

# Current Biology

## Prefrontal cortex represents heuristics that shape choice bias and its integration into future behavior

### Highlights

- Highly trained monkeys exhibit slow and fast behavioral biases.
- Both bias types are represented in prestimulus activity of prefrontal cortex cells.
- That bias representation allows predicting upcoming choice.
- Initially, bias shifts evidence accumulation to be later dynamically integrated.

### Authors

Gabriela Mochol, Roozbeh Kiani,  
Rubén Moreno-Bote

### Correspondence

[gabriela@mochol.com](mailto:gabriela@mochol.com)

### In Brief

Using a decision task Mochol et al. demonstrate that highly trained monkeys develop choice biases fluctuating over slow and fast timescales. Bias information is present in prefrontal cortex activity and enables to predict animal choice prior to stimulus. Bias initially acts as an offset to the decision variable and is dynamically accumulated later.



## Article

# Prefrontal cortex represents heuristics that shape choice bias and its integration into future behavior

Gabriela Mochol,<sup>1,6,\*</sup> Roozbeh Kiani,<sup>2,3,4,5</sup> and Rubén Moreno-Bote<sup>1,5</sup><sup>1</sup>Center for Brain and Cognition and Department of Information and Communications Technologies, Pompeu Fabra University, Barcelona, Spain<sup>2</sup>Center for Neural Science, New York University, New York, NY 10003, USA<sup>3</sup>Neuroscience Institute, NYU Langone Medical Center, New York, NY 10016, USA<sup>4</sup>Department of Psychology, New York University, New York, NY 10003, USA<sup>5</sup>These authors contributed equally<sup>6</sup>Lead contact\*Correspondence: [gabriela@mochol.com](mailto:gabriela@mochol.com)<https://doi.org/10.1016/j.cub.2021.01.068>

## SUMMARY

Goal-directed behavior requires integrating sensory information with prior knowledge about the environment. Behavioral biases that arise from these priors could increase positive outcomes when the priors match the true structure of the environment, but mismatches also happen frequently and could cause unfavorable outcomes. Biases that reduce gains and fail to vanish with training indicate fundamental suboptimalities arising from ingrained heuristics of the brain. Here, we report systematic, gain-reducing choice biases in highly trained monkeys performing a motion direction discrimination task where only the current stimulus is behaviorally relevant. The monkey's bias fluctuated at two distinct time scales: slow, spanning tens to hundreds of trials, and fast, arising from choices and outcomes of the most recent trials. Our findings enabled single trial prediction of biases, which influenced the choice especially on trials with weak stimuli. The pre-stimulus activity of neuronal ensembles in the monkey prearcuate gyrus represented these biases as an offset along the decision axis in the state space. This offset persisted throughout the stimulus viewing period, when sensory information was integrated, leading to a biased choice. The pre-stimulus representation of history-dependent bias was functionally indistinguishable from the neural representation of upcoming choice before stimulus onset, validating our model of single-trial biases and suggesting that pre-stimulus representation of choice could be fully defined by biases inferred from behavioral history. Our results indicate that the prearcuate gyrus reflects intrinsic heuristics that compute bias signals, as well as the mechanisms that integrate them into the oculomotor decision-making process.

## INTRODUCTION

Choice biases are prevalent.<sup>1</sup> Biases that reflect imbalanced priors or reward expectations in the environment are advantageous as they could improve the speed or overall gain of our choices.<sup>2–8</sup> However, biases could also hinder performance,<sup>9</sup> especially when they arise from heuristics that do not truly capture the environment or task structure. It is often hypothesized that the history of past stimuli, actions, and outcomes inform these heuristics for sequential choices, where subjects perform a series of similar decisions.<sup>5,9–22</sup> However, the computations involved in these heuristics, specifically what constitutes past history (e.g., time scales) and how it influences future choices (e.g., changes in the starting point of the decision-making process<sup>19,23</sup> or dynamic bias signals<sup>4,20</sup>) is not well understood. Further, the neural representation of history dependent biases (e.g., locus and neural code) and their integration in the decision-making process remain underexplored.

Addressing these questions is especially impactful in a task in which heuristic biases are not rewarding, and thus there is no task incentive to develop them. Biases that increase reward rate encourage alteration of decision strategies based on task structure.<sup>4,5,10,11</sup> The underlying computations and neural mechanisms may, therefore, not generalize to conditions where biases are non-rewarding and not induced by the experimenter. Investigating non-rewarding biases that persist despite training provides an opportunity to study the neural computations underlying choice bias free from experimenter-induced factors and offers an opportunity to compare the findings with those of past studies<sup>4,5,10,11,17</sup>

Addressing our questions greatly benefits from single trial quantification of the magnitude of bias, as well as recording from neural ensembles in brain regions that represent both the bias and the decision-making process. Single trial quantification of the magnitude of bias necessitates development of behavioral models that can accurately predict the bias on individual trials.

Although many studies attempted to do so,<sup>16,23,24</sup> there are few comprehensive models that achieve sufficient accuracy, often because they ignore one or more key factors that shape the bias. Additionally, past studies on the neural representation of bias focused largely on single neuron activity,<sup>4,13,16–19</sup> whereas accurate characterization of the moment-by-moment fluctuation of the state of the neural population requires simultaneous recording of many neurons.<sup>24–26</sup> Finally, past studies focused largely on finding a neural representation of bias and rarely explored how the bias is integrated in the decision-making process. Here, we overcome these challenges for the first time by developing a comprehensive framework that determines the magnitude of bias on individual trials, characterizes bias representation by the prefrontal neural population, and determines the computational mechanism for integrating the bias in the decision-making process.

We recorded simultaneously from large populations of prearcuate gyrus neurons while monkeys performed a direction discrimination task, designed such that past history was unrelated to the present stimulus, making history-dependent biases suboptimal. Nonetheless and despite extensive training, monkeys showed detectable biases that fluctuated throughout and across experimental sessions. Their choice biases stemmed from two sources: fast biases shaped by actions and outcomes of past trials, especially the most recent one, and slow biases fluctuating over tens to hundreds of trials. Our single-trial quantification of bias enabled us to show that the total bias, as well as its two sources, was represented in the population activity of pre-arcuate gyrus neurons prior to the stimulus onset. The same pre-stimulus neural responses were also predictive of the upcoming choice. These neuronal representations of bias and choice were well aligned in the activity state space, suggesting that the pre-stimulus choice prediction was achieved through the representation of bias, which itself reflected past choices and feedbacks. Finally, we demonstrate that the bias influenced the decision-making process as an initial offset in the accumulation of evidence, followed by a dynamic bias signal throughout the stimulus presentation, pointing at the computational mechanism for the integration of bias in the decision-making process.

## RESULTS

The present study is the reanalysis of previously published data in the context of new scientific question.<sup>26,27</sup> Using a 96-channel multi-electrode array, we recorded neural population activity from the prearcuate gyrus (PAG, area 8Ar; 169–250 single and multi-units in each session; median = 220), while monkeys ( $n = 2$ ) performed a direction discrimination task (9 and 7 sessions for monkeys 1 and 2, respectively<sup>26,27</sup>). Each trial began with the monkey fixating on a central fixation point on the screen, following by the appearance of two targets (T1 and T2; Figure 1A), and a circular patch of random dot kinematogram.<sup>28</sup> The percentage of coherently moving dots (coherence or motion strength) and the net motion direction varied randomly from trial to trial. The motion stimulus was shown for 800 ms and was followed by a delay period. The monkey reported motion direction at the end of the delay period with a saccadic eye movement to the corresponding targets. We use signed motion coherence<sup>19,28,29</sup> to jointly represent the stimulus strength and

direction with a single variable (positive for motion toward T1 and negative for motion toward T2).

### Highly trained monkeys exhibit slow and fast choice fluctuations

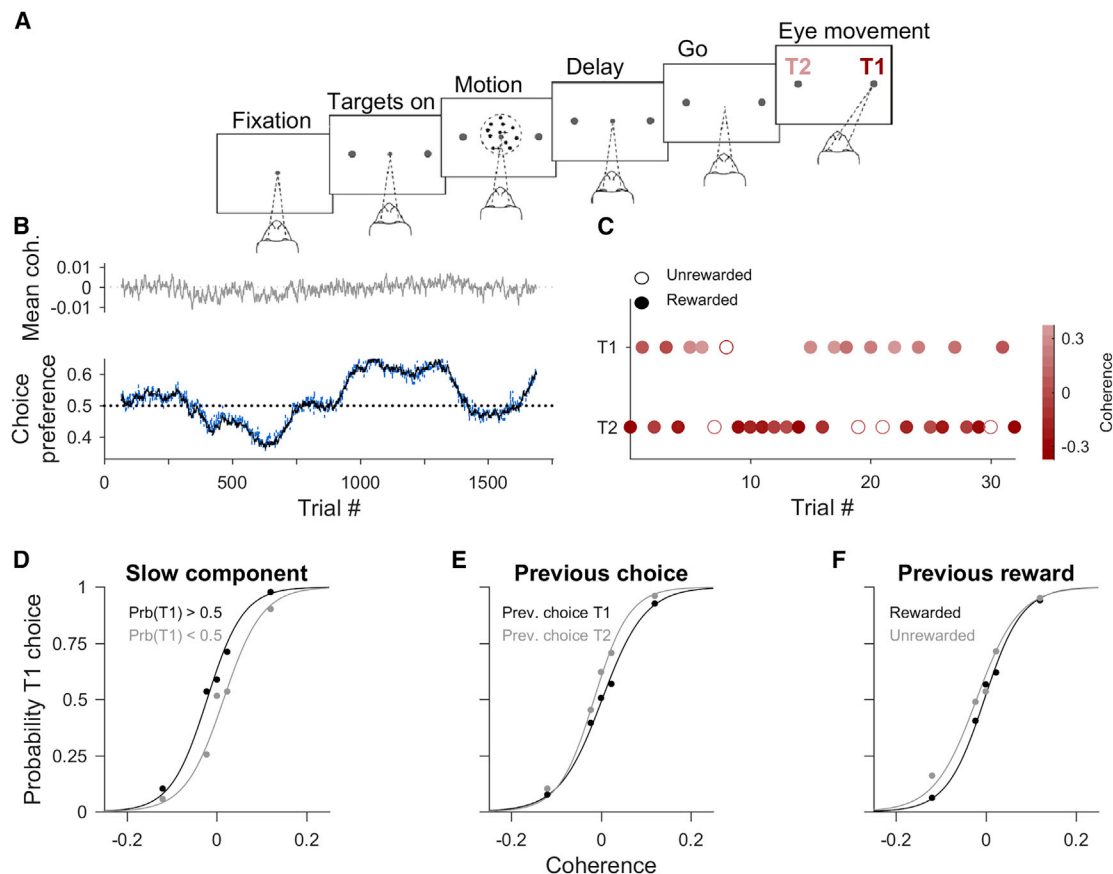
Monkeys were extensively trained in the task and showed stable performance prior to neural recordings. The average fraction of correct choices was equal to  $0.78 \pm 0.01$  and  $0.74 \pm 0.008$  for monkeys 1 and 2, respectively. The mean sensitivity threshold (coherence strength corresponding to 0.75 accuracy) equaled  $0.08 \pm 0.012$  and  $0.11 \pm 0.019$ . Both monkeys barely exhibited lapses; the lapse rate (1- accuracy for the strongest stimulus) was close to zero ( $0.002 \pm 0.0007$  and  $0.028 \pm 0.007$ ). Despite their extensive training and trial-to-trial independence of stimulus conditions, both monkeys demonstrated slow fluctuations in their choice preference, where monkeys chose one target more frequently than the other for tens to hundreds of trials before reversing their preference (Figure 1B; black line at the bottom). These fluctuations were spontaneous and could not be explained by fluctuations of motion direction across trials because motion direction was largely balanced in those periods (Figure 1B, top). To further ensure that the slow choice preference fluctuation did not merely reflect random fluctuations arising from spurious unbalance of motion directions and coherence, we repeated our analysis and replicated our results by subsampling trials to equalize the number of trials with stimuli moving toward T1 or T2 for each coherence (Figures 1B, S1A, and S1B blue line; see STAR Methods, Equation 2). Across sessions, the correlation coefficient between slow choice preference fluctuations and signed motion coherence was weak and not statistically significant (mean  $\pm$  SEM  $0.02 \pm 0.04$ , permutation test  $p$ -value = 0.52). Additionally, the mean auto-correlogram of slow choice preference fluctuations calculated on coherence-balanced trial history showed a statistically significant broad central peak (Figure S1; one-sided permutation test,  $p$ -value < 0.05), indicating the presence of slow fluctuations of response preference, irrespective of fluctuations of stimulus statistics.

