# Results

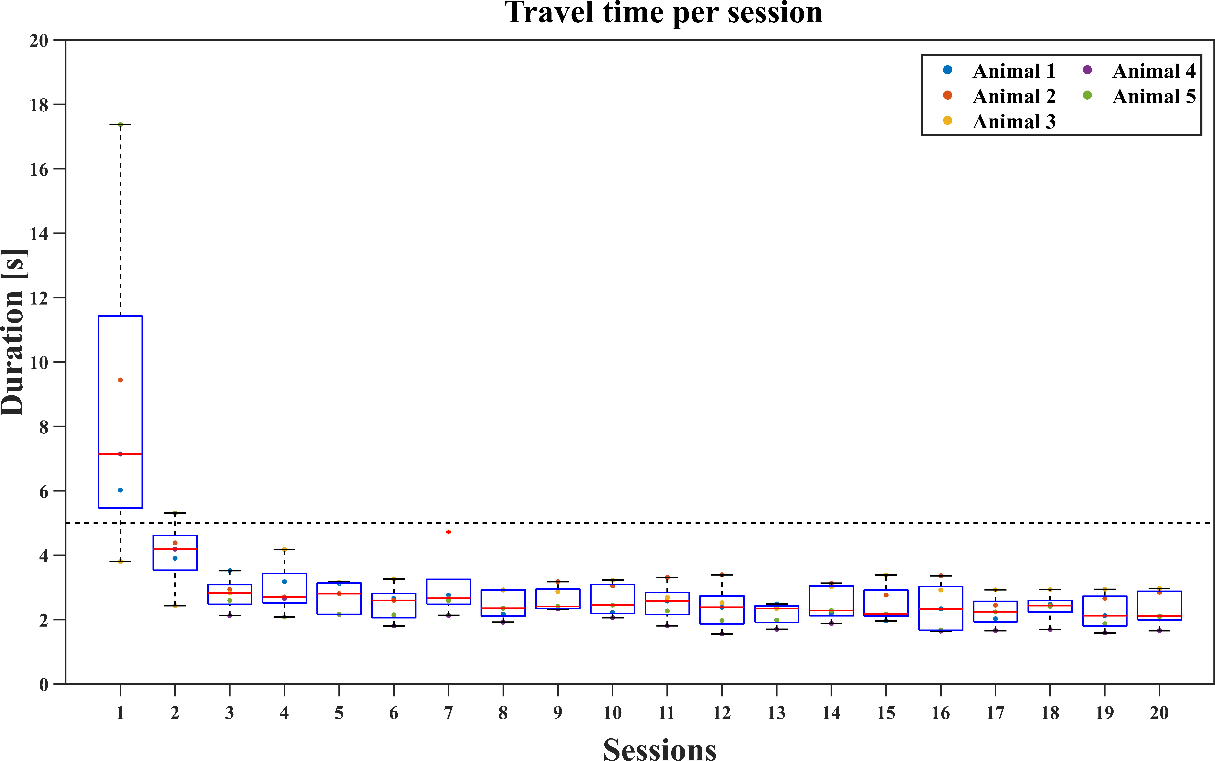
The purpose of this study was to investigate the neuronal underpinnings of attentional resource allocation in a probabilistic foraging paradigm using Mongolian gerbils. Specifically, we aimed to determine whether the frontal cortex, a region known to play a pivotal role in such decision-making processes in humans, also contributes to decision boundaries based on reward probabilities in rodents. To this end, we recorded neuronal data from the frontal cortex of the gerbils using multichannel electrodes as they engaged in a foraging task where food rewards at a particular spout depleted over time, prompting a decision to switch to another spout promising more food. This task represents a typical exploration-exploitation trade-off experiment.

