Supplementary Materials for

Hierarchical reasoning by neural circuits in the frontal cortex

Morteza Sarafyazd and Mehrdad Jazayeri*

*Corresponding author. Email: mjaz@mit.edu

Published 17 May 2019, *Science* **364**, eaav8911 (2019)
DOI: 10.1126/science.aav8911

This PDF file includes:

Materials and Methods

Inferred rule experiment

A trial was initiated when the animal fixated a gray central fixation spot (diameter = 0.5 deg). After a random delay (400-800 ms, uniform hazard), a blue and a red visual target (diameter = 1 deg) were presented above and below the fixation point (distance = 7 deg). The animal had to report its belief about the rule by making a saccade to one of the targets within 1000 ms of their appearance (blue for C_1 and red for C_2). After the rule report epoch, the animal had to shift its gaze back to the fixation point to initiate the interval discrimination epoch (trial was aborted if re-fixation took longer than 500 ms). The trial proceeded regardless of whether the rule was reported correctly. After another random delay (400-900 ms, uniform hazard), two 50-ms flashes (diameter = 0.75 deg) were presented sequentially, a flash at the center of the screen followed by a flash to the left or right of the fixation point (distance = 9 deg). These two flashes demarcated the sample interval, t_s, which was drawn randomly from a 9-point discrete uniform distribution ranging between 530 and 1170 ms. In less than 700 ms after the second flash, the animal had to make a prosaccade toward the second flash or an antisaccade away from it to report whether its estimate of t_s was shorter or longer than a criterion of 850 ms, which corresponds to the mean of t_s distribution. We limited the reaction time to 700 ms to motivate the monkey to control the preparatory prepotent response inhibition dynamically during the interval measurement epoch. At the end of each trial, reward was provided (with a green visual feedback at the target location) if both the rule report in the first epoch and the interval discrimination in the second epoch were correct. No visual feedback or reward was given in error trials. When t_s was equal to the criterion (850 ms) in which case there was no correct answer for the interval discrimination, the animal got reward on a random 50% of the trials if it reported the rule correctly. The trial was aborted if the animal broke fixation before the second flash.

Instructed rule experiment

This task was identical to the rule inference experiment task with the only difference that the rule was cued explicitly and therefore the animal did not need to infer the rule based on history of errors. To specify the rule, we presented a red or blue disc around the fixation spot during the first 3-5 trials after every rule switch and 50-80% of trials until the next rule switch. The color of the disc served as an explicit cue for the current rule. To make sure the structure of the experiments was identical, in the instructed rule experiment, similar to the inferred rule experiment, animals had to first report the rule. On trials that the disc was not presented, the animal could infer it from memory of the previous trial since all rule switches were explicitly cued. The colored disc around the fixation spot disappeared immediately after the animal made a saccade to report the rule so that the interval discrimination epoch of the two tasks were identical.

Model for behavior

We developed a model with two hierarchically organized stages to capture behavior. The model is illustrated schematically in Fig. S3 and S4. The first stage captures the psychometric function describing the probability of Anti (Pr(Anti)) within a trial as a function of rule and t_d in that trial (Fig. S3). The second stage captures the probability of

switch (Pr(Sw)) across trials as a function of feedback, t_d and rule in the preceding trials (Fig. S4). 1 and 2 refer to the parameters associated with the first and second stages, respectively. We use simple letter names for either experimentally controlled variables (e.g., t_s) or observable behavioral variables (e.g., action: A), and letter names with superscript (^) for behaviorally unobservable variables (e.g., internal belief about rule, $\hat{C}$, or internal estimate of discriminant interval, $\hat{t}_d$) and intended actions (e.g., intended pro- or antisaccade, $\hat{A}$).

Modeling the first level of the decision hierarchy

We assumed that the observer makes a noisy measure of t_s ($\hat{t}_s$). Consistent with scalar variability of timing *(93, 94)*, we modeled the noise by a zero-mean Gaussian whose standard deviation scales with t_s with the constant of proportionality w_m, as follows:

$$p(\hat{t}_s|t_s, w_m) \sim \mathcal{N}(t_s, w_m t_s)$$

The observer compares $\hat{t}_s$ to an internal criterion, $\hat{t}_c$ ($\hat{t}_d = \hat{t}_s$-$\hat{t}_c$), to categorize the interval as "Short" or "Long". We treated $\hat{t}_c$ as a parameter that was fixed throughout each behavioral session but was allowed to change across sessions. Moreover, we included a lapse rate, Γ_t to accommodate trials in which the animal categorized t_s randomly (50% "Short" and 50% "Long"). With this formulation, the probability of an intended antisaccade can be written as follows:

$$p(\hat{A} = \text{Anti}|\hat{C}, t_s) = \frac{\Gamma_t}{2} + (1 - \Gamma_t) \int_{\hat{t}_s} p(\hat{A} = \text{Anti}|\hat{C}, \hat{t}_s, \hat{t}_c) p(\hat{t}_s|t_s, w_m) d\hat{t}_s$$

where the first term in the integrant is determined by task contingencies as follows:

$$p(\hat{A} = \text{Anti}|\hat{C}, \hat{t}_s, \hat{t}_c) = \begin{cases} 1, & \text{if } \hat{C} = C_1 \text{ and } \hat{t}_s > \hat{t}_c \\ 0, & \text{if } \hat{C} = C_1 \text{ and } \hat{t}_s < \hat{t}_c \\ 0, & \text{if } \hat{C} = C_2 \text{ and } \hat{t}_s > \hat{t}_c \\ 1, & \text{if } \hat{C} = C_2 \text{ and } \hat{t}_s < \hat{t}_c \end{cases}$$

We assumed that the observer uses $\hat{A}$ to generate a Pro or Anti action (A). In decision making studies, typically there is no need to distinguish between intended and executed actions. In our work, this distinction is warranted because animals occasionally fail to execute an intended antisaccade *(39)* when the inhibition to suppress the competing prepotent prosaccade is not sufficiently strong *(25, 39, 40)*. To account for this asymmetry, the model was augmented with a parameter, Γ_{Anti}, to account for the small percentage of trials in which an intended antisaccade resulted in an erroneous prosaccade. The full model thus used 4 parameters ($\Psi_1 = \{w_m, \hat{t}_c, \Gamma_t, \Gamma_{Anti}\}$) to capture the probability of antisaccades as a function of t_s and C:

$$p(A = \text{Anti}|\hat{C}, t_s) = (1 - \Gamma_{\text{Anti}}) \left[\frac{\Gamma_t}{2} + (1 - \Gamma_t) \int_{\hat{t}_s} p(\hat{A} = \text{Anti}|\hat{C}, \hat{t}_s, \hat{t}_c) p(\hat{t}_s|t_s, w_m) d\hat{t}_s \right]$$

Finally, we assumed that actions across trials were conditionally independent and used the sum of the logarithm of individual conditional probabilities to derive the maximum likelihood estimate (MLE) of the model parameters:

$$\Psi_1(\text{MLE}) = \arg\max_{\Psi} \left[\sum_{i=1}^{N_{\text{Anti}}} \log p(A^i = \text{Anti}|\hat{C}^i, \hat{t}_s^i, \Psi) + \sum_{i=1}^{N_{\text{Pro}}} \log(1 - p(A^i = \text{Anti}|\hat{C}^i, \hat{t}_s^i, \Psi)) \right]$$

where indices i and j count over N_{Anti} antisaccade and N_{Pro} prosaccade trials, respectively. We fit this model to behavioral data in individual sessions of both the inferred rule (subjective rule) and instructed rule experiments.

Modeling the second level of the decision hierarchy

As a first step, we developed an ideal observer model (IOM) whose switch behavior on trial k is governed by posterior odds of rule change compared to no rule change (3), denoted Q_{IOM}^k. If we index the last rewarded trials by k_0-1, Q_{IOM}^k has to be computed based on the consecutive error trials $k_0,\ldots,k$ (denote $[k_0{:}k]$) following the last rewarded trials:

$$Q_{IOM}^k = \frac{Pr(C^{k+1} \neq \hat{C}^k | t_s^{[k_0:k]}, \hat{C}^{[k_0:k]})}{Pr(C^{k+1} = \hat{C}^k | t_s^{[k_0:k]}, \hat{C}^{[k_0:k]})}$$

The denominator is the probability of no rule change despite negative feedback across all $[k_0{:}k]$ trials. For every trial i, the probability of being incorrect without a rule change is the product of the probability that no objective rule change occurred multiplied by the probability that the interval was discriminated incorrectly. If we denote the probability of correctly discriminating t_s in trial i by y^i, and the probability of objective covert rule change in trial i by its hazard function, denoted λ^i, the probability of an incorrect discrimination in the absence of rule change becomes $(1\text{-}\lambda^i)(1\text{-}y^i)$. The denominator therefore is the product of this probability across all $[k_0{:}k]$ trials. Using a similar logic, the numerator is the sum of multiple probabilities including the probability that the rule change occurred on trial k_0, that it occurred on trial k_0+1 but not on trial k_0, and the all the other terms until the last terms, which corresponds to a rule change on trial k with no rule change in the preceding $[k_0{:}k\text{-}1]$ trials. Accordingly, Q_{IOM}^k can be written as follows:

$$Q_{IOM}^k = \frac{\lambda^{k_0} + \sum_{i=k_0+1}^{k} \lambda^i \prod_{j=k_0}^{i-1} (1 - \lambda^j)(1 - y^j)}{\prod_{i=k_0}^{k} (1 - \lambda^i)(1 - y^i)}$$

Since y^i is a measure of expected accuracy (i.e., the probability of being correct), we quantified it based on our model of the subjective psychometric function after marginalizing over t_s (which the ideal observer does not know). For λ^i, to avoid over-parameterizing the model, we did not attempt to model the animal's subjective hazard

function and instead, assumed a constant hazard rate consistent with our experimentally imposed block transition scheme. With y^i and λ^i specified, we can fully derive Q^k_{IOM} with no additional free parameters.