On a finer timescale, we also observed that the monkey's choices were influenced by recent choices and outcomes. As illustrated by the example trial sequence in Figure 1C, the monkey tended to choose the opposite target after error trials in this session. Fluctuations of both slow and fast choice preference were reflected as a shift in the psychometric curve (see STAR Methods, Equation 1) when it was calculated conditioned on the direction of slow choice preference (Figure 1D), previous choice (Figure 1E), or recent outcome (Figure 1F).

### Improvement of choice prediction accuracy with slow and fast choice fluctuations

To quantify how monkeys' decisions were affected by the slow and fast fluctuations of choice preference, we measured whether and how much they would improve the prediction accuracy of upcoming choice beyond that given by motion stimulus alone.

We built a logistic regression model to predict choices based on three variables: stimulus coherence, slow choice preference fluctuation, and fast choice preference fluctuations expressed as a combination of previous choice and reward (see STAR Methods, Equations 2–4). The cross-validated model prediction accuracy was assessed using a *leave-one-out* procedure. To

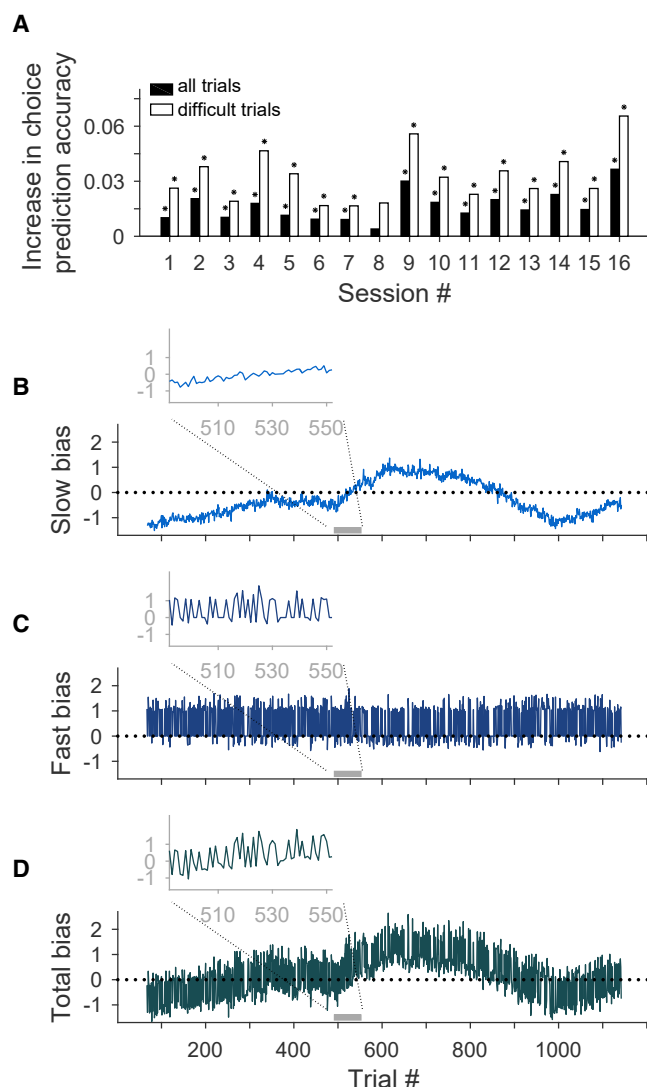


**Figure 1. Well-trained monkeys in a random dots direction-discrimination task show slow and fast fluctuations in their preferred choice**  
(A) Task design. After the monkey acquired a central fixation point, the patch of randomly moving dots appeared on the screen for 800 ms. The fraction of dots moving coherently in a given direction defined the trial difficulty. The motion was followed by a delay period with variable length, after which the monkey indicated its choice by making a saccade towards one of the two targets (T1 or T2).  
(B) Top – average of signed motion coherence. Dashed line indicates 0% coherence, and positive and negative values indicate motion toward T1 and T2, respectively. Bottom – black line shows the fraction of trials in which the monkey chose T1. The dashed blue line also shows the choice preference towards T1 but calculated after balancing coherence in a given window. All curves calculated in a 130 trial running window.  
(C) The monkey's choices in a sample sequence of 30 trials. Color intensity indicates signed motion coherence direction, with positive values showing motion toward T1 and negative values motion toward T2.  
(D–F) Psychometric curves from one experimental session (the same session as in B and C) computed conditioned on the monkey's slow choice preference (D), previous choice (E), and previous reward (F). Dots indicate actual data and the solid lines are maximum likelihood fits of logistic functions (see STAR Methods). See also Figure S1.

measure the role of choice preference fluctuations in determining choices, we compared the prediction accuracy with a baseline obtained by fitting the model using shuffled fast and slow choice preferences across trials. Because shuffling destroys the statistical relationship between current choices and choice preference fluctuations, the comparison with the baseline isolates improvement of the prediction accuracy conferred by the latter. The mean prediction accuracy improvement across all sessions was small but significantly larger than zero ( $0.016 \pm 0.002$ , one sample t-test,  $p$ -value =  $10^{-6}$ ; Figure 2A; black bars; see Table S2 with results for two monkeys separately). Importantly, when we focused only on difficult trials where the stimulus is less informative and biases could have a larger influence on choice, the improvement of choice prediction accuracy doubled (mean  $\pm$  SEM:  $0.032 \pm 0.004$ ; one sample t-test,  $p$ -value =  $10^{-7}$ ; Figure 2A; open bars; see Table S2 with

results for two monkeys separately). Consistently, the improvement tested only on easy trials was not different from zero (mean  $\pm$  SEM:  $0.0001 \pm 0.0002$ ; one sample t-test,  $p$ -value = 0.69; see Table S2 with results for two monkeys separately), which is expected because prediction accuracies based on stimulus strength alone are already close to ceiling. Overall, biases had a tangible effect on upcoming choices especially for more difficult decisions.

It is possible to use similar choice prediction models to quantify the temporal extent of fast and slow choice preference fluctuations. For fast choice preference fluctuations, information about choice and reward of the previous trial significantly improved choice prediction accuracy (Equation 5 in STAR Methods, mean difference tested on difficult trials equaled  $0.015 \pm 0.003$ ; one sample t-test,  $p$ -value =  $3 \times 10^{-5}$ ). However, including information about more distant past (from two to five



**Figure 2. Predicting the monkey's choices becomes more accurate by using fluctuations of behaviorally defined choice preference**

(A) Improvement of choice prediction accuracy measured as the difference of the accuracy of a logistic regression model with fast and slow choice preference fluctuations and a reference model in which choice preferences were shuffled across trials. Both models contained stimulus strength (signed motion coherence) as a regressor in addition to the bias terms. Improvement in choice prediction accuracy was higher when computed for difficult trials (white bars, compare to black bars for all trials). Each pair of bars represents a single session. (\* one-sided permutation test;  $p$ -value < 0.05)

(B–D) Traces of slow (B, blue), fast (C, navy blue), and total (summed fast and slow, D, green) biases across a sample experimental session. On each trial, biases were computed from the choice preference fluctuations multiplied by their corresponding weights from the logistic regression model. Insets zoom in on a sequence of 60 trials in the session (the grey bar on the x-axis) for better visualization of the dynamics of different types of biases. Note the different time scale between slow and fast biases but their similar contribution to the total bias in terms of their magnitudes.

See also Figure S2, Tables S1, and S2.

trials back) did no improve the prediction accuracy further (paired t-test on prediction accuracies,  $p$ -value > 0.32; Figure S2A), indicating that in our task, fast choice fluctuations

were shaped in a time scale that was no longer than one single trial in the past.

We studied how the size of the trials window used to calculate slow choice preference improved performance of a model that included both slow and fast choice preference fluctuations compared to a model in which only fast choice preference fluctuations was used (in both models coherence was also used as a regressor). If slow fluctuations were defined using trial windows of less than 130 trials, there was not statistically measurable effect (Figure S2B; Equation 4, paired t-test on prediction accuracies,  $p$ -value > 0.063). However, when slow fluctuations were calculated in larger windows (130–400 trials), there was a statistically significant increase in prediction accuracy (paired t-test,  $p$ -value < 0.05 for 19 cases; Figure S2B). Therefore, we conclude that slow choice preference fluctuations estimated in a window of at least 130 trials and fast choice fluctuations capturing the immediately preceding choice and reward were the only reliable and measurable variables that predicted choice apart from the stimulus.

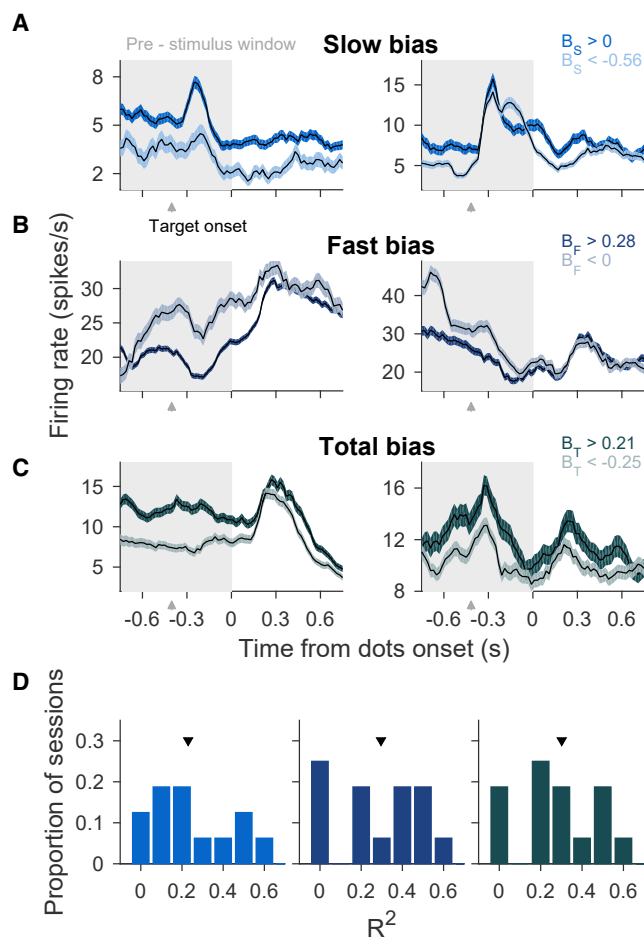
An important question is to know the relative strength of the effect of fast and slow choice preference fluctuations. To compare their effects, we first expressed both types of fluctuations in the same units (log-odd units) by using the logistic regression model described above. Thus, “slow bias” (Figure 2B) was defined as the product of slow choice preference fluctuation and its corresponding weight in the model plus the model offset (see STAR Methods). Similarly, “fast bias” (Figure 2C) was defined as the product of fast choice preference fluctuation and its model weights (see STAR Methods). To compare the relative strength of slow and fast biases on choice prediction, we calculated the mean ratio of the absolute value of each divided by the sums of absolute values of both biases and stimulus strength in log-odds space (see STAR Methods, Equation 6). We found that both fast and slow biases were effective in shaping the choice but, on average, the fast bias had a lower impact on choice (mean  $\pm$  SEM  $0.18 \pm 0.01$ ) compared to the slow bias (mean  $\pm$  SEM  $0.25 \pm 0.01$ ; difference of fast and slow =  $-0.07 \pm 0.016$ , one sample t-test  $p$ -value = 0.0006).

