Title: Dynamic task-belief is an integral part of decision-making

Authors: Cheng Xue, Lily E. Kramer, Marlene R. Cohen

Affiliations: Department of Neuroscience and Center for Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, PA

Summary: Natural decisions involve two seemingly separable processes: inferring the relevant task (task-belief) and performing the believed-relevant task. The assumed separability has led to the traditional practice of studying task-switching and perceptual decision-making individually. Here, we used a novel paradigm to manipulate macaque monkeys' task-belief, and demonstrated inextricable neuronal links between flexible task-belief and perceptual decision-making that substantially impact behavior. We showed that in animals, but not artificial networks trained to make perceptual decisions under fluctuating task-belief, stronger task-belief is associated with better perception. Correspondingly, recordings from neuronal populations in monkey cortical areas 7a and V1 revealed that stronger task-belief is associated with better discriminability of the believed-relevant but not the believed-irrelevant feature. Perception also impacts belief updating: noise fluctuations in V1 help explain how task-belief is updated. Our results demonstrate that complex tasks and multi-area recordings can reveal fundamentally new principles of how biology affects behavior in health and disease.

Body Text: Humans and animals make countless decisions every day that affect their well-being or even survival. In the laboratory, decision-making has typically been studied by observing behaviors and neuronal activity while subjects perform simple, well-understood sensory-motor integration tasks (Gold and Shadlen, 2007; Kable and Glimcher, 2009; Uchida, Kepecs and Mainen, 2006). But real-life decisions usually need to contend with a more important problem even before making perceptual judgements: inferring the relevant task to solve in a certain situation (i.e., task-belief). Task-beliefs allow decision-makers to focus on a relevant subset of the huge amount of information in natural environments, and task-beliefs are flexibly adapted as the environment evolves (task-switching(Monsell, 2003)). Flexibly adapting task-belief is critical but difficult: the inability to appropriately respond to changing conditions is a debilitating symptom of disorders including autism, dementia, and substance abuse (Brady, Gray and Tolliver, 2011; Thapar *et al.*, 2016; Dickstein *et al.*, 2007).

Typically, task-belief is assumed to be a separate functional module that occurs before, and independent of, perceptual decision-making. In this view, the task-belief module (possibly involving parietal, prefrontal, and cingulate cortical areas (Stoet and Snyder, 2009; Buschman *et al.*, 2012; Kamigaki, Fukushima and Miyashita, 2009; Sarafyazd and Jazayeri, 2019; Bartolo and Averbeck, 2020)) identifies the relevant task and then the perception module (involving sensory areas such as visual cortex) performs perceptual judgements on the chosen task (Purcell and Kiani, 2016a; Sarafyazd and Jazayeri, 2019; Mante *et al.*, 2013) (Figure S1A, upper panel). We trained monkeys on a two-feature discrimination task that allows rhesus monkeys to report both their task-belief, and their perceptual judgment on the relevant feature. By decoding task- and perception-related signals from simultaneously recorded neuronal populations in parietal cortical area 7a and visual cortical area V1, we demonstrated that: 1) trial-by-trial fluctuations in task-belief strength

correlate with perceptual performance and the fidelity of task-relevant information encoding in visual cortex; and 2) neuronal variability in visual cortex, even those unrelated to the visual input, plays an important role in updating task-belief (Figure S1A, lower panel). These results revealed a very integrated neuronal system that serves as the common substrate for flexible task-belief and decision-making.

As neuroscience and cognitive science research start to shift focus from simplified functions in isolation towards more complex and natural behavior, our work has implications for methodology, basic science, and clinical translations at the time of this paradigm shift. 1) Methodologically, our work showcased how dynamic cognitive states (e.g. task-belief) in natural behavior can be systematically manipulated and measured by large-scale neuronal recording from multiple brain areas during complex cognitive tasks. 2) Scientifically, we demonstrated inextricable links between task-belief and perceptual decisions, which would be neglected if either process is studied alone. It is therefore wrong to assume that all complex cognitions can be understood by investigating their components separately. 3) In translational research, a wide variety of neuropsychiatric disorders lead to malfunctions in task-switching and decision-making. The neuronal link between the two processes suggests that potential treatments might do well to target neurotransmitters that mediate communication between brain areas.