Due to its history-dependent nature, direct use of IOM to characterize animal's behavior is not feasible (i.e., would need far more data that we can collect). As an approximation, we developed a simpler confidence-based model (CBM). The behavior of CBM is governed by a graded latent variable representing trial-by-trial evidence for a covert switch, which we denote by X_Σ. In this model, X_Σ is incremented after each error with a magnitude that depends on the previous trial's expected accuracy, and leads to a switch when X_Σ breaches a threshold (θ). To match CBM as closely as possible to IOM, we assumed that X_Σ on each trial has a Gaussian distribution. We set the mean and standard deviation (denoted by $\mu_{X\Sigma}$ and $\sigma_{X\Sigma}$) of X_Σ such that Q^k_{CBM} (probability that X_Σ is above θ over the probability it is below θ) was equal to Q^k_{IOM} for 1B-Er. Moreover, we assumed that switch evidence is subject to diffusive noise across trials; i.e., that variance grows linearly with the number of error trials (i.e., $\sigma^2_{X\Sigma}$ for N errors = $N\,\sigma_{X\Sigma}$ for 1 error). We also considered a nonlinear relationship between variance and number of trials (Fig. S5D-F) but the simple diffusive (linear) model was able to capture animals' switch behavior more accurately. Note that in the main manuscript, we denote latent representation of subjective switch evidence as $\hat{X}_\Sigma$, however, here we use X_Σ (which is the actual switch evidence) for simplicity.

$$p(X_\Sigma) \sim \mathcal{N}(\mu_{X_\Sigma}, \sigma_{X_\Sigma})$$

$$Q^k_{CBM} = \frac{P(X_\Sigma \geq \theta)}{P(X_\Sigma < \theta)} = \frac{\int_\theta^{\inf} N(\mu_{X_\Sigma},\ \sigma_{X_\Sigma}) dX_\Sigma}{1 - \int_\theta^{\inf} N(\mu_{X_\Sigma},\ \sigma_{X_\Sigma}) dX_\Sigma} = Q^k_{IOM}$$

Rearranging this equation, we can write:

$$\int_{-\inf}^{\theta} \mathcal{N}(\mu_{X_\Sigma}, \sigma_{X_\Sigma}) dX_\Sigma = 1 - \frac{Q^k_{IOM}}{1 + Q^k_{IOM}}$$

which can be solved analytically and numerically to derive $\mu_{X\Sigma}$ in terms of Q^k_{IOM} and $\sigma_{X\Sigma}$. This formulation enabled CBM to closely approximate IOM. To fit CBM to animal's behavior, we made two simplifying assumptions: $X_\Sigma=0$ after a rewarded trial, and $\theta=1$. These are inconsequential because their effect on behavior can be fully subsumed by $\mu_{X\Sigma}$ and $\sigma_{X\Sigma}$.

To account for the deviation of each animal's behavior from the ideal observer model, we augmented $\mu_{X\Sigma}$ by a less-than-one multiplicative 'perseverance factor', α. This factor reduces the magnitude of increments in X_Σ and causes an overall drop in the probability of switch. The full model thus has three parameters, $\mu_{X\Sigma}$, $\sigma_{X\Sigma}$, and α. However because of the relationship between $\mu_{X\Sigma}$, $\sigma_{X\Sigma}$, and the known Q^k_{IOM}, $\mu_{X\Sigma}$ is not a free parameter. Therefore, our model for rule inference had only two free parameters

($\Psi_2 = \{\sigma_{X\Sigma}, \alpha\}$). In this formulation, the probability of switch ($X_{y/n}$=1) can be simplified as follows:

$$p(X_{y/n} = \mathrm{Sw}|\Psi_2) = \int_1^{\inf} \mathcal{N}(\alpha\mu_{X_\Sigma}, \sigma_{X_\Sigma})dX_\Sigma$$

Finally, we assumed that switch decisions were conditionally independent across trials and used the sum of the logarithm of individual conditional probabilities to derive the maximum likelihood estimate (MLE) of the model parameters. The model fits were based on data in the 1B-Er and 2B-Er trials.

$$\Psi_2(\mathrm{MLE}) = \arg\max_{\Psi} \left[\sum_{i=1}^{N_{\mathrm{Sw}}} \log p(X_{y/n} = \mathrm{Sw}|\Psi) + \sum_{j=1}^{N_{\mathrm{No\ Sw}}} \log(1 - p(X_{y/n} = \mathrm{Sw}|\Psi)) \right]$$

We performed several additional analyses to validate our assumptions about CBM and to verify that it was a suitable model: (1) *In-silico* model lesioning: we assessed the necessity of $\sigma_{X\Sigma}$ and α by testing the capacity of the model to capture the observed switch behavior when either of those parameters were fixed. Results in Fig. S5G-L show two examples, one in which $\sigma_{X\Sigma}$ was set to 0.5, and α was fitted, and another when α was set to 1 and $\sigma_{X\Sigma}$ was fitted. However, our conclusions were general and applied to other fixed values of $\sigma_{X\Sigma}$ and α so long as the fixed values were not identical to those recovered by the full model. (2) Model identification: we applied the model fitting procedure to simulated data generated by CBM to ensure that parameters could be recovered. Results are shown in Fig. S5C. (3) 2-fold cross-validation: to ensure that the success of the model was not due to overfitting, we repeatedly fit the model to a random half of each dataset and test the model on the remaining half. Results are shown in Fig. S5B. (4) Predictive validity: we simulated the model fit to the subjective psychometric function in the presence of experimentally imposed covert rule switches, and asked whether it could predict the observed objective psychometric function without additional parameterization and/or fitting. Results are shown in Fig. 1F.

Model free analysis of switch behavior

We also evaluated whether switch probability depended on trial difficulty (indexed by $|t_d|$) and the number of consecutive trials using a simple logistic regression:

$$P(X_{y/n}^{i+1} = \mathrm{Sw}|t_d^i, n\mathrm{B-Er}) = \frac{1}{1+e^{-(\beta_0 + \beta_{|td|}|t_d^i| + \beta_{n\mathrm{B-Er}}[n\mathrm{B-Er}])}} \qquad \text{Equation 1}$$

where $P(X_{y/n}^{i+1})$ represents the probability of switch in trial $i+1$, t_d^i, the discriminant interval in trial i, and nB-Er, the number of consecutive error before trial $i+1$. The parameters of the logistic model (β_0, $\beta_{|td|}$, $\beta_{n\text{B-Er}}$) capture the baseline, dependence on $|t_d^i|$, and dependence on nB-Er of $P(X_{y/n}^{i+1})$, respectively. We used MLE to fit the model to behavior, and used fits to $\beta_{|td|}$ and $\beta_{n\text{B-Er}}$ to assess the statistical significance of the effect of $|t_d^i|$ and nB-Er on switch probability, respectively.

Analysis of single neurons

We estimated firing rates of each neuron per condition by computing average spike counts in 25 ms time bins followed by convolution with a Gaussian boxcar filter (σ =50 ms). For each neuron, we compared average firing rates during the inter-trial-interval (ITI) after rewarded and unrewarded trials (Fig. 2). We also compared responses for both rewarded and unrewarded trials further broken down according to whether the preceding trial was relatively "easy" or "difficult. We defined "easy" and "difficult" as trial in which $|t_d|\geq 160$ ms and $|t_d|<160$ ms, respectively. Collapsing the full range of t_d to a small number of categories was necessary to overcome the limited number of error trials per sessions, but the results were qualitatively the same regardless of the criterion for defining easy versus difficult trials (we also did all the analyses for $|t_d|\geq 80$ ms and $|t_d|<80$ ms). We also compared average firing rates as a function of number of preceding consecutive errors either across all values of t_d or after further dividing trials of the last error according to trial difficulty ($|t_d|\geq 160$ ms versus $|t_d|<160$ ms).

Retrospective and prospective information in DMFC and ACC

We used a time-dependent linear regression model to examine the time course of the representations of trial outcome (Rw/Er) and switch behavior ($X_{y/n}$) across individual neurons in DMFC and ACC. We z-scored the firing rates of each neuron across all trials and times ($r(t)$), and formed a regression model that aimed to explain the firing rates as a function of trial outcome, $X_{y/n}$, and other experimental and behavioral variables as follows:

$$r(t) = \beta_{outcome} I_{Rw/Er} + \beta_{Sw} I_{y/n} + \beta_{Target} I_{left/right} + \beta_A I_{Pro/Anti} + \beta_{td} t_d + \beta_{Rule} I_{C1/C2} + \beta_0$$

Equation 2

In this regression, $I_{Rw/Er}$ was set to 1 (Er) or 0 (Rw), $I_{y/n}$ was set to 0 (no switch) or 1 (switch), $I_{left/right}$, to -1 or 1 depending on the saccade direction, $I_{Pro/Anti}$, to -1 or 1 depending on the animal's action, $I_{C1/C2}$ to -1 or 1 depending on rule, and t_d was the experimentally imposed variable. Solving this equation yielded $\beta_{outcome}$ and β_{Sw} as a function of time for each neurons. Since our interest was to evaluate the relative strength of sensitivity across neurons and areas, we converted $\beta_{outcome}$ to a selectivity index by 1) taking the absolute value of each $\beta_{outcome}$ and 2) normalizing all $\beta_{outcome}$ at each time point by the average $\beta_{outcome}$ across all neurons, areas, and time during ITI. We did the same for β_{Sw}.

$$SI_{outcome} = \frac{|\beta_{outcome}|}{\langle |\beta_{outcome}| \rangle_{\text{neurons, areas}}}$$

$$SI_{\text{Sw}} = \frac{|\beta_{\text{Sw}}|}{\langle |\beta_{\text{Sw}}| \rangle_{\text{neurons, areas}}}$$

Finally, we averaged the selectivity index associated with $\beta_{outcome}$ and β_{Sw} across significantly sensitive neurons in each area and compared the time course of these two selectivity indices across the two areas (Fig. 3).

Sensitivity of population activity in DMFC and ACC to X_Σ and $X_{y/n}$

We employed a targeted dimensionality reduction (63) to assess the encoding of information about X_Σ and $X_{y/n}$ across the population in DMFC and ACC. This approach has been carefully described before (63). Here, we briefly describe the procedure. First, we z-scored firing rates of each neuron across all trials and times, and formed a regression model that aimed to explain the entire vector of firing rates across the population (Z) at each point in time as a function of various experimental and behavioral variables as follows:

$$Z = \beta_{outcome}\, I_{Rw/Er} + \beta_{X\Sigma}\, X_\Sigma + \beta_{Sw}\, I_{y/n} + \beta_{Target}\, I_{left/right} + \beta_A\, I_{Pro/Anti} + \beta_{td}\, t_d + \beta_{Rule}\, I_{C1/C2} + \beta_0$$

Equation 3

The regressors are the same as described for Equation 2, with the addition of X_Σ which was inferred on a trial-by-trial basis from the fit of the confidence-based model to behavior. Second, we solved the regression model to derive the regression coefficients. Third, we denoised the coefficients by finding their projection in the subspace spanned by the first 10 principal components (PCs) of data that explained ~70% (or more) of the total variance in both areas and animals. To compute the PCs, we applied PCA to data resorted to a matrix of dimensions $[N_{units}, N_C \times T]$ (N_{units}: number of neurons, N_C: number of conditions, T: number of time bins). Fourth, for each regressor, we found the time at which the coefficient has the largest norm (i.e., the encoding of the corresponding variable of interest was strongest). Fifth, we orthogonalized the coefficients. To do so, we applied QR-decomposition to a matrix whose columns were the coefficients, and used the columns of the matrix Q as the new orthogonalized coefficients.