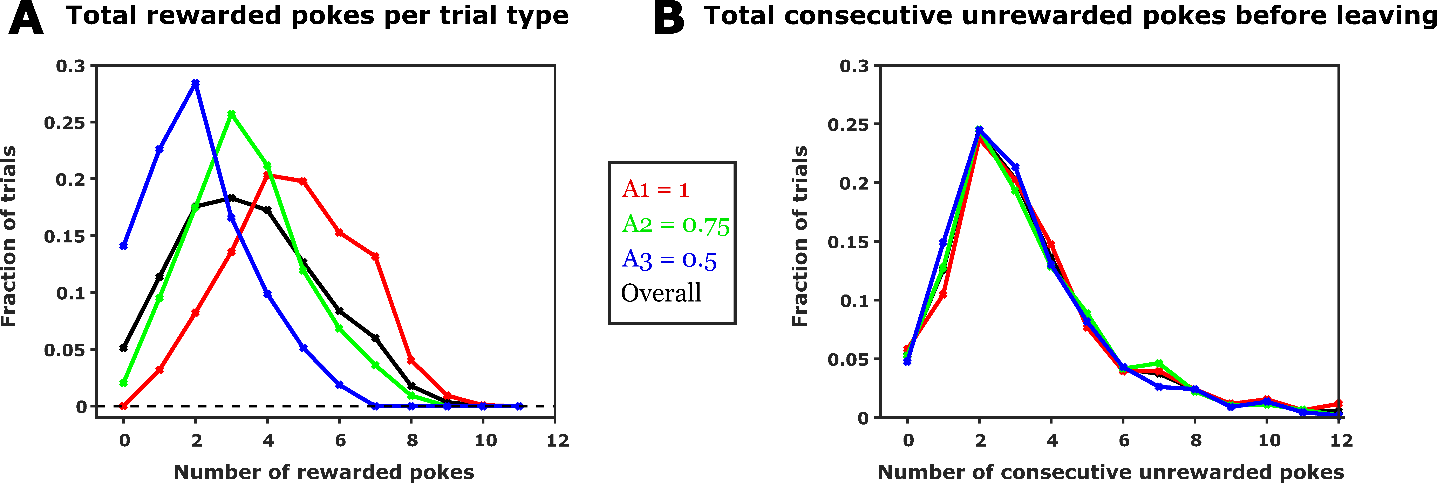
## Behavioural analysis

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

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). Lesser and consistent travel time indicate that the animal is goal-directed and not randomly exploring. The consistent and lower travel time indicates that after each trial, the animals were directed towards the other spout and not randomly exploring the cage (Fig.2A). Based on this, only sessions from 6 to 20 were considered for further behavioural and electrophysiological analyses.