Perception during task-switching

Task-belief is an internal state that continually changes. Even with experimenters' best attempts to keep task-belief constant (with fixed stimuli, explicit instructions, and task statistics), internal belief states still have uncontrolled fluctuations (Purcell and Kiani, 2016b; Ebitz, Tu and Hayden, 2020; Cohen and Maunsell, 2011a; Cohen and Maunsell, 2011b), some with effects on

visual cortical activity and perceptual performance that is supposedly confined to the perception module (Cohen and Maunsell, 2011a; Cohen and Maunsell, 2011b; Monsell, 2003). These results suggest that beliefs and perception interact in complex ways. The biggest barrier to understanding such interactions is estimating task-belief during each decision, which is by definition internal and continually changing.

To address this challenge, we devised a novel two-feature discrimination task to assess perception and belief simultaneously. We trained animals to discriminate either the spatial location or spatial frequency of two Gabor patches presented in series. The animals indicated both the subjective belief about which feature was task-relevant and the corresponding perceptual judgment by making a saccade to one of four targets (Figure 1A, upper left panel). They were rewarded only when both task-belief and perception were correct (Figure 1A, right panel). The relevant feature was not cued and switched with a low probability from one trial to the next (Figure 1A, lower left panel). This design provides rich and easy measurements and manipulations of subjects' behavior during dynamic belief-based decision-making. Meanwhile, we recorded from groups of neurons from which we could decode information about both visual features the monkeys discriminated (in visual cortical area V1) and task-belief (in parietal area 7a(Kamigaki, Fukushima and Miyashita, 2009; Stoet and Snyder, 2004)) (Figure 1B). Together, these measurements provide a unique window into belief updating and perceptual decision-making on every trial.

After training, the animals successfully discriminated the feature change they believed to be relevant, and largely ignored the feature believed to be irrelevant (Figure 1C). The animals also effectively updated their belief according to the evolving task requirements, switching tasks only a couple of trials after the (uncued) task changes occurred (Figure 1D). The number of trials the

animals took to notice task changes was close to optimal given their perceptual sensitivity (Figure 1E).