Next, we used the orthogonalized coefficients, $\beta_{X\Sigma}$ and β_{Sw}, to examine how DMFC and ACC encoded X_Σ and $X_{y/n}$ at the population level. We first describe the procedure of $\beta_{X\Sigma}$: 1) we found the projection of data onto $\beta_{X\Sigma}$, which we refer to as $Z_{X\Sigma}$; 2) we divided $Z_{X\Sigma}$ into two groups, one containing all the trials that did not lead to a switch and another for all the trials that did lead to a switch; 3) we divided trials in each group to 6 overlapping bins depending on the value of X_Σ (i.e., low X_Σ, intermediate X_Σ, and high X_Σ). The bin and overlapping window were chosen so that the number of trials in each bin were nearly equal. 4) we used d' as a distance metric to quantify the separation between the distributions of $Z_{X\Sigma}$ in each bin to the distribution of $Z_{X\Sigma}$ associated with rewarded trials (i.e., when X_Σ=0). This distance metric was used to assess the sensitivity of DMFC and ACC to X_Σ. To examine the statistics, the population analysis method was cross-validated by resampling dataset into equally-sized training and testing sets (bootstrapping method, 100-times). The procedure for β_{Sw} was identical except that we used the projections of data onto β_{Sw} (instead of $\beta_{X\Sigma}$), which we refer to as Z_{Sw}. Analysis of Z_{Sw} with respect to X_Σ enables us to assess the extent to which DMFC and/or ACC encoded $X_{y/n}$ as a binary variable (i.e., independent from X_Σ). In addition, as a control analysis, we conditioned the same subsets of trials on anti/pro-saccades, and separately on left/right saccades, and the projections into $\beta_{X\Sigma}$ were measured to quantify the action dependency of X_Σ neural state (Fig. S10).

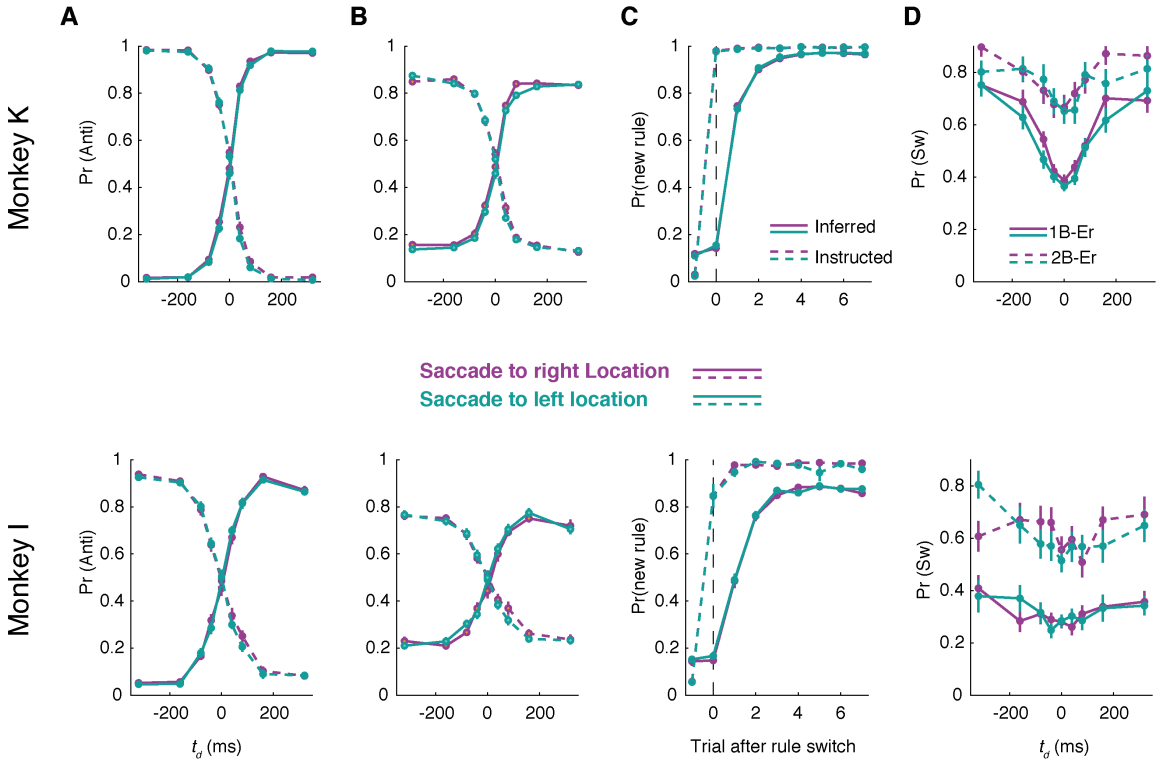

Fig. S1. Switch behavior of the two monkeys as a function of saccade location. (A-D) Results are shown with the same ordering and format as in Fig. 1E-H. Behavior is largely independent of the direction of saccade. There were no significant difference between corresponding points on each plot (point-by-point Wilcoxon rank sum test; $P > 0.05$).

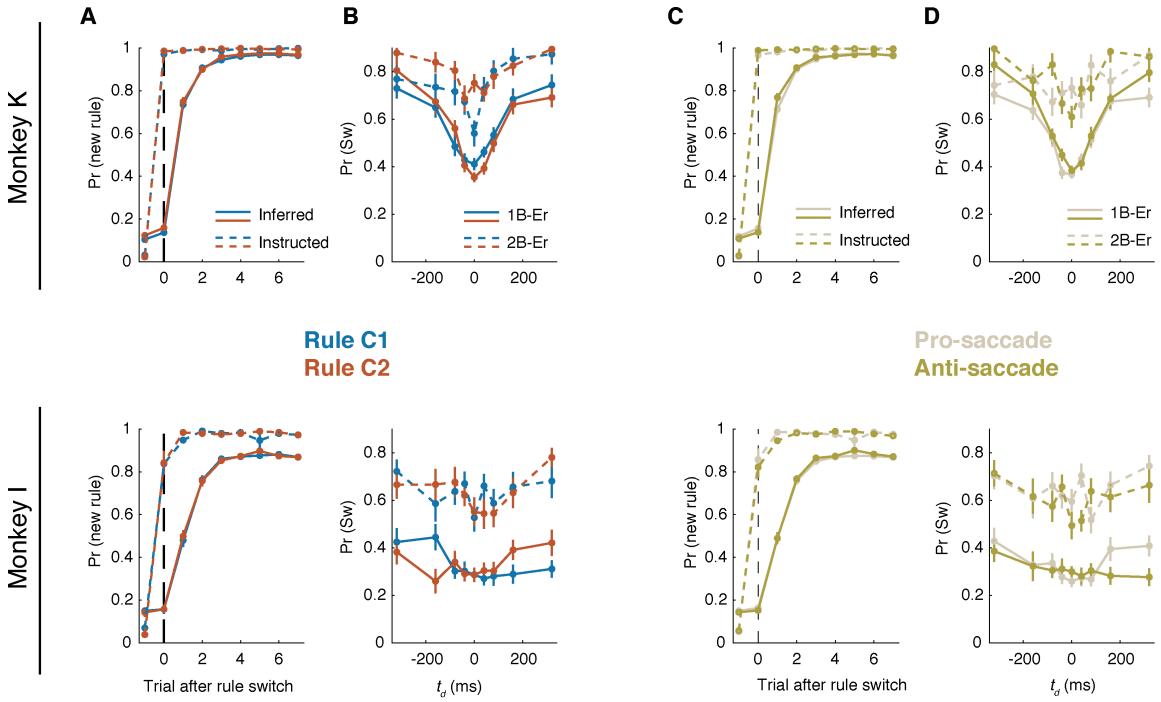

Fig. S2. Switch behavior of the two monkeys as a function of rule (C_1 versus C_2) and action (Pro versus Anti). Results are shown with the same ordering and format as in Fig. 1G-H. Behavior is largely independent of the direction of saccade. There were no significant difference between corresponding points on each plot (point-by-point Wilcoxon rank sum test, $P > 0.05$).

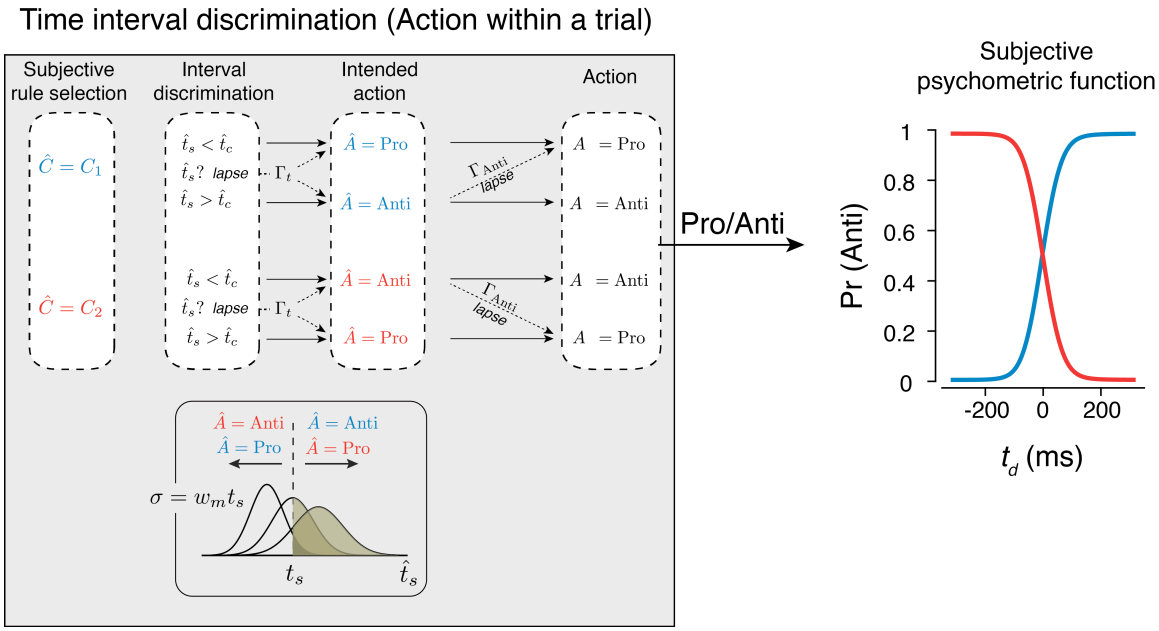

Fig. S3. **Description of the observer model of time interval discrimination. The observer model makes a noisy measurement of the sample interval, t_s, denoted $\hat{t}_s$. The measurement noise has a zero-mean Gaussian distribution and its standard deviation (σ) scales with t_s with constant of proportionality, w_m ($\sigma = w_m t_s$). The model determines the intended action, $\hat{A}$ based on rule $\hat{C}$ ($\hat{C} = C_1$ *or* $\hat{C} = C_2$) and $\hat{t}_s$ relative to an internal criterion ($\hat{t}_c$). When $\hat{C} = C_1$ and $\hat{t}_s > \hat{t}_c$, $\hat{A}$, is an antisaccade (Anti), and when $\hat{C} = C_1$ and $\hat{t}_s < \hat{t}_c$, $\hat{A}$, is a prosaccade (Pro). For $\hat{C} = C_2$, these response contingencies are reversed. The model also allows for a percentage of 'lapse' trials (Γ_t), in which $\hat{A}$ is randomly assigned to Pro or Anti (shown as $\hat{t}_s$?). The model's action, A usually follows from $\hat{A}$, but it also allows for a percentage of additional lapses parametrized by Γ_{Anti} in which $A = Pro$ despite $\hat{A} = \text{Anti}$ due to a failure to inhibit the prepotent prosaccade response. In sum, the model has 4 parameters ($\Psi_1 = \{w_m, \hat{t}_c, \Gamma_t, \Gamma_{Anti}\}$) to capture the probability of antisaccades as a function of t_s and $\hat{C}$. The two psychometric functions on the right are examples of the model's behavior for the two rules shown in terms of proportion of antisaccades (Pr(Anti)) as a function of the discriminant interval, t_d.**