To capture the overall effect of slow and fast biases on behavior, we defined “total bias” (Figure 2D) on each trial as the sum of slow and fast biases. The total bias in log-odds units corresponds to the single-trial quantification of bias that is central for our analysis (for comparison, we provide Table S1 where total bias is expressed in units of coherence). On average, the contribution of total bias to the monkey's choices was  $0.25 \pm 0.01$  in units of log-odds. For comparison, it was more than twice smaller than the contribution of motion stimulus alone (mean  $\pm$  SEM  $0.58 \pm 0.02$ ), reflecting the fact that monkeys made their decision largely based on the presented stimuli, but they were also slightly impacted by their current bias.

### Impact of slow and fast biases on monkey's reward rate

Since our task was designed such that the stimulus sequence across trials did not have temporal correlations, the existence of behavioral biases described above can only impair the monkey's performance. As expected, periods with larger total bias correlated with periods with lower animal accuracy (fraction of correct choices; Pearson correlation coefficient  $-0.08 \pm 0.03$ , permutation test,  $p$ -value = 0.001). However, the reduction in accuracy was very small ( $0.007 \pm 0.003$ ; one sample t-test,





**Figure 3. Pre-stimulus responses of prearcuate gyrus neurons carry information about behavioral biases**

(A–C) Peristimulus time histograms (PSTHs) of example cells (left, monkey 1, and right, monkey 2) averaged across trials with high positive (dark) or negative (light) slow (A, blue), fast (B, navy blue), or total (C, green) biases. Specific bias ranges for which trials were averaged (see legends) corresponded to the 0.35 (negative) and 0.65 (positive) quantiles of the bias distribution for all sessions pulled together. Shaded areas correspond to SEM. On the x-axis, zero refers to stimulus onset. Target onset time is indicated by an arrow. Spikes were counted in 100 ms moving windows in steps of 20 ms.

(D) Histograms of the coefficients of determination (cross-validated  $R^2$ ) of a linear regression model fitted to predict slow (left, blue), fast (middle, navy blue), or total (right, green) biases from the pre-stimulus population activity of prearcuate gyrus cells ( $T = 800$  ms; grey areas in A–C). Arrows indicate mean across sessions ( $0.23 \pm 0.05$  for slow;  $0.29 \pm 0.06$  for fast and  $0.30 \pm 0.05$  for total biases respectively; mean  $\pm$  SEM). See also Figure S3.

$p$ -value = 0.02), suggesting that monkeys may not have noticed the adverse effect of biases on their performance or they did not find enough incentive to fully abolish them.

### Representation of bias in the pre-stimulus neural responses of PAG

Since the slow and fast biases influenced the monkey's choice (Figure 2A), the information about them should be present in brain regions involved in the decision-making process. An area of interest could be PAG, as it provides easy access to a large

scale recordings and its neural responses are related to the accumulation of evidence.<sup>26,30,31</sup> Also, since the observed behavioral biases had a history dependent component, there should be neurons sensitive to the bias even prior to the stimulus presentation. Such a tuning is illustrated in Figures 3A–3C, where we show responses of example neurons representing slow (Figure 3A), fast (Figure 3B), and total (Figure 3C) biases.

To investigate whether these biases were represented in the responses of PAG neural population, we used a linear model in which each type of bias (slow, fast, or total) was regressed against pre-stimulus activity of simultaneously recorded PAG neurons (Equation 8; spike counts calculated in 800 ms window prior to stimulus onset). The analyses were performed using first few PCA components of the neural population responses that collectively explained 50% of the variance (Equation 7). Here, we used PCA to reduce overfitting,<sup>31,32</sup> but qualitatively similar results were obtained without dimensionality reduction. All three biases were significantly represented in the pre-stimulus responses of PAG population (Figure 3D; mean of cross-validated  $R^2$  across sessions was equal to  $0.23 \pm 0.05$ ,  $0.29 \pm 0.06$  and  $0.30 \pm 0.05$  for  $B_S$ ,  $B_F$  and  $B_T$  respectively, one-sided permutation test,  $p$ -value = 0.001 in all three cases).

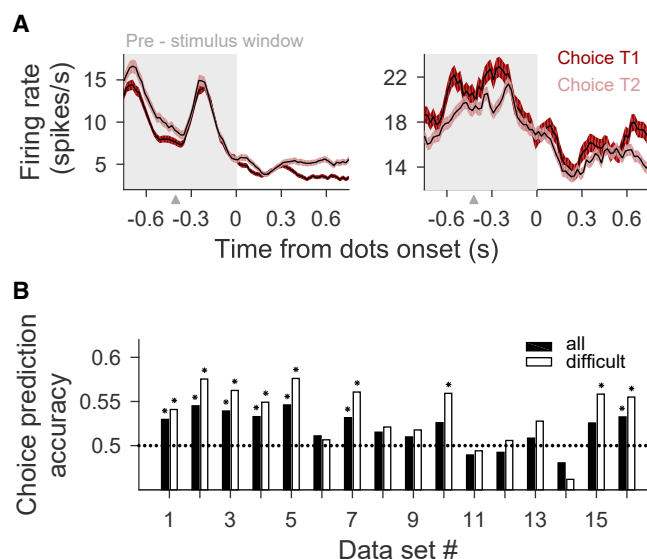
Consistent with a representation of the fast bias, PAG activity prior to the stimulus onset represented previous choice and reward, which, together, defined the fast bias. Examples of four neurons for which firing rates were modulated by previous choice or reward are shown in Figure S3 (A and C, respectively). In all cases, the response modulation toward previous choice or recent outcome persisted through the pre-stimulus period.

Fitting a logistic regression model to the firing rate of the population of simultaneously recorded PAG neurons (Equation 9 for  $n < 0$ ) revealed that previous choices could be decoded up to three trials back in the past (Figure S3B; means  $\pm$  SEMs:  $0.52 \pm 0.005$ ;  $p$ -value = 0.003 for  $n = -3$ ;  $0.53 \pm 0.006$ ;  $p$ -value = 0.0003 for  $n = -2$  and  $0.81 \pm 0.02$ ;  $p$ -value  $< 10^{-11}$  for  $n = -1$ , one-sided t-test; see Table S2 with results for two monkeys separately). A similar model (Equation 9 for  $n < 0$ ) could predict the outcome (reward) of the preceding trial (Figure S3D; mean  $\pm$  SEM  $0.81 \pm 0.01$ ; one-sided paired t-test,  $p$ -value 0.0001; see Table S2 with results for two monkeys separately), but not further back (see STAR Methods).

### Predicting choices from pre-stimulus neural responses of PAG

Given that slow and fast biases influenced monkeys' decisions (see Figure 2A) and that they were represented in the pre-stimulus PAG activity (see Figure 3), we asked whether pre-stimulus activity was also predictive of monkeys' upcoming choices. Figure 4A shows two example units. One of the units (Figure 4A, left) had a higher firing rate for choosing the T2 target and the other unit had a higher firing rate for the opposite target (Figure 4A, right). Importantly, both units represented the upcoming choice even prior to the stimulus presentation (grey areas), matching the representation of bias in the neural population.

To quantify this activity modulation at the population level, we fit a logistic regression model to predict upcoming choices using the PCA-dimensionality-reduced PAG population responses in the 800 ms before stimulus onset (Equation 9 for  $n = 0$ ). For many sessions (44%; 7 out of 16), the cross-validated prediction accuracy



**Figure 4. Pre-stimulus activity of prearcuate gyrus neurons predicts upcoming choice**

(A) Mean firing rate of two cells (from two monkeys) averaged across trials with T1 (dark) or T2 (light) choices. Shaded area corresponds to SEM. Arrow on the x-axis indicates target onset. Firing rates were calculated in a 100 ms windows moved in steps of 20 ms.

(B) Cross-validated prediction accuracy of a logistic regression model for predicting upcoming choice from the pre-stimulus activity of simultaneously recorded neurons in prearcuate gyrus (window size, 800 ms; grey area in A, PCA dimensionality reduction). Accuracy was higher when assessed for difficult trials only (white bars). Choice prediction accuracy was above chance level (0.5, dotted line) for seven or nine sessions when calculated for all or difficult trials respectively (\*, one-sided permutation test;  $p$ -value < 0.05). Mean across session was equal ( $0.52 \pm 0.005$ ,  $p$ -value = 0.001 for all trials;  $0.54 \pm 0.008$   $p$ -value = 0.001 for difficult trials mean  $\pm$  SEM  $p$ -values from one-sided permutation test).

See also Figure S3 and Table S2.

was above chance level (0.5) (Figure 4B, black bars; mean across sessions,  $0.52 \pm 0.005$ , one-sided permutation test  $p$ -value = 0.001; see Table S2 with results for two monkeys separately). Similar to the behavioral model, the cross-validated prediction accuracy was higher when we focused on difficult trials (white bars;  $0.54 \pm 0.008$  one-sided permutation test,  $p$ -value = 0.001; see Table S2 with results for two monkeys separately), and not significantly different from chance level for easy trials (mean  $\pm$  SEM across sessions  $0.5 \pm 0.004$  one-sided permutation test,  $p$ -value = 0.31; see Table S2 with results for two monkeys separately).

### Prediction of choices based on pre-stimulus neural activity is due to the representation of slow and fast biases

A key question is whether choice predictive neural responses prior to stimulus onset are due to the representation of the fast and slow biases that we have defined behaviorally. One possibility is that encoding of our behaviorally defined biases fully explains the representation of choice prior to the stimulus onset, which would imply that total bias and choice representations are “aligned” in neuronal activity space. Alternatively, choice predictive neural responses could arise from factors not fully captured by fast and slow biases, which would cause

misalignments between choice and total bias representations (see STAR Methods, Equation 13). To differentiate these two possibilities, we asked if the neural representation of biases was as predictive of behavior as the neural representation of choice.