**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.

### Animals make inference-based decisions

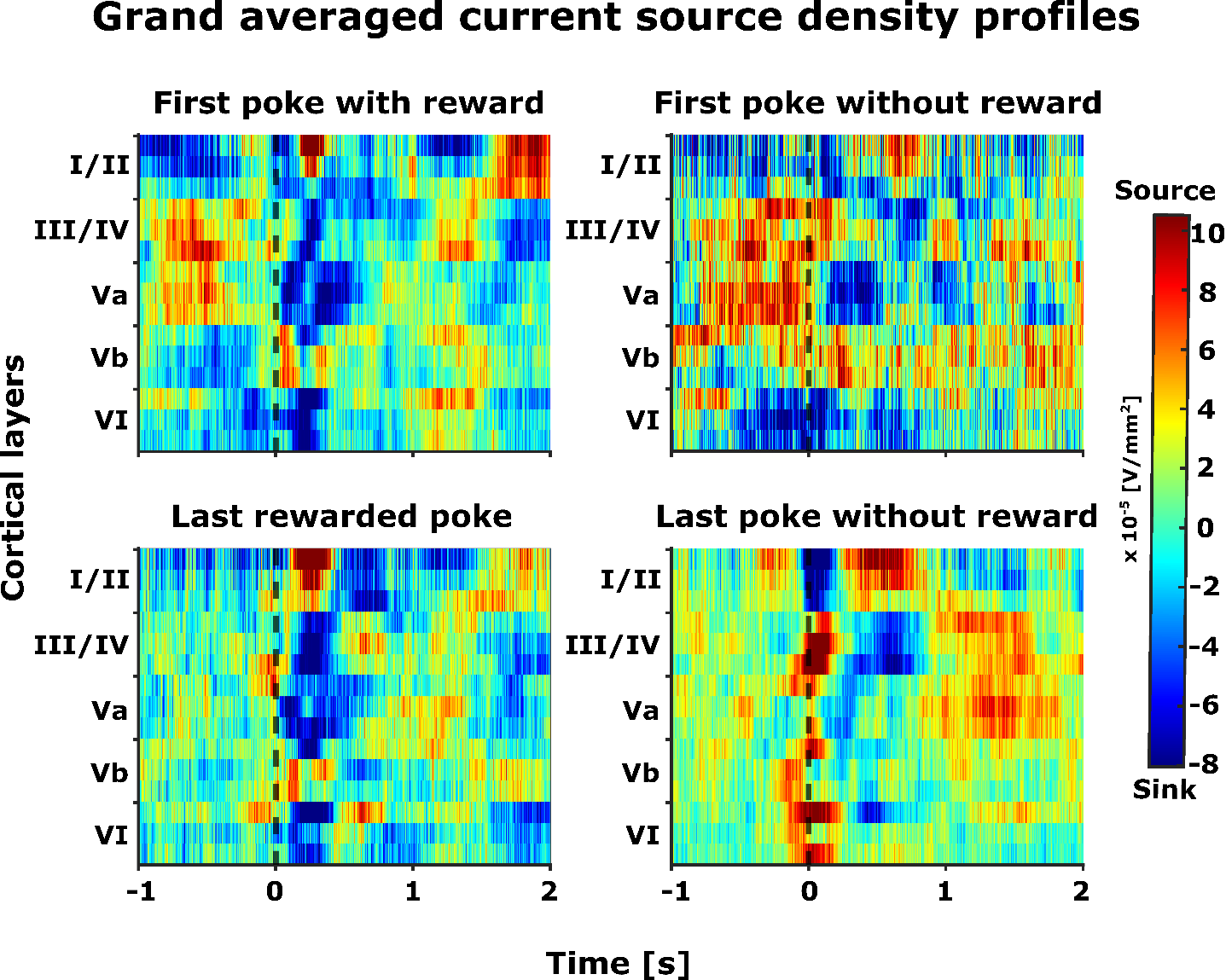
To evaluate how well the animal learnt to make decisions during the exploitation-exploration dilemma, we focussed on the time point where the animal decided to shift from exploitation to exploration. This is because every trial begins with exploiting the current spout for rewards and slowly as the rewards gets exhausted, the important decision is to decide when to leave the current spout to explore the other. This can be interpreted by the consecutive number of unrewarded pokes the animal makes before leaving the current spout.

Following the experimental design (Fig.1C), Fig.2B shows that the distribution of total rewards shifts towards right for trials starting with higher initial reward probability. This indicates that the if the animal makes same number of pokes in each trial, it will receive a greater number of rewards for trials starting with higher initial reward probability. On the contrary, this distinction between trials starting with different starting reward probabilities gets lost when it comes to the consecutive unrewarded pokes made before switching (Fig.2C) 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 consistency in decision making may reflect the possibility that the animals form an inference about the hidden reward structure based on their learning of the task.

## Distinct spatiotemporal activity patterns in the frontal field A

To confirm if the target region FrA encodes distinct activity patterns that reflects the animal’s behaviour, four different time points that are unique and represents crucial stages during the foraging session were selected based on the animal’s performance (Fig1B and Fig2). These are first poke (rewarded and unrewarded), last rewarded poke and the last poke. The time between first poke until last rewarded poke can be considered as the exploitation phase because during this phase, even after experiencing unrewarded pokes, the animal insists on staying on the same side expecting more rewards. The last rewarded poke is a critical point as it represents the end of exploitation phase and beginning of transition towards exploration. This is because after the last rewarded poke, the animal slowly starts to alter its expectation and move towards exploration phase which is best represented by the last poke in a trial.

The current source density (CSD) profile of the aforementioned time points shows distinct spatiotemporal neural activity within frontal field A that may encode different features such as the poke related motor activity and the ensuing reward related information (Fig.3). Assuming that the decision-making process to shift or stay from the current spout is temporally linked towards the end of each poke, the computed CSD signals represent -1 to +2 seconds from the end of the poke (black dashed line, t=0). This enables us to compare the spatiotemporal frontal neural activity between different time points (i.e., different pokes) during the decision-making phase and identify distinct neural signatures. Further, the differential reward related activity patterns between rewarded and unrewarded pokes helped identify and distinguish infragranular layers from superficial layers in the laminar recordings and perform channel-layer specification.