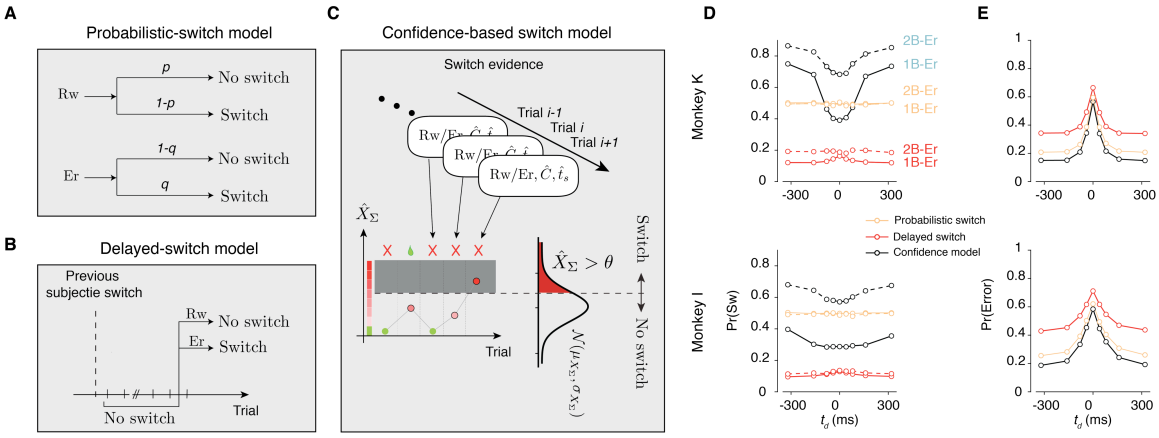

Fig. S4. Alternative models for switching behavior. (A) Probabilistic switch model. This model implements a variant of the win-stay lose-switch strategy in which the animal switches the rule with different fixed probabilities (p and q) depending on trial outcome (Rw: rewarded trial, Er: error trial). (B) Delayed switch model. This model switches after an error only after a certain number of trial since the last successful switch. The model emulates an observer that exploits knowledge about the block-structure in the experiment. (C) Confidence-based switch model. This model computes a graded variable, $\hat{X}_\Sigma$, whose value is determined by the outcome (Rw/Er) and difficulty (indexed by $\hat{t}_s$) of preceding trials. $\hat{X}_\Sigma$ is reset to zero after each trial with positive outcome, and is incremented by a sample of Gaussian distribution after error trials. The mean and standard deviation of the distribution ($\mu_{X\Sigma}$ and $\sigma_{X\Sigma}$) are set so that the model's behavior approximates the behavior of an ideal observer (see Methods). When the updated $\hat{X}_\Sigma$ breaches a threshold, θ, the model switches its belief about the rule ($\hat{C}$). The model also uses a perseveration factor (α) to account for the deviation of each animal's behavior from optimality (not shown in C but see Methods). Since $\mu_{X\Sigma}$ was determined by the ideal observer, the behavior of this stage of the model was fully explained by two parameters ($\Psi_2 = \{\sigma_{X\Sigma}, \alpha\}$). (D) Switch probability (Pr(sw)) based on simulations of three models. Each model was constructed such that it matched as closely as possible to each animal's subjective psychometric function. The data from the probabilistic switch model was generated assuming no switch after rewarded trials and 50% switch after error trials irrespective of confidence. The data for the delay switch model was generated by assuming that no switch for the first 10 trials (consistent with the experimentally imposed hazard) after each successful switch, and 50% switch for each error trial afterwards. The data for the confidence-based switch model are the same as those shown in the main manuscript. For each model, the results are shown for 1-back error (1B-Er; solid) and 2-back error (2B-Er; dashed) conditions. (E) Percentage error as a function of the discriminant interval, t_d, for each of the three models with the same specifications as in D. The Confidence-based switch model has the lowest probability of error in comparison to the the other two models.

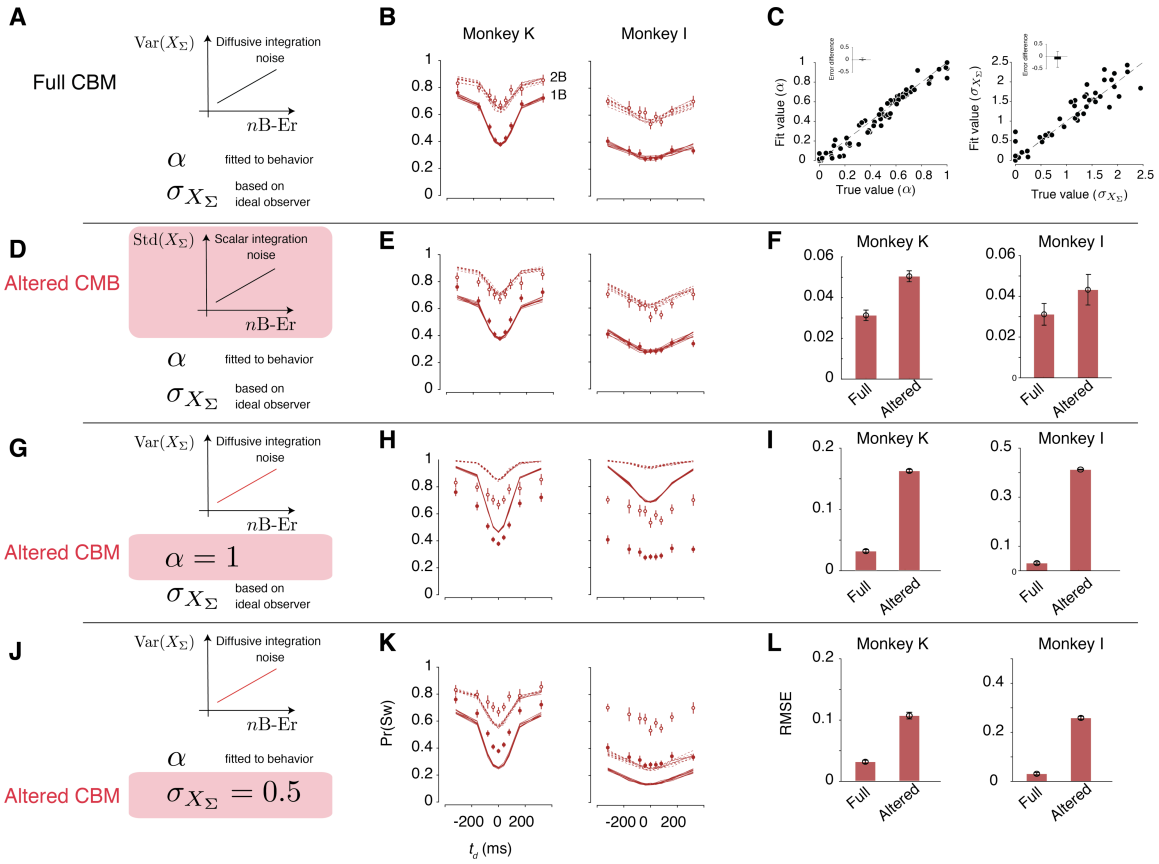

Fig. S5. Cross-validation, identification and lesion studies of the confidence-based belief update model. (A) Full CBM model with two parameters α and $\sigma_{X\Sigma}$. (B) Validation of the full CBM model. Fit of CBM (line) to switch behavior of two monkeys (symbols) in the same format as Fig. 1H. The model fits were cross-validated and repeated 10 times for random train and validation set. (C) Model identifiability. Fits of CBM to data generated from CBM with random choices of α and $\sigma_{X\Sigma}$. Inset: Difference between fitted and actual value across simulations. There is not significant difference between the fit and true values ($P > 0.05$). (D) Altered CBM in which the standard deviation (instead of variance) of noise increases with consecutive trials. The altered component (the one that differs from the full CBM) is highlighted by the shaded box. (E) Same as panel B for alternate model in D. (F) Comparison of the full and alternate CBM by using RMSE across 10 cross-validated fits to behavioral data (error bar: SEM). (G) Altered CBM in which α is set to 1 (no perseverance) while $\sigma_{X\Sigma}$ is fitted to behavior. (H,I) Same as E,F for the altered CBM in G. (J-L) Same as panels (G-I) for an altered CBM in which $\sigma_{X\Sigma}$ is set to 0.5 (instead of being inferred from the ideal observer) while α is fit.

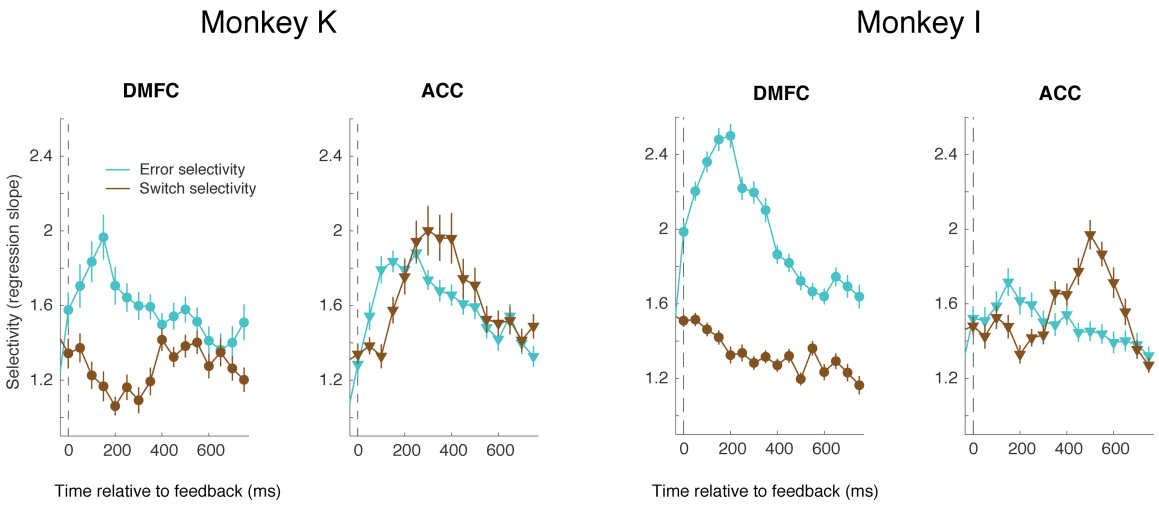

Fig. S6. Retrospective and prospective computations in DMFC and ACC during ITI, separately for the two animals. Results are shown in the same format as in Fig. 3A,B, separately for the two animals.

