted for the selected time intervals (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 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 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 (Fig.5A, brown poke), with reduced yet notable early activity beyond 100 ms, to the exploration phase (Fig.5A, green poke), 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 (Fig.5A – shift towards exploration).



**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).

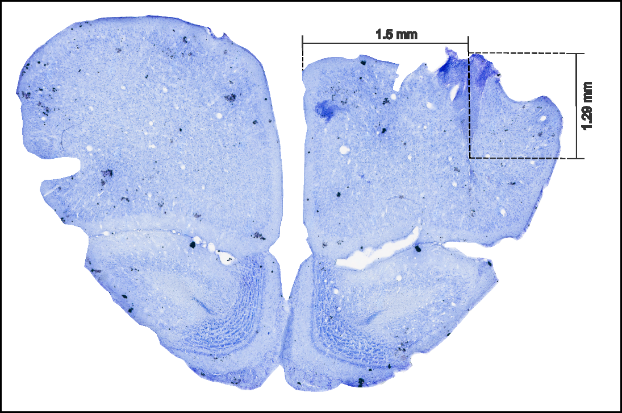
## Layer specific spatiotemporal activity patterns in the frontal field A

### 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.

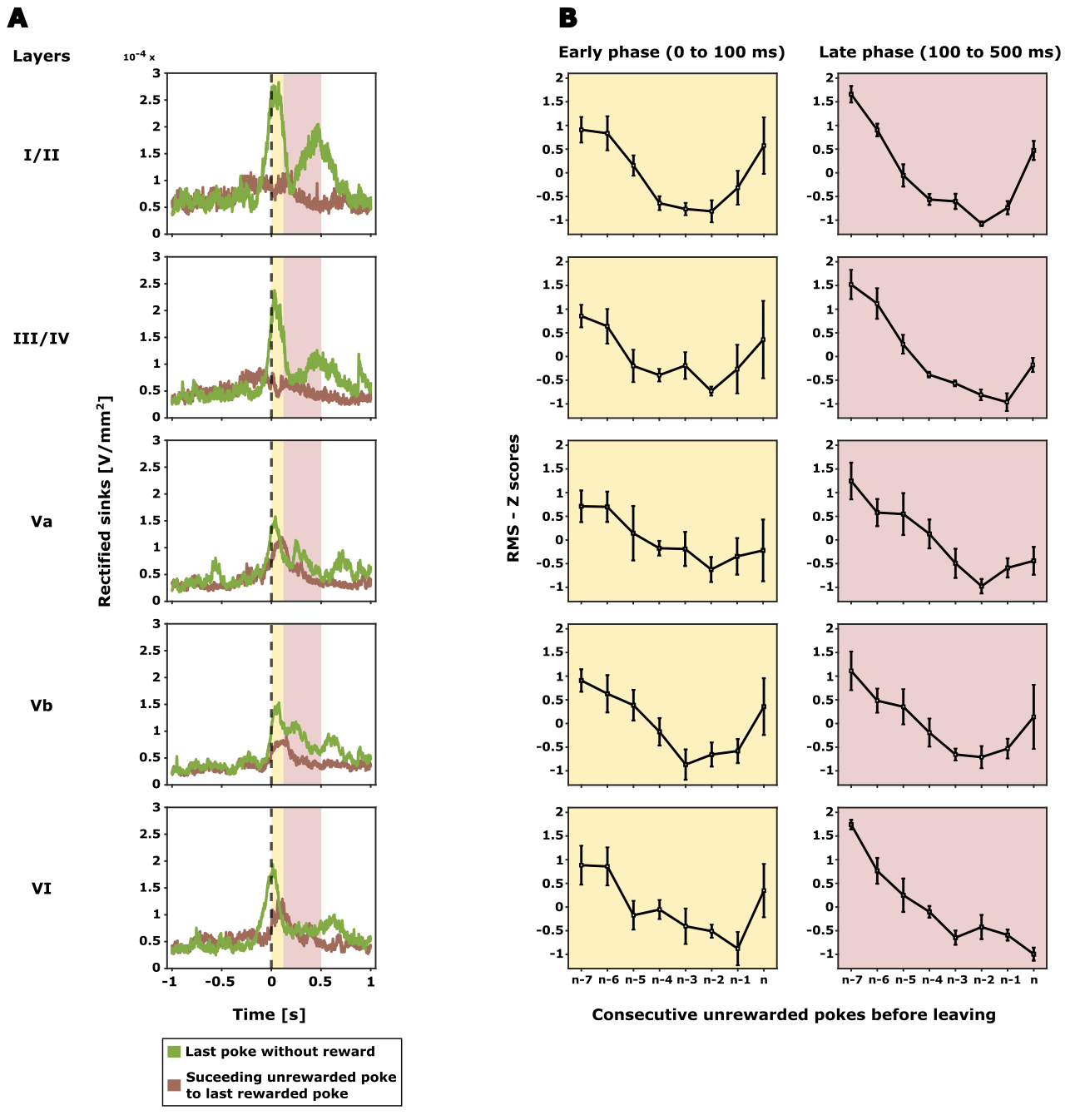


**Figure 1: Sample histology image of the frontal field A (FrA).** Sample histology image from Frontal field A (FrA), taken 4.85 mm anterior to the Bregma. The electrode track is evident at 1.5 mm lateral to the Bregma, reaching a cortical depth of approximately 1.29 mm

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

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.