RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

# Hierarchical reasoning by neural circuits in the frontal cortex

Morteza Sarafyazd and Mehrdad Jazayeri*

INTRODUCTION: Research on the neurobiology of decision-making has emerged as a fertile ground for integrating cognitive, systems, computational, and, more recently, circuit and molecular neuroscience. However, examinations of the underlying neural mechanisms have been largely limited to categorizing stimuli under uncertainty or choosing among volatile rewards. To realize the broader impact of this integration, we need to understand the neural underpinnings of decision-making in more sophisticated behavioral paradigms that demand cognitive reasoning and characterize the computational principles that underlie such reasoning.

RATIONALE: Cognitive reasoning often involves making hierarchically organized decisions. For example, imagine you want to prepare a dish you once enjoyed at a restaurant. You try an online recipe, but the outcome falls short of expectations. You ask yourself, "Is it me or is this the wrong recipe?" Depending on your confidence in your cooking skills, you may try the recipe a few more times, but if the results remain unsatisfactory, you may switch to another recipe. Behavioral studies have shown that humans reason about their failures by assessing their confidence after one or more attempts. However, the neural computations supporting this high-level reasoning strategy are not understood. We sought to characterize these computations in the frontal cortex of nonhuman primates.

RESULTS: We trained monkeys to perform a task comprised of two hierarchically organized decisions. In their first decision, monkeys had to choose between two stimulus-response contingency rules that alternated covertly throughout the experiment. Subsequently, monkeys had to make a perceptual judgment about a stimulus and respond according to the underlying contingency rule. In this task, making

the wrong choice in either decision could lead to an error. Therefore, to correctly infer the cause of the error, one has to reason hierarchically and ask, "Did the rule change, or did I make a perceptual error?" We found that monkeys, like humans, relied on their confidence to decide whether to attribute errors to
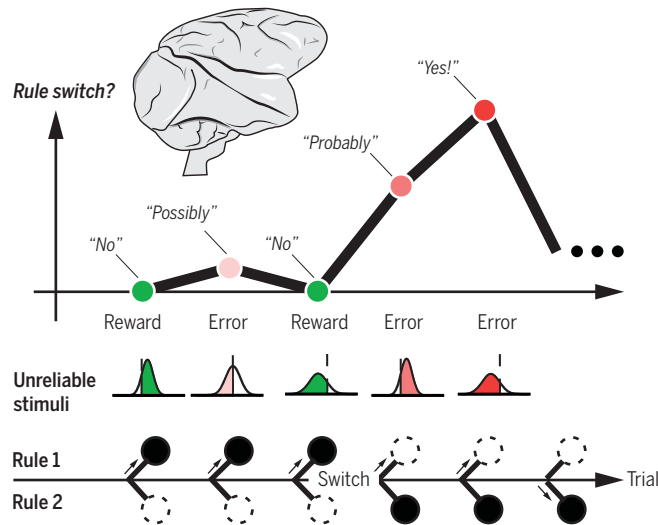
Cognitive error reasoning in the frontal cortex. In a hierarchical reasoning task comprised of two alternating rules (bottom), animals inferred covert rule switches by monitoring the outcome of their perceptual decisions about unreliable stimuli (middle). In nonhuman primates, this cognitive capacity was supported by circuit-level interactions in the frontal cortex computing the belief about rule switches on the basis of the outcome of the preceding trials (top).

themselves or to covert rule switches. They treated each failure as evidence for a covert rule switch but did so rationally by updating their belief about the underlying rule depending on their level of confidence in their perceptual judgments across multiple trials.

To assess the animals' behavior rigorously, we developed a model of hierarchical decision-making that was composed of two processes, one supporting perceptual decisions within each trial and another supporting decisions about covert rule switches across trials. The model was able to capture the animals' be-

havior accurately and provided a quantitative account of how the belief about covert rule switches was updated.

Next, we sought to characterize how neural computations in the frontal cortex could provide a substrate for representing and updating the belief about rule switches. We focused on anterior cingulate cortex (ACC) and dorsomedial frontal cortex (DMFC), which have been implicated in performance monitoring, adaptive reasoning, and strategic decision-making. Electrophysiological recordings indicated that neural activity in both areas reflected the animals' belief about the rule on the basis of the outcome of the preceding trials. A detailed comparison of the nature of the signals in the two areas revealed that only ACC had a strong correlate of the animals' decisions about rule switches. Further probing of these circuits using causal tools revealed that ACC operates downstream of DMFC, integrates trial-outcome information, and drives decisions about when a rule switch might have occurred.

CONCLUSION: Our behavioral results reveal that monkeys, like humans, can reason hierarchically and make rational decisions that rely on evidence at multiple time scales. This opens the possibility for a detailed examination of the neurobiology of hierarchical reasoning, which is a central theme in cognitive neuroscience. We were able to build on previous foundational work on models of decision-making to create a unified framework for understanding the computational principles of hierarchical reasoning. In addition, our neural recording and perturbation experiments revealed a distributed and hierarchically organized neural circuit in the frontal cortex, including DMFC and ACC, that is functionally responsible for hierarchical reasoning about errors. Confidence-based updating of beliefs in uncertain environments is an integral part of human cognition, and our discovery of its underlying computational principles and neural mechanisms is likely to help bridge the gap between research in cognitive and systems neuroscience. ∎

RESEARCH ARTICLE

NEUROSCIENCE

Hierarchical reasoning by neural circuits in the frontal cortex

Morteza Sarafyazd and Mehrdad Jazayeri*

Humans process information hierarchically. In the presence of hierarchies, sources of failures are ambiguous. Humans resolve this ambiguity by assessing their confidence after one or more attempts. To understand the neural basis of this reasoning strategy, we recorded from dorsomedial frontal cortex (DMFC) and anterior cingulate cortex (ACC) of monkeys in a task in which negative outcomes were caused either by misjudging the stimulus or by a covert switch between two stimulus-response contingency rules. We found that both areas harbored a representation of evidence supporting a rule switch. Additional perturbation experiments revealed that ACC functioned downstream of DMFC and was directly and specifically involved in inferring covert rule switches. These results reveal the computational principles of hierarchical reasoning, as implemented by cortical circuits.

A hallmark of cognition is the ability to process information hierarchically. Consider the deliberations of a doctor helping a patient with equivocal symptoms. The doctor has to choose a judicious diagnostic test, interpret the results, prescribe a suitable medicine on the basis of test results, and, finally, evaluate the outcome. The hierarchical nature of this decision process makes failures ambiguous. When facing an unfavorable outcome, the doctor may question the dosage, the medicine, the test results, or the suitability of the test. Resolving this ambiguity demands a sophisticated causal inference strategy. Although human capacity to make such inferences is well established (*1–7*), the link between the key computational principles and the underlying neurobiology is not well understood.

When decisions are organized hierarchically, causal inference about errors demands two critical computations. First, one has to compute a graded expectation of potential outcomes—also known as confidence—depending on the quality of evidence. Decades of work have provided strong evidence that humans and animals compute confidence over their choices (*4*, *8–15*) and use it to improve subsequent decisions (*3*, *16*, *17*). Second, one must monitor performance at multiple time scales to tease apart proximal versus higher-order causes of a failure (e.g., wrong choice of drug versus wrong assumption about the disease). Numerous experiments have found strong error-dependent signals in the dorsomedial frontal cortex (DMFC) and anterior cingulate cortex (ACC) consistent with performance monitoring in a variety of instrumental and conditioning tasks (*18–27*). An important observation has been that both cortical areas carry

performance monitoring signals and that ACC harbors representations of reward on a longer time scale that could be used to regulate strategic exploratory behavior in nonstationary environments (*5*, *28–38*).