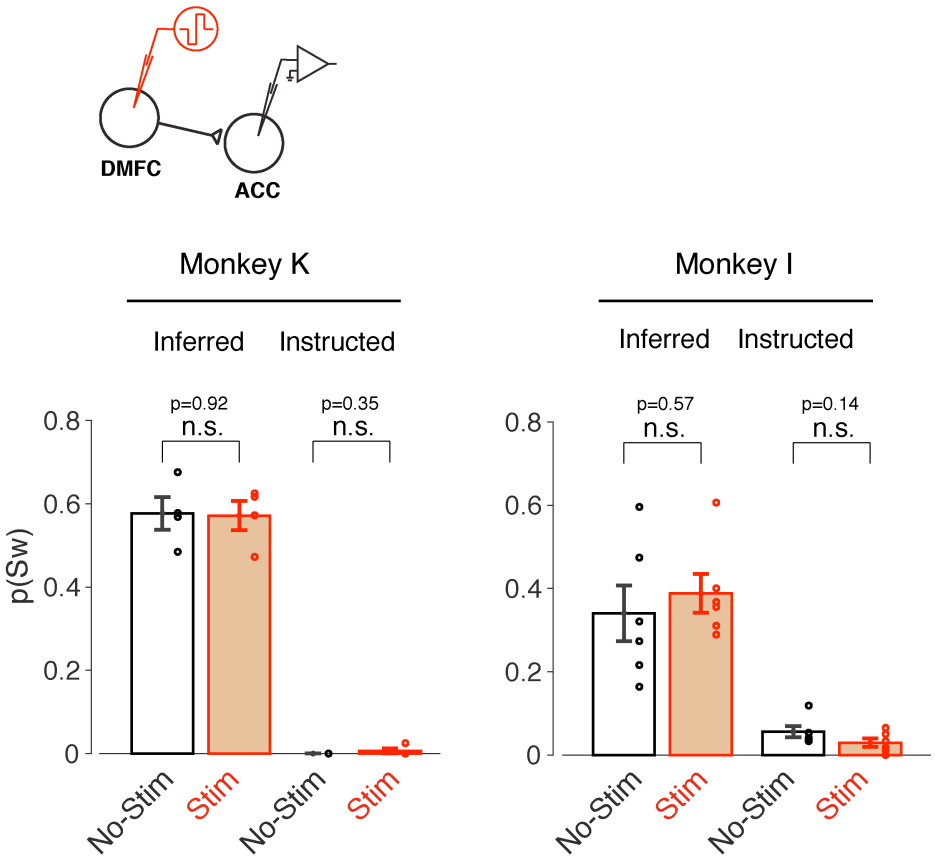

Fig. S7. No effect of DMFC microstimulation on switch probability. The ordinate shows average switch probability (Pr(Sw)) across all sample intervals after 1-back error trials separately for trials when DMFC was and was not stimulated. There was no significant change in Pr(Sw) in the inferred or instructed experiments as a result of DMFC microstimulation (n.s.: $P > 0.05$). For the instructed experiment, Pr(Sw) is only shown for trials where the subjective switch was not cued.

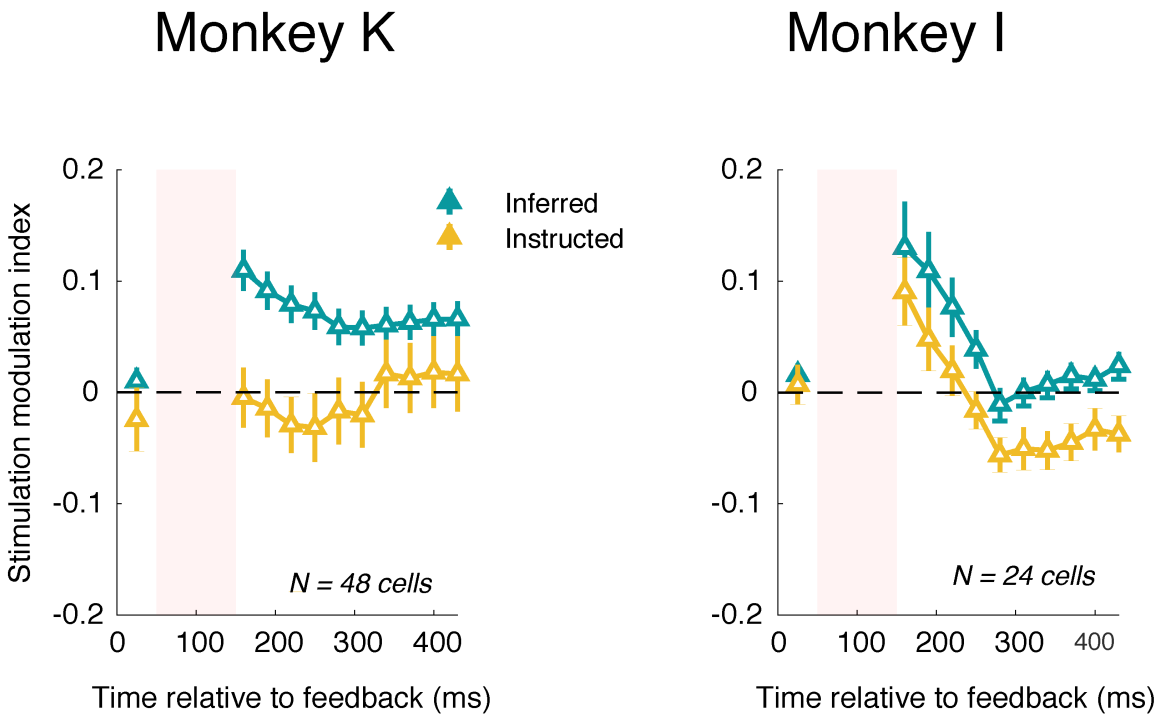

Fig. S8. Modulation of ACC activity due to DMFC microstimulation shown for the two animals separately. Results are shown in the same format as Fig. 3D. We examine the statistical significance of the stimulation modulation index in the inferred and instructed experiments independently. In each case, we compared the distribution of firing rates in the stimulated trials to a null distribution created by random labeling of trials as stimulated and unstimulated. We used a shuffling procedure to avoid making arbitrary assumptions about the sampling distributions. In the inferred experiment, DMFC microstimulation significantly increased the stimulation modulation index in ACC within 150-300 ms relative to feedback (Monkey K: mean=0.091, $P<0.001$; Monkey I: mean=0.102, $P<0.01$;). In contrast, there was no significant change in ACC activity in the same window after DMFC microstimulation in the instructed experiment (Wilcoxon rank sum test, $P>0.05$). Since the results for the instructed experiment were not significant, we did not perform a direct statistical comparison of the inferred and instructed experiments. Although the results for both animals were statistically significant, we wish to highlight the difference between the two animals as this difference may provide stronger support for our hypothesis. The animal for which the effect of stimulation in the instructed task was stronger ("monkey I"), was less flexible at alternating between the inferred and instructed tasks. This is evident from a careful examination of the behavior. In monkey I, who occasionally did not switch following an instructed switch (compare yellow traces in Fig. 1G), DMFC stimulation influenced ACC during the instructed experiment. In contrast, in monkey K who followed the rule instruction nearly perfectly (Fig. 1G), the DMFC-stimulation had nearly no effect on ACC in the instructed experiment. In other words, the stimulation effects, when considered in relation to each animal's behavior, seem to suggest that when rule inference via feedback is at play, DMFC stimulation has a more prominent effect on ACC, which is consistent with our hypothesis.

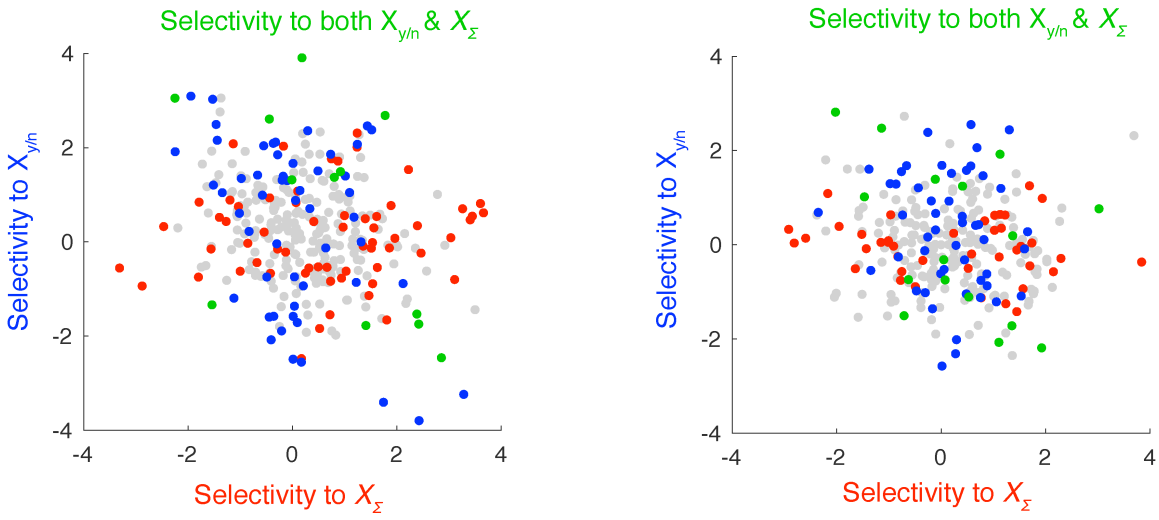

Fig. S9. Mixed representation of $\hat{X}_\Sigma$ and $X_{y/n}$ at the level of individual neurons in ACC. The selectivity of each neuron to switch evidence and switch choice was measured using linear regression (see Methods). Colored dots represent neurons with significant selectivity (permutation test, 100 times, $P < 0.01$) to $\hat{X}_\Sigma$ (red) and $X_{y/n}$ (blue) or both (green).

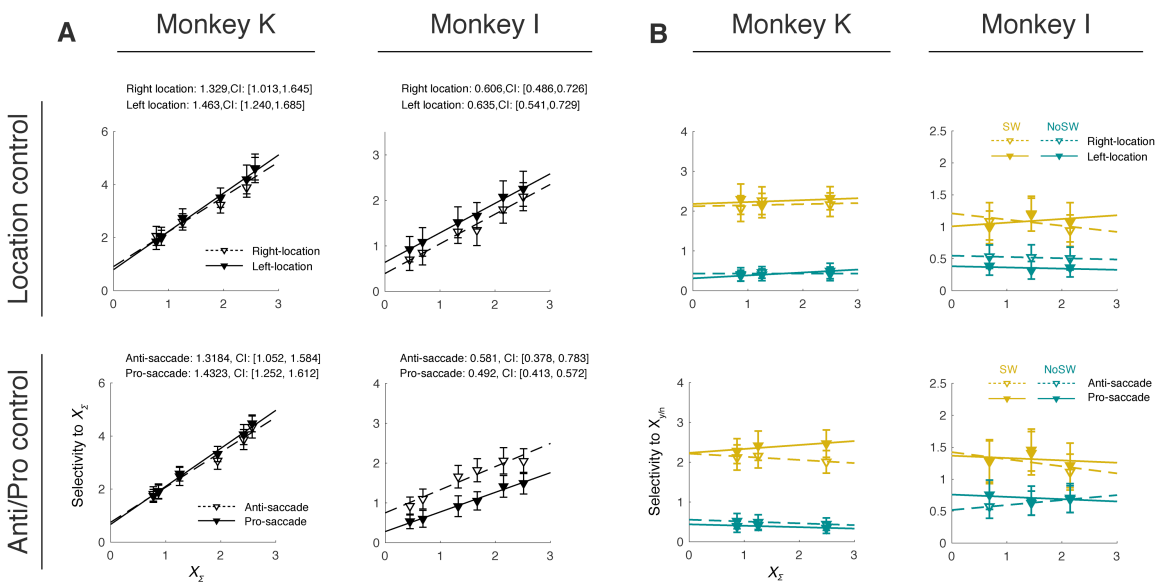

Fig. S10. Sensitivity of ACC to switch evidence was robust to nuisance parameters. (A) After fitting the regression model to ACC (Fig. 4), we separated trials based on saccade direction (left versus right) and response type (prosaccade versus antisaccade) to characterize the extent to which the encoding of switch evidence (X_Σ) was dependent on these nuisance variables. Results indicated that representation of X_Σ in ACC was present for both saccade directions and both response types. All regression slopes were significantly larger than zero ($P < 0.01$). (B) Same as panel (A) for the switch choice selectivity ($X_{y/n}$). Results indicated that representation of switch choice in ACC was present for both saccade directions and both response types ($P < 0.01$).

