Sequential sampling from memory underlies action selection during abstract decision-making

S. Shushruth and Michael N. Shadlen Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY

INTRODUCTION

The facility to assign abstract labels to sensory percepts belies the complexity of the underlying neural computations that support this ability. Abstract labels assigned to sensory percepts have two defining features that make them indispensable for higher order cognition. First, abstract labels enable parsing the vast array of sensory information into behaviorally useful categories (Tenenbaum et al 2011). Second, abstract labels endow immense flexibility to the process of sensorimotor transformation (Chafee & Crowe 2012). For example, one can assign the abstract labels "rightward" and "leftward" to consolidate motion perceived to the right or left, irrespective of its precise direction or motion strength. These abstract labels then allow for the implementation of flexible action plans such as "press a red button if you see rightward motion". Despite the centrality of sensory abstraction to human cognition, neither its underlying neural basis nor its evolutionary antecedents in nonhuman primates are well understood.

The extent to which nonhuman primates can assign abstract labels to sensory percepts and exploit them to be flexible in their actions is unclear. The process of abstraction, by definition, unyokes the sensory evaluation processes from the process of acting on the sensory information. However, multiple lines of research suggest that in monkeys, the process of sensory evaluation is intimately coupled to the actions that can result from the evaluative process (Cisek 2007, Cisek & Kalaska 2010, Shadlen et al 2008). This framework, wherein cognitive processes are embodied in terms of the motor actions they afford, is supported by the patterns of neural activity found in association and premotor cortices of monkeys (Cisek & Kalaska 2005, Klaes et al 2011, Kubanek & Snyder 2015, Shadlen & Kiani 2013). But monkeys can be trained to decide on properties of sensory stimuli even when unaware of the exact motor action that will be required of them to report their decision (Bennur & Gold 2011, Freedman & Assad 2006, Genovesio et al 2009, Gold & Shadlen 2003, Wang et al 2019). Many of these studies have focused on the neural representation of abstract perceptual labels. However two important aspects of abstract perceptual decisions are still not well understood – (i) what is the decision making process for assigning abstract labels to perceptually ambiguous stimuli and (ii) how are stored decisions flexibly converted to motor actions.

To investigate these aspects of abstraction, we trained two monkeys to decide on the net direction of stochastic motion stimuli and associate two possible directions of motion with two colors. The monkeys reported the direction of motion by making an eye movement to the target of the associated color. The locations of the targets were randomized across trials and revealed only after the cessation of the motion stimulus. To perform the task well, monkeys needed to integrate motion information in the stimulus over time to make an abstract decision about the direction of motion. This imposition allowed for the investigation of how an abstract perceptual decision is formed when the action-choices associated with the decision are unknown. And since the decision making phase is unyoked from the motor planning phase, this task allowed for the investigation of how that abstract decision is converted into a motor action.

Surprisingly, we found that the two aspects of abstract decision making — evidence evaluation and action selection — that our task was supposed to unyoke were, in fact, coupled. The accuracy of the monkeys' decisions suggested that the monkeys did learn to integrate evidence towards an abstract decision about the direction of motion. However, the process of deliberation about the quality of evidence transpired during the action selection epoch instead of the evidence evaluation epoch. Consistent with this behavioral finding, neural activity in the

sensorimotor association area LIP during the action selection epoch reflected the instantiation of a sequential sampling process. Our results show how the ability to assign abstract labels to perceptual stimuli could have developed upon the underlying motor-intention centric architecture of the primate brain.

RESULTS

A task that entails decision-making in the abstract



FIGURE 1

We trained two monkeys (one female AN and one male SM) to decide whether the net direction of a random dot motion (RDM) stimulus was to the right or left, and report it by making an eye movement to a cyan or yellow target (**Fig. 1, top**). The associations between directions of motion and target colors were counterbalanced across animals. The two targets appeared after a short delay (200-333 ms) following the termination of the motion stimulus. The locations of the two targets were randomized across trials. Thus the monkeys had to make a decision about the rightwardness/ leftwardness of the RDM without knowing the actual eye movement that would be required of them to report the decision.