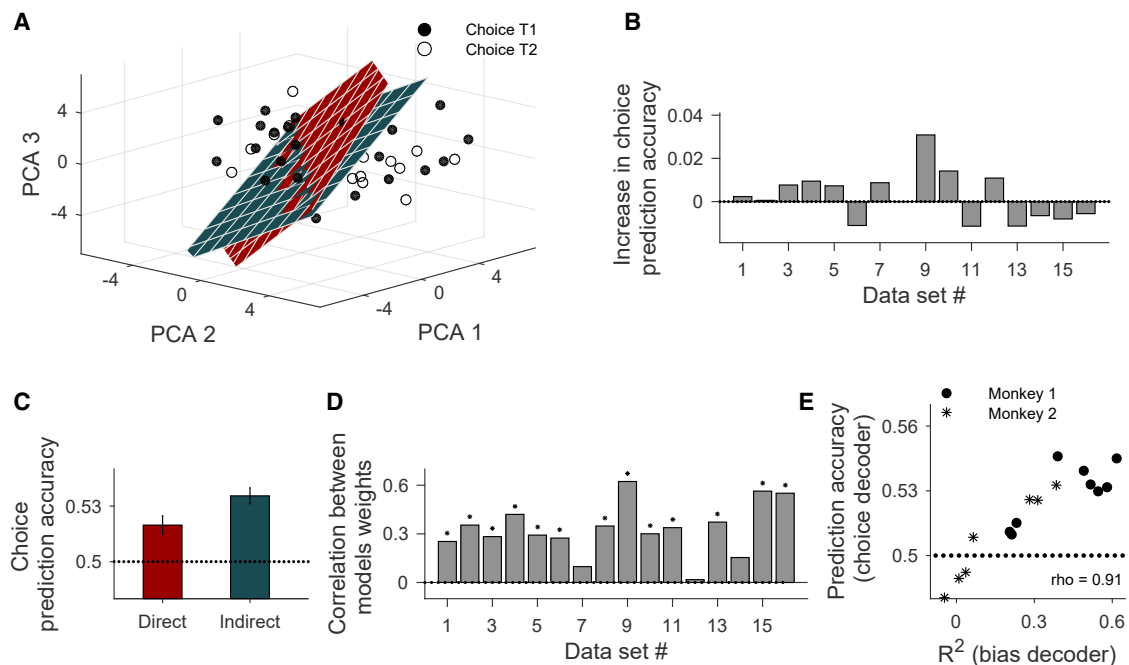
For each trial, we used the remaining trials in the session to find the best hyperplanes that explained the choice (Equation 9) and total bias (Equation 8) based on pre-stimulus responses (Figure 5A). Then, we calculated the distance of pre-stimulus responses of the left-out trial from those two hyperplanes ( $d_{\text{choice}}$  and  $d_{\text{bias}}$ ). If the pre-stimulus choice and bias representations were aligned, predicting the upcoming choice based on  $d_{\text{bias}}$  would be as accurate as using both  $d_{\text{choice}}$  and  $d_{\text{bias}}$  (Equations 10 and 11). Indeed, this was what we observed. Across sessions, the difference in predicted accuracy was negligible (Figure 5B; mean  $\pm$  SEM,  $0.002 \pm 0.003$ ) and not significant (paired t-test,  $p$ -value = 0.42), suggesting that the neural representation of total bias was functionally indistinguishable from the neural representation of choice prior to stimulus onset. Consistent with these results, predicting the choice based on the sign of  $d_{\text{bias}}$  or  $d_{\text{choice}}$  was comparable, with slightly better accuracies for  $d_{\text{bias}}$  (mean difference:  $0.02 \pm 0.004$ , paired t-test  $p$ -value = 0.001; see Table S2 with results for two monkeys separately); Figure 5C; mean prediction accuracy based on choice hyperplane,  $0.52 \pm 0.005$ ; one-sided permutation test  $p$ -value = 0.001; based on bias hyperplane  $0.54 \pm 0.004$ ; one-sided permutation test  $p$ -value = 0.001; see Table S2 with results for two monkeys separately), further supporting our conclusion.

Additional insight about choice predictive neural responses prior to stimulus onset is gained from comparing the geometry of choice and bias decoder hyperplanes. Vectors in high-dimensional spaces tend to be orthogonal.<sup>33</sup> However, if the representation of choice and bias are functionally aligned, one would expect that the angle between the norms of their respective hyperplanes is less than 90 deg. Indeed, the mean angle was equal  $71 \pm 2.5$  deg. and was significantly smaller than 90 deg. (mean  $\pm$  SEM, permutation test  $p$ -value = 0.001; for extended analysis see Figure S4). Additionally, we found that the weight vectors that defined the norm of the choice and bias hyperplanes ( $\beta_i$  in Equation 8 and  $\alpha_i$  in Equation 9) were positively correlated and the correlation coefficients were significant for the majority of sessions (Figure 5D;  $n = 13$ , one-sided permutation test  $p$ -value < 0.05; across session mean  $\pm$  SEM,  $0.33 \pm 0.04$ ; one-sided permutation test  $p$ -value = 0.001), supporting the notion that the hyperplane norms were not orthogonal.

Further supporting the alignment hypothesis, we found that sessions with stronger representation of the total bias (higher cross-validated  $R^2$ ) in pre-stimulus activity of PAG also had a higher cross-validated choice prediction accuracy (Figure 5E; Pearson correlation coefficient 0.91  $p$ -value =  $9 \times 10^{-7}$ ). These results suggest that across-session variability in choice predictive power could be explained by the across-session variability in the representation of total bias.

### The integration of bias into the accumulation of evidence

Given the presence of the bias signal prior to stimulus presentation, the question arises of whether and how this bias impacts the accumulation of the sensory evidence during the stimulus-



**Figure 5. Choice predictive power of pre-stimulus PAG neural responses is related to the representation of total bias**

(A) The total bias (green) and choice (red) decoders, corresponding to two discriminant hyperplanes that split neuronal state space in two regions for T1/T2 choices or positive and negative total bias, were trained using the same pre-stimulus activity of PAG neuronal populations. The panels show example data points (40 trials) where each dot represents the pre-stimulus population firing rates of a trial projected on the first three PCA dimensions. Filled circles correspond to T1 choices and hollow circles to T2 choices. Although the bias decoder has been trained to predict biases, it can be used “indirectly” to predict choices because the sign of the bias indicates tendency toward a choice in each trial (“indirect method”). In contrast, the “direct method” uses the choice decoder to predict choices. (B) Difference in choice prediction accuracy between a model including the distances from neuronal activity to the choice and bias hyperplanes as predictors, and a model with the distance to the bias hyperplane as the only predictor. The mean prediction accuracy difference between the two models was equal  $0.002 \pm 0.003$  and not significant (paired t-test,  $p$ -value = 0.42).

(C) Experimental data. Choice prediction accuracies from the direct and indirect methods were comparable, suggesting that the bias and choice decoders are aligned as illustrated in (A). As in (B), choice decoding was based on PAG activity from an 800 ms window before stimulus onset (0.5 chance level marked by dotted line; mean  $\pm$  SEM calculated across sessions, PCA dimensionality reduction).

(D) Consistent with an alignment of the bias and choice hyperplanes, the correlation coefficient between weights of choice and bias decoder hyperplanes (y-axis) was significantly positive for most of the sessions (\* one-sided permutation test,  $p$ -value < 0.05).

(E) Across sessions choice prediction accuracy (direct method) correlated with the strength of the total bias representation (defined as the  $R^2$  of the total bias linear regression model;  $corr = 0.91$ ,  $p$ -value =  $9 \times 10^{-7}$ ). Each point represents a single session (dots or stars for Monkey 1 or 2, respectively).

See also Figure S4 and Table S2.

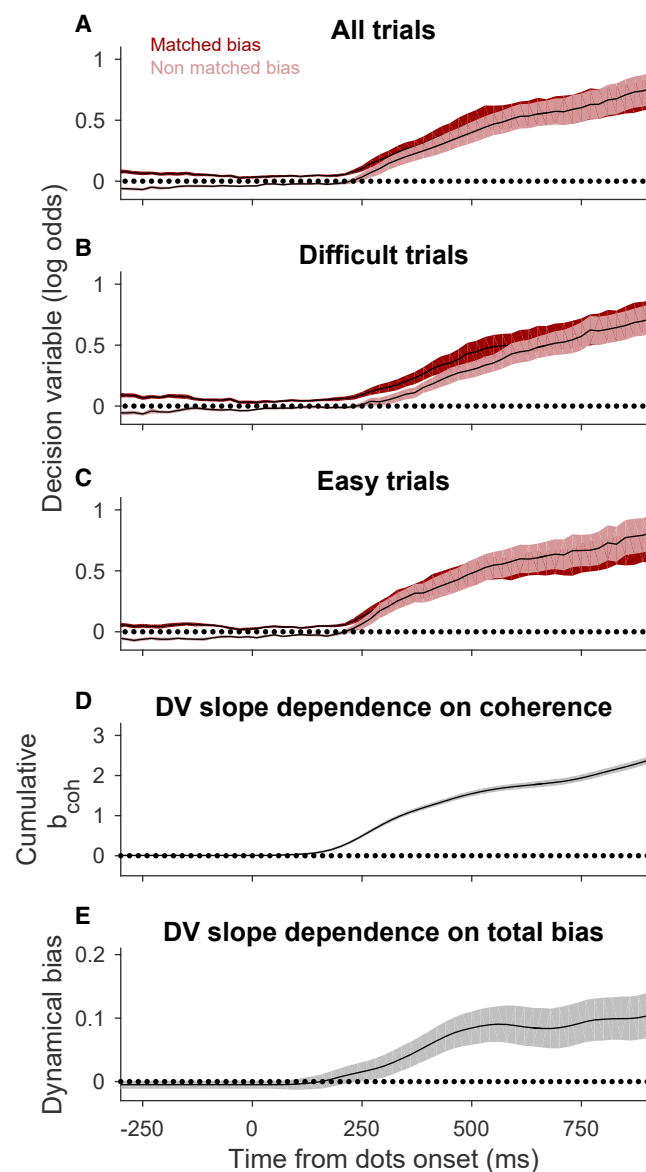
viewing period. Our result about the alignment of bias and choice decoders before stimulus onset suggests that the bias could be implemented as an initial offset in baseline activity before accumulation of sensory information begins in each trial. This initial offset is best visible in the neuronal activity axes where the decision is encoded (orthogonal to the choice decoder hyperplane shown in Figure 5A). Thus, we plotted how the decision variable (projection of neuronal activity onto the choice axes) evolved over time and how its slope depended on coherence and bias (Figure 6, see STAR Methods Equations 14, 15).

We found, consistent with our expectations, that there was an offset in the decision variable before stimulus onset. The offset was positive when the bias favored the final choice (Figure 6A, dark curve) and negative when the bias was against the final choice (light curve). This offset was roughly constant for the whole duration before stimulus onset and persisted during the first few hundreds of milliseconds of stimulus viewing period, when the decision was formed. Interestingly, the offset lasted longer for more difficult stimuli (Figure 6B), where

monkeys integrated the sensory evidence longer. However, toward the end of the stimulus presentation, the offset vanished because the monkey had likely reached a decision on the majority of trials. Our results suggest that the initial bias was integrated into the decision-making process and contributed to the formation of the choice.

The initial offset and final convergence of the decision variables for positive (Figure 6A, dark curve) and negative (light curve) biases are consistent with predictions of bounded evidence-accumulation models for the decision-making process.<sup>34–36</sup> The gradual dynamics of the decision-variables and larger and longer-lasting effects of offset for weaker motion stimuli (compare Figures 6B and 6C) is compatible with these models too. This is because the accumulation of sensory evidence is slower for weaker stimuli, and the decision variable takes longer to hit the decision bound. This slower ramping lets the bias-induced offset (light and dark red) survive longer during stimulus viewing. In contrast, sensory evidence accumulates quickly on easy trials, causing accelerated convergence of the decision





**Figure 6. Total bias acts as an initial offset of the decision variable that is dynamically accumulated afterwards during the stimulus presentation**

(A–C) Instantaneous decision variable (DV) from the logistic regression model of choice. The DV is averaged across trials in which the total bias was aligned with the monkey's choice (dark red, matched) or trials in which the total bias was against the final choice (light red, non-matched). Different panels show the average DVs for all (A), difficult (B), or easy (C) trials. The DVs were calculated based on 100ms of dimensionality-reduced neural responses centered at each time (PCA axes for dimensionality reduction were the same as those used in earlier figures and calculated for the 800ms window before stimulus onset). The analyses window was moved in steps of 20 ms. Shading indicates SEM across sessions.