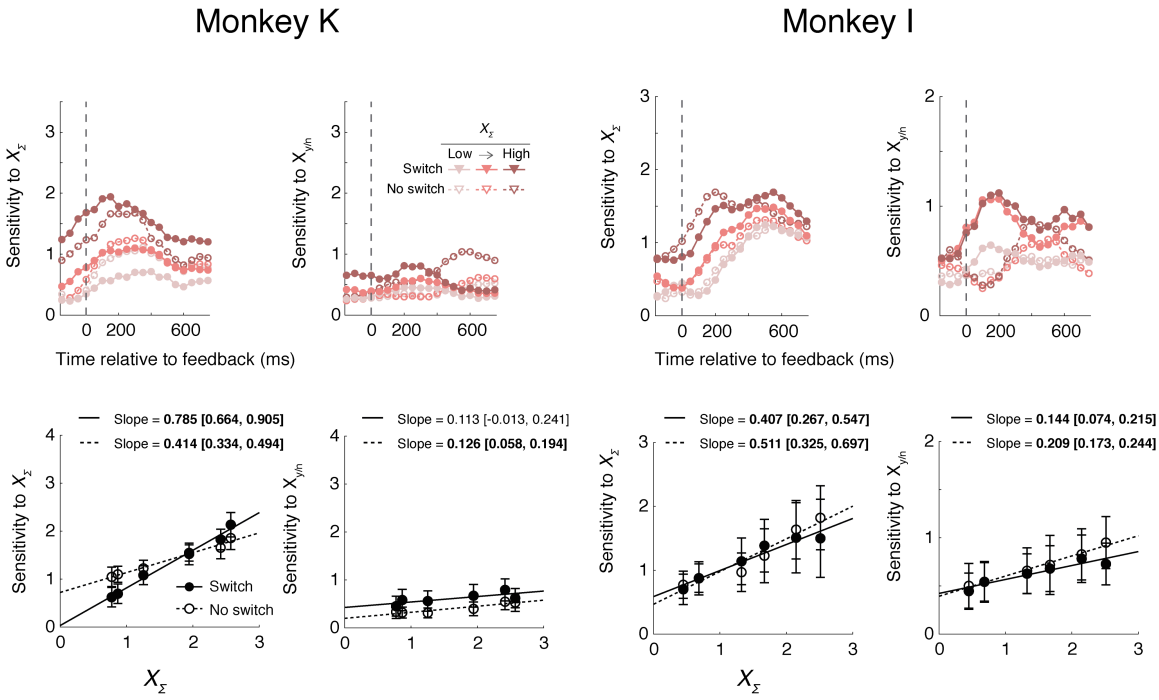

Fig. S11. Representations of switch evidence and switch choice in DMFC. Results are shown in the same format as in Fig. 4A,B, separately for the two animals. For the bottom panel, the slope of the regression line and the corresponding confidence intervals are reported. Regression slopes that were significantly different from zero are in bold.

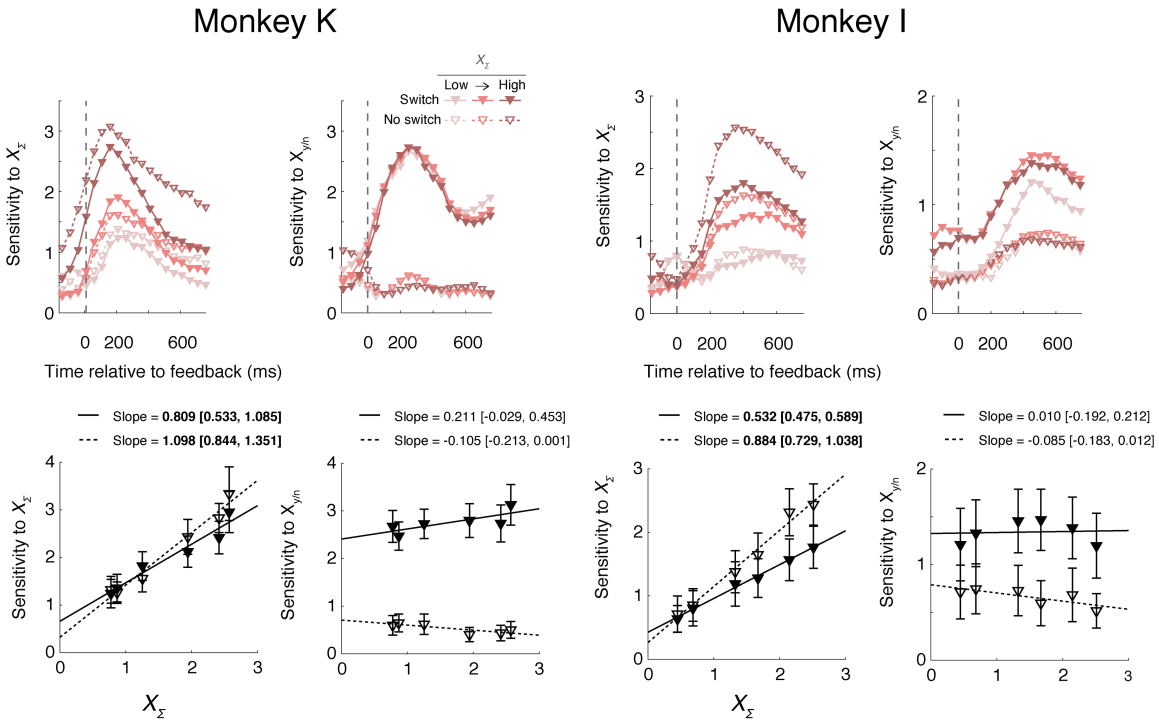

Fig. S12. Representations of switch evidence and switch choice in ACC. Results are shown in the same format as in Fig. 4A,B, separately for the two animals. For the bottom panel, the slope of the regression line and the corresponding confidence intervals are reported (slope [confidence interval]). Regression slopes that were significantly different from zero are in bold. Note that in ACC, selectivity to switch is largely independent of X_Σ (bottom row, second and fourth panels from left), and that neural responses strongly predict animals' switch behavior in the next trial (Table below shows the baseline of the regression line as a function of $X_{y/n}$ for the two animals). These results support our hypothesis that the decision to switch (i.e., threshold crossing) is likely made in ACC.

	$X_{y/n}$ (Sw)	Sensitivity to $X_{y/n}$ b_0 [CI]
Monkey K	Yes No	**2.408, [1.977, 2.840]** **0.707, [0.514, 0.900]**
Monkey I	Yes No	**1.325, [0.993, 1.657]** **0.788, [0.627, 0.950]**

* CI: 95% Confidence interval. Significant estimates ($P < 0.01$) are in bold.

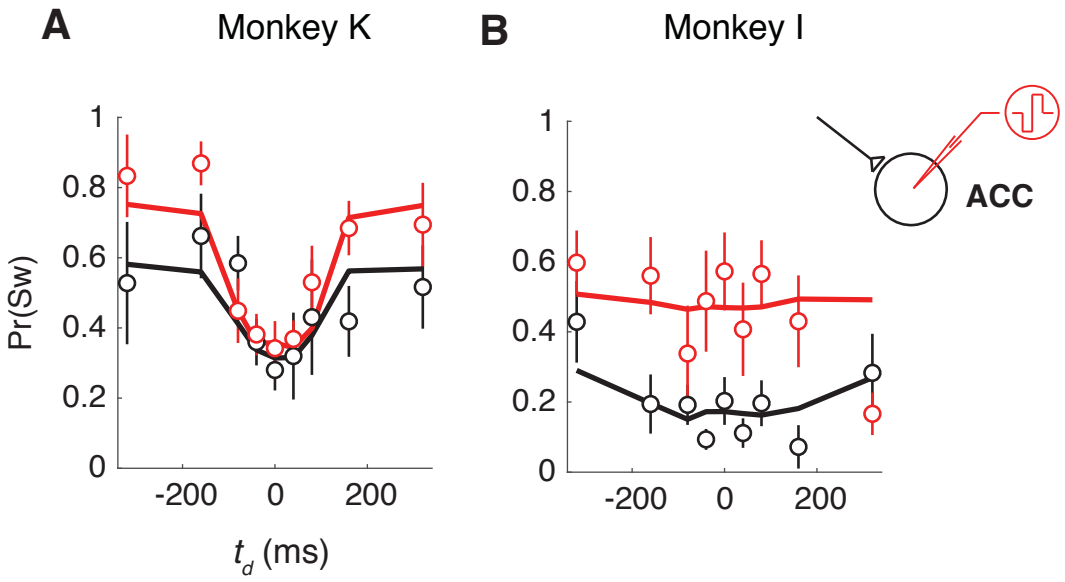

Fig. S13. The effect of ACC microstimulation on switch probability of the two animals separately. The effect of ACC microstimulation (red) compared to stimulation (black) is shown as a function of t_d in the preceding trial for monkey K (A) and monkey I (B). Stimulation was performed after a randomly selected 50% of 1-back error trials. The lines represent the fit of the confidence-based model (see Methods) to each animal for each condition.

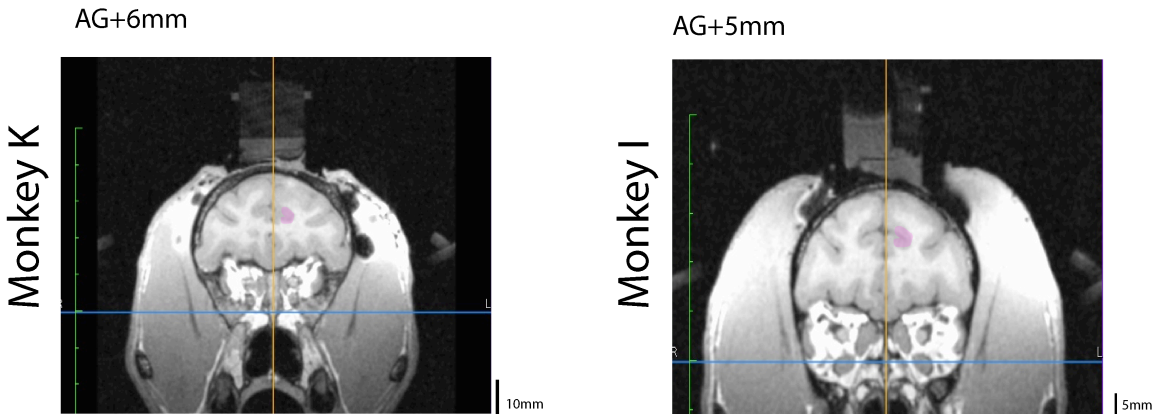

Fig. S14. Coronal section containing ACC from the structural MRI scan of the two animals. Magenta shows the recording site in ACC (AG: anterior to the genu of arcuate sulcus).