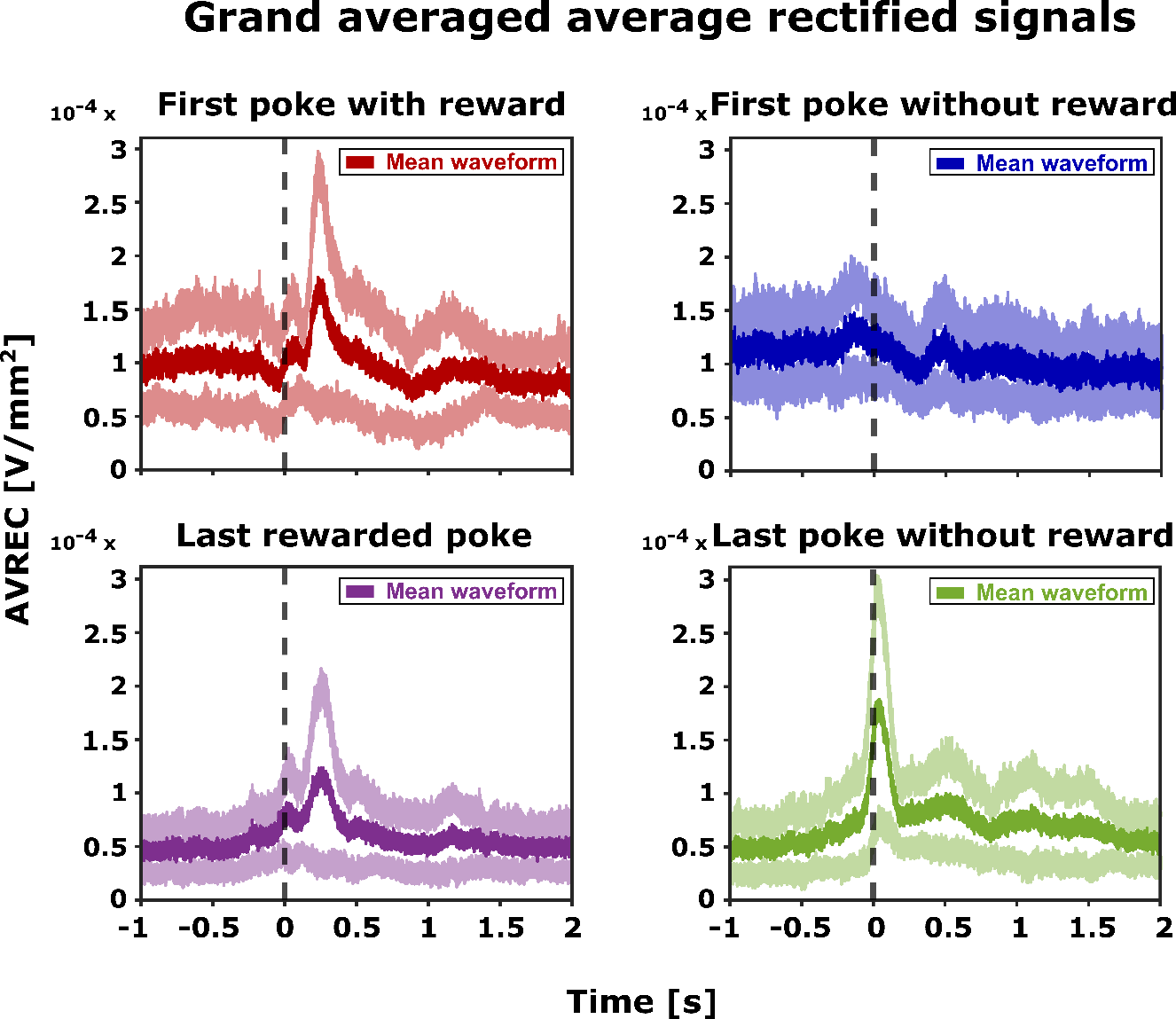


**Figure 3: Grand averaged current source density (CSD) profiles (n=5)** **– Distinct motor and reward related spatiotemporal neural activity in frontal field A**. The selected epochsrepresent -1 to +2 seconds from the end of the poke (black dashed line, t=0). The selected time interval was taken for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