(D and E) Cumulative  $b_{coh}$  (D) and cumulative (dynamic)  $b_{bias}$  (E) from a linear model in which the slope of the trial by trial DV was estimated as a weighted sum of coherence and total bias locally in time (see STAR Methods). Shaded area represents 95% confidence interval (bootstrap procedure,  $n = 1000$ ). Both models' weights for coherence and total bias were accumulated only from periods of significant  $b_{coh}$  and were not normalized by the number of points used in their calculation. DV slope was fitted in 200 ms window (10

variables and leaving minimal room for the bias-induced offset to influence the final decision.

Another interesting feature of the neurally inferred decision variable is that it continues to rise after the convergence of positive and negative bias curves in Figure 6. Because our model is designed to predict the choice (Equation 14), the magnitude of the model decision variable is influenced by any factor that improves its accuracy. Those include neural responses that may represent only the final choice but not necessarily the decision-making process that leads to the choice. Such choice-related responses have been shown to emerge in motor-planning regions toward the end of the stimulus viewing period in the dots task,<sup>37</sup> and could be responsible for additional rise of the model decision variable after 700 ms from stimulus onset when the initial bias is no longer represented.

An initial DV offset is not the only way bias can influence the evidence accumulation process. Favoring one choice over another can be also achieved by asymmetries in the rate in which stimulus information is accumulated.<sup>4,20</sup> To test this, we first fitted the slope of the trial by trial DV traces locally in time by using a sliding window of 200 ms to avoid the potential saturating effects of DV due to the decision threshold.<sup>4</sup> Next, we asked how this time-resolved slope depended on stimulus coherence and total bias by expressing it as a weighted sum of both quantities (Figures 6D, 6E, and S5; Equation 15). As expected, the DV's slope significantly depended on coherence shortly after stimulus onset (Figures 6D and S5A). Restricting ourselves to this time period, we also found that the total bias influenced the rate of evidence accumulation, but its impact was weak compared to that of the stimulus strength itself (Figures 6E and S5B).

## DISCUSSION

We have studied the dynamics and neuronal representation of biases in highly trained monkeys performing a direction discrimination task while recording simultaneous responses of hundreds of neurons in the prefrontal cortex. Despite trial-by-trial independence of the stimulus direction, monkeys exhibited weak but measurable suboptimal behavioral biases. Observed biases emerged at two distinct time scales. The slow bias, reflecting the monkey's preference towards one of the targets, fluctuated at a time scale spanning tens to hundreds of trials. The fast bias was shaped by the choice and outcome of only the preceding trial. Together, these biases improved prediction accuracy of upcoming choices beyond that given by stimulus alone. As expected, this increase was higher on trials with weak stimuli. Further, we found that pre-stimulus population activity of prearcuate gyrus represented the fast and slow biases. The same activity was also predictive of the monkey's upcoming choices. Critically, the axes that represented bias and choice in the neural population state space were similar; suggesting that choice-prediction power of pre-stimulus prefrontal activity was largely due

points) and moved with 20 ms resolution. Analysis was based on pooled data from all sessions. Note that the accumulated effect of bias (E) is approximately one order of magnitude smaller than the accumulated effect of coherence (D). See also Figure S5.

to the representation of the fast and slow biases. Conditioned on behavioral biases, we demonstrated that biases were incorporated into the decision-making process as an offset of baseline activity that was dynamically accumulated throughout the integration of sensory evidence during motion stimulus presentation.

To make optimal decisions, one should take into account all available and relevant information. This idea is expressed in Bayesian decision theory where choices depends on both the current sensory information and prior expectations (for review, see<sup>38</sup>). Such prior expectations might reflect for example the previously learned statistics of the stable environment and as such ease correct decisions especially when sensory evidence is ambiguous.<sup>4,39</sup> In contrast, when environment is unpredictable, prior expectations do not provide additional useful information into decision at hand but instead might lead to suboptimal behavioral biases that potentially reduce accuracy. In our task, prior history was uninformative, and thus all biases that we observed despite extensive monkey's training could not arise from a reward optimization strategies. As expected, the observed biases were weak, but nevertheless could be reliable measured on a trial-by-trial basis. That gave us a unique opportunity to explore innate mechanisms that shape history-dependent biases that arise spontaneously and are unrelated to any imposed task strategy.

In perceptual and value based decision making tasks one particular type of biases refers to the fact that current choices can depend on previous history of stimuli, rewards, and choices even in the conditions where such dependence is not relevant for the task, so-called sequential biases.<sup>5,9,10,12–14,16–19</sup> Such history dependences can span from one trial up to several trials in the past.<sup>5,9,12,16,17,23,40</sup> Only a couple of studies have reported the presence of separate time scales in sequential biases. One such example refers to the study where monkeys performed perceptual decision task and during training exhibited intrinsic slow (tens of trials) and fast (previous few trials) choice sequential biases.<sup>23</sup> As the slow bias component decreased with training, it is likely that it was a byproduct of learning. Fast and slow bias components have been shown to depend also on sensory components of the previous history.<sup>9</sup> Our results add to that body of research and constitute a rather unique example of such biases in highly trained primates performing unstructured tasks in which past trials are irrelevant for choice at hand. As it has been suggested, the existence of such biases might be a byproduct of a priori adaptive mechanisms that take advantage of stability of natural environment to make faster and more accurate decisions.<sup>38</sup> As such, biases can reflect different learning mechanisms with a slow process responsible for tracking the changes in the statistics of a non-stationary environment.<sup>41,42</sup> However, in our task biases, while having an impact on choice, they have very little impact on task performance. Whereas fast bias might reflect fast learning mechanism described in above mentioned study, it is unlikely that slow bias is a consequence of a slow learning process as it does not relate to stimulus or reward statistics. Instead, it might rather reflect slow fluctuation in the internal cognitive state of the animal as recently demonstrated.<sup>43</sup>

In many studies, fast sequential biases depend on past choices and rewards from more than one trial back.<sup>5,9,10,12–14,16–19,44</sup> One of the differences between the above studies and our results is the fact that, to define the fast bias, past choices and outcomes

were jointly used. Each of those pre-trial variables, when treated independently, had lower effect on upcoming choices, probably due to the high level of training of our animals. The level of expertise could also be a reason why fast biases were restricted to the most recent past choice and reward. Nevertheless, to track the sources for those differences, it would be important to jointly analyze training and steady-state behaviors to determine how time scales evolve with learning.

Given that history-dependent biases impact choices, they should be represented in the form of choice-predictive neuronal activity in brain regions that represent both the bias and the decision-making process. Such choice predictability has been described in the pre-stimulus activity of LIP neurons in the seminal work of Shadlen and Newsome.<sup>19</sup> However, as the authors did not investigate history-dependent biases, it was unclear what the source of this choice predictability was. Several studies investigated whether choice predictability reflects past history of stimuli, choices, and outcomes,<sup>9,13,16,17,19,21,45–48</sup> Although recently it has been demonstrated that choice-predictive signals in visual cortex could be partially accounted for by previous history,<sup>16</sup> the remaining unexplained residuals probably reflected unmeasured biases at the behavioral level that are nevertheless measurable at the neuronal level. At the behavioral level, our results go beyond these results by providing a single-trial quantification of bias that seems to exhaust all biases that are linearly decodable from neuronal population activity in PAG. That allowed us to demonstrate that choice predictive power of prefrontal cortex pre-stimulus activity can be explained by history dependent biases. While it is still possible that other biases exist in the monkeys' behavior, their temporality or non-linearity make them hard to detect and measure. Similar results to the ones reported here could potentially also hold in other brain areas, such as lateral intraparietal area (LIP), superior colliculus (SC), frontal eye field (FEF), or dorsolateral prefrontal cortex (dlPFC)<sup>4,19,30,49,50</sup> where neuronal activity has been shown to reflect accumulation of evidence. Therefore, although it is possible that PAG is not the only brain area where an integration of bias and stimulus signals holds, the details about how this integration works as reported in the present study remains to be seen in other candidate brain areas.

Several models have proposed how biases might mechanistically combine with sensory information to form the final decision.<sup>51–55</sup> One of the predictions is that biases can act either as an offset or as a change in the slope of sensory information accumulation<sup>20,36,53,54</sup> (but see<sup>56</sup>). Previous experimental work provided evidence in favor of the offset hypothesis.<sup>19,23</sup> For instance, conditioning neuronal responses on the final choice Shadlen and Newsome demonstrated an offset in pre-stimulus LIP neuronal activity that persisted during the stimulus presentation period.<sup>19</sup> However, since the choice is a combination of bias and stimulus evidence integration, whether similar offsets in accumulation of evidence would be observed when directly conditioning on bias has remained unsolved. Here, we provide this missing evidence and demonstrate an offset in pre-stimulus neuronal activity between trials in which behavioral bias matched or mismatched final decision. This bias is later dynamically accumulated during stimulus presentation.

To conclude, we have provided neuronal evidence that behavioral biases and choices are represented in the same neuronal circuits along similar directions of activity state space. The implications of these results can be multifarious. For instance, in a speculative vein, the fact that biases are directly incorporated into the decision process as an offset, just as veridical information would do, could speak about why it is so difficult to eliminate deleterious biases from our daily life behavior,<sup>57,58</sup> and it is in line with current work on decision making proposing bottlenecks in sensory<sup>59</sup> and value-based processing.<sup>60</sup>

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animal model
- METHOD DETAILS
  - Behavioral tasks
  - Neural recording
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Behavioral data analysis
  - Modeling behavior
  - Neuronal data analysis

## SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cub.2021.01.068>.

## ACKNOWLEDGEMENTS

We would like to thank to Mohamad Saleh Esteki, William Newsome, Ramon Nogueira, Gouki Okazawa, and Jacob Yates for fruitful discussions about the project. We would also like to acknowledge the following funding agencies: G.M. is supported by IJCI-2014-21937 from MINECO (Spain); R.K. is supported by the Simons Collaboration on the Global Brain (542997), McKnight Scholar Award, Pew Scholarship in the Biomedical Sciences, and National Institutes of Mental Health (R01 MH109180-01); R.M.-B. is supported by BFU2017-85936-P from MINECO (Spain), the Howard Hughes Medical Institute (HHMI; ref 55008742), an ICREA Academia award, and the Bial Foundation (grant number 117/18).