Parameter	Inferred rule experiment		Instructed rule experiment	
	Monkey K	**Monkey I**	**Monkey K**	**Monkey I**
w_m	0.0581±0.0042	0.1024±0.0107	0.0596 ±0.0043	0.1083±0.0093
$\hat{t}_c$ (ms)	850.6±129.7 (C_1) 846.8±129.1 (C_2)	853.3±164.2 (C_1) 846.0±162.8 (C_2)	848.4±129.4 (C_1) 856.0±130.5 (C_2)	856.6±164.8 (C_1) 838.7±161.4 (C_2)
Γ_{Anti}	0.0084±0.0013	0.0108±0.0021	0.0080 ±0.0012	0.0095±0.0018
Γ_t	0.0122±0.0019	0.0150±0.0029	0.0096±0.0015	0.0240 ±0.0046

Table S1. Model parameters for the psychometric functions (M±SEM). Since the responses were asymmetric (prosaccade versus antisaccade), we fit the $\hat{t}_c$ for the two rules (C_1 and C_2) separately. However, results indicated that there was not significant difference in $\hat{t}_c$ between the two rules ($P > 0.05$). There was no significant difference between fitted model parameters in inferred and instructed rule experiments ($P > 0.05$).

Parameters	Monkey K	Monkey I
α	0.72±0.03	0.44±0.07
$\sigma_{X\Sigma}$	1.56±0.30	1.09±0.11

Table. S2. Model parameters for switch probability (M±SEM)

	Monkey K		**Monkey I**	
Area	**DMFC**	**ACC**	**DMFC**	**ACC**
Stereotactic coordinates	ML 2-6mm AG 0-6 mm	ML 2-5mm AG 0-8 mm	ML 2-6mm AG 0-7 mm	ML 2-5mm AG 0-8 mm
Recording sessions/trials	18/32148	16/27463	14/18008	17/24672
Stimulation sessions/trials	4/5738	4/6362	6/8685	8/10062

Table S3. DMFC/ACC electrophysiology. ML: mediolateral relative to midline; AG: Anterior-posterior relative to the genu of arcuate sulcus. DMFC stimulation pertains to experiments in which we simultaneously recorded from ACC. Fig. S14 shows a coronal section of the two animals' frontal cortex containing a sample recording site in ACC.

References

1. M. Donoso, A. G. E. Collins, E. Koechlin, Foundations of human reasoning in the prefrontal cortex. *Science* **344**, 1481–1486 (2014). doi:10.1126/science.1252254 Medline

2. C. M. Glaze, A. L. S. Filipowicz, J. W. Kable, V. Balasubramanian, J. I. Gold, A bias–variance trade-off governs individual differences in on-line learning in an unpredictable environment. *Nat. Hum. Behav.* **2**, 213–224 (2018). doi:10.1038/s41562-018-0297-4

3. B. A. Purcell, R. Kiani, Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E4531–E4540 (2016). doi:10.1073/pnas.1524685113 Medline

4. A. Zylberberg, D. M. Wolpert, M. N. Shadlen, Counterfactual reasoning underlies the learning of priors in decision making. *Neuron* **99**, 1083–1097.e6 (2018). doi:10.1016/j.neuron.2018.07.035 Medline

5. N. Kolling, J. Scholl, A. Chekroud, H. A. Trier, M. F. S. Rushworth, Prospection, perseverance, and insight in sequential behavior. *Neuron* **99**, 1069–1082.e7 (2018). doi:10.1016/j.neuron.2018.08.018 Medline

6. Q. J. M. Huys, N. Lally, P. Faulkner, N. Eshel, E. Seifritz, S. J. Gershman, P. Dayan, J. P. Roiser, Interplay of approximate planning strategies. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3098–3103 (2015). doi:10.1073/pnas.1414219112 Medline

7. C. Summerfield, T. E. Behrens, E. Koechlin, Perceptual classification in a rapidly changing environment. *Neuron* **71**, 725–736 (2011). doi:10.1016/j.neuron.2011.06.022 Medline

8. R. Kiani, M. N. Shadlen, Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* **324**, 759–764 (2009). doi:10.1126/science.1169405 Medline

9. A. Kepecs, N. Uchida, H. A. Zariwala, Z. F. Mainen, Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231 (2008). doi:10.1038/nature07200 Medline

10. S. M. Fleming, H. C. Lau, How to measure metacognition. *Front. Hum. Neurosci.* **8**, 443 (2014). doi:10.3389/fnhum.2014.00443 Medline

11. P. G. Middlebrooks, M. A. Sommer, Neuronal correlates of metacognition in primate frontal cortex. *Neuron* **75**, 517–530 (2012). doi:10.1016/j.neuron.2012.05.028 Medline

12. T. J. Pleskac, J. R. Busemeyer, Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychol. Rev.* **117**, 864–901 (2010). doi:10.1037/a0019737 Medline

13. R. Ratcliff, J. J. Starns, Modeling confidence judgments, response times, and multiple choices in decision making: Recognition memory and motion discrimination. *Psychol. Rev.* **120**, 697–719 (2013). doi:10.1037/a0033152 Medline

14. J. Drugowitsch, R. Moreno-Bote, A. Pouget, Relation between belief and performance in perceptual decision making. *PLOS ONE* **9**, e96511 (2014). doi:10.1371/journal.pone.0096511 Medline

15. Y. Komura, A. Nikkuni, N. Hirashima, T. Uetake, A. Miyamoto, Responses of pulvinar

neurons reflect a subject's confidence in visual categorization. *Nat. Neurosci.* **16**, 749–755 (2013). doi:10.1038/nn.3393 Medline

16. R. van den Berg, A. Zylberberg, R. Kiani, M. N. Shadlen, D. M. Wolpert, Confidence is the bridge between multi-stage decisions. *Curr. Biol.* **26**, 3157–3168 (2016). doi:10.1016/j.cub.2016.10.021 Medline

17. A. Lak, G. M. Costa, E. Romberg, A. A. Koulakov, Z. F. Mainen, A. Kepecs, Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron* **84**, 190–201 (2014). doi:10.1016/j.neuron.2014.08.039 Medline

18. S. Ito, V. Stuphorn, J. W. Brown, J. D. Schall, Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* **302**, 120–122 (2003). doi:10.1126/science.1087847 Medline

19. J. D. Schall, V. Stuphorn, J. W. Brown, Monitoring and control of action by the frontal lobes. *Neuron* **36**, 309–322 (2002). doi:10.1016/S0896-6273(02)00964-9 Medline

20. J. W. Brown, T. S. Braver, Learned predictions of error likelihood in the anterior cingulate cortex. *Science* **307**, 1118–1121 (2005). doi:10.1126/science.1105783 Medline

21. C. Amiez, J.-P. Joseph, E. Procyk, Anterior cingulate error-related activity is modulated by predicted reward. *Eur. J. Neurosci.* **21**, 3447–3452 (2005). doi:10.1111/j.1460-9568.2005.04170.x Medline

22. A. W. MacDonald III, J. D. Cohen, V. A. Stenger, C. S. Carter, Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* **288**, 1835–1838 (2000). doi:10.1126/science.288.5472.1835 Medline

23. H. Seo, D. Lee, Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* **27**, 8366–8377 (2007). doi:10.1523/JNEUROSCI.2369-07.2007 Medline

24. M. Shidara, B. J. Richmond, Anterior cingulate: Single neuronal signals related to degree of reward expectancy. *Science* **296**, 1709–1711 (2002). doi:10.1126/science.1069504 Medline

25. M. Schlag-Rey, N. Amador, H. Sanchez, J. Schlag, Antisaccade performance predicted by neuronal activity in the supplementary eye field. *Nature* **390**, 398–401 (1997). doi:10.1038/37114 Medline

26. F. Bonini, B. Burle, C. Liégeois-Chauvel, J. Régis, P. Chauvel, F. Vidal, Action monitoring and medial frontal cortex: Leading role of supplementary motor area. *Science* **343**, 888–891 (2014). doi:10.1126/science.1247412 Medline

27. S. W. Kennerley, T. E. J. Behrens, J. D. Wallis, Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat. Neurosci.* **14**, 1581–1589 (2011). doi:10.1038/nn.2961 Medline

28. B. Y. Hayden, J. M. Pearson, M. L. Platt, Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* **14**, 933–939 (2011). doi:10.1038/nn.2856 Medline

29. M. M. Botvinick, T. S. Braver, D. M. Barch, C. S. Carter, J. D. Cohen, Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652 (2001). doi:10.1037/0033-

295X.108.3.624 Medline

30. J. G. Kerns, J. D. Cohen, A. W. MacDonald III, R. Y. Cho, V. A. Stenger, C. S. Carter, Anterior cingulate conflict monitoring and adjustments in control. *Science* **303**, 1023–1026 (2004). doi:10.1126/science.1089910 Medline

31. A. Shenhav, M. M. Botvinick, J. D. Cohen, The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217–240 (2013). doi:10.1016/j.neuron.2013.07.007 Medline

32. K. Johnston, H. M. Levin, M. J. Koval, S. Everling, Top-down control-signal dynamics in anterior cingulate and prefrontal cortex neurons following task switching. *Neuron* **53**, 453–462 (2007). doi:10.1016/j.neuron.2006.12.023 Medline

33. S. W. Kennerley, M. E. Walton, T. E. J. Behrens, M. J. Buckley, M. F. S. Rushworth, Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947 (2006). doi:10.1038/nn1724 Medline

34. J. Wang, D. Narain, E. A. Hosseini, M. Jazayeri, Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018). doi:10.1038/s41593-017-0028-6 Medline

35. E. D. Remington, D. Narain, E. A. Hosseini, M. Jazayeri, Flexible sensorimotor computations through rapid reconfiguration of cortical dynamics. *Neuron* **98**, 1005–1019.e5 (2018). doi:10.1016/j.neuron.2018.05.020 Medline

36. X. Lu, M. Matsuzawa, O. Hikosaka, A neural correlate of oculomotor sequences in supplementary eye field. *Neuron* **34**, 317–325 (2002). doi:10.1016/S0896-6273(02)00657-8 Medline

37. J. D. Murray, A. Bernacchia, D. J. Freedman, R. Romo, J. D. Wallis, X. Cai, C. Padoa-Schioppa, T. Pasternak, H. Seo, D. Lee, X.-J. Wang, A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014). doi:10.1038/nn.3862 Medline

38. T. E. J. Behrens, M. W. Woolrich, M. E. Walton, M. F. S. Rushworth, Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007). doi:10.1038/nn1954 Medline

39. S. Everling, M. C. Dorris, D. P. Munoz, Reflex suppression in the anti-saccade task is dependent on prestimulus neural processes. *J. Neurophysiol.* **80**, 1584–1589 (1998). doi:10.1152/jn.1998.80.3.1584 Medline

40. N. Amador, M. Schlag-Rey, J. Schlag, Primate antisaccades. I. Behavioral characteristics. *J. Neurophysiol.* **80**, 1775–1786 (1998). doi:10.1152/jn.1998.80.4.1775 Medline