However, the neural substrates and mechanisms that allow humans and animals to compute and integrate confidence about low-level decisions to make strategic adjustments to higher-level decisions are not known. To tackle this problem, we designed a hierarchical decision-making task for monkeys in which the rule relating the sensory evidence to behavioral response changed covertly throughout the experiment so that the animals had to compare outcomes with expected outcomes across multiple trials to infer rule changes (Fig. 1A). Behavioral results indicated that monkeys, like humans, make such causal inferences using their level of confidence. Concurrent neural recordings revealed that both DMFC and ACC carry signals related to performance monitoring, with ACC playing a key causal role in making inferences about covert rule changes on longer time scales.

A behavioral task for hierarchical reasoning

In a volatile environment with two alternating response contingency rules (C), the animals had to discriminate a sample interval, t_s (Fig. 1B). The interval varied between 530 and 1170 ms and was demarcated by two flashes, one at the fixation point and one in the periphery (Fig. 1C). In rule 1 (C_1), the animals had to look toward the second flash (prosaccade) when t_s was shorter than the median interval, 850 ms, and away from it (antisaccade) when t_s was longer than 850 ms (Fig. 1B). In rule 2 (C_2), response contingencies were reversed: prosaccade ("Pro") for "Long" and an antisaccade ("Anti") for "Short." Because generating antisaccades requires inhibition of a prepotent prosaccade response (*25*, *39*, *40*), we

refer to C_1 and C_2 as "late inhibition" and "early inhibition," respectively (Fig. 1B). The two rules alternated in a blocked fashion, and the length of each block was a minimum of 10 trials plus a sample from a geometric distribution with a mean of 6 trials (Fig. 1C). In each trial, the animals had to make two hierarchically organized decisions. First, they had to report their belief about the current rule by making a saccade to one of two colored targets (blue for C_1 and red for C_2). Subsequently, t_s was presented and the animals had to make a pro- or antisaccade depending on t_s and the rule. In the main experiment, which we refer to as the "inferred rule experiment," rules were not cued, and rule switches occurred covertly. Therefore, the animals had to infer the rule based on the outcome of previous trials. The animals also performed a control "instructed rule experiment" in which rule switches were explicitly cued. To receive reward, the animals had to correctly report the rule and correctly discriminate t_s (Fig. 1C).

To assess the animals' performance, we measured the proportion of antisaccades [Pr(Anti)] as a function of the sample interval. As a matter of convenience, we will express the interval in each trial relative to the criterion, which we refer to as the discriminant interval, $t_d = t_s - 850$ ms. In the instructed rule experiment, Pr(Anti) increased lawfully with t_d for the late inhibition rule (Fig. 1D, blue, and table S1) and decreased for the early inhibition rule (Fig. 1D, red, and table S1). In the inferred rule experiment, we analyzed responses both in terms of the subjective rule (i.e., the rule reported by the animal; Fig. 1E) and in terms of the experimentally imposed objective rule (Fig. 1F). The animals' performance with respect to the subjective rule was not statistically different from that of the instructed rule experiment (table S1), verifying that the animals understood the hierarchical structure of the task and followed the response contingencies according to their belief about the rule. Errors, when measured with respect to the objective rule, increased by a factor of 1.76 ± 0.04 and 1.62 ± 0.05 in monkeys K and I, respectively (Wilcoxon rank sum test, $P < 0.001$). This is a straightforward consequence of the incorrect rule reports after covert rule switches.

Next, we focused on the animals' decision about the rule (Fig. 1, G and H). In the instructed rule experiment, the animals switched their selection of rule immediately after every objective rule switch, indicating that they successfully learned to follow instructions (Fig. 1G, yellow). In the inferred rule experiment, rule switches were covert and thus had to be inferred from the pattern of feedbacks. A rational observer would make the following considerations: (i) Any rewarded trial indicates that the rule was selected correctly and no rule switch is needed; (ii) after a negative outcome, there is a chance that a covert rule switch might have occurred, and the chances increase after multiple consecutive negative outcomes; and (iii) when t_d is far from the median value (i.e., trials in which interval discrimination is relatively easy), the

Department of Brain & Cognitive Sciences, McGovern Institute for Brain Research, MIT, Cambridge, MA, USA.
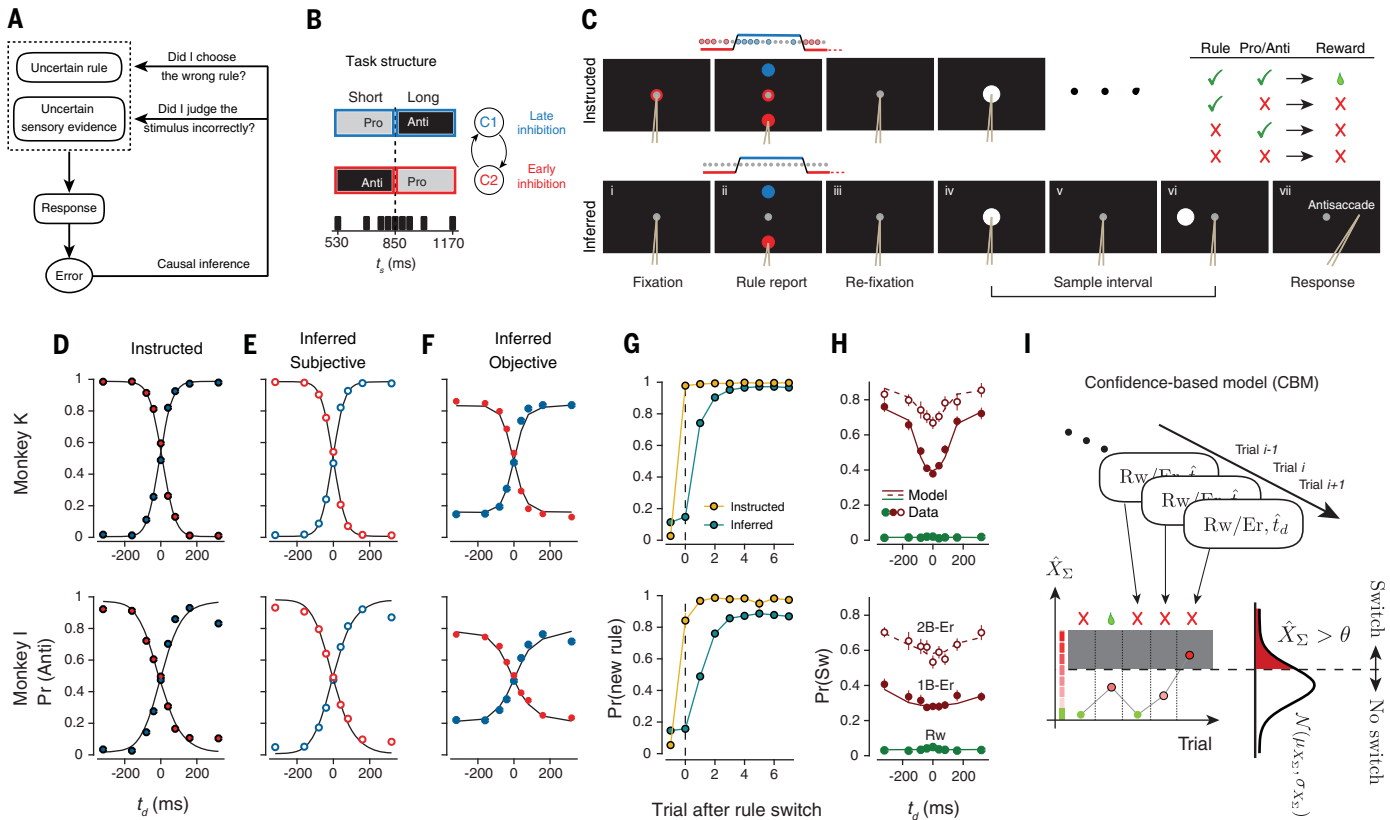*Corresponding author. Email: mjaz@mit.edu