## AUTHOR CONTRIBUTIONS

G.M., R.K. and R.M.B. designed the project and analyses; R.K. provided the data (adopted from<sup>27</sup>); G.M. analyzed the data; G.M., R.K. and R.M.B. wrote the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 23, 2020

Revised: October 1, 2020

Accepted: January 20, 2021

Published: February 26, 2021

## REFERENCES

1. Gardner, J.L. (2019). Optimality and heuristics in perceptual neuroscience. *Nat. Neurosci.* 22, 514–523.
2. Averbeck, B.B., Sohn, J.W., and Lee, D. (2006). Activity in prefrontal cortex during dynamic selection of action sequences. *Nat. Neurosci.* 9, 276–282.
3. Fan, Y., Gold, J.I., and Ding, L. (2018). Ongoing, rational calibration of reward-driven perceptual biases. *eLife* 7, <https://doi.org/10.7554/eLife.36018>.
4. Hanks, T.D., Mazurek, M.E., Kiani, R., Hopp, E., and Shadlen, M.N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J. Neurosci.* 31, 6339–6352.
5. Hermoso-Mendizabal, A., Hyafil, A., Rueda-Orozco, P.E., Jaramillo, S., Robbe, D., and de la Rocha, J. (2020). Response outcomes gate the impact of expectations on perceptual decisions. *Nat. Commun.* 11, <https://doi.org/10.1038/s41467-020-14824-w>.
6. Hesselmann, G., Kell, C.A., Eger, E., and Kleinschmidt, A. (2008). Spontaneous local variations in ongoing neural activity bias perceptual decisions. *Proc. Natl. Acad. Sci. USA* 105, 10984–10989.
7. Moreno-Bote, R., Shpiro, A., Rinzel, J., and Rubin, N. (2008). Bi-stable depth ordering of superimposed moving gratings. *J. Vis.* 8, 1–13.
8. Moreno-Bote, R., Knill, D.C., and Pouget, A. (2011). Bayesian sampling in visual perception. *Proc. Natl. Acad. Sci. USA* 108, 12491–12496.
9. Akrami, A., Kopec, C.D., Diamond, M.E., and Brody, C.D. (2018). Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature* 554, 368–372.
10. Abrahamyan, A., Silva, L.L., Dakin, S.C., Carandini, M., and Gardner, J.L. (2016). Adaptable history biases in human perceptual decisions. *Proc. Natl. Acad. Sci. USA* 113, E3548–E3557.
11. Braun, A., Urai, A.E., and Donner, T.H. (2018). Adaptive History Biases Result from Confidence-Weighted Accumulation of past Choices. *J. Neurosci.* 38, 2418–2429.
12. Busse, L., Ayaz, A., Dhruv, N.T., Katzner, S., Saleem, A.B., Schölvinck, M.L., Zaharia, A.D., and Carandini, M. (2011). The detection of visual contrast in the behaving mouse. *J. Neurosci.* 31, 11351–11361.
13. Eskandar, E.N., and Assad, J.A. (1999). Dissociation of visual, motor and predictive signals in parietal cortex during visual guidance. *Nat. Neurosci.* 2, 88–93.
14. Fritsche, M., Mostert, P., and de Lange, F.P. (2017). Opposite Effects of Recent History on Perception and Decision. *Curr. Biol.* 27, 590–595.
15. Hwang, E.J., Dahlen, J.E., Mukundan, M., and Komiyama, T. (2017). History-based action selection bias in posterior parietal cortex. *Nat. Commun.* 8, 1242.
16. Lueckmann, J.-M., Macke, J.H., and Nienborg, H. (2018). Can Serial Dependencies in Choices and Neural Activity Explain Choice Probabilities? *J. Neurosci.* 38, 3495–3506.
17. Nogueira, R., Abolafia, J.M., Drugowitsch, J., Balaguer-Ballester, E., Sanchez-Vives, M.V., and Moreno-Bote, R. (2017). Lateral orbitofrontal cortex anticipates choices and integrates prior with current information. *Nat. Commun.* 8, 14823.
18. Padoa-Schioppa, C. (2013). Neuronal origins of choice variability in economic decisions. *Neuron* 80, 1322–1336.
19. Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936.
20. Urai, A.E., de Gee, J.W., Tsetsos, K., and Donner, T.H. (2019). Choice history biases subsequent evidence accumulation. *eLife* 8, e46331.
21. Wyart, V., and Tallon-Baudry, C. (2009). How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. *J. Neurosci.* 29, 8715–8725.
22. Barbosa, J., Stein, H., Martinez, R.L., Galan-Gadea, A., Li, S., Dalmau, J., Adam, K.C.S., Valls-Solé, J., Constantinidis, C., and Compte, A. (2020). Interplay between persistent activity and activity-silent dynamics in the

- prefrontal cortex underlies serial biases in working memory. *Nat. Neurosci.* 23, 1016–1024.
23. Gold, J.I., Law, C.-T., Connolly, P., and Bennur, S. (2008). The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. *J. Neurophysiol.* 100, 2653–2668.
24. Jasper, A.I., Tanabe, S., and Kohn, A. (2019). Predicting Perceptual Decisions Using Visual Cortical Population Responses and Choice History. *J. Neurosci.* 39, 6714–6727.
25. Arandia-Romero, I., Nogueira, R., Mochol, G., and Moreno-Bote, R. (2017). What can neuronal populations tell us about cognition? *Curr. Opin. Neurobiol.* 46, 48–57.
26. Kiani, R., Cueva, C.J., Reppas, J.B., and Newsome, W.T. (2014). Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. *Curr. Biol.* 24, 1542–1547.
27. Kiani, R., Cueva, C.J., Reppas, J.B., Peixoto, D., Ryu, S.I., and Newsome, W.T. (2015). Natural grouping of neural responses reveals spatially segregated clusters in prearcuate cortex. *Neuron* 85, 1359–1373.
28. Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745–4765.
29. Kiani, R., Hanks, T.D., and Shadlen, M.N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017–3029.
30. Kim, J.-N., and Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* 2, 176–185.
31. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503, 78–84.
32. Yu, B.M., Cunningham, J.P., Santhanam, G., Ryu, S.I., Shenoy, K.V., and Sahani, M. (2009). Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *J. Neurophysiol.* 102, 614–635.
33. Hall, P., Marron, J.S., and Neeman, A. (2005). Geometric representation of high dimension, low sample size data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* 67, 427–444.
34. Link, S.W. (1992). The wave theory of difference and similarity (Lawrence Erlbaum Associates, Inc).
35. Ratcliff, R., and Smith, P.L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* 111, 333–367.
36. Shadlen, M.N., and Kiani, R. (2013). Decision making as a window on cognition. *Neuron* 80, 791–806.
37. Peixoto, D., Kiani, R., Chandrasekaran, C., Ryu, S.I., Shenoy, K.V., and Newsome, W.T. (2018). Population dynamics of choice representation in dorsal premotor and primary motor cortex. *bioRxiv*. <https://doi.org/10.1101/283960>.
38. Summerfield, C., and de Lange, F.P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* 15, 745–756.
39. Rao, V., DeAngelis, G.C., and Snyder, L.H. (2012). Neural correlates of prior expectations of motion in the lateral intraparietal and middle temporal areas. *J. Neurosci.* 32, 10063–10074.
40. Barbosa, J., and Compte, A. (2020). Build-up of serial dependence in color working memory. *Sci. Rep.* 10, 10959.
41. Igaya, K., Ahmadian, Y., Sugrue, L.P., Corrado, G.S., Loewenstein, Y., Newsome, W.T., and Fusi, S. (2019). Deviation from the matching law reflects an optimal strategy involving learning over multiple timescales. *Nat. Commun.* 10, 1466.
42. Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.
43. Cowley, B.R., Snyder, A.C., Acar, K., Williamson, R.C., Yu, B.M., and Smith, M.A. (2020). Slow Drift of Neural Activity as a Signature of Impulsivity in Macaque Visual and Prefrontal Cortex. *Neuron* 108, 551–567.e8.
44. Fründ, I., Wichmann, F.A., and Macke, J.H. (2014). Quantifying the effect of intertrial dependence on perceptual decisions. *J. Vis.* 14, <https://doi.org/10.1167/14.7.9>.
45. Akaishi, R., Umeda, K., Nagase, A., and Sakai, K. (2014). Autonomous mechanism of internal choice estimate underlies decision inertia. *Neuron* 81, 195–206.
46. Bonaiuto, J.J., Berker, A., and Bestmann, S. (2016). Response repetition biases in human perceptual decisions are explained by activity decay in competitive attractor models. *eLife* 5, e20047.
47. Fischer, J., and Whitney, D. (2014). Serial dependence in visual perception. *Nat. Neurosci.* 17, 738–743.
48. Williams, Z.M., Elfar, J.C., Eskandar, E.N., Toth, L.J., and Assad, J.A. (2003). Parietal activity and the perceived direction of ambiguous apparent motion. *Nat. Neurosci.* 6, 616–623.
49. Horwitz, G.D., and Newsome, W.T. (1999). Separate signals for target selection and movement specification in the superior colliculus. *Science* 284, 1158–1161.
50. Horwitz, G.D., and Newsome, W.T. (2001). Target selection for saccadic eye movements: prelude activity in the superior colliculus during a direction-discrimination task. *J. Neurophysiol.* 86, 2543–2558.
51. Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J.D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765.
52. Drugowitsch, J., and Pouget, A. (2012). Probabilistic vs. non-probabilistic approaches to the neurobiology of perceptual decision-making. *Curr. Opin. Neurobiol.* 22, 963–969.
53. Drugowitsch, J., Mendonça, A.G., Mainen, Z.F., and Pouget, A. (2019). Learning optimal decisions with confidence. *Proc. Natl. Acad. Sci. USA* 116, 24872–24880.
54. Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.
55. Rustichini, A., and Padoa-Schioppa, C. (2015). A neuro-computational model of economic decisions. *J. Neurophysiol.* 114, 1382–1398.
56. Sohn, H., Narain, D., Meirhaeghe, N., and Jazayeri, M. (2019). Bayesian Computation through Cortical Latent Dynamics. *Neuron* 103, 934–947.e5.
57. Gigerenzer, G. (2008). Gut feelings: the intelligence of the unconscious (Penguin Books).
58. Kahneman, D. (2011). Thinking, fast and slow, First Edition (Farrar, Straus and Giroux).
59. Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., and Pouget, A. (2014). Information-limiting correlations. *Nat. Neurosci.* 17, 1410–1417.
60. Hayden, B.Y., and Moreno-Bote, R. (2018). A neuronal theory of sequential economic choice. *Brain Neurosci. Adv.* 2, 0.1177/2398212818766675.
61. Hastie, T., Tibshirani, R., and Friedman, J. (2001). The Elements of Statistical Learning, Second Edition (Springer-Verlag).

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Bacterial and Virus Strains		
Biological Samples		
Chemicals, Peptides, and Recombinant Proteins		
Critical Commercial Assays		
Deposited Data		
Raw and analyzed data	This paper	<a href="https://www.cns.nyu.edu/kianilab/Datasets.html">https://www.cns.nyu.edu/kianilab/Datasets.html</a>
Experimental Models: Cell Lines		
Experimental Models: Organisms/Strains		
Oligonucleotides		
Recombinant DNA		
Software and Algorithms		
MATLAB	MathWorks	N/A
Custom made code specific to the study	This paper	<a href="https://www.cns.nyu.edu/kianilab/Datasets.html">https://www.cns.nyu.edu/kianilab/Datasets.html</a>
Other		

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Gabriela Mochol ([gabriela@mochol.com](mailto:gabriela@mochol.com)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

Data and code allowing replicating results reported in the manuscript can be found here:  
<https://www.cns.nyu.edu/kianilab/Datasets.html>.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

#### Animal model

We recorded extracellular activity from populations of neurons in the prearcuate gyrus (PAG) of two adult male macaque monkeys performing a direction discrimination task. All training, surgery, and recording procedures conformed to the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by Stanford University Animal Care and Use Committee (protocol number is 9720).

### METHOD DETAILS

#### Behavioral tasks

The direction discrimination task is illustrated in [Figure 1A](#). Each trial began when a fixation point (FP; 0.3° diameter) appeared at the center of the monitor. The monkey was required to fixate within  $\pm 1.5^\circ$  of the FP. Afterwards two targets (T1 and T2) appeared. In 11 sessions, the targets were placed on opposite sides of the screen. In the remaining five sessions, both targets were placed contralateral to the recorded hemisphere. After a short delay (400 ms for 15 data sets and 500–1500 ms for one data set, median 876 ms), a patch of randomly moving dots was shown at the center of the screen for 800 ms. The fraction of coherently moving dots (stimulus



strength or coherence) defined the difficulty of a given trial.<sup>19,28,29</sup> The motion direction and strength were chosen randomly on each trial from a set of predefined values. The coherent motion direction could be toward one target or the other. Coherence ranged from 0 to 0.8. We use signed motion coherence,  $C$ , to specify motion direction and coherence using one number, where positive values indicate motion toward T1 and negative values motion toward T2. The coherence range was tailored for each monkey to obtain the full range of performance accuracy from the chance level to nearly perfect. The stimulus was followed by a delay period of variable duration (302–1478 ms, median = 758 ms) randomly selected on each trial. At the end of the delay period, FP disappeared (Go cue) and the monkey had to report the perceived direction of motion by making a saccade towards the corresponding target and maintaining gaze on the target until the trial outcome was revealed. Correct choices were rewarded with a drop of juice. The mean inter-trial interval (the time between revealing an outcome and the next fixation) was equal to  $1882 \pm 43$  ms (mean  $\pm$  SEM across all trials pooled from 16 sessions; median = 1269 ms). For the zero coherence trials, choices were rewarded randomly with a probability of 0.5. Throughout the session the eye position was monitored at 1 kHz with a scleral search coil (CNC Engineering, Seattle).