41. M. M. Botvinick, Y. Niv, A. C. Barto, Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition* **113**, 262–280 (2009). doi:10.1016/j.cognition.2008.08.011 Medline

42. D. Badre, Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn. Sci.* **12**, 193–200 (2008). doi:10.1016/j.tics.2008.02.004 Medline

43. M. M. Botvinick, Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci.*

12, 201–208 (2008). doi:10.1016/j.tics.2008.02.009 Medline

44. D. W. Schneider, G. D. Logan, Hierarchical control of cognitive processes: Switching tasks in sequences. *J. Exp. Psychol. Gen.* **135**, 623–640 (2006). doi:10.1037/0096-3445.135.4.623 Medline

45. E. Koechlin, C. Ody, F. Kouneiher, The architecture of cognitive control in the human prefrontal cortex. *Science* **302**, 1181–1185 (2003). doi:10.1126/science.1088545 Medline

46. T. D. Kim, M. Kabir, J. I. Gold, Coupled decision processes update and maintain saccadic priors in a dynamic environment. *J. Neurosci.* **37**, 3632–3645 (2017). doi:10.1523/JNEUROSCI.3078-16.2017 Medline

47. A. Shenhav, J. D. Cohen, M. M. Botvinick, Dorsal anterior cingulate cortex and the value of control. *Nat. Neurosci.* **19**, 1286–1291 (2016). doi:10.1038/nn.4384 Medline

48. C. H. Donahue, H. Seo, D. Lee, Cortical signals for rewarded actions and strategic exploration. *Neuron* **80**, 223–234 (2013). doi:10.1016/j.neuron.2013.07.040 Medline

49. K. Iigaya, M. S. Fonseca, M. Murakami, Z. F. Mainen, P. Dayan, An effect of serotonergic stimulation on learning rates for rewards apparent after long intertrial intervals. *Nat. Commun.* **9**, 2477 (2018). doi:10.1038/s41467-018-04840-2 Medline

50. D. J. Barraclough, M. L. Conroy, D. Lee, Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* **7**, 404–410 (2004). doi:10.1038/nn1209 Medline

51. C. H. Donahue, D. Lee, Dynamic routing of task-relevant signals for decision making in dorsolateral prefrontal cortex. *Nat. Neurosci.* **18**, 295–301 (2015). doi:10.1038/nn.3918 Medline

52. B. Massi, C. H. Donahue, D. Lee, Volatility facilitates value updating in the prefrontal cortex. *Neuron* **99**, 598–608.e4 (2018). doi:10.1016/j.neuron.2018.06.033 Medline

53. M. R. Nassar, R. C. Wilson, B. Heasly, J. I. Gold, An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J. Neurosci.* **30**, 12366–12378 (2010). doi:10.1523/JNEUROSCI.0822-10.2010 Medline

54. K. A. Hadland, M. F. S. Rushworth, D. Gaffan, R. E. Passingham, The anterior cingulate and reward-guided selection of actions. *J. Neurophysiol.* **89**, 1161–1164 (2003). doi:10.1152/jn.00634.2002 Medline

55. W. H. Alexander, J. W. Brown, Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* **14**, 1338–1344 (2011). doi:10.1038/nn.2921 Medline

56. B. Y. Hayden, M. L. Platt, Neurons in anterior cingulate cortex multiplex information about reward and action. *J. Neurosci.* **30**, 3339–3346 (2010). doi:10.1523/JNEUROSCI.4874-09.2010 Medline

57. N. S. Narayanan, M. Laubach, Top-down control of motor cortex ensembles by dorsomedial prefrontal cortex. *Neuron* **52**, 921–931 (2006). doi:10.1016/j.neuron.2006.10.021 Medline

58. K. Shima, J. Tanji, Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* **282**, 1335–1338 (1998). doi:10.1126/science.282.5392.1335 Medline

59. V. Stuphorn, T. L. Taylor, J. D. Schall, Performance monitoring by the supplementary eye field. *Nature* **408**, 857–860 (2000). doi:10.1038/35048576 Medline

60. S. R. Heilbronner, B. Y. Hayden, Dorsal anterior cingulate cortex: A bottom-up view. *Annu. Rev. Neurosci.* **39**, 149–170 (2016). doi:10.1146/annurev-neuro-070815-013952 Medline

61. M. Rigotti, O. Barak, M. R. Warden, X.-J. Wang, N. D. Daw, E. K. Miller, S. Fusi, The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013). doi:10.1038/nature12160 Medline

62. M. L. R. Meister, J. A. Hennig, A. C. Huk, Signal multiplexing and single-neuron computations in lateral intraparietal area during decision-making. *J. Neurosci.* **33**, 2254–2267 (2013). doi:10.1523/JNEUROSCI.2984-12.2013 Medline

63. V. Mante, D. Sussillo, K. V. Shenoy, W. T. Newsome, Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013). doi:10.1038/nature12742 Medline

64. M. T. Kaufman, M. M. Churchland, S. I. Ryu, K. V. Shenoy, Cortical activity in the null space: Permitting preparation without movement. *Nat. Neurosci.* **17**, 440–448 (2014). doi:10.1038/nn.3643 Medline

65. D. Raposo, M. T. Kaufman, A. K. Churchland, A category-free neural population supports evolving demands during decision-making. *Nat. Neurosci.* **17**, 1784–1792 (2014). doi:10.1038/nn.3865 Medline

66. L. T. Hunt, B. Y. Hayden, A distributed, hierarchical and recurrent framework for reward-based choice. *Nat. Rev. Neurosci.* **18**, 172–182 (2017). doi:10.1038/nrn.2017.7 Medline

67. C. B. Holroyd, M. G. H. Coles, The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* **109**, 679–709 (2002). doi:10.1037/0033-295X.109.4.679 Medline

68. V. Stuphorn, J. D. Schall, Executive control of countermanding saccades by the supplementary eye field. *Nat. Neurosci.* **9**, 925–931 (2006). doi:10.1038/nn1714 Medline

69. N. S. Narayanan, J. F. Cavanagh, M. J. Frank, M. Laubach, Common medial frontal mechanisms of adaptive control in humans and rodents. *Nat. Neurosci.* **16**, 1888–1895 (2013). doi:10.1038/nn.3549 Medline

70. H. Gemba, K. Sasaki, V. B. Brooks, 'Error' potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. *Neurosci. Lett.* **70**, 223–227 (1986). doi:10.1016/0304-3940(86)90467-2 Medline

71. H. Niki, M. Watanabe, Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Res.* **171**, 213–224 (1979). doi:10.1016/0006-8993(79)90328-7 Medline

72. C. Shen, S. Ardid, D. Kaping, S. Westendorff, S. Everling, T. Womelsdorf, Anterior cingulate cortex cells identify process-specific errors of attentional control prior to transient prefrontal-cingulate inhibition. *Cereb. Cortex* **25**, 2213–2228 (2015). doi:10.1093/cercor/bhu028 Medline

73. M. Ullsperger, D. Y. von Cramon, Subprocesses of performance monitoring: A dissociation of error processing and response competition revealed by event-related fMRI and ERPs.

Neuroimage **14**, 1387–1401 (2001). doi:10.1006/nimg.2001.0935 Medline

74. M. N. Shadlen, W. T. Newsome, Motion perception: Seeing and deciding. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 628–633 (1996). doi:10.1073/pnas.93.2.628 Medline

75. J. D. Roitman, M. N. Shadlen, Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* **22**, 9475–9489 (2002). doi:10.1523/JNEUROSCI.22-21-09475.2002 Medline

76. A. Hernández, V. Nácher, R. Luna, A. Zainos, L. Lemus, M. Alvarez, Y. Vázquez, L. Camarillo, R. Romo, Decoding a perceptual decision process across cortex. *Neuron* **66**, 300–314 (2010). doi:10.1016/j.neuron.2010.03.031 Medline

77. A. K. Churchland, R. Kiani, M. N. Shadlen, Decision-making with multiple alternatives. *Nat. Neurosci.* **11**, 693–702 (2008). doi:10.1038/nn.2123 Medline

78. G. D. Horwitz, W. T. Newsome, Separate signals for target selection and movement specification in the superior colliculus. *Science* **284**, 1158–1161 (1999). doi:10.1126/science.284.5417.1158 Medline

79. T. D. Hanks, C. D. Kopec, B. W. Brunton, C. A. Duan, J. C. Erlich, C. D. Brody, Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**, 220–223 (2015). doi:10.1038/nature14066 Medline

80. P. Znamenskiy, A. M. Zador, Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* **497**, 482–485 (2013). doi:10.1038/nature12077 Medline

81. C.-T. Law, J. I. Gold, Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* **11**, 505–513 (2008). doi:10.1038/nn2070 Medline

82. M. K. Mian, S. A. Sheth, S. R. Patel, K. Spiliopoulos, E. N. Eskandar, Z. M. Williams, Encoding of rules by neurons in the human dorsolateral prefrontal cortex. *Cereb. Cortex* **24**, 807–816 (2014). doi:10.1093/cercor/bhs361 Medline

83. N. P. Rougier, D. C. Noelle, T. S. Braver, J. D. Cohen, R. C. O'Reilly, Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7338–7343 (2005). doi:10.1073/pnas.0502455102 Medline

84. M. F. S. Rushworth, K. A. Hadland, D. Gaffan, R. E. Passingham, The effect of cingulate cortex lesions on task switching and working memory. *J. Cogn. Neurosci.* **15**, 338–353 (2003). doi:10.1162/089892903321593072 Medline

85. E. K. Miller, J. D. Cohen, An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001). doi:10.1146/annurev.neuro.24.1.167 Medline

86. R. E. Passingham, *The Frontal Lobes and Voluntary Action* (Oxford Psychology Series, issue 21, Oxford Univ. Press, 1993).

87. W. F. Asaad, G. Rainer, E. K. Miller, Task-specific neural activity in the primate prefrontal cortex. *J. Neurophysiol.* **84**, 451–459 (2000). doi:10.1152/jn.2000.84.1.451 Medline

88. N. Zarr, J. W. Brown, Hierarchical error representation in medial prefrontal cortex. *Neuroimage* **124** (part A), 238–247 (2016). doi:10.1016/j.neuroimage.2015.08.063 Medline

89. C. Amiez, J. P. Joseph, E. Procyk, Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex* **16**, 1040–1055 (2006). doi:10.1093/cercor/bhj046 Medline

90. Z. M. Williams, G. Bush, S. L. Rauch, G. R. Cosgrove, E. N. Eskandar, Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nat. Neurosci.* **7**, 1370–1375 (2004). doi:10.1038/nn1354 Medline

91. R. Quilodran, M. Rothé, E. Procyk, Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* **57**, 314–325 (2008). doi:10.1016/j.neuron.2007.11.031 Medline

92. C. M. Glaze, J. W. Kable, J. I. Gold, Normative evidence accumulation in unpredictable environments. *eLife* **4**, e08825 (2015). doi:10.7554/eLife.08825 Medline

93. C. Malapani, S. Fairhurst, Scalar timing in animals and humans. *Learn. Motiv.* **33**, 156–176 (2002). doi:10.1006/lmot.2001.1105

94. J. Gibbon, Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* **84**, 279–325 (1977). doi:10.1037/0033-295X.84.3.279