Both monkeys were naïve to the RDM stimulus when they began training on the task. This ensured that the monkeys had not formed an association between the direction of motion and a provisional motor plan. This is in contrast to previous studies with RDM stimuli (Bennur & Gold 2011, Gold & Shadlen 2003) in which the monkeys were initially trained to associate leftward motion with an eye movement to the left and rightward motion with an eye movement to the right. In many previous studies (Bennur & Gold 2011, Freedman & Assad 2006, Wallis et al 2001) in which monkeys performed sensory evaluation separately from motor planning, the potential motor actions available to the monkeys to report their decision were limited to two. This restricted action set allowed for the possibility that the monkeys formed a provisional mapping of a sensory percept with one of the two motor plans and then either implemented that plan or its reversal depending on which of the two actions were required in the reporting epoch. In our task, to preclude such strategies, we increased the possible motor actions available to the monkey. The two targets could appear in any of 8 or 12 configurations and the set of configurations varied across sessions.

Since one of our goals was to investigate how decisions are converted to motor actions, the monkeys were allowed to report their decision as soon as the targets were presented. For reasons that we explain below, Monkey SM was also trained on a variation of the task in which an additional waiting time was imposed after the appearance of the targets (**Fig. 1, bottom**). Hereon we will refer to the two versions of the task as 'go' and 'wait' respectively and denote them as subscripts to the monkey names (AN_{go}, SM_{go} and SM_{wait}).

Monkeys can learn to integrate evidence towards an abstract percept

The abstract decision making task proved to be challenging for the monkeys to learn. Monkey AN_{go} needed 170 sessions to reach a psychophysical threshold of 15% or less. Monkey SM plateaued at a psychophysical threshold of 20% even after 180 sessions (~700 trials per session on average) while trained on the 'go' version of the task. This monkey was then trained on the 'wait' version of the task for a further 60 sessions until the psychophysical thresholds decreased and stabilized at around 11%.



FIGURE 2

At the end of training, both monkeys were significantly above chance at the lowest nonzero motion coherence (**Fig. 2 A-B**). While the monkeys never achieved perfect performance at the highest motion strengths (overall error at highest motion strength: 9% for Monkey AN_{go} and 6% for Monkey SM_{wait}), these lapse rates are considerably lower than those reported in previous studies with similar abstract direction discrimination paradigms (> 12% in (Gold & Shadlen 2003) and >15% in (Bennur & Gold 2011). In monkeys performing the same direction discrimination task with known mapping between motion direction and actions, the lapses are usually <1%. This marked increase in lapses emphasizes the difficulty for monkeys to make perceptual decisions in the abstract. Note that nearly all previous studies of perceptual discrimination in macaques with unyoked sensory evaluation and motor planning epochs have reported significant lapse rates {Fitzgerald, 2011 #311, Freedman & Assad 2006, Genovesio et al 2009, Wallis et al 2001, Wang et al 2019). To determine whether the monkeys utilized information from prolonged periods of time to drive their decision, we computed the leverage of momentary fluctuations in motion energy during 0% coherence trials on the monkeys' choices (**Fig. 2 C-D**). Motion energy fluctuations had a significant impact on choice for hundreds of milliseconds in both monkeys (357 ms for Monkey AN_{go} and 261 ms for Monkey SM_{wait}). This observation alone does not confirm that the monkeys integrated information over this whole duration during each trial (Stine et al 2020). For example, the monkeys could have integrated information over shorter time durations during each trial, but used different time epochs across trials. However, we will provide other behavioral and neural data in the following sections to support the thesis that monkeys integrated motion information over time during a trial.

The conversion of an abstract decision to an action is a deliberative process

The behavioral task is structured to separate the decision-making epoch from the action selection epoch. The natural expectation is that the monkey makes a decision about the rightwardness or leftwardness of motion (or equivalently, to choose the cyan or the yellow target) while viewing the motion stimulus (**Fig. 3A**, **Strategy 1**). If so, the action selection epoch would involve a simple translation of the categorical decision into an eye movement. Once the targets appear, the reporting time should be short and consistent, perhaps modulated by a few tens of ms by confidence or reward expectation (Schall & Thompson 1999, Tanaka et al 2015). Alternatively, the monkey might store the experienced stream of evidence in memory during the motion presentation epoch and consult this memory to make a decision about which of the two targets to make an eye movement to (**Fig. 3A, Strategy 2**).