Fig. 1. Hierarchical causal inference, behavioral task, performance, and model. (**A**) Causal inference when the task involves two hierarchically organized decisions. The observer has to infer the stimulus-response contingency rules and judge the sensory evidence. Because both are uncertain, when a decision leads to an error (i.e., negative feedback), the observer has to decide whether to attribute the error to an incorrect decision about the rule or an incorrect judgment about the stimulus. (**B**) Task contingencies. The sample interval (t_s) varied between 530 and 1170 ms and was designated Short or Long depending on whether it was shorter or longer than 850 ms. For the first rule (C_1), the animal had to make a prosaccade (Pro) when $t_s < 850$ ms and an antisaccade (Anti) when $t_s > 850$ ms. For the second rule (C_2), the response contingencies were reversed. Response contingency rules were volatile and switched throughout the experiment (arrows). (**C**) Experimental conditions and trial structure. The order of events in every trial was as follows: (i) fixation: the animal had to fixate a central spot; (ii) decision about the rule: the animal had to report its belief about the underlying rule by making a saccade to one of two targets presented above and below the fixation spot corresponding to two rules (blue, C_1; red, C_2); (iii) refixation: choice targets were removed, cueing the animal to immediately refixate; (iv to vi) t_s presentation: two flashes, one around the fixation point (iv), and one to the left or right of the fixation point (vi), separated by t_s were presented (v); (vii) response: depending on its belief about the rule and the duration of t_s, the animal had to either make a prosaccade (Pro) toward the second flash or an antisaccade (Anti) away from it. Pro and Anti responses by the animal were followed by reward if both rule and interval discrimination were correct. The animal performed two variants of the experiment. In the instructed rule experiment, a colored cue around the fixation spot indicated the correct rule. This rule cue was only provided during the initial fixation before the rule decision was made. In the inferred rule experiment, no cue was provided regarding the rule, and the animal had to infer the rule from the pattern of errors and its own expected accuracy. As shown above the

"rule report" screen in (C), rule switches occurred in a blocked fashion. In the instructed rule experiment, the rule was cued after each switch and intermittently during the block. In the inferred rule experiment, there was no external information about the rule (gray cue). (Top right) The animal could be correct (green checkmark) or incorrect (red cross) about either the rule or the action (Pro or Anti). Reward was provided only if both were correct. (**D**) The proportion of Anti [Pr(Anti)] as a function of interval ($t_d = t_s − 850$ ms) for the instructed rule experiment for the two monkeys. Lines show model fits (see methods). Blue, C_1; red, C_2. (**E**) Same as (D) for the inferred rule experiment, plotted with respect to the rule reported by the animal (i.e., "subjective"). (**F**) Same as (E), plotted with respect to the experimentally imposed rule (i.e., "objective"). (**G**) Probability of choosing a new rule by the animal [P(new rule)] after an objective rule switch for the instructed (yellow) and inferred (turquoise) experiments. (**H**) Proportion of subjective rule switches [Pr(Sw)] as a function of t_d after reward (Rw, green), one-back error (1B-Er, filled red circle), and two-back error (2B-Er, open red circle) after a rewarded trial. Lines show model fits (see methods). (**I**) Confidence-based switch model. This model updates the value of switch evidence (different levels of red), X_Σ, on the basis of the outcome of the previous trials (Rw/Er, reward versus error), and the animal's belief about t_s ($\hat{t}_s$). After rewarded trials (green drop), X_Σ is reset to zero. After each error (red cross), X_Σ is incremented. We modeled X_Σ as a Gaussian distribution whose mean and standard deviation (μ_{X_Σ} and σ_{X_Σ}) as a function of t_s and the number of consecutive errors were set such that the model's behavior approximates the behavior of an ideal observer (see methods). When X_Σ breaches a threshold (dashed line, θ), the model switches the rule. Modeling X_Σ as a Gaussian distribution, the probability of switch at any trial can be related to the area under the distribution beyond the threshold (red region under the Gaussian). In (D) to (H), error bars (SEM across sessions) are included, but for most points, the bars are not visible because they are smaller than the symbol size.

error is more likely to be due to a covert rule switch than incorrect timing. Qualitatively, the animals' pattern of subjective rule switches was consistent with all these predictions. For example, the probability of choosing a new rule was low in the first trial after an objective rule switch and increased monotonically afterward (Fig. 1G, turquoise). Similarly, the probability of subjective rule switches [Pr(Sw)] was small after rewarded trials (Fig. 1H; Rw, green), increased after one error (Fig. 1H; 1B-Er, filled red), and further increased after two consecutive errors (Fig. 1H; 2B-Er, open red). Moreover, Pr(Sw) after both one and two errors seemed to increase systematically with t_d (Fig. 1H). To assess these observations quantitatively, we used logistic regression to measure the effect of trial difficulty and number of consecutive errors on Pr(Sw). The regression coefficients (β) associated with both trial difficulty (indexed by $|t_d|$) and the number of consecutive errors (indexed by nB-Er) were positive (monkey K: $\beta_{|td|} = 5.78 \pm 0.37$, $P < 10^{-16}$ and $\beta_{nB\text{-}Er} = 0.91 \pm 0.07$, $P < 10^{-16}$; monkey I: $\beta_{|td|} = 1.73 \pm 0.26$, $P < 10^{-8}$ and $\beta_{nB\text{-}Er} = 1.17 \pm 0.09$, $P < 10^{-8}$). These characteristics were present for both saccade directions (fig. S1), both rules, and both response types (fig. S2). Therefore, we concluded that the animals (i) updated their belief about covert rule switches on the basis of their trial-by-trial confidence in their interval judgments and (ii) accumulated evidence across consecutive errors. These findings demonstrate that monkeys, like humans (3, 4, 16, 41–47), are capable of adopting a sophisticated causal inference strategy in a hierarchical decision task.

A computational model for hierarchical reasoning

We tested various classes of models to infer the relevant latent variables that guide the animals' behavior (figs. S3 and S4). In one model, we assumed that the animals implemented a probabilistic switching behavior, which is a generalization of the well-known win-stay lose-switch strategy (48). According to this model, the agent switches the rule with fixed probabilities depending on trial outcome without regard to trial difficulty and/or the number of consecutive errors (fig. S4A). Although variants of this model have successfully captured the behavior of monkeys and rodents in a number of simple tasks (49–52), it failed to explain the behavioral characteristics of monkeys during hierarchical reasoning (fig. S4D) and was unable to reach the animals' level of performance (fig. S4E). In another model, we assumed that monkeys learned the hazard rate of environmental switches and delayed their subjective switches accordingly (fig. S4B). This model also failed to capture the animals' switch behavior (fig. S4D) and level of performance (fig. S4E).

We also considered other models of hierarchical reasoning that were used to explain how humans decompose tasks into hierarchies (6), track higher-order parameters (i.e., hyperparameters) and environmental states (2, 4, 7, 53), and reduce failures by comparing new, old, and counterfactual strategies (1). Although these models

have successfully captured human behavior in a range of hierarchical decision-making tasks, they could not straightforwardly be adapted to our experiment because they were not intended to account for failures caused by misjudgments of unreliable stimuli, which is a central component of our task.