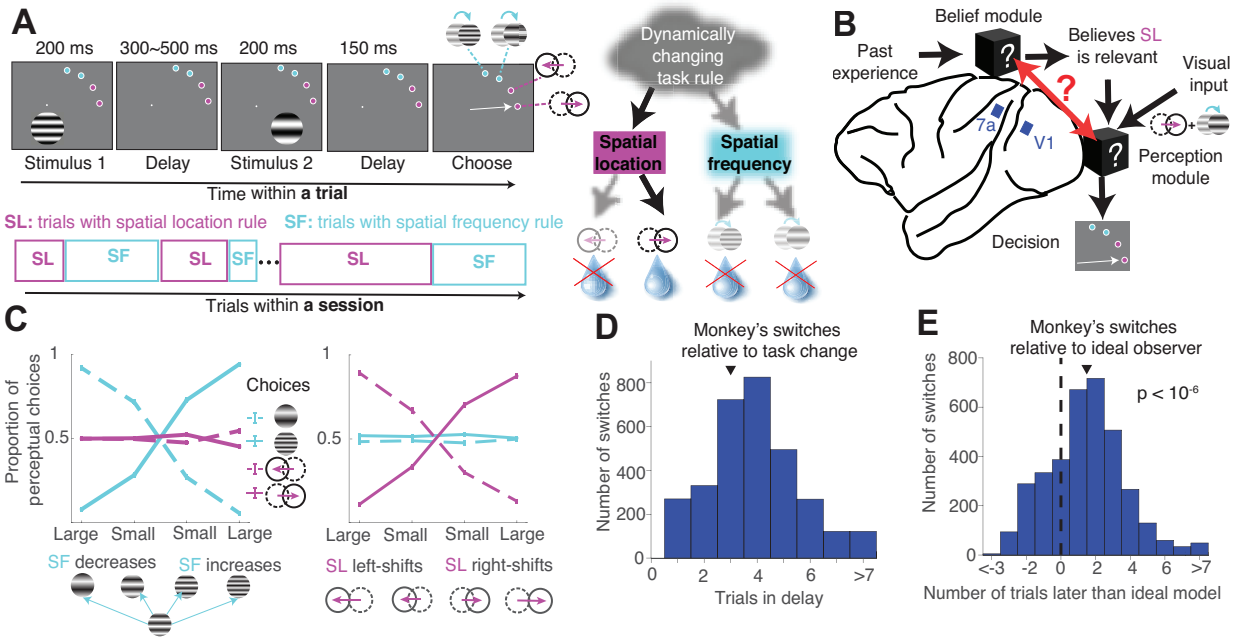

Figure 1. Behavioral paradigm and electrophysiological recording. (**A**) Schematic of the two-interval, two-feature discrimination task with stochastic task switching. On each trial, monkeys discriminate the difference in either spatial frequency or spatial location between two subsequent Gabor stimuli; and are rewarded for correctly reporting the sign of the change in the relevant feature. The relevant feature is uncued and changes with 2.5% probability on each trial. The monkeys indicate their perceptual decision and the feature believed to be relevant by making a saccade to one of four choice targets. (**B**) Belief-based decisions could potentially be solved by independent hierarchical modules that compute belief and perception (black boxes). We simultaneously recorded population activity from one representative brain region for each module (7a and V1 respectively, blue squares show approximate implant locations) to test the hypothesis that these modules are non-independent (red arrow). (**C**) Psychometric curves showing the monkeys' perceptual choice proportion as a function of spatial frequency (left panel) and spatial location (right panel) differences. The flat curves for the irrelevant feature show that animals successfully ignored irrelevant visual information. (**D**) Distribution of number of trials it took the monkeys to adapt to task changes across experimental sessions (mean 3.1 trials). (**E**) Distribution of number of trials the monkey took to adapt to the task change relative to an ideal observer model (see STAR

Methods). Positive values refer to occasions where monkeys were slower than the model; negative values indicate that the monkeys were accidentally faster (mean 1.5 trials).

Perception covaries with belief strength

Our behavioral results demonstrate that dynamic task-belief strength affect the accuracy of perceptual decision-making. By design, the animals' perceptual choices (as opposed to task choices) are informed by stimulus information within each trial and should ideally be independent from trial history. However, across experimental sessions (focusing on subsets of trials with the same stimulus conditions), the animals had better perceptual performance (i.e., perceptual accuracy of whichever task the animal chose to perform) after rewarded trials (which reinforced task-belief) than after unrewarded trials (which would weaken the monkeys' task-belief) (Figure 2A upper panel, Figure S2A). This association between strong task-belief and improved perception persisted even when taking task-switch cost and task-set inertia (Alport, Styles and Hsieh, 1994) into account (Figure S2C-D). Furthermore, perceptual performance was even worse when the previous unrewarded trial contained stronger perceptual evidence against task-belief, such as a non-reward following an obvious change in a believed-relevant feature (Figure S2E). Correspondingly, the monkeys' perceptual choices were strongly related to V1 activity following rewarded than unrewarded trials (Figure S2B).

The observation that perceptual performance is different following rewarded than unrewarded trials is broadly consistent with either of two classes of mechanisms. This behavioral interaction could be caused by the underlying biology: there could be an interaction between the neurons that encode task-belief strength and those that encode the relevant or irrelevant visual features. Alternately, the behavioral could be caused purely by the details of the task structure: the monkey

133 has no way of knowing whether a lack of reward indicates an error in perception, task-belief, or
134 both. If this ambiguity introduces general uncertainty about the logic of the experiment, all aspects
135 of the behavior might simply be worse following negative feedback.

136 To differentiate the two hypotheses, we trained a recurrent neural networks to perform this
137 task. Like the monkeys, the recurrent neural networks have information about both the visual
138 stimuli on the current trial and the recent trial history, and need to infer the relationship between
139 those inputs and the correct behavior. If task-uncertainty induced by negative feedback worsens
140 behavior, this should affect the model as well as the monkey. If, on the other hand, worse
141 perception following non-rewards is induced by a neuronal interaction between the biological
142 mechanisms mediating perception and task-belief, the model may be unaffected.

143 The behavior of the trained recurrent neural networks bears many similarities to the monkeys'
144 behavior. Like the monkeys, the network can distinguish relevant information from irrelevant
145 information (Figure S3B) and can efficiently detect task changes and adjust behavior accordingly
146 (Figure S3C).