### Neural recording

We recorded extracellular activity of a population of PAG neurons using 96 channel microelectrode arrays (Blackrock Microsystems, Salt Lake City; electrode length = 1.5 mm; spacing = 0.4 mm; impedance  $\sim 0.5$  M $\Omega$ ), while monkeys performed the behavioral task. The electrode array was implanted anterior to the concavity of the arcuate sulcus and posterior to the tip of principal sulcus. Neural signals were saved online with 30 kHz sampling rate and spike waveforms were sorted offline (Plexon Inc., Dallas). Recording artifacts simultaneously occurring in a large number of channels were removed using customized algorithms. We identified 169–250 single and multi-units in each session (median = 220). We use the term “units” to refer to both well-isolated single neurons and multi-units. The data sets analyzed in the present study included 9 and 7 recording sessions from monkeys 1 and 2, respectively. These sessions were chosen based on three factors: large number of trials per session ( $>1000$ ), high quality of recordings, and large number of units. Although the position of the electrode array could not be changed by the experimenter after implantation, the recorded units could change from one session to another, presumably due to small movements of cortex relative to the array. The analyzed data were published previously<sup>26,27</sup> in the context of different scientific questions.

Due to the length of the experimental sessions and the large size of the datasets, they were saved as multiple small files, each containing data from an experimental block of more than a hundred trials. These data files were concatenated offline at the end of the session. For 11 of the recording sessions, online spike detection thresholds varied in different blocks of the experiment, causing non-stationary baseline firing rates across the session. To make certain that our analyses were not affected by this non-stationarity, we z-scored firing rates in each block after removing the first and last three trials in the block (median number of removed trials per session, 18, range, 6–30). The z-scored firing rates were concatenated across the session and used in the session by session analyses explained below. Similar but noisier results were obtained when the blocks were analyzed without z-scoring.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Behavioral data analysis

Psychometric curves were defined as the fraction of trials in which the monkey chose target T1 for a given signed motion coherence,  $C$ , for each session. We fit a logistic function to the psychometric curve:

$$\text{Logit}[P_{T1}(C)] = \beta + \alpha C \quad (1)$$

where  $P_{T1}$  is the probability of choosing target T1, and  $\alpha$  and  $\beta$  are model parameters, describing sensitivity and overall bias, respectively. We used a maximum-likelihood fitting procedure for Equation 1 and all subsequent models in the paper.

We use animal accuracy calculated for unsigned coherence in each session to define difficult and easy motion strengths. Difficult motion strengths were those associated with lower than 75% accuracy. Easy motion strengths were those associated with accuracy equal or higher than 75%.

### Modeling behavior

We hypothesized (and confirmed below) that behavior is influenced by past history of choices at two different time scales: long timescale changes in target preference varying over tens to hundreds of trials, and short timescale preference shaped by action and reward history in a few previous trials.

We defined the slow timescale fluctuations of choice preference based on the frequency of choosing targets T1 and T2:

$$\psi_s(i) = \frac{N_{T1}(i)}{N_{T1}(i) + N_{T2}(i)} - \frac{1}{2} \quad (2)$$

where  $N_{T1}(i)$  and  $N_{T2}(i)$  correspond to the number of T1 or T2 choices, respectively.  $N_{T1}(i)$  and  $N_{T2}(i)$  were computed for a group of trials around  $i$ :  $[i-W/2, i-2] \cup [i+1, i+W/2]$ . Trials  $i$  and  $i-1$  were excluded from calculation of  $\psi_s(i)$  to avoid confounds and any overlap with the variables that will be used to define the fast timescale fluctuations of choice preference (see below). We tested various trial windows,  $W$ , as explained below. To ensure that  $\psi_s(i)$  reflected spontaneously generated choice preference and not random fluctuations in the stimulus history, we subsampled trials in the analysis window to balance the number of trials for each signed coherence. Excluding trials after the  $i$ th trial — making the definition of  $\psi_s(i)$  causal — did not qualitatively change our results.

We defined fast time-scale changes of choice preference based on immediately preceding trials. Past studies suggest that the outcome of decisions before trial  $i$  can influence the choice on current trials  $i^{5,10,12}$ . We use three indicator variables to define different combinations of choice and outcome for each preceding trial. For trial  $i-n$ , where  $i$  indicates the current trial and  $n$  indicates how many trials back in the past are considered, the vector of indicator variables are:

$$\boldsymbol{\psi}_i^{-n}(i) = \begin{cases} (1 \ 0 \ 0) & \text{if } T1^- \\ (0 \ 1 \ 0) & \text{if } T1^+ \\ (0 \ 0 \ 1) & \text{if } T2^- \\ (0 \ 0 \ 0) & \text{if } T2^+ \end{cases} \quad (3)$$

where T1 and T2 indicate the two choices, and +/- indicate the two possible outcomes (rewarded/unrewarded) of trial  $i-n$ . Here and further in the text we use bold symbol notation to denote vectors.

To test the effect of fast and slow timescale choice preference on the monkey's behavior, we used a logistic regression model<sup>61</sup> that included both the motion strength and direction from the current  $i^{\text{th}}$  trial (signed coherence  $C$ ), together with the slow and fast choice preference fluctuations:

$$\text{Logit}[P_{T1}(i)] = b_0 + b_1 C(i) + \mathbf{b}_2 (\boldsymbol{\psi}_i^{-1}(i))^T + b_3 \psi_s(i) \quad (4)$$

where  $P_{T1}(i)$  is the probability of a T1 choice in trial  $i$ . The constant term  $b_0$  captures the overall monkey's preference towards one of the targets (constant within a given session), and thus it is added into the definition of the slow bias below. The predictor  $\boldsymbol{\psi}_i^{-1}(i)$  defines the fast choice preference fluctuations, which are shaped by choice and reward on the previous trial (Equation 3), and  $\psi_s(i)$  is the slow choice preference fluctuation that varies at time scales much larger than the immediately experienced history (Equation 2). Here the model weights  $b_0$ ,  $b_1$  and  $b_3$  are constants and  $\mathbf{b}_2$  is a row vector composed of three constants.

The model in Equation 4 was cross-validated using a leave-one-out procedure: the model was fit to all trials except for a held-out trial and its preceding trial, which was used for estimating the fast choice preference for the held-out trial. We used the model parameters to predict the probability of a T1 choice in the held-out trial. The procedure was repeated for all trials in the experiment. For each training set we balanced the number of T1 and T2 choices by randomly removing trials corresponding to the surplus choice. We considered that the model prediction was correct if it gave a higher probability to the target chosen by the monkey. The overall prediction accuracy of the model was calculated for all trials in a session, or sub-groups of easy and difficult trials.

We used the model described in Equation 4 to estimate fast and slow biases on each trial of each session. The "fast bias" at trial  $i$  was calculated as  $B_f(i) = \mathbf{b}_2 (\boldsymbol{\psi}_i^{-1}(i))^T$ . The "slow bias" was calculated as  $B_s(i) = b_0 + b_3 \psi_s(i)$ . We defined the "total bias" as the sum of the fast and slow biases,  $B_t(i) = B_f(i) + B_s(i)$ . The advantage of defining fast and slow bias effects on the current choice using the parameters obtained by the logistic regression model is that all variables are measured using log-odds, thereby allowing a direct comparison between the strength of the three variables on choices. For comparison, for each session we calculated total bias in coherence units as  $B_t(i)/b_1(i)$  (see Table S1).

The significance of the effects of slow and fast choice preference fluctuations on choices was assessed by shuffling  $\boldsymbol{\psi}_i^{-n}(i)$  and  $\psi_s(i)$  independently across trials and fitting Equation 4 to this shuffled data. We repeated this process 1000 times to calculate the distribution of prediction accuracy for the shuffled data, corresponding to the null hypothesis distribution. As in our shuffling procedure we kept the relation between the monkey's choice and motion coherence intact, the null hypothesis distribution was centered on the baseline prediction accuracy given solely by the stimulus direction and strength. We subtracted that mean from the model prediction accuracy to calculate the accuracy improvement conferred by adding the monkey's fast and slow biases into the logistic regression model. The  $p$ -value for the significance of this improvement was calculated as the fraction of shuffles for which the prediction accuracy was higher than or equal to the prediction accuracy for the unshuffled data (one-tailed permutation test). To test whether the mean improvement across all the behavioral sessions was different from zero, we used two-sided, paired  $t$ -test between the predicted and mean shuffled accuracies across sessions.

To select the time scale of the slow fluctuations of choice preference,  $\psi_s(i)$ , we tested a wide range of  $W$  from 20 to 500 trials in steps of 10 trials. A wide range of window sizes provided choice prediction accuracies significantly higher than a model without slow choice preference. We chose the shortest window in a consecutive set of significant window sizes ( $W = 130$ ; Figure S2). This window was used for calculation of slow choice preference in all subsequent analyses. We would like to emphasize that a 130 trial window was somehow an arbitrary choice and any other sufficiently larger window would lead to qualitatively comparable results in Figure S2.

The time scale of fast bias was chosen using a similar procedure by progressively including  $\boldsymbol{\psi}_i^{-2}(i)$  to  $\boldsymbol{\psi}_i^{-5}(i)$  in the model:

$$\text{Logit}[P_{T1}(i)] = b_0 + b_1 C(i) + \sum_{n=1}^j \mathbf{b}_{n+1}^T \boldsymbol{\psi}_i^{-n}(i) \quad (5)$$

where  $j = \{1, \dots, 5\}$ . Because choice prediction accuracy did not show tangible improvement by these extensions, we limited our definition of fast bias to the immediately preceding trial, as in Equation 4.

To compare the strength of each bias on the decision, we used the model in Equation 4 to define the motion coherence in log-odds space of choice as  $M_c(i) = b_1 C(i)$ . Next for each session separately, we defined the total choice predictive power in  $i^{\text{th}}$  trial as  $E_t(i) = |M_c(i)| + |B_f(i)| + |B_s(i)|$ . The impact of each bias ( $l_x$ ) on the decision, (where  $x \in \{s, f, t\}$  stands for slow, fast or total bias) was defined as:

$$I_x = \frac{1}{N} \sum_{i=1}^N \frac{|B_x(i)|}{E_t(i)} \quad (6)$$

To evaluate whether variations of total bias,  $B_t$ , correlates with the monkey's accuracy, we measured average accuracy in windows of 130 trials. Accuracy was based on a subsampled group of trials to balance motion coherences. We used 130 trials to match the time window in which total bias was calculated. Next we calculated Pearson correlation coefficient between absolute value of total bias and accuracy. To assess the significance of Pearson correlation coefficients, we used a two – sided permutation test by randomly shuffling rewarded trials and recomputing accuracy and its correlation with absolute value of total bias ( $n = 1000$ ). To measure the reduction of accuracy (and reward) caused by the total bias, we calculated the mean difference of accuracy in trials with low and high total bias in each session. Trials were labeled as low (or high) bias, if the absolute value of the total bias was within the smallest (or largest) tertiles of the distribution of absolute total bias across all sessions.