To develop a suitable model for our experiment, we first formulated the problem in terms of the behavior of an ideal observer, similar to a recent behavioral study in humans (3). The ideal observer computes the posterior probability of a rule switch by integrating information about (i) the expected accuracy of trial-by-trial interval judgments and (ii) the outcome of preceding trials (see methods). We then created a simplified confidence-based model that computes the evidence for or against a covert rule switch by a single graded latent variable, X_Σ, whose value as a function of task difficulty (indexed by t_d) and the number of consecutive errors was inferred from the ideal observer model (see methods). After updating the value of X_Σ, the model sets a binary switch decision, $X_{y/n}$, to "switch" or "no switch" depending on whether X_Σ is larger or smaller than a threshold, θ (Fig. 1I). To capture both animals' behavior quantitatively, we had to augment the model with a perseveration factor. This factor enabled us to account for the deviation of each animal's behavior from the ideal observer model that computes switch evidence optimally (see methods). The model captured the characteristic dependence of switch behavior on trial difficulty and number of consecutive errors for both animals (Fig. 1H, and fig. S4D; see tables S1 and S2 for model parameters). Further interrogation of the model using cross-validation, parameter identification, and in silico lesioning indicated that the key parameters were both necessary and identifiable (fig. S5). We also verified the predictive validity of the model by confirming that simulations of the model fitted to the subjective psychometric function (Fig. 1E) in the presence of experimentally imposed covert switches were able to capture the objective psychometric function without additional fitting or parameterization (Fig. 1F). On the basis of the success of this model and the failure of alternatives in capturing the animals' switch behavior (Fig. 1H), we hypothesized that the computational logic of the underlying neural circuitry could be understood in terms X_Σ and $X_{y/n}$.

Electrophysiology

Previous work has established a central role for the DMFC and dorsal ACC in monitoring and predicting outcomes (18–25), using outcomes to regulate actions (48, 54–58) and strategic decisions (5, 28–33). Therefore, we recorded neural activity in DMFC—comprising supplementary eye field, dorsal supplementary motor area (i.e., excluding the medial bank), and presupplementary motor area—and ACC (stereotactic coordinates in table S3) while the animals performed the task. Because our main focus was to understand how the animal used decision outcomes to disambiguate errors, we focused our analyses of

neural activity during the intertrial interval (ITI) after the trial outcome was revealed.

As a first step, we characterized individual neurons in terms of their sensitivity to trial outcome, trial difficulty, and the number of consecutive errors (Fig. 2). In both areas, a large proportion of neurons responded differently depending on trial outcome (DMFC: Fig. 2, A to C; ACC: Fig. 2, J to L). Many of the neurons that signaled error trials were differentially modulated depending on trial difficulty (i.e., magnitude of t_d). This difficulty-dependent modulation was present at the level of single neurons (Fig. 2, D, E, M, and N) and across the population (Fig. 2, F and O), as evident from a comparison of firing rates associated with relatively easy ($|t_d| \geq 160$ ms) and difficult ($|t_d| < 160$ ms) trials. Moreover, the firing rate of many error-modulated neurons was modulated depending on the number of preceding errors (Fig. 2, G to I and P to R). The sensitivity of DMFC and ACC neurons to error, difficulty, and trial history corroborates previous findings in a variety of sensorimotor, cognitive, and economic decision-making tasks (18, 59) and suggests that these areas might serve as the neural substrate for making causal inferences about errors in our hierarchical decision task.

Integration of DMFC error-related signals in ACC

Rule inference during ITI depends on two sequential computations: a retrospective computation to evaluate the outcome of the previous trial (i.e., reward versus error) and a subsequent prospective computation to decide whether or not to switch the rule in the upcoming trial. Previous work found that ACC encodes errors with longer latencies than DMFC (18) and suggested that it may act as a "storage buffer" (60), tracking task-relevant variables across trials (24, 28, 33). Following these suggestions, we analyzed the sensitivity of neurons in DMFC and ACC during the ITI to (i) error versus reward in the previous trial and (ii) switch versus no switch in the next trial. Our prediction was that DMFC would exhibit strong and early error-modulated activity but would be less sensitive to switch behavior in the next trial. For ACC, on the other hand, we expected neurons to show a longer latency modulation reflecting subsequent rule switches. To test these predictions, we used a simple linear regression model that aimed to explain modulations of spiking activity in each neuron in terms of two indicators: one specifying whether the animal made an error in the previous trial and one specifying whether the animal switched the rule in the next trial (see methods).

We found that both DMFC and ACC signals during ITI were modulated depending on both error in the previous trial and the animal's switch behavior in the next trial (Fig. 3, A and B). DMFC selectivity (i.e., regression slope, see methods) averaged across neurons that were significantly modulated with respect to error (permutation test for regression slope for each neuron, $P < 0.01$) exhibited strong sensitivity early during ITI. The peak sensitivity occurred at a latency of 145 ms
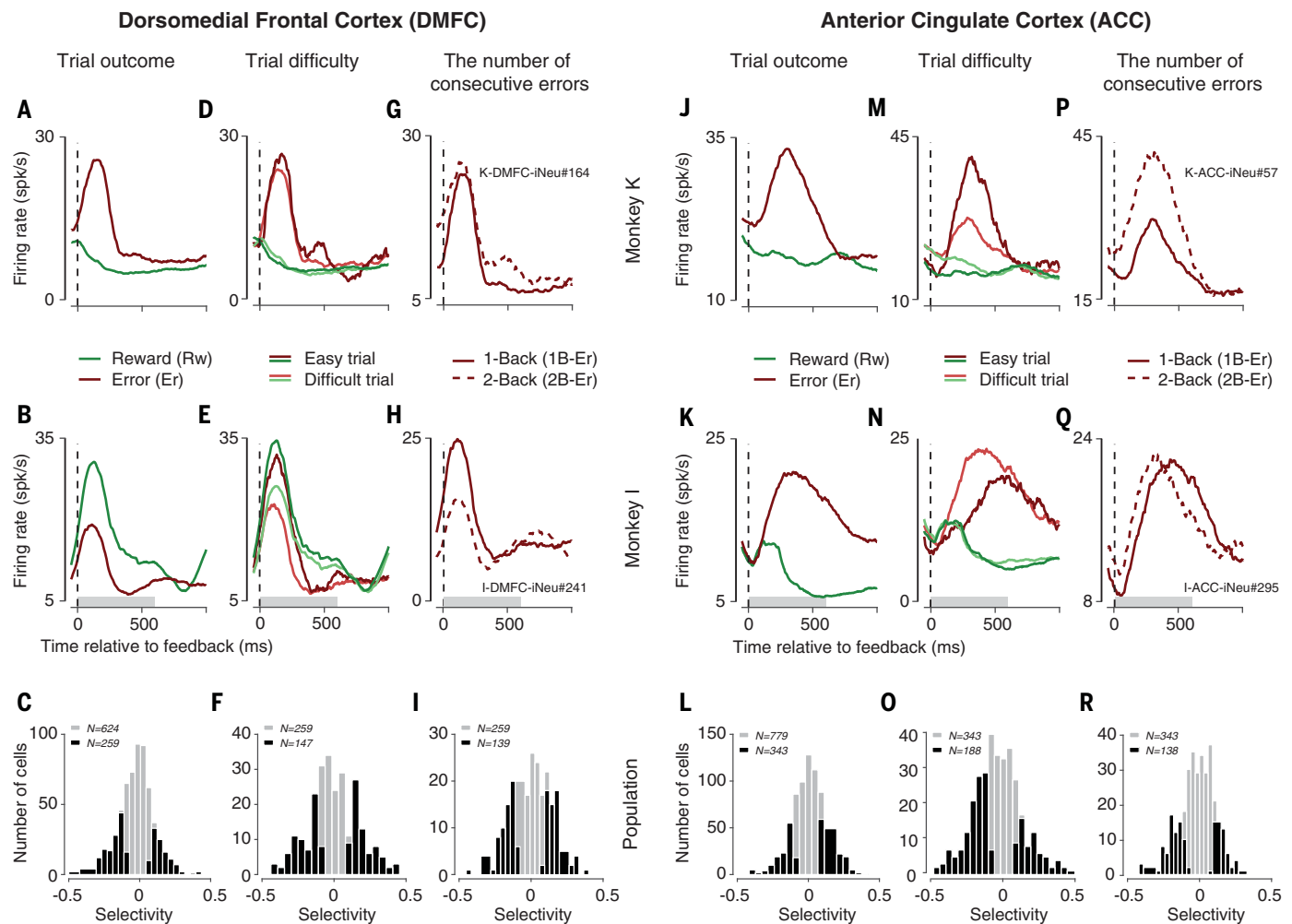
Dorsomedial Frontal Cortex (DMFC)