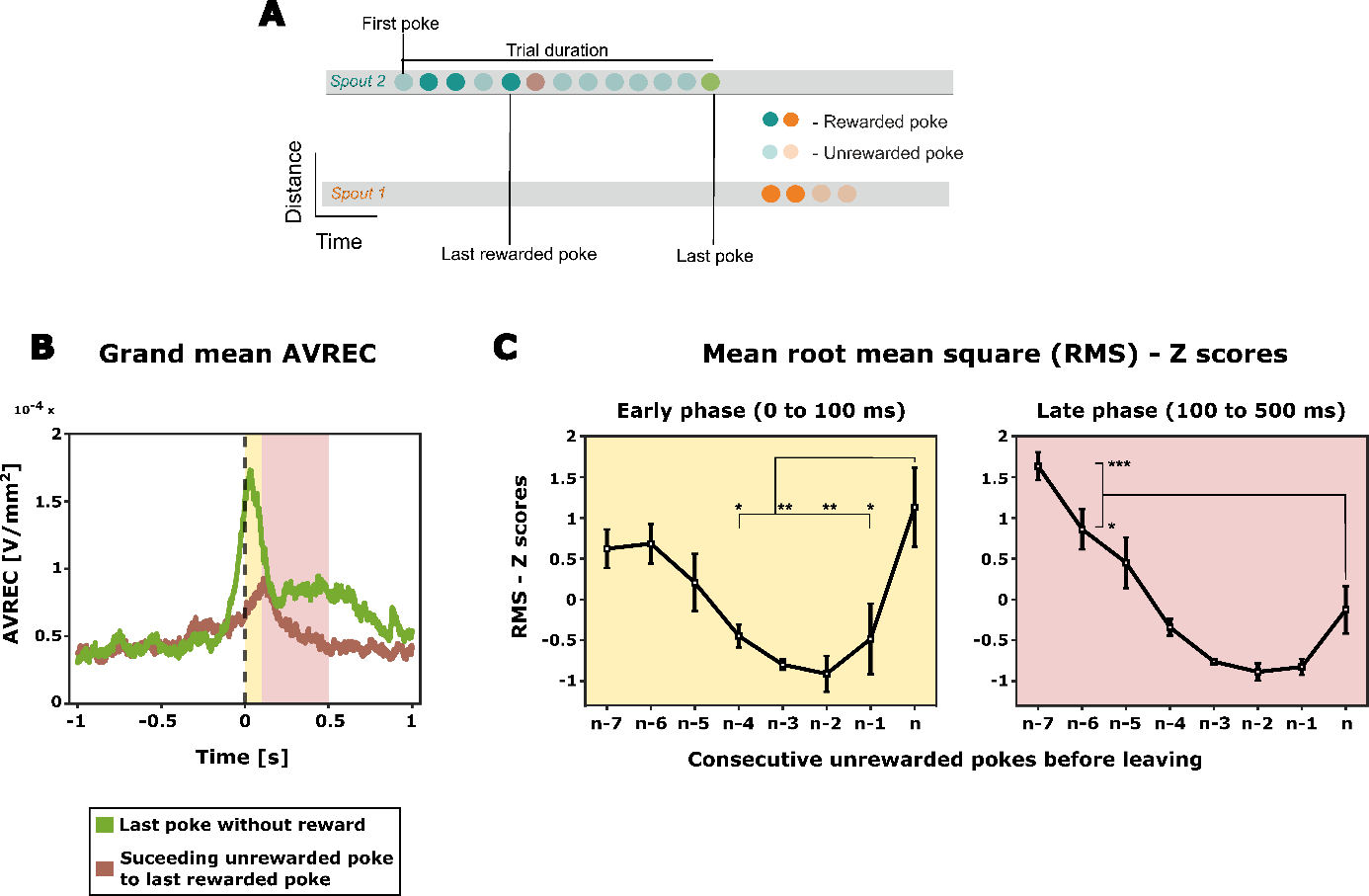
In order to visualize the overall frontal cortical activity, the CSD signals were rectified and averaged across the laminar electrodes to lose the spatial information. The average rectified signals (AVREC) from the CSD profiles show distinct overall activity patterns between different pokes (Fig.4). Mainly, all the pokes except first unrewarded poke show a bimodal waveform with an early peak just after the end of the poke (0-100 ms) and a second late peak after 100 ms from the end of the poke. The first unrewarded poke also shows a bimodal distribution but the early peak appears before the end of the poke (-100 – 0 ms). Further, the overall activity pattern of rewarded pokes (first rewarded and last rewarded pokes) is similar to each other and is different from the unrewarded pokes (first unrewarded and last poke). Finally, the last poke, which acts as a start point of exploration, shows a heightened early peak in amplitude (<100 ms) that is distinct from rest of the 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)** **– Overall frontal cortical activity shows distinct motor and reward related signal**. Mean average rectified waveform (dark) along with its standard error (light) is plotted for selected time intervals (epochs). The selected epochsrepresent -1 to +2 seconds from the end of the poke (t=0). AVREC were 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.