Surprisingly, we found that the behavior of both monkeys were consistent with implementations of the second strategy. The action selection epoch either involved (in Monkey AN_{go}) or necessitated (in Monkey SM_{wait}) a prolonged period of deliberation spanning hundreds of milliseconds. Given that only the two colored targets were visible during the action selection epoch, this deliberation is likely in service of a decision about which saccadic motor plan to execute . The behavior of the two monkeys provided complementary insights into the deliberative process and we will now describe them separately.

Monkey ANgo:

The time taken by Monkey AN_{go} to report its decision (reaction time after the go-cue, hereafter *go-RT*) depended strongly on the strength of the motion presented in the previous epoch (Fig. 3B). Note that, irrespective of the strength of the motion experienced beforehand, the action selection epoch only involved deciding which of two highly distinguishable colored targets to make a saccade to. Yet, the average go-RT for correct decisions ranged from 440 to 771 ms, depending on the motion strength. The go-RTs were 146 to 176 ms slower on trials with the weakest motion strength compared to trials with the strongest motion strength (4% vs. 64% coherence — 440 vs. 616 ms for rightward motion and 624 vs. 771 ms for leftward motion). The prolongation of go-RTs is unlikely to be due to difficulty in distinguishing the colors of the targets. The range of go-RTs emitted by Monkey AN_{ao} spanned >2-4 times the mean go-RT of ~190 ms previously reported in monkeys discriminating between distinctly colored targets (Tanaka et al 2015). The prolongation is also unlikely to be a reflection of the animal's confidence in its decision as the range of go-RTs in monkey AN_{ao} is an order of magnitude larger than expected were it a function of the animal's confidence in its decision. For example, in monkeys performing a similar direction discrimination task with fixed target locations, the range of mean go-RTs as a function of stimulus strength is <20 ms (Gold & Shadlen 2003).

We hypothesized that the act of choosing between the colored targets was a deliberative process involving a sequential sampling of the evidence experienced beforehand (**Fig. 3A**, **Strategy 2**). We modeled the mean go-RTs as the terminations of drift-diffusion processes arising from sequential sampling of motion evidence from memory. This model provided a good fit for the goRTs observed following trials of different motion strengths (**Fig. 3B**). In the model, the time taken for the termination of a sampling process and the probability of the process resolving to an accurate decision are coupled. Thus, the parameters from the fit to go-RT furnish predictions of the monkey's accuracy as a function of motion strength (Shadlen & Kiani 2013); (**Fig. 3C**). These

predictions were surprisingly close to a logistic model fit to the actual choices of the monkey. We quantified the quality of the prediction by comparing it to predictions obtained from random perturbations of the mean go-RT. We used a conservative approach in which we preserved the actual go-RT at the highest and zero motion strengths and imposed an orderly dependence of the other random go-RTs on motion strength (**Fig. 3C inset**). Even with this approach, small perturbations of the go-RT (6%) were sufficient to produce substantially poorer predictions (p<0.001). The fidelity of the predictions supports the hypothesis that the prolonged go-RTs reflect a bounded sequential sampling of information from memory leading to the rendering of a decision.

A more elaborate drift-diffusion model fit to the observed distribution of goRTs provided additional insights into the action selection process. We fit the model to go-RTs of all trials and derived an estimate of the average decision time for each motion strength. For 0% coherence motion the estimated integration time was 228±184 ms (mean±SD). This estimate obtained from goRTs independently corroborated the finding from the motion kernel analysis (see Fig. 2C) that the monkey integrated motion information over hundreds of milliseconds to render its decision.

Monkey SM_{wait}:

In the 'go' version of the task, Monkey SM failed to show behavioral signatures of deliberative decision making. Despite being trained for longer than Monkey AN_{go}, the psychophysical thresholds plateaued at ~20% (**Fig. 3D**, **green**). Note that the lapse rates at the highest motion strengths were comparable to that of Monkey AN_{go}, making it unlikely that the low sensitivity was due to confusion in learning the mapping between motion direction and target color. The motion energy fluctuations during the 0% coherence motion trials had only weak and transient leverage over choices (**Fig. 3E**), suggesting that the monkey failed to integrate motion information over time. The failure to integrate information could be the result of a strategy wherein the monkey uses only the last few frames of motion information before target onset to drive its decision. To investigate this possibility, we varied motion strength within trials in six sessions. The

motion strength started out at 0% coherence and either stepped up or exponential built up to one of the two highest motion strengths. The time of the step or the slope of the build-up was varied across trials. Restricting strong motion evidence to only the end of each trial significantly worsened the monkey's accuracy, with performance dropping to chance levels even for the highest motion strength. Thus the monkey was using motion information from a short time window at the beginning of each trial to base its decisions.