Anterior Cingulate Cortex (ACC)

Fig. 2. DMFC and ACC selectivity patterns after feedback. (**A** and **B**) Average firing rate of two example DMFC neurons in monkeys K (A) and I (B) for rewarded (green) and unrewarded (red) trials relative to the time of feedback (dashed line). The gray bar in (B) represents a 600-ms window used for analysis in (C), (F), and (I). K-DMFC-iNeu#164 indicates neuron #164 in DMFC of animal K. spk/s, spikes per second. (**C**) Histogram of selectivity to trial outcome (reward versus no reward) across DMFC neurons. Selectivity was computed using receiver operating characteristic (ROC) analysis based on spike counts within a 600-ms window after feedback [gray bar in (A) and (B)]. Black corresponds to neurons with significant selectivity (259/624; permutation test, 100 times, $P < 0.05$). (**D** and **E**) Average firing rate of the same neurons in (A) and (B), respectively, sorted with respect to

relatively "easy" (dark red and green, $|t_d| \geq 160$ ms) and "difficult" (light red and green, $|t_d| < 160$ ms) trials. (**F**) Histogram of selectivity to trial difficulty (easy versus difficult) across DMFC neurons using an ROC analysis similar to that used in (C). Black corresponds to neurons that were significantly modulated by trial difficulty (147/259; permutation test, 100 times, $P < 0.05$). (**G** and **H**) Average firing rate of the same neurons shown in (A) and (B), respectively, after one error (solid line, 1B-Er) and two consecutive errors (dashed line, 2B-Er) following a rewarded trial. (**I**) Histogram of selectivity to consecutive errors (1B-Er versus 2B-Er) across DMFC neurons using an ROC analysis similar to that used in (C). Black corresponds to neurons that were significantly modulated with respect to the number of consecutive errors (139/259; permutation test, 100 times, $P < 0.05$). (**J** to **R**) Same as (A) to (I) for ACC.

and was highly significant (peak sensitivity = 2.48, Wilcoxon rank sum test, $P < 0.001$). ACC was also strongly modulated with respect to trial outcome early during ITI (peak latency = 187 ms, peak sensitivity = 2.21, $P < 0.001$). Additionally, ACC exhibited a strong and late modulation that forecasted the animal's switch behavior in the subsequent trial (peak latency = 415 ms, peak sensitivity = 2.1, $P < 0.001$). DMFC was also modulated by the animal's switch behavior, but these modulations were significantly weaker than ACC (selectivity ratio of DMFC to ACC = 0.74, $P < 0.001$) and lacked the distinctive temporal structure evident in ACC that was indicative of a switch-related computation. Consistent with our

hypothesis, these results suggest that DMFC signals errors and that ACC receives error signals and helps the animal decide whether and when to switch (fig. S6 shows the same results for the two monkeys separately).

To validate this hypothesis, we performed an experiment to test whether the late responses in ACC were sensitive to early signals in DMFC. In a random 50% of error trials, we used electrical microstimulation to alter DMFC activity within a window of 50 to 150 ms in ITI while recording the ensuing spiking activity in ACC. Our choice of the time window of microstimulation was informed by our analysis of the temporal profile of sensitivity of DMFC to error versus reward

(Fig. 3A), and the stimulation current was weak (see methods) to avoid changes in the animal's overt behavior (e.g., evoking saccades). Consistent with our hypothesis, microstimulation of DMFC changed the poststimulation spiking activity of single neurons in ACC (Fig. 3C). To summarize this effect across the population of ACC neurons, we used a modulation index to quantify the relative increase in firing rate after microstimulation. This analysis revealed that ACC neurons were strongly and significantly modulated as a result of DMFC microstimulation (Wilcoxon rank sum test, $P < 0.01$; Fig. 3D). Importantly, we verified that DMFC microstimulation had no statistically significant effect on behavior, ensuring
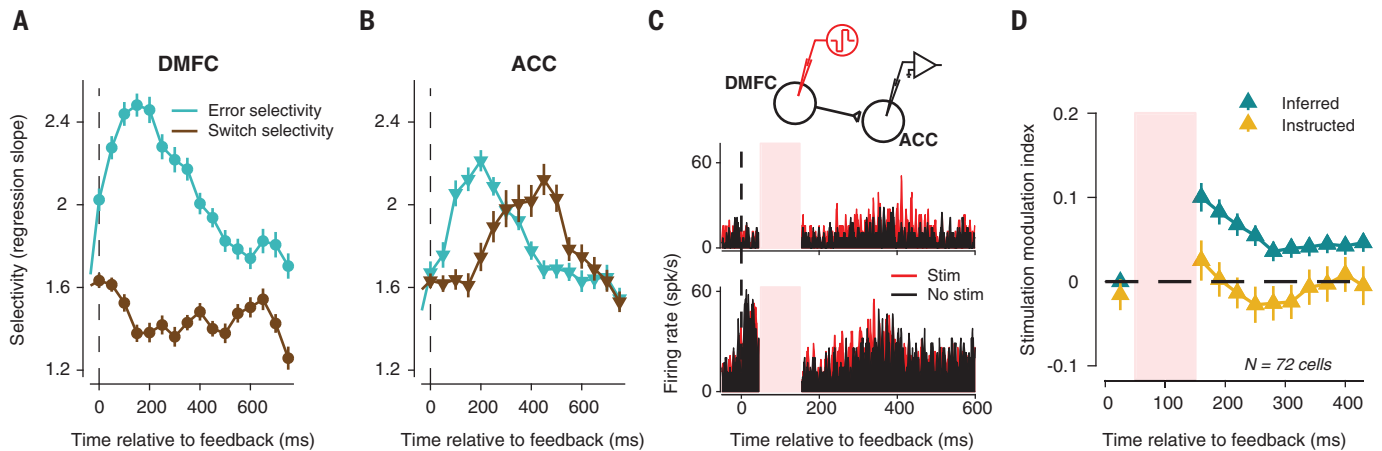
Fig. 3. Retrospective and prospective computations in DMFC and ACC during ITI. (**A**) Time course of DMFC selectivity to error in the previous trial and to rule switches in the next trial. The blue trace shows the time course of selectivity to error versus reward averaged across neurons with significant error-modulated activity. The brown trace shows the time course of selectivity to switch versus no switch averaged across DMFC neurons with significant switch-predictive activity. (**B**) Same as (A) for ACC neurons. (**C**) Modulation of spiking activity in example ACC neurons due to weak electrical microstimulation of DMFC. DMFC microstimulation early during ITI (pink-shaded region, 50 to 150 ms) was followed by an increase in the firing rate of ACC neurons (red) compared with no stimulation (black). (**D**) Modulation of ACC activity due to DMFC microstimulation during the inferred and instructed rule experiments. For each ACC neuron, we computed a microstimulation modulation index quantifying the relative increase in firing rate as a result of microstimulation. Denoting the firing rate in the stimulated and nonsimulated trials by r_1 and r_2 (measured by a sliding window of 150 ms), respectively, the modulation index was computed as $(r_1 - r_2)/(r_1 + r_2)$. The averaged modulation index was then computed across neurons before and after the stimulation window separately for the inferred (green) and instructed (yellow) rule experiments. Stimulation sensitivity index was significantly increased within 150 to 300 ms relative to feedback in the inferred rule experiment, but not for the instructed rule experiment (Wilcoxon rank sum test, $P < 0.01$). In (A), (B), and (D), error bars indicate SEM.

that modulation of ACC activity was not due to an indirect change in switch probability (fig. S7).

