Summary 4
Debarpita Dash

Traversing Marr levels for value decisions: In value-based decision-making, the brain follows Marr's three levels of analysis. At the computational level, the brain's goal is to choose the option with the highest subjective value by weighing risks and rewards. At the algorithmic level, it assigns value using methods like reinforcement learning or simple decision rules (heuristics). At the implementation level, this process happens in brain regions.

Framework for decision making: It involves identifying possible actions (representation), assessing their value based on internal and external factors (valuation), choosing the best option (action selection), evaluating the outcome, and updating future decisions through experience (learning)

Does the LIP area encode decision-related computations?

Cued Saccade Task: Monkeys fixated on a central stimulus while two peripheral targets appeared. A color change cued the correct target, with reward size and selection probability varying across blocks. LIP neurons fired more for larger expected rewards, with activity peaking early and declining later. Stronger firing was observed when a movement had a higher reward probability, but firing did not correlate with movement parameters.

Free-Choice Task: Monkeys freely choose between two targets with changing reward values. LIP neurons encoded expected gain, showing higher activity for higher rewards. Frequent choices of a movement correlated with increased firing, supporting the idea that LIP encodes subjective value. Early in trials, LIP activity reflected expected rewards, while later it aligned with movement selection. As reward differences increased, LIP firing changed linearly, consistently predicting the chosen target. Notably, the gain ratio plot remained stable despite absolute reward variations, reinforcing that LIP encodes subjective value rather than just sensory input or movement execution.

Common Currency for Value in the Brain: Researchers have found that a subregion of the ventromedial prefrontal cortex (VMPFC) /orbitofrontal cortex (OFC) encodes subjective value independent of sensory attributes. Using fMRI, participants were shown rewards like money, juice, or other incentives, and this brain region consistently activated, regardless of the reward type.

Behavioral Evidence: The Becker-Degroot-Marschak (BDM) Auction: It provides behavioral evidence for a common currency system. In this task, participants bid money on rewards like food or gifts, revealing their willingness to pay (WTP). Neural activity in the VMPFC strongly correlated with WTP, reinforcing the idea that this region encodes subjective value in a comparable way across different reward types.

Reflex Actions: Reflexes are extremely fast, automatic responses shaped by evolution. They do not involve new learning but are driven by evolutionary value they exist because they increase survival chances. For example, pulling your hand away from fire is an instinctive protective reflex, bypassing conscious thought for speed.

Pavlovian Conditioning: It occurs when an organism associates a stimulus with an outcome, leading to automatic responses. Example :Imagine you are camping in a tiger reserve. You hear rustling leaves, which could indicate a predator. Even if no tiger appears, your brain has learned to associate rustling with danger, prompting you to pack up and leave. This is a Pavlovian response, even though your action does not directly change the outcome (whether or not a tiger is present).However, Pavlovian learning is limited associations can be unlearned if they become disadvantageous (e.g., a dog can stop salivating if the bell is no longer followed by food).It does not allow learning entirely new behaviors in response to a conditioned stimulus (e.g., you cannot make a dog bark instead of salivating when it hears a bell).

Goal-Directed Behavior: It involves evaluating all available options and selecting the one with the highest value. This system adapts to changes in the environment and is ideal for making flexible, rational decisions.

Stimulus devaluation: You crave chocolate, so you go to the store and buy one.If you eat too much chocolate in one day, the next time someone offers you another piece, you might reject it because you are satiated.

Habit Formation: If an action is repeated frequently enough, it shifts from goal-directed to habitual learning. If you go to the store every day to buy chocolate, over time, the action becomes automatic.Even if you are not craving chocolate, you may still go to the store and buy it simply out of habit. Habitual behavior is less sensitive to changes in reward value. If chocolate is no longer enjoyable, a habitual individual may still buy it, unlike a goal-directed individual who would stop.

Neural Circuits for Reflexes, Pavlovian Learning, and Habitual vs. Goal-Directed Actions

Dopamine, produced in the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc), interacts with different brain circuits depending on the type of learning. Goal-directed learning relies on the dorsomedial striatum (Dm) for reward evaluation, the medial prefrontal cortex (mPFC) for planning, and the orbitofrontal cortex (OFC) for tracking reward value. Habitual learning, driven by the dorsolateral striatum (Dl), stores automatic behaviors, which disappear if this region is disrupted. Pavlovian learning involves the amygdala (BLA & CEA) for emotional processing and the ventral striatum (VS) for reinforcing stimulus-reward associations. These systems operate simultaneously, often competing—goal-directed circuits can override habits, enabling flexibility, while deeply ingrained habits may override rational decision-making.