147 But unlike the monkeys, the network showed no dependence of perceptual performance on the
148 reward history (Figure 2A), suggesting that biological interactions cause this dependence in our
149 monkeys. The difference between monkey and model behavior cannot be explained by the network
150 having mastered the task better than the monkey: the network model never displayed significant
151 dependence of perceptual performance on past feedback during any stage of training (Figure 2B).
152 These results demonstrate that better perceptual performance after rewarded than after unrewarded
153 trials is not a necessary aspect of performing this task. Instead, they suggest that the difference

between monkeys and models reflects the biological realities of how brains calculate task-belief and perception.

We therefore investigated potential neuronal interactions between task-belief strength and the representation of relevant and irrelevant features in visual cortex in V1. To test this hypothesis on a trial-by-trial basis, we leveraged the fact that neuronal populations in parietal area 7a encode task-related information (Stoet and Snyder, 2004; Kamigaki, Fukushima and Miyashita, 2009). We decoded this continuous measure of the animals' task-belief on each trial (Figure 2C). Consistent with the idea that rewards reinforce beliefs, the animals' task choice was better classifiable after a rewarded than an unrewarded trial (Figure S1C). Decisions to switch tasks were associated with a dynamic change in decoded task-belief away from the old task and toward the new task (Figure 2D).

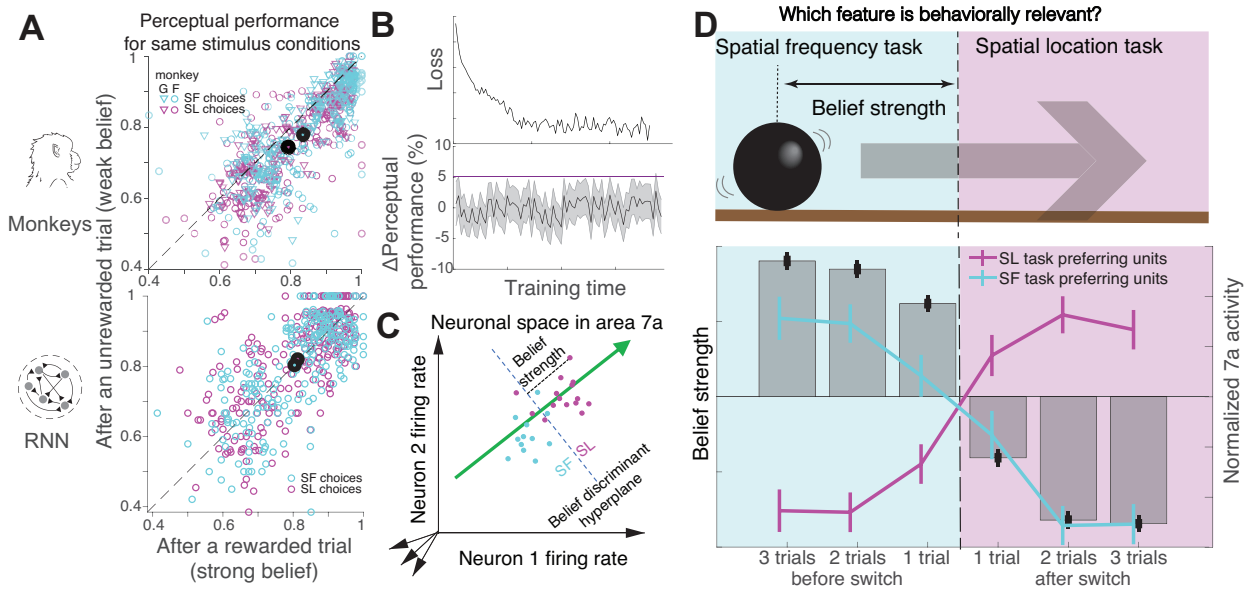

Figure 2. Behavioral and neuronal measures of belief strength. (**A**) Comparing perceptual performance following rewarded trials (abscissa) and unrewarded trials (ordinate) for monkeys (upper panel) and the recurrent network model (lower panel). Each point