To ensure that the communication channel from DMFC to ACC played a task-dependent functional role and was not due to some nonspecific microstimulation-induced corticocortical interactions, we additionally examined the effect of DMFC microstimulation on ACC in the instructed rule experiment using an identical protocol (same experimental sessions, same time window, same current levels, etc.). The key difference in the instructed rule experiment compared with the inferred rule experiment was that rule switches were instructed and the animals did not have to rely on errors in the previous trial to decide whether to switch. In other words, errors did not inform rule switches, and, therefore, there was no need to engage the communication channel from DMFC to ACC to integrate early error-related signals. Accordingly, we found that the same microstimulation protocol had no statistically significant effect on ACC activity in the instructed rule experiment (Fig. 3D and fig. S8, for the individual animals). These results provide evidence that ACC integrates error-related signals in DMFC selectively when there is a need to evaluate errors with respect to a hierarchy of potential causes and decide whether to switch.

ACC represents cumulative switch evidence and drives switching behavior

Both the task-selective effect of DMFC microstimulation on ACC (Fig. 3, C and D) and the strong switch-predictive signals in ACC (Fig. 3B) motivated a hypothesis that ACC may be involved in computing the decision to switch. For

the hypothesis to hold, ACC must harbor both a graded representation of the evidence supporting a switch and a binary signal predicting future switches. To test this hypothesis, we examined ACC signals in the context of our model of behavior. The model formalized subjective switches in terms of a binary variable, $X_{y/n}$, whose state (switch or no switch) was set on the basis of the value of cumulative switch evidence relative to a threshold (Fig. 1I). The cumulative switch evidence, in turn, was captured by a single graded latent variable, $\hat{X}_\Sigma$, whose value absorbed the dependence of behavior on both trial difficulty (e.g., larger for easier trials) and the number of consecutive errors.

We performed a regression analysis on individual ACC neurons to investigate whether $\hat{X}_\Sigma$ and $X_{y/n}$ were encoded by distinct subpopulations of neurons or were mixed across the population. Consistent with numerous recent studies in various higher cortical areas (*35*, *61*–*65*), this analysis indicated that $\hat{X}_\Sigma$ and $X_{y/n}$ were mixed across the population (fig. S9). Therefore, we used a recently developed targeted dimensionality reduction technique (*63*) that allowed us to tease apart signals encoding $\hat{X}_\Sigma$ and $X_{y/n}$ independently across the population (see methods).

We found a strong representation of $\hat{X}_\Sigma$ in ACC population activity (Fig. 4A, top). We quantified the sensitivity of ACC to $\hat{X}_\Sigma$ by measuring the slope of the regression line relating $\hat{X}_\Sigma$ to ACC population activity associated with $\hat{X}_\Sigma$ based on firing rates 200 to 400 ms after feedback (Fig. 4A, bottom). The regression slope was significantly positive (* indicates significance

at $P < 0.05$) for both animals independent of the binary choice about rule switch {monkey K: 0.809*, confidence interval (CI): [0.533, 1.085] when $X_{y/n} = 1$ and 1.098*, CI: [0.844, 1.351] when $X_{y/n} = 0$; monkey I: 0.532*, CI: [0.475, 0.589] when $X_{y/n} = 1$ and 0.884*, CI: [0.729, 1.038] when $X_{y/n} = 0$}. This observation suggests that ACC either computed or received inputs that encode $\hat{X}_\Sigma$.

ACC also had a clear representation of $X_{y/n}$, as evidenced by the convergence of the population activity associated with $X_{y/n}$ on to one of two distinct states depending on whether the animal would switch in the next trial or not (Fig. 4B, top). We used linear regression to quantify the degree to which ACC responses 200 to 400 ms after feedback were modulated by $X_{y/n}$ (Fig. 4B, bottom). The baseline for the regression was significantly larger when the animal switched in the next trial (monkey K: 2.408*, CI: [1.977, 2.840]; monkey I: 1.325*, CI: [0.993, 1.657]) compared with when it did not (monkey K: 0.707*, CI: [0.514, 0.900]; monkey I: 0.788*, CI: [0.627, 0.950]).

All the analyses relating ACC activity to $\hat{X}_\Sigma$ and $X_{y/n}$ were performed on cross-validated data and were not the result of an overfitted regression model. Moreover, the encoding of $\hat{X}_\Sigma$ and $X_{y/n}$ in ACC was robust to various nuisance parameters (fig. S10), including saccade direction and response type (Pro versus Anti). These results provided evidence in support of the hypothesis that ACC may be responsible for converting the cumulative switch evidence, $\hat{X}_\Sigma$, to the binary decision to switch, $X_{y/n}$. (Figs. S11 and S12 show the results of the same analysis for DMFC
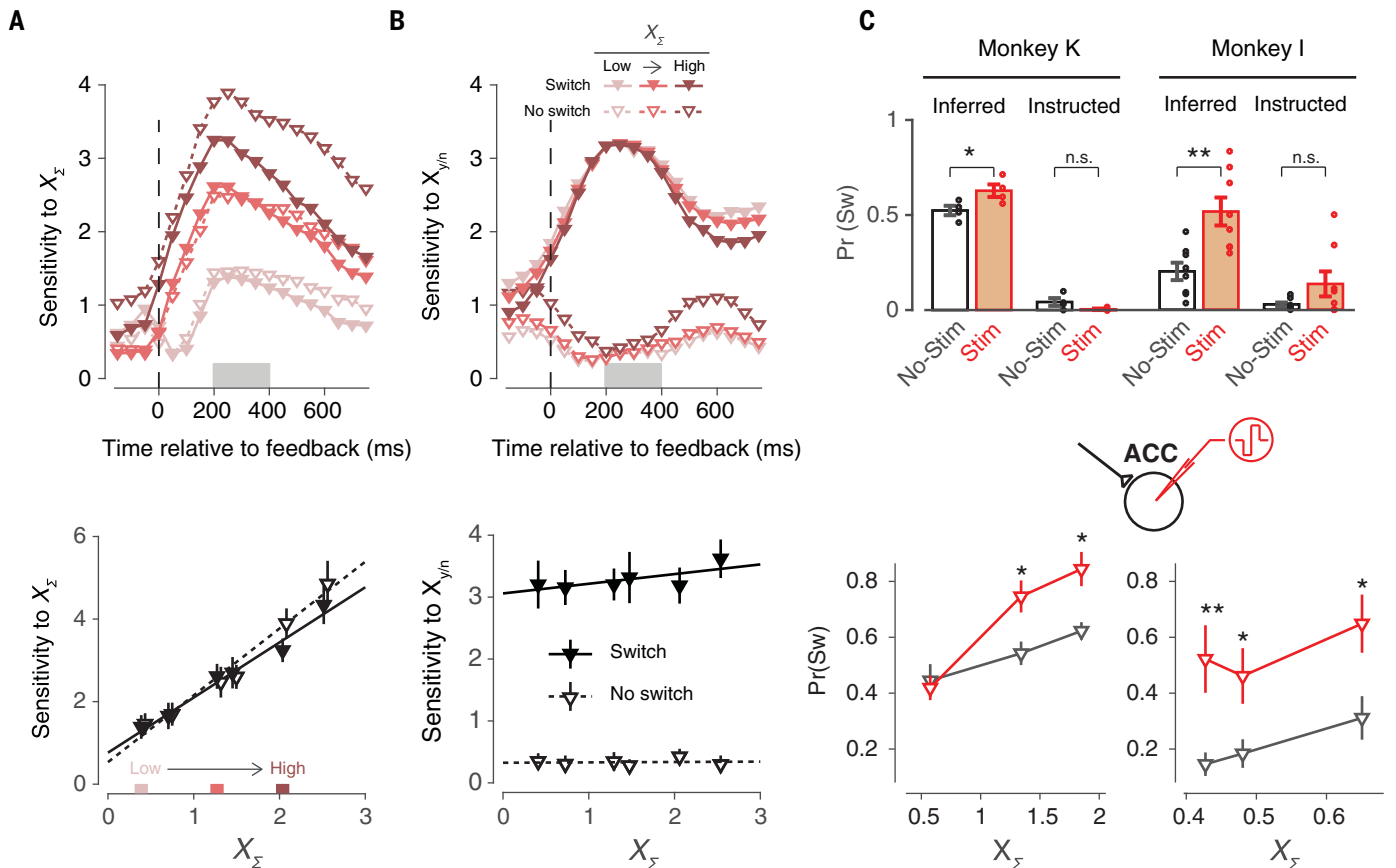
Fig. 4. Representation of switch evidence and causal manipulation of switch behavior in ACC. (**A**) Encoding of the graded switch evidence (X_Σ) by population activity in ACC. Using targeted dimensionality reduction, we identified the pattern of population activity in ACC that encoded X_Σ, inferred from the animals' behavior. We then derived a sensitivity index (d′) to quantify the distance between activity associated with different levels of X_Σ (low to high, shown in three colors) and activity associated with rewarded trials (i.e., $X_\Sigma = 0$). The sensitivity is shown separately for trials that led to a switch (filled symbols) and those that did not (open symbols). The top plot shows sensitivity as a function of time, and the bottom plot shows the response tuning to X_Σ within a time window of 200 to 400 ms after feedback (gray region in the top plot). Results for both panels correspond to cross-validated data (see methods). (**B**) Encoding of the binary switch variable ($X_{y/n}$) by population activity in ACC. Applying the same analysis technique as in (A) revealed that ACC activity patterns that were sensitive to $X_{y/n}$ were organized in a binary fashion (i.e., switch versus no switch) regardless of the value of X_Σ (cross-validated; see methods). (**C**) (Top) The effect of ACC electrical microstimulation within the 200 to 400 ms relative to feedback on switch probability [Pr(Sw)] in the inferred and instructed task. Open circles represent individual sessions. Electrical stimulation was applied on ~50% of error trials (red, Stim), and the effect of Pr(Sw) was compared with the nonstimulated trials (black, No-Stim). In both animals, ACC microstimulation led to an increased Pr(Sw) in the inferred task (Wilcoxon rank sum test, *P < 0.05, **P < 0.01). By contrast, switches in the instructed task did not increase after ACC microstimulation (n.s., not significant, P > 0.05). For the instructed task, Pr(Sw) is only shown for trials where the subjective switch was not cued. (Bottom) The effect of ACC stimulation in the inferred rule experiment on Pr(Sw) as a function of X_Σ of 1B-Er trials (Wilcoxon rank sum test, *P < 0.05, **P < 0.01). In (A) to (C), error bars indicate SEM.