## Shifts in frontal activity patterns: Exploitation to exploration



**Figure 5: Shift from exploitation to exploration**. **A** – A representation of an example trial consisting of a series of pokes highlighting the first unrewarded poke after last poke (brown) and the last poke that is unrewarded (green). **B** – shows the distinct activation pattern of grand AVREC of first unrewarded poke after last rewarded poke (brown) and the last poke without reward (green). Based on the grand AVREC data (A), two distinct time intervals (epochs) were chosen for RMS computation: early phase (0 – 100 ms, yellow), and late phase (100 – 500 ms, light pink). **C** – The AVREC – RMS Z score for unrewarded pokes between last rewarded poke and the last poke (n) before disengaging from the current spout (in this figure, we show a scenario of 7 consecutive unrewarded pokes where n-7 is the first unrewarded poke after last rewarded poke (brown) while n represents the last poke (unrewarded, green)). One way ANOVA with Bonferroni correction was performed to find differences between the pokes (\* p < 0.05, \*\* p < 0.01, \*\* p<0.001)

### What makes the last poke the last one?

We have confirmed that the frontal field A encodes distinct activity patterns for different pokes (Fig.3 and Fig.4 – first poke (rewarded and unrewarded), last rewarded poke and last poke). Specifically, the last poke shows a distinct neural activity pattern compared to the rest of the pokes (Fig.4). This distinct neural activity pattern of the last poke. does it pertain only to the last poke in a trial? If yes, how this distinct neural activity pattern emerges towards the last poke? For this, we looked into the shift in frontal activity pattern from exploitation to exploration phases (Fig.5A).

Analyzing this shift in activity helps to understand the crucial decision-making part i.e., how the animal decides the last poke in a trial? This is essential as the animals may encounter multiple consecutive unrewarded pokes after the last rewarded poke (Fig.5A). Hence, the last poke also being an unrewarded poke needs to be distinguished from its preceding unrewarded pokes. To this end, we compared the overall frontal activity of last unrewarded poke (Fig5A – green poke) with the first unrewarded poke after the last rewarded poke (Fig5A – brown poke) revealing its distinct activity patterns (Fig.5B). While the last unrewarded poke has a bimodal distribution with a sharp peak in frontal activity immediately after the end of the poke (< 100 ms from the end of the poke) followed by a persistent activity pattern later (> 250 ms from the end of the poke), the first unrewarded poke succeeding the last rewarded poke has only a single prominent peak after the end of the poke (<100 ms) that is smaller in amplitude compared to its counterpart.

As the two prominent peaks observed in the last poke lies within 500 ms from the end of the poke, we narrowed our focus into this time window to see if this is the crucial window where the decision to shift to the other spout happens. Further due to its bimodal nature, two distinct phases were selected within this 500 ms i.e., early and late phase to separate the two peaks. Early phase represents the first 100 ms (Fig.5B, yellow region) capturing the sharp frontal activity immediately after the poke and the late phase represents 100-500 ms (Fig.5B, light pink region) after the end of the poke which captures the persistent frontal activity after the first sharp peak.

Fig.5B reveals that when the animal is still at the end of exploitation phase (Fig.5A – brown peak; first unrewarded poke after last rewarded poke), less activity is observed in the early phase but a small peak followed by an extended persistent activity is observed after 100 ms following the nose poke. Conversely, at the onset of the exploration phase (last poke), heightened frontal activity is evident within 100 ms after the nose poke (early phase). This shift in activity pattern makes the last poke unique and hints that the early phase could be a crucial window where the primary decision to explore may get encoded.