## Neuronal data analysis

### Firing rate and dimensionality reduction

In the present study we focused on the activity of PAG units before stimulus appearance, the period of time when the effects of choice biases could be most easily detected as stimulus cannot yet affect neuronal responses<sup>19</sup>. For each recorded unit,  $m$ , the firing rate (the number of spikes per unit of time) at trial  $i$ ,  $r_m(i)$ , was computed in an 800 ms window that terminated 10 ms before dots onset.

Because of the large number of recorded units and limited number of trials, models that use neural responses to predict bias or choice are prone to overfitting. To diminish overfitting, we reduced dimensionality of the neuronal population activity using principal component analysis (PCA). The projection of the neural responses on the  $j^{\text{th}}$  principal component in the  $i^{\text{th}}$  trial is defined as

$$\tilde{r}_j(i) = \sum_{m=1}^M \lambda_m^j r_m(i) \quad (7)$$

where  $\lambda_m(j)$  is the  $j^{\text{th}}$  PCA coefficient for the  $m^{\text{th}}$  unit and  $M$  is the number of simultaneously recorded units. For each recording session, we used the lowest number of PCA components, denoted  $J$ , which explained at least 50% of the total variance (range 38–52, median 48 components). The 50% cutoff provided a good balance between reducing overfitting (increasing prediction accuracy) and maintaining task-related variance of neural responses. Qualitatively similar results were obtained for different variance cutoffs or for the raw data. Principal components were calculated across all trials.

### Decoding biases—linear regression models

A linear regression model was used to investigate whether PAG population activity represents fast, slow and total biases. We regressed any of these biases with the first  $J$  principal components of population activity as

$$B_x(i) = \beta_0 + \sum_{j=1}^J \beta_j \tilde{r}_j(i) + \varepsilon \quad (8)$$

where “x” stands for {s, f, t} for slow ( $B_s$ ), fast ( $B_f$ ), or total bias ( $B_t$ ), respectively, and  $\varepsilon$  is a Gaussian noise term.  $J$  corresponds to the number of principal components that explain 50% of the total neural response variance, as explained above. Bias variables were z-scored using all trials in the session. The model was cross-validated using a leave-one-out procedure. The significance of the cross-validated  $R^2$  was assessed based on a permutation test ( $n = 1000$  random shuffles of  $B_x$ ; the one-tailed  $p$ -value was computed as the fraction of shuffles leading to a  $R^2$  higher than the one obtained from unshuffled  $B_x$ ).

### Decoding Choice - logistic regression models

We used a logistic regression model to predict the monkey's  $n^{\text{th}}$  past choice, future choice, or outcome, based on population activity of the units before stimulus onset on the current trial.

$$\text{Logit}[P_x(i+n)] = \alpha_0 + \sum_{j=1}^J \alpha_j \tilde{r}_j(i) \quad (9)$$

where  $i$  is the current trial,  $n \in \{0, \pm 1, \pm 2, -3, -4, -5\}$  and “x” stands for T1 for choice prediction and ‘+’ for outcome prediction (reward or not). Positive and negative  $n$  indicate trials after and before  $i$ , respectively.

For the choice decoder, we balanced the training set for T1 and T2 choices to make chance level equal to 0.5 for the model prediction accuracy. Similar to the behavioral model in Equation 4, for each repetition/cross-validation, we randomly removed trials corresponding to the surplus choice to have equal numbers of T1 and T2 choices. For the outcome decoder, we did not balance rewarded and unrewarded responses because errors comprised only a small fraction of trials (16%–29% across sessions) and balancing the number of rewarded and unrewarded trials led to exclusion of more than half of the trials in each session. Rather than dropping trials, we calculated the chance level for predicting trial outcome as the monkey's overall reward rate (the fraction of the rewarded trials). Both models (choices or outcomes) were cross-validated using a leave-one-out procedure. For the model predicting upcoming choice ( $n = 0$ ) the significance of the model prediction accuracy was assessed based on a permutation test similar to the ones described earlier ( $n = 1000$  random shuffles of choices or outcomes). For the remaining models we tested if the mean prediction accuracy calculated across sessions was significantly higher than chance level (one – sided t – test). For the

outcome decoder due to slight model overfitting we observed significant decoding accuracy below chance for ( $n \neq 0$ ) using a left – sided t – test. The effect disappeared when we trained the model on data with a larger dimensionality reduction.

### Alignment of choice and total bias decoders

The bias and choice decoders (Equations 8 and 9) provide distances of pre-stimulus responses from the discriminant hyperplanes that best explain the monkey's choice or total bias ( $d_{choice}$  and  $d_{bias}$ ). If the pre-stimulus choice and bias representations were aligned, predicting the upcoming choice based on  $d_{bias}$  would be as accurate as using both  $d_{choice}$  and  $d_{bias}$ . That is,  $d_{choice}$  would not provide additional information for predicting the choice beyond what is provided by  $d_{bias}$ . To test this hypothesis we trained and compared two logistic regression models:

$$\text{Logit}[P_{T1}(i)] = a_1 d_{bias}(i) \quad (10)$$

and

$$\text{Logit}[P_{T1}(i)] = \hat{a}_1 d_{bias}(i) + \hat{a}_2 d_{choice}(i) \quad (11)$$

Alternatively, one could test for the alignment of the neural representations of choice and bias by directly comparing the two discriminant hyperplanes. If our definition of total bias and its neural representation provide a complete account of the representation of choice prior to stimulus onset, one would expect parallel hyperplanes in Equations 8 and 9, and thereby strong correlations in the weight vectors that determine the norm of hyperplanes ( $\vec{\alpha}$  and  $\vec{\beta}$ ). In contrast, if our definition of total bias is an incomplete account of the factors that predict the choice prior to stimulus onset, and if those factors have distinct neural representations from our total bias, the choice hyperplane would not align with our total bias hyperplane. Specifically, if we assume that factors beyond our total bias add up to make a new bias term,  $\vartheta$ , that has a neural representation captured by  $\vartheta(i) = \gamma_0 + \sum_{j=1}^J \gamma_j \tilde{r}_j(i)$ , we can update Equation 8 to include all biases

$$\tilde{B}_t(i) = (\beta_0 + \gamma_0) + \sum_{j=1}^J \beta_j \tilde{r}_j(i) + \sum_{j=1}^J \gamma_j \tilde{r}_j(i) + \varepsilon \quad (12)$$

where  $\tilde{B}_t$  is the corrected bias term that includes both our slow and fast biases and additional factors that we may have failed to identify behaviorally in this paper. Because the combination of all possible bias terms is what enables the prediction of the upcoming choice based on neural responses prior to stimulus onset ( $\text{Logit}[P_{T1}(i)] = \tilde{B}_t(i)$ ), we can combine Equations 9 and 12 to write

$$\sum_{j=1}^J \alpha_j \tilde{r}_j(i) = (\beta_0 - \alpha_0) + \sum_{j=1}^J \beta_j \tilde{r}_j(i) + \sum_{j=1}^J \gamma_j \tilde{r}_j(i) + \varepsilon \quad (13)$$

Equation 13 clarifies that in the presence of additional bias factors not captured by our definition of total bias, the choice hyperplane would not need to align well with our total bias hyperplane. Alternatively, if our definition of total bias is complete within the precision conferred by our dataset, the third term on the right-hand side of Equation 13 would be negligible and the choice and total bias hyperplanes would align well. We test for the alignment of hyperplanes by calculating the correlation of the weight vectors that determine their norms ( $\vec{\alpha}$  and  $\vec{\beta}$ ), and by calculating the angle of the two vectors. Since both hyperplanes lay in a highly dimensional space, angles calculated are from two random hyperplanes are biased towards 90 deg (that is, they are biased to be orthogonal). To account for this bias, we also computed the angle between the two hyperplanes using only from two up to 38 top PCA dimensions used for the previous analysis. Additionally, we checked choice prediction accuracy and total bias representation in a reduced space.

In addition to the analyses above, we compared the accuracy of predicting choices using the neural representation of our total bias or the neural representation of choice prior to stimulus onset. For simplicity, we relied on the sign of  $d_{choice}$  and  $d_{bias}$ . Based on Equation 9, positive values of  $d_{choice}$  mean a higher probability of choosing T1, while negative values mean a higher probability of choosing T2. Similarly, positive and negative  $d_{bias}$  suggest leaning toward T1 and T2, respectively.

### Choice decoder during stimulus presentation

To explore the dynamics of choice prediction accuracy based on neural responses, we extended Equation 9 to include time:

$$\text{Logit}[P_{T1}(i, t)] = \alpha_0(t) + \sum_{j=1}^J \alpha_j(t) \tilde{r}_j(i, t) \quad (14)$$

where  $i$  is the current trial and  $t$  is the center of time window used for the analysis. We used a 100-ms sliding window that moved from 800 ms before to 1100 ms after stimulus onset in steps of 20ms. The projection of the neural responses on the  $j^{th}$  principal component in the  $i^{th}$  trial  $\tilde{r}_j(i, t)$  was defined using the same PCA coefficients used in Equation 7 and based on the 800 ms activity prior to stimulus onset. Similar to the model in Equation 9, we used a leave-one-out cross-validation and balanced choices in the training sets.

The model decision variable (DV) given by the right-hand side of the above equation is the distance of population neural responses from a linear discriminant hyperplane separating T1 and T2 choices. To investigate how behavioral bias interacts with the DV, we

calculated the mean DV conditional on trials where the pre-stimulus total bias favored the choice finally made by the animal (matched trials) or on trials where the total bias was against the final choice (non-matched). To plot the average of DV for each group of trials, we flipped the sign of the DV for T2 choice trials before averaging.

To investigate how the rate of accumulation of evidences depends on coherence and total bias, we first fit the trial by trial DV slope ( $S_{DV}$ ) using a polynomial fit of order one (linear model; matlab *polyfit* function). The fit was done locally in time based on a 200 ms window (10 consecutive time points) moved with 20 ms (1 point) resolution and performed for each trial separately (DV was given by the right side of the Equation 14). Next, we fitted the linear model in which  $S_{DV}$  was expressed as a weighted combination of coherence  $C$  and total bias  $B_t$ :

$$S_{DV}(i, t) = b_0(t) + b_{coh}(t)C(i) + b_{bias}(t)B_t(i) + \varepsilon \quad (15)$$

The model weights were fitted based on all trials pooled from 16 session ( $n = 19715$  trials). The confidence intervals of model weights were obtained from bootstrap, by randomly subsampling trials ( $n = 1000$  random samples with replacement) and refitting the model (Equation 15). We used the central 95% of the bootstrapped weight's distribution as confidence intervals. The statistical significance of the weights was assessed from permuted data where we randomly shuffled SDV across trial dimension but keeping the temporal dimension unchanged and re-fitting the model in Equation 15. We used a two-sided permutation test ( $n = 1000$  shuffles).

To investigate the dynamical accumulation of the total bias during integration of sensory evidence we focused only on the time interval in which  $b_{coh}$  was statistically significant ( $p$ -value  $< 0.05$ ; two-tailed permutation test). Outside that interval, we set  $b_{coh}$  and  $b_{bias}$  to zero and computed their cumulative sum across time (matlab function *cumsum*). The cumulative sum is an unnormalized estimate of the overall slope of DV as a function of coherence and bias that takes into account the non-linear effects that introduces the presence of a hard decision boundary. The procedure was repeated for bootstrapped data to obtain the 95% confidence interval.

We emphasize that as DV rises to the decision threshold non-linearly and reaches it at a different times points depending on stimulus strength, it is better to estimate DV's slope instantaneously and then accumulate the local slopes than performing a single fit using a large window following stimulus onset to obtain a single slope. As a side effect, the estimates of the linear model weights (Equation 15), especially for the total bias, were noisy (Figure S5).

All mathematical analysis were performed using *Matlab* (MathWorks). All statistical tests were two-sided if not stated otherwise.