and ACC, respectively, separately for the two animals.)

As a final test for our hypothesis, we used electrical microstimulation to perturb ACC during the window of 200 to 400 ms in which the sensitivity of the population activity to $\hat{X}_\Sigma$ and $X_{y/n}$ was evident (Fig. 4, A and B). We reasoned that if the animals relied on ACC to decide when to switch, perturbation of signals in ACC in this window would interfere with this computation and would result in a measurable difference in the observed switch probability. Consistent with this prediction, we found that switch probabilities increased significantly in the random subset of error trials after which ACC was electrically stimulated (Fig. 4C and fig. S13). This experiment provided evidence that ACC plays a functional

role in making causal inferences about covert rule switches. Although the effect of microstimulation was strong and significant in both animals, in one animal the stimulation enhanced the sensitivity of the animal to preceding errors, as evidenced by an increase in slope (Fig. 4C, bottom left), whereas in the other animal, the effect was an overall increase the switch probability, evident as a change in baseline (Fig. 4C, bottom right).

To ensure that this effect was indeed related to ACC supporting the animal's attempt to infer covert rule switches, we additionally tested the effect of ACC microstimulation in the instructed rule experiment. Because switches in this control experiment were instructed and not based on cumulative switch evidence, we predicted that

ACC perturbation should have minimal effect on the animal's switch probability. Results of this experiment supported our prediction. ACC microstimulation in the instructed rule task had nearly no effect on switch probability. It is possible that the animals made inferences about the rule in both tasks, but this inference was overridden in the instructed task by the reliable visual cue. The positive effect of microstimulation in the inferred rule task coupled with the negative effect in the instructed task further substantiated a central role for ACC in making causal inferences about errors in a hierarchical setting.

Discussion

Our results provide an understanding of the computational principles and neurobiological

underpinnings of adaptive decision-making when the agent has to reason about errors in the presence of multiple hierarchically organized sources of uncertainty. In terms of behavior, we were able to establish that macaque monkeys are capable of implementing a causal inference strategy with a level of sophistication that is comparable to humans (*3*). Both monkeys treated each negative outcome as evidence for a covert rule switch but did so rationally by taking into account that (i) errors after easy trials (i.e., higher expectation of reward) were more likely due to a rule switch than a timing error and (ii) repeated errors provided stronger evidence for a rule switch. In terms of neurobiology, we found that the animals' behavior was supported by distributed and hierarchically organized neural circuits in the frontal cortex (*45, 66*). Specifically, we found that DMFC and ACC, two areas that have been implicated in error monitoring (*18, 26, 67–73*) and the control of adaptive behavior (*5, 28–36*), carried signals relevant for causal inference. ACC, however, seemed to function downstream of DMFC and played a direct role in integrating evidence and computing whether and when to attribute failures to a covert rule switch.

Both DMFC and ACC were modulated retrospectively by the outcome of the low-level decisions, which is consistent with previous findings that suggest a role for these areas in performance monitoring (*18–21, 23, 25*). In other words, both areas had a representation of switch evidence. However, the hierarchical nature of our task allowed us to uncover a distinct and longer latency form of modulation in ACC that was predictive of the animal's subsequent switch behavior. This result, coupled with an experiment involving simultaneous DMFC perturbation and ACC recording, enabled us to reveal the sequential nature of computations in DMFC and ACC and provided a coherent explanation of their distinct functional link to behavior, their different activation latencies, and their circuit-level interactions. In particular, our results suggest that the long latency signals in ACC result from an integration of early error-related modulations in DMFC, much like what is seen in well-established first-order perceptual decision-making tasks between sensory and higher-order association and premotor areas (*74–81*). An important outstanding question that our study motivates is the problem of maintaining a desired behavioral strategy between switches, which may involve other brain areas such as the dorsolateral prefrontal cortex (*82–88*).

ACC is modulated by reward history (*23, 89*) and reward expectations (*24*) and is linked to reward-dependent control of adaptive behavior in many situations, including movement selection (*58, 90*), decision-making under risk (*33*), foraging (*28*), and regulation of exploration versus exploitation (*91*). Our work extends the role of ACC in adaptive control to the general class of cognitive tasks that demand hierarchical reasoning about errors. As we found in our work, and others, in other behavioral settings (*3, 16, 92*), reasoning about errors can be captured by a model comprised of two key computational variables: (i) a latent switch evidence variable that accumulates confidence-dependent error information and (ii) a latent threshold that controls a binary decision to switch depending on the strength of the cumulative switch evidence. We found that ACC harbors distinct activity patterns associated with both the cumulative switch evidence and the binary decision to switch that were mixed at the level of single neurons but dissociated across the population. Moreover, we found that perturbation of ACC interfered with the animal's rationality during causal reasoning about errors. Together, these results suggest that ACC uses graded evidence derived from errors in low-level processes in a decision hierarchy to select between longer-term behavioral strategies associated with higher levels of the hierarchy.