Sign-Tracking vs. Goal-Tracking: How Animals Differ in Learning

Goal-trackers focus on obtaining the actual reward, such as a rat pressing a lever for chocolate and immediately collecting it. In contrast, sign-trackers fixate on cues rather than rewards; some rats become obsessed with the lever itself, repeatedly interacting with it instead of retrieving the chocolate. This behavior is driven by dopamine reinforcement, where the brain overvalued the cue rather than the actual goal.

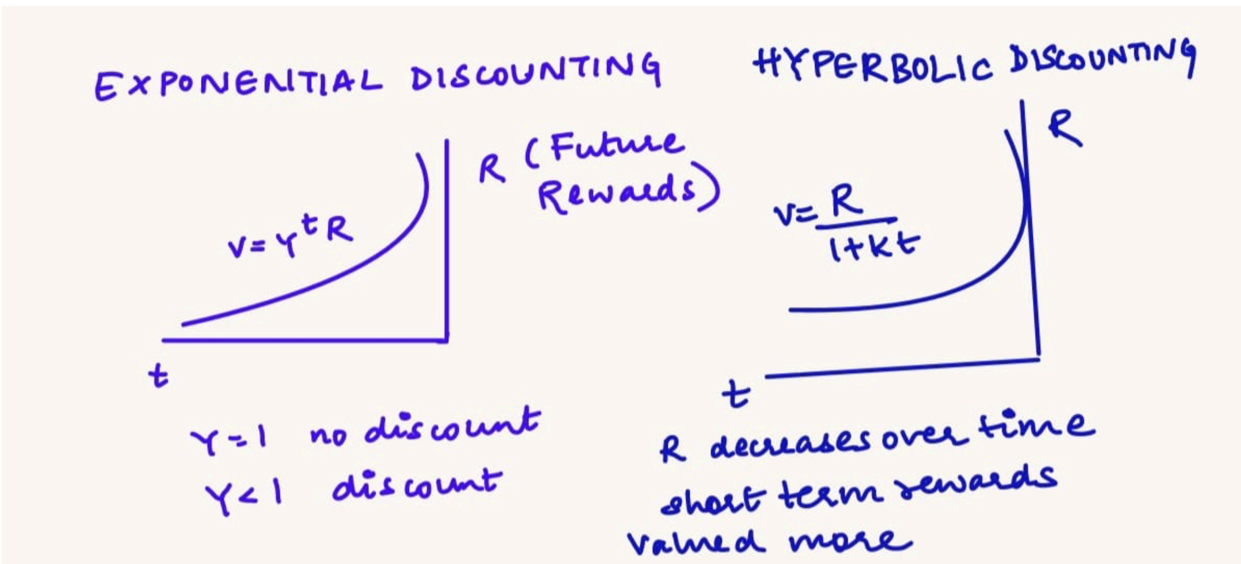
Marmosets as a Model for Studying Decision-Making: Their brains have a smooth surface (no gyri or sulci), making them ideal for studying neural circuits.

Does preference remain the same over time?

People's preferences generally remain stable when decision parameters are unchanged. However, in intertemporal choices, time influences valuation through temporal discounting, where distant rewards are perceived as less valuable.

Given \$50 now vs. \$100 in 12 months, most choose the immediate reward, implying high discounting of future gains.However, for \$50 in 24 months vs. \$100 in 36 months, people often pick the larger reward, showing reduced discounting when both options are delayed.

This contradicts exponential discounting, where preferences should remain proportional over time, suggesting that decision-making follows a hyperbolic rather than strict exponential decay model.



High discounters (high k) prefer immediate rewards, while low discounters (low k) wait for larger ones. A study of 20,000 participants found a 1000-fold variation in discount rates, influenced by genetics (12%) and environment (trust, scarcity, socioeconomic status).

Landmark studies show that while individual neurons don't follow hyperbolic patterns, behavior at the aggregate level does.A mix of neurons with different discounting profiles can collectively produce hyperbolic discounting.