Unlike Monkey AN_{go}, Monkey SM_{go} did not spontaneously develop a strategy of deliberation during the action selection epoch. The go-RTs were short and did not show a dependence on previously experienced motion strength (mean±sd: 192±25 ms). We surmised that the inability of the monkey to integrate motion evidence stemmed from this lack of deliberation after the targets appeared. We therefore enforced a wait time after the onset of the targets in the hope that this manipulation would allow the monkey to perform evidence integration, despite the fact that no additional information from vision is available in this epoch. This simple modification greatly improved the monkey's sensitivity. The psychophysical thresholds dropped from 20% to 11% motion coherence (Fig. 3D, green vs. blue traces). This two-fold improvement in sensitivity implied that the monkey now used four times the number of independent samples to form its decision. This was reflected in the time window over which motion information influenced decisions, which increased from 40 ms to 261 ms (cf. Fig. 3E and Fig. 2D), Thus, the imposition of a wait after the onset of the targets enabled the monkey to perform integration of motion evidence, even though the evidence was acquired well before this epoch.

The behavioral data from both monkeys provides complementary evidence that deliberation during the action selection epoch is necessary for integrating previously observed motion information. This integration appears to be instantiated as a sequential sampling of information from memory. We investigated whether the responses of saccade preparatory neurons in the parietal cortex represents such a process.

The evolution of neural responses in LIP during action selection reflect previously presented evidence

FIGURE 4



To understand the neural basis of how remembered motion information directs the choice of an appropriate eye movement, we recorded from single units in area LIP, a parietal association area involved in the planning of eye movements and attention. We recorded from 60 neurons (29 in Monkey AN_{go} and 31 in Monkey SM_{wait}) that showed spatially selective saccade preparatory response fields (RF) during a delayed saccade task. During the recording sessions, the target locations were still randomized from trial to trial, but the RF locations were sampled at a higher rate (Each colored target appeared in the RF on 33% and 28% of the trials for Monkey AN_{go} and Monkey SM_{wait} respectively).

During the motion viewing epoch the neural responses showed weak (Monkey AN_{go}) to absent (Monkey SM_{wait}) dependence on motion strength (**Fig. 4A,B**). In the action selection epoch, the appearance of a colored target in the RF initially elicited a strong visual response at a

latency of 80 ms (**Fig. 4C-F**). The evolution of the response after that reflected both the strength and direction of the RDM stimulus that had been presented in the previous epoch. To visualize the relationship between the neuronal response and the previously presented RDM stimulus, we removed the sensory component of the response by subtracting the mean response across all trials for each color target (insets in **Fig. 4C-F**). This revealed that the strength of the motion stimulus affected the rate of rise of responses in the action selection epoch.

FIGURE 5



To quantify the effect of the motion strength on the slope of the response, we first determined the earliest times at which the responses diverged at the strongest motion strength for the two directions of motion. We computed an ROC metric from the spike rates in 60 ms bins starting from the onset of the target (**Fig. 5A,B**). The trials were sorted by the direction of motion of the RDM presented earlier in the trial. For Monkey AN_{go}, the earliest of consecutive time bins that showed statistical significance was 170 ms, but for Monkey SM_{wait}, this was 100 ms. We

considered these respective time points to be the start of the neural correlates of the memory dependent deliberation process for each monkey.

We quantified the effect of the strength of the previously presented RDM on the rate of increase of neural responses ("buildup rate") after target onset. We estimated the slope for each neuron at each coherence from the mean detrended response in 20 ms time bins. For Monkey AN_{go}, we computed buildup rates starting from 170 ms after stimulus onset, but discarding spikes occurring within 200 ms of saccade onset in each trial (to remove any pre-saccadic bursts). For Monkey SM_{wait}, buildup rates were computed starting from 100 ms (putative start of buildup from latency analysis) to 350 ms (a 250 ms window of motion information determined from the motion energy analysis shown in Fig. 2D).