Materials and methods

All experimental procedures conformed to the guidelines of the National Institutes of Health and were approved by the Committee of Animal Care at the Massachusetts Institute of Technology. Experiments involved two awake, behaving monkeys (species: *Macaca mulatta*; ID: K and I; weight: 7.5 and 10.5 kg; age: 4 and 5 years old, respectively). Animals were head-restrained and seated comfortably in a dark and quiet room and viewed stimuli on an Acer H236HL LCD monitor (23 inch; refresh rate: 60 Hz; resolution: 1920 by 1080). All reported stimulus presentation times had to be rounded to a multiple of the frame duration (16.67 ms). Eye movements were registered by an infrared camera and sampled at 1 kHz (Eyelink 1000, SR Research Ltd., Ontario, Canada). The MWorks software package (http://mworks-project.org) was used to present stimuli and to register eye position. A photodiode was used for registering the timing of events during stimulus presentation.

Neurophysiology recordings were made by 24- or 36-channel laminar probes with 100-μm interelectrode spacing (V-probe, Plexon Inc.) through a biocompatible cranial implant whose position was determined by stereotaxic coordinates and structural MRI scan of the two animals. Extracellular signals were bandpass filtered (300 Hz to 6 kHz) and digitized (sampling rate: 30 kHz; resolution: 16 bit) using a digital Intan headstage (Intan Technologies, http://intantech.com/), and the data were collected using the OpenEphys system (www.open-ephys.org/). Single or multi-unit action potential waveforms were detected and sorted offline using MKsort (https://github.com/ripple-neuro/mksort). Analysis of both behavioral and spiking data was performed using custom MATLAB code (Mathworks, MA).

Electrophysiology

We recorded neural activity in (i) DMFC, comprising supplementary eye field, dorsal supplementary motor area (i.e., excluding the medial bank), and presupplementary motor area; and (ii) ACC. Recording sites and number of sessions and trials are reported in table S3.

ACC electrical microstimulation

On a random 50% of error trials, we stimulated ACC in a 200- to 400-ms time window after the saccade reached to target. We generated biphasic current pulses (amplitude: 40 to 60 μm peak-to-peak; pulse duration: 300 μs; frequency: 200 Hz; total stimulation duration: 200 ms) using a commercial stimulator (CereStim 96, Blackrock Microsystems) and injected current using a custom bipolar system with two tungsten microelectrodes (10 to 50 kilohm, Microprobes), one serving as a current source and another as a dedicated return. To cover the region of interest in ACC, we placed the current source just below the border of gray-white matter in ACC margin and the return electrode, ~3 mm above and ~2 mm lateral to the source electrode. The number of stimulation sessions and trials are reported in table S1. To have an appropriate control for these stimulation experiments, both the inferred and instructed experiments were tested on every stimulation session.

Simultaneous electrical microstimulation of DMFC and recording in ACC

In a random 50% of error trials, we stimulated DMFC in a 50- to 150-ms time window after the saccade reached to target and recorded the effect of stimulation on ACC afterward. We injected current (amplitude: 20 to 30 μm peak-to-peak; pulse duration: 300 μs; frequency: 200 Hz; total stimulation duration: 100 ms) through a tungsten electrode placed ~2 mm deep just below the border of gray-white matter in DMFC and used a guide tube piercing the dura as the return, while recording simultaneously with a separate V-probe. The stimulation current was weak to make sure the behavior would not change. This was particularly important because stimulation of DMFC could evoke saccades (e.g., around supplementary eye fields), and such explicit change in behavior would complicate the interpretation of any observed changes in ACC activity. The number of stimulation sessions and trials are reported in table S3. To have an appropriate control for these stimulation experiments, both the inferred and instructed experiments were tested on every stimulation session.

REFERENCES AND NOTES

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; pmid: 24876345
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; pmid: 27432960
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; pmid: 30122376
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; pmid: 30189202

6. Q. J. M. Huys et al., Interplay of approximate planning strategies. Proc. Natl. Acad. Sci. U.S.A. 112, 3098–3103 (2015). doi: 10.1073/pnas.1414219112; pmid: 25675480

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; pmid: 21867887

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; pmid: 19423820

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; pmid: 18690210

10. S. M. Fleming, H. C. Lau, How to measure metacognition. Front. Hum. Neurosci. 8, 443 (2014). doi: 10.3389/fnhum.2014.00443; pmid: 25076880

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; pmid: 22884334

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; pmid: 20658856

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; pmid: 23915088

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; pmid: 24816801

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; pmid: 23666179

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; pmid: 27866891

17. A. Lak et al., Orbitofrontal cortex is required for optimal waiting based on decision confidence. Neuron 84, 190–201 (2014). doi: 10.1016/j.neuron.2014.08.039; pmid: 25242219

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; pmid: 14526085

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; pmid: 12383784

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; pmid: 15718473

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; pmid: 16026482

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; pmid: 10846167

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; pmid: 17670983

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; pmid: 12040201

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; pmid: 9389478

26. F. Bonini et al., Action monitoring and medial frontal cortex: Leading role of supplementary motor area. Science 343, 888–891 (2014). doi: 10.1126/science.1247412; pmid: 24558161

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; pmid: 22037498

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; pmid: 21642973

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; pmid: 11488380

30. J. G. Kerns et al., Anterior cingulate conflict monitoring and adjustments in control. Science 303, 1023–1026 (2004). doi: 10.1126/science.1089910; pmid: 14963333

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; pmid: 23889930

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; pmid: 17270740

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; pmid: 16783368

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; pmid: 29203897

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; pmid: 29879384

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; pmid: 11970861

37. J. D. Murray et al., A hierarchy of intrinsic timescales across primate cortex. Nat. Neurosci. 17, 1661–1663 (2014). doi: 10.1038/nn.3862; pmid: 25383900

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; pmid: 17676057

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; pmid: 9744965

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; pmid: 9772238

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; pmid: 18926527

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; pmid: 18403252

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; pmid: 18420448

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; pmid: 17087577

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; pmid: 14615530

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; pmid: 28242793

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; pmid: 27669989

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; pmid: 24012280

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; pmid: 29946069

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; pmid: 15004564

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; pmid: 25581364

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; pmid: 30033151

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; pmid: 20844132

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; pmid: 12574489

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; pmid: 21926982

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; pmid: 20203193

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; pmid: 17145511

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; pmid: 9812901

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; pmid: 11130724

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; pmid: 27090954

61. M. Rigotti et al., The importance of mixed selectivity in complex cognitive tasks. Nature 497, 585–590 (2013). doi: 10.1038/nature12160; pmid: 23685452

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; pmid: 23392657

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; pmid: 24201281

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; pmid: 24487233

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; pmid: 25383902

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; pmid: 28209978

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; pmid: 12374324

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; pmid: 16732274

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; pmid: 24141310

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; pmid: 3774229

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; pmid: 111772

72. C. Shen *et al.*, 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; pmid: 24591526

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; pmid: 11707094

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; pmid: 8570606

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; pmid: 12417672

76. A. Hernández *et al.*, Decoding a perceptual decision process across cortex. *Neuron* **66**, 300–314 (2010). doi: 10.1016/j.neuron.2010.03.031; pmid: 20435005

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; pmid: 18488024

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; pmid: 10325224

79. T. D. Hanks *et al.*, Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**, 220–223 (2015). doi: 10.1038/nature14066; pmid: 25600270

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; pmid: 23636333

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; pmid: 18327253

82. M. K. Mian *et al.*, Encoding of rules by neurons in the human dorsolateral prefrontal cortex. *Cereb. Cortex* **24**, 807–816 (2014). doi: 10.1093/cercor/bhs361; pmid: 23172774

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; pmid: 15883365

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; pmid: 12729487

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; pmid: 11283309

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; pmid: 10899218

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; pmid: 26343320

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; pmid: 16207931

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; pmid: 15558064

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; pmid: 18215627

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; pmid: 26322383

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/364/6441/eaav8911/suppl/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S3
References (*93*, *94*)