# 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). Our study illuminated the gerbil’s reliance on a more sophisticated decision-making rather than rigid foraging rules. This was evidenced by their consistent behavior in the face of diminishing rewards, which indicates an adherence to Giving Up Time (GUT) rule or similar heuristics. Furthermore, the distinct neural patterns within the FrA suggest an intricate encoding of motor actions, as well as reward expectation and evaluation, pointing to a complex cognitive process that integrates past experiences with immediate action-outcome assessments. Additionally, our findings on the layer-specific activity within the FrA hint at a nuanced, layer-dependent processing mechanism, potentially underpinning the adaptive decision-making observed in these animals.

## Inference-Bound Decision-Making in Gerbils

In the world of foraging, the 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 et al. in 1976, 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. Thus, they face uncertainty: 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 (Results Fig) indicate that they did not conform to fixed-time or fixed-number 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 such unpredictable settings, 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 – 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 this rule. Regardless of the spout's starting probability or the number of rewards accumulated, gerbils exhibited a consistent behavior of leaving after a set number of unrewarded pokes (results fig). This consistency suggests a reliance on the 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 saligns with observed behaviours in other rodents, primates, and humans, where adaptive strategies are formed through experience (to be cited). 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 suggests the presence of an intricate neural circuitry in the gerbils, facilitating a decision-making strategy that is both adaptive and informed by experience.

## 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 gerbils' engagement in a probabilistic foraging task. 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 are integral to the gerbils' 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. 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 milliseconds later, which aligns with the actual receipt of the food pellet and thus likely signifies the evaluation of the reward (refer to AVREC figure).

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. This divergence indicates a neural encoding of prediction error rather than reward receipt. 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.

Furthermore, the notably heightened amplitude in the early peak of the last unrewarded poke suggests an integration of expectation and cognitive demands for planning the switch to a new foraging location. This amplified neural activity during the early phase of the last poke distinguishes it from others and underscores a decision-making process within the FrA that goes beyond simple stimulus-response patterns, involving an anticipation of outcomes shaped by experience—an essential element underlying inference-driven behavior.

## The Dynamic Transition of Expectation 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 reveals 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 Results figure), 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 Results figure) is characterized by a sharp increase in activity within 100 ms post-poke. This heightened early-phase activity suggests an overlap of outcome expectation with the motor aspects of the foraging action itself, indicating that the neural coding of the last unrewarded poke, which signals the shift to exploration, is distinct from previous pokes. Besides, it hints that the early phase could be the crucial phase where the primary signal to explore may get encoded.

Delving into a quantitative assessment, we took an example where the animal made seven unrewarded pokes before leaving to study the evolution of the FrA activity. The RMS analysis of individual pokes illustrates the non-linear evolution of FrA activity, forming a U-shaped curve as the trial progresses from exploitation to exploration. 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-7th poke). This is characterized by a dip in activity 100 ms pose-poke where the reward receipt is usually encoded (brow waveform, Fig). As unrewarded pokes accumulate (n-6 to n-2nd poke), the animal enters a state of uncertainty, diminishing overall frontal activity and shifting attention from the anticipated reward to the act of nose poking—a more informative action for updating expectations. Eventually, a new expectation sets in—the anticipation of no reward—which, upon confirmation through continued unrewarded pokes (n-2 to nth poke), leads to an increase in activity and the subsequent spout switch.

This nuanced modulation of expectation suggests a sophisticated underlying neural mechanism at play, where layer-specific processing in the FrA could be critical. The transition from a certain expectation (reward) through uncertainty to a new certainty (no reward) for the same action implies differential engagement of cortical layers, potentially reflecting the layered architecture's role in integrating past experiences with present-moment decision-making.

## Functional Implications of Layer-Specific Activity During Foraging

In the pursuit of understanding foraging strategies, the laminar current source density (CSD) recordings from the frontal region A (FrA) have provided a detailed map of the local layer-specific micro-circuitry activity. The gerbils' foraging behavior has allowed us to discern how the brain allocates 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 orchestrated within the FrA in a layer-dependent manner, as revealed in our results (Figure 6).

The transition from exploitation to exploration is fundamentally an attentional reallocation challenge. 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 (references), 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 and Kirchmar, 2015, 2017), which are implicated in updating working memory content through reward-prediction error signaling—integral to 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. 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 animals' 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 gerbils' dynamic shift in search strategies during foraging behavior.