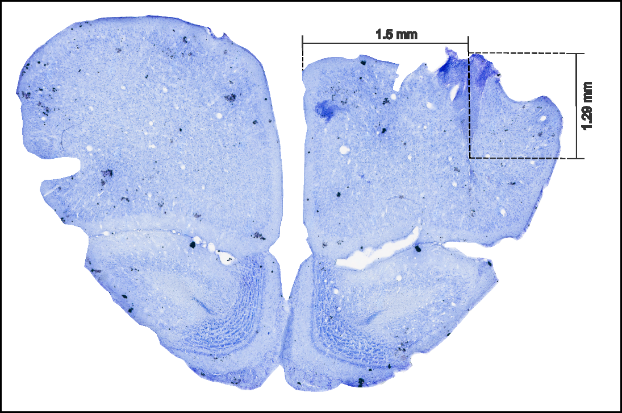
### The evolution of frontal activity from exploitation to exploration

To get a deeper understanding of the evolution of change in frontal activity patterns from exploitation to exploration, we focussed on the individual pokes. Root mean square of overall frontal activity was chosen to obtain a single quantitative measure that represents the average temporal activity of each poke in the selected time window of interest and can be compared between different pokes. We computed and compared the root mean square (RMS) of the AVREC signal for all the unrewarded pokes between last rewarded poke and last poke (Fig.5B and 5C) and z-normalized it within each animal. Over the transition from exploitation to exploration (Fig.5B, n-7 to nth poke), during the early phase, the overall frontal activity initially decreased (n-6 to n-2) and then increased just before the animal decides to leave the spout (n-2 to nth poke) (One way ANOVA with Bonferroni correction, \* p < 0.05, \*\* p < 0.01, \*\* p<0.001). However, during the late phase, its frontal activity constantly decreases with a slight increase before leaving (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).

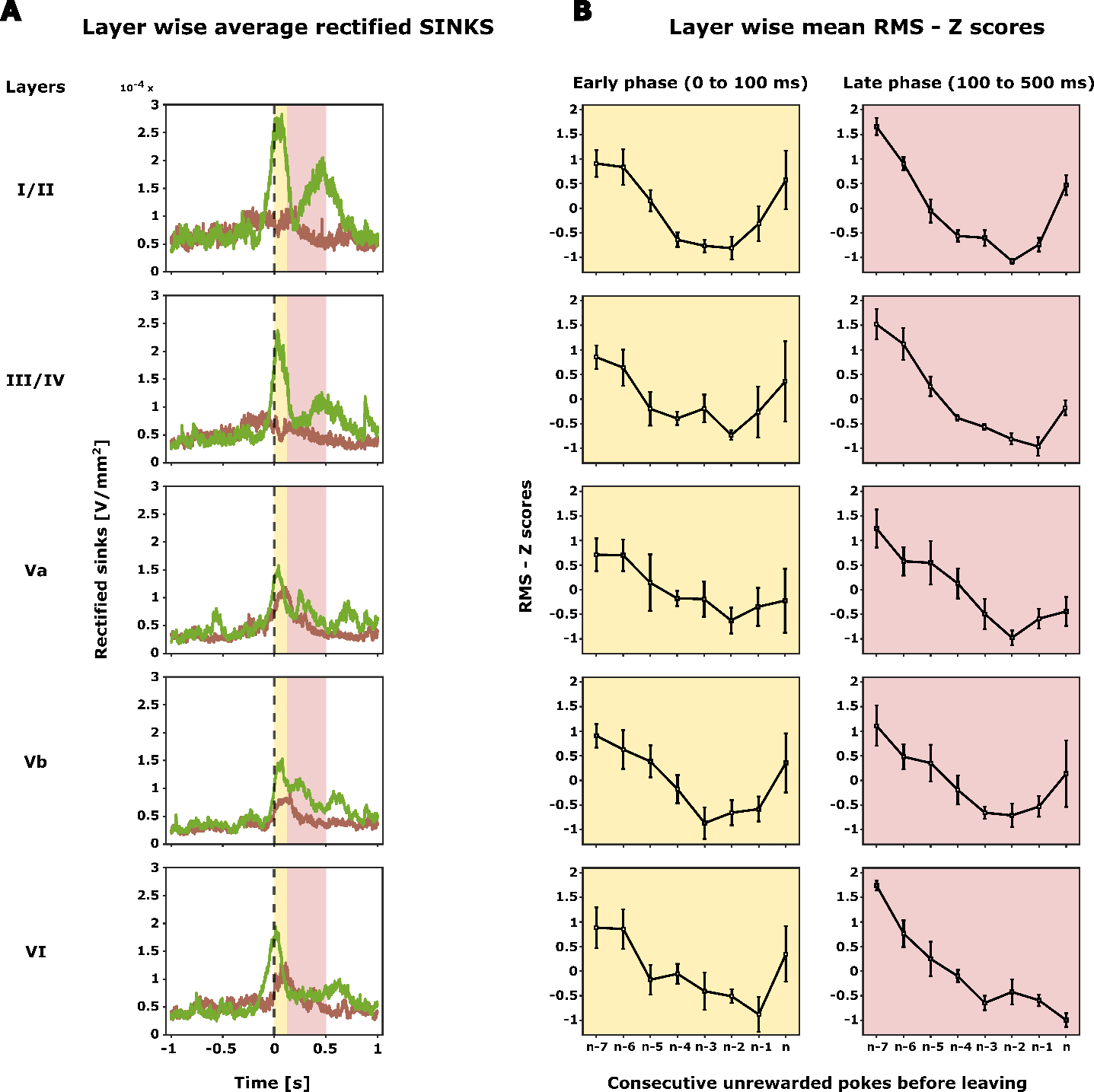
## Layer specific motor and reward related activity patterns in the frontal field