represents one stimulus condition of an experimental session, and we compute perceptual performance based on the subjectively chosen task, regardless of whether that task-belief was correct. Upper panel, the monkeys' perceptual performance is better following rewarded trials than unrewarded trials. The distribution lies significantly below the unity line ($p<10^{-6}$ for both monkeys and both features), showing lower perceptual performances following a non-rewarded trial than following a rewarded trial, with the same perceptual difficulty. Lower panel, the perceptual performance of the artificial network does not significantly depend on feedback history ($p>0.05$ for both features). (**B**) The difference between the monkey and network model is not explained by different extent of training. The upper panel shows that the loss function decreased during time, indicating the gradual learning process of the model on the task. In the lower panel, the black line indicates the network model's difference between perceptual performance following rewarded and unrewarded trials at each point of training. The red line indicates the corresponding difference in the monkeys' perceptual performance, which always lies outside the 90% confidence interval for the model (gray shading). (**C**) In a high dimensional neuronal space expanded by the activity of 7a units during the delay period, we find the best hyperplane to discriminate the task the animal performed on the trial. We define our single-trial neuronal measure of belief strength as the Euclidean distance from 7a population activity on each trial to the hyperplane. (**D**) Belief strength is schematized as the distance from a rolling ball to a boundary. for trials leading up to the animals' decision to switch tasks, the average belief strength decreased monotonically, changed sign right at the point the monkey decided to switch tasks and recovered as the new task-belief was reinforced (histograms in bottom panel). Normalized activity of task-selective 7a units tracked the same dynamics as decoded belief around task switches (lines in bottom panel). Error bars indicate standard errors.

Similarly, we estimated trial-by-trial feature discriminability using V1 population responses in the corresponding feature encoding dimensions (Figure 3A). As expected, when comparing trials using the same stimuli (i.e. same difficulty), trials with larger relevant feature discriminability yield better perceptual performance (Figure S1E-F). For each task-belief and stimulus condition, we look for potential correlation between belief strength measured from area 7a and perceptual discriminability measured in V1 (Figure 3B). Despite the fact that the resulting correlation is based on few trials and only a few dozen neurons across two very weakly connected areas (Markov *et al.*, 2014), there is a positive correlation between belief and the encoding of the feature that is believed to be relevant, but not when the feature is believed to be irrelevant (Figure 3C-D). Together, our results indicate that belief-based decision making is an integrated system rather than

a separable two-stage computation (first the categorical task-belief, then the corresponding perceptual decision).