The average buildup rates across the neural population varied as a function of motion strength and motion direction for each monkey, when either of the colored targets was in the RF (Fig. 5 C-F). We tested whether these buildup rates scaled with coherence across the population in each stimulus configuration by fitting a linear model regressing the buildup rates against signed coherence. For Monkey AN_{go}, when either target was in the RF, buildup rates were linearly dependent on motion strength when the preferred direction had been previously presented (Cyan target in RF: buildup=2.8 spikes per s² per % coherence, p=10⁻⁴; Yellow in RF: buildup=1.3 spikes per s² per % coherence, p=0.005). If the non-preferred direction had been presented, then the responses showed a significant inverse dependence on coherence only when the Yellow target was in the RF (buildup=-1.7 spikes per s² per % coherence, p=0.001). For Monkey SM_{wait}, the linear dependence was seen in all four conditions (Cyan target in RF, Fig. 3D: buildup=1 spikes per s² per % coherence, p=0.046 when motion favored the target and buildup=-1.8, p=0.002 when motion was against; Yellow in RF, Fig. 3F: buildup=1.2, p=0.01 when motion favored the target and buildup= -0.9, p=0.037 when motion was against). The fact that the strength of the motion presented in an earlier epoch affects the rate of change of neural responses in the target selection epoch suggests a sustained influence of the motion information stored in memory

REFERENCES:

- Bennur S, Gold JI. 2011. Distinct representations of a perceptual decision and the associated oculomotor plan in the monkey lateral intraparietal area. *J Neurosci* 31: 913-21
- Chafee MV, Crowe DA. 2012. Thinking in spatial terms: decoupling spatial representation from sensorimotor control in monkey posterior parietal areas 7a and LIP. *Front Integr Neurosci* 6: 112
- Cisek P. 2007. Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos Trans R Soc Lond B Biol Sci* 362: 1585-99
- Cisek P, Kalaska JF. 2005. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* 45: 801-14
- Cisek P, Kalaska JF. 2010. Neural mechanisms for interacting with a world full of action choices. *Annu Rev Neurosci* 33: 269-98
- Freedman DJ, Assad JA. 2006. Experience-dependent representation of visual categories in parietal cortex. *Nature* 443: 85-8
- Genovesio A, Tsujimoto S, Wise SP. 2009. Feature- and order-based timing representations in the frontal cortex. *Neuron* 63: 254-66
- Gold JI, Shadlen MN. 2003. The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *The Journal of neuroscience* 23: 632-51
- Klaes C, Westendorff S, Chakrabarti S, Gail A. 2011. Choosing goals, not rules: deciding among rulebased action plans. *Neuron* 70: 536-48
- Kubanek J, Snyder LH. 2015. Reward-based decision signals in parietal cortex are partially embodied. *J Neurosci* 35: 4869-81
- Schall JD, Thompson KG. 1999. Neural selection and control of visually guided eye movements. *Annu Rev Neurosci* 22: 241-59
- Shadlen MN, Kiani R. 2013. Decision making as a window on cognition. Neuron 80: 791-806
- Shadlen MN, Kiani R, Hanks TD, Churchland AK. 2008. An intentional framework. *Better than conscious? Decision-making, the human mind, and implications for institutions*: 71-101
- Stine GM, Zylberberg A, Ditterich J, Shadlen MN. 2020. Differentiating between integration and nonintegration strategies in perceptual decision making. *Elife* 9
- Tanaka T, Nishida S, Ogawa T. 2015. Different target-discrimination times can be followed by the same saccade-initiation timing in different stimulus conditions during visual searches. *J Neurophysiol* 114: 366-80
- Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND. 2011. How to grow a mind: statistics, structure, and abstraction. *Science* 331: 1279-85
- Wallis JD, Anderson KC, Miller EK. 2001. Single neurons in prefrontal cortex encode abstract rules. *Nature* 411: 953-6
- Wang M, Montanede C, Chandrasekaran C, Peixoto D, Shenoy KV, Kalaska JF. 2019. Macaque dorsal premotor cortex exhibits decision-related activity only when specific stimulus-response associations are known. *Nat Commun* 10: 1793