In order to accurately map neuronal activity throughout all cortical layers, it is essential to ensure that the laminar depth electrode is positioned perpendicularly to the cortical surface. Histological verification is a critical step in confirming that the electrode spans the entire cortical thickness, thereby enabling a comprehensive analysis of the layer-specific activity within the behavioural paradigm. The histological examination involved post-experiment tissue processing and sectioning to visualize the electrode track and confirm its orientation relative to the cortical layers. By correlating the electrode's position with the known cytoarchitecture of the frontal cortex, one could infer that the recordings obtained indeed reflect the activity from the full depth of the cortex. This verification process is pivotal for the interpretation of our electrophysiological data, as it ensures that the recorded signals are representative of the neuronal dynamics across the cortical layers, which is hypothesized to be crucial for attentional resource allocation during decision-making tasks (reference).

Once all the foraging sessions were completed, the animals were sacrificed and the brain slices were taken and Nissl stained to confirm the location of implanted electrode. According to the Gerbil brain atlas (Radtke-Schuller et al., 2016), the target frontal region A (FrA) is at 4.65 to 5 mm anterior and 1.5 mm lateral to the bregma. Fig.1 shows a sample histology slice from one animal taken at 4.85 mm anterior and 1.5 mm lateral to the bregma indicating that the electrode is placed in the FrA. The electrode trace can be clearly identified with a maximum cortical depth of approximately 1.29mm. This is just one qualitative example of an ongoing quantitative analysis being done on other set of animals which is not the primary focus of this thesis.



**Figure 1: Sample histology image of the frontal field A (FrA).** The histology slice is taken 4.85 mm anterior to Bregma. The electrode location from the staining can be seen 1.5 mm lateral with a cortical depth of 1.29 mm approximately.



**Figure 6: Layer specific frontal motor and reward related activity**. **A** – The grand averaged rectified sinks (n=5) was computed for all the identified layers from the grand CSD profile (Fig.3). The selected epochsrepresent -1 to +2 seconds from the end of the poke (t = 0). The selected time interval was taken for first unrewarded poke after last rewarded poke (brown) and last poke without reward (green). Based on the averaged rectified sinks (A), two distinct time intervals (epochs) were chosen for RMS computation: early phase (0 – 100 ms, yellow), and late phase (100 – 500 ms, light pink). **B** – The layer wise Z scores of root mean square (RMS) was calculated from the average rectified sinks for unrewarded pokes between last rewarded poke and the last poke (n) before disengaging from the current spout. (Similar to Fig.5B, in this figure, we show a scenario of 7 consecutive unrewarded pokes where n-7 is the first unrewarded poke after last rewarded poke while n represents the last poke (unrewarded)).

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. However, at the beginning of the exploration phase (towards the last poke (n)), 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.