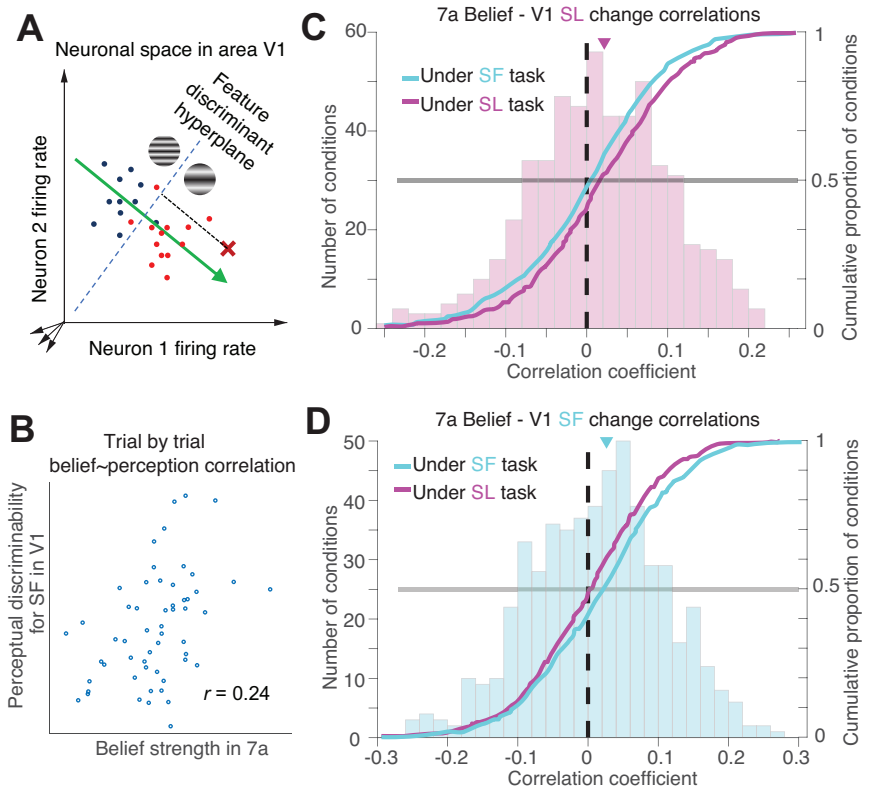

Figure 3. Belief and perception are linked on a trial-by-trial basis. (**A**) Using a procedure similar to that described in Figure 2B, we define the perceptual discriminability of each stimulus feature change on each trial as the Euclidean distance from V1 population activity to the hyperplane that best classifies the stimulus change of that feature (e.g., higher vs. lower spatial frequency). (**B**) Trial-by-trial comparison between belief strength (abscissae, decoded from 7a) and perceptual discriminability (ordinates, decoded from V1) for an example stimulus/task condition. If belief decisions and perceptual decisions are implemented by separate functional modules of the brain, then internal fluctuations of the two systems should have no correlation. (**C**) The belief- spatial location discriminability correlation is positive when spatial location is believed to be relevant (histogram and magenta cumulative distribution curve, $p=4\times10^{-6}$), but not when it is believed to be irrelevant (cyan cumulative distribution curve, $p>0.05$). The two distributions are significantly different (Wilcoxon rank sum test, $p=0.014$). (**D**) Similarly, belief- spatial frequency discriminability is significantly positive when spatial frequency is believed to be relevant

(p=0.0015) but not when it is believed to be irrelevant (p>0.05). The two distributions are significantly different (Wilcoxon rank sum test, p=0.03).

V1 fluctuations affect belief updating

In addition to the trial-by-trial interaction between belief and perception, two pieces of evidence demonstrate that perception on the previous trial affects how task-belief is updated for the upcoming trial. First, trial information beyond reward outcome affects how task-belief will be updated. On average, the monkeys were more likely to switch tasks after they missed rewards on trials with big changes in the stimulus feature believed to be relevant (Figure S4C). This reliance of belief updating on vision is captured by our ideal observer model which optimally updates belief to changes in the environment based on historical reward, stimulus, and choices (see Methods and Figure S4A). The ideal observer model consistently predicts the animals' behavior better than an alternative strategy in which every unrewarded trial affects belief independent of visual and choice experience (Figure S4B). These results demonstrate that, consistent with findings from studies with similar task structure (Purcell and Kiani, 2016a; Sarafyazd and Jazayeri, 2019), confidence in historical choices inform belief updating.

Second, even when trial conditions (stimulus, choice, and reward) were identical, there is a trial-to-trial relationship between fluctuations in the representation of visual stimuli and belief updating. We captured uncontrolled fluctuations in perception by fitting estimates of each feature from V1 (Figure 3A) using logistic regression (Figure 4A, similar to the strategy in (Peixoto *et al.*, 2021)). We used this population neurometric curve to estimate the monkey's confidence about each perceptual choice and related that confidence estimate to the animals' task switching decisions (Figure 4B). The resulting model predicts the animals' task switching decisions better than an alternative model that does not incorporate trial-by-trial variability in V1 (Figure S4E).

Furthermore, if we shuffle V1 responses among trials with identical trial conditions (Figure 4C), the model's switch prediction performance suffers significantly (Figure 4D-E). This difference likely reflects confidence fluctuations in past visual discrimination, since for identical trial conditions, the monkeys were more likely to switch tasks after they missed rewards on trials with larger relevant feature discriminability estimated from V1 (Figure S4D). Together, these results demonstrate that trial-to-trial fluctuations in perception affect belief updating on the subsequent trial, even though these fluctuations provide no benefit for estimating the relevant feature.

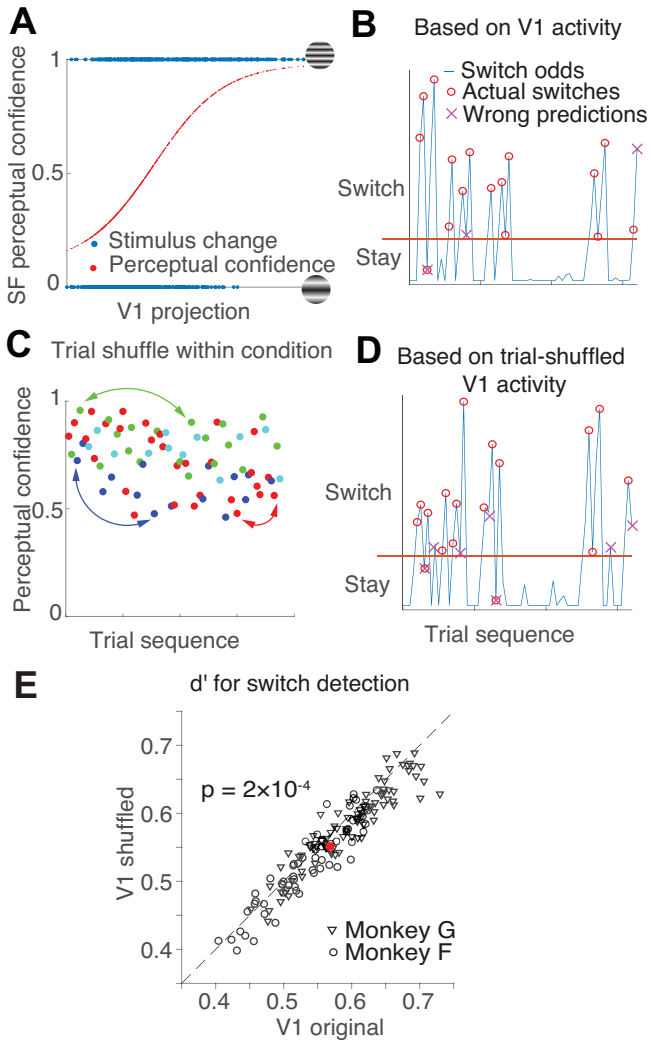

Figure **4. Trial to trial variability in visual cortex affects belief updating.** (**A**) Example neurometric curve showing the ability of a decoder to discriminate spatial frequency changes

from the population of recorded V1 neurons using logistic regression on the perceptual discriminability of spatial frequency (as in Figure 3A). (**B**) Based on perceptual confidence on each trial (estimated from V1 population activity), a normative model determines whether the subject should switch tasks given the trial history (see STAR Methods). (**C**) Based on the V1 projections to the relevant feature subspace on each trial, we estimate from the neurometric curve, which represents the probability the monkey's behavioral choice is correct (i.e. perceptual confidence, (Hangya, Sanders and Kepecs, 2016)). In the trial-shuffle analysis, we randomly switch the confidence within trials with the same conditions (dots with same color). (**D**) Model predictions after trial-shuffle, conventions as in (b). (**e**) Trial-to-trial variability in V1 is related to belief. The model's ability to predict whether the monkey would switch tasks is better using the actual than trial-shuffled V1 activity ($p=2\times10^{-4}$). Each data point here represents an experimental session, and its coordinates show the sensitivity index (*d'*) of switch prediction for the model based on original (x-axis) or trial-shuffled V1 activity (y-axis).

Discussion

Our findings demonstrate that there is no such thing as a standalone perceptual decision-making process: every aspect of perceptual decision-making is profoundly integrated with the dynamic belief states that dominate natural behavior. The relationship between task-belief and perception is bidirectional.

First, we demonstrated that fluctuating task-beliefs affect decisions. Using a combination of multi-neuron, multi-area physiology, complex but controlled behavior, and hypothesis-driven dimensionality reduction, we demonstrated that perception and task-belief are intimately intertwined such that weak task-beliefs are associated with poor perception of task-relevant information. This suggests that fluctuation in belief strength, instead of reflecting a homogenous process such as arousal, has specific effects on the believed relevant information only. This task-belief specific modulation may be a critical mechanism that underlie the process that limits reinforcement learning to the relevant feature (Niv *et al.*, 2015). The specific plasticity mechanism underlying such flexible modulation remains to be revealed, but it may involve a systematic change in the correlation structure of V1 caused by feedback from higher brain regions such as 7a (Bondy,

Haefner and Cumming, 2018). On the other hand, it will also be interesting to determine whether fluctuations in other types of belief (e.g. those reviewed in (Ma and Jazayeri, 2014)) interact with decision making in different ways.

Second, we demonstrated that fluctuations in perceptual-decision related neuronal activity affect belief updating. This aspect has largely been missed by studies of cognitive flexibility that use unambiguous stimuli(Bartolo and Averbeck, 2020; Botvinick and Braver, 2015). Using perceptually challenging tasks not only is more realistic, but also offers powerful tools to investigate, manipulate, and quantitatively understand the specific neuronal process governing perception and belief updating. Our study demonstrates that incorporating these more natural behavioral features into well-controlled laboratory studies leads to important new insights. In this regard, our findings are merely a starting point of a wide range of potential scientific explorations that will shed light on cognitive flexibility, belief-based decision-making, and learning.

The idea that belief-based decision-making relies on the inextricable link between dynamic task-beliefs and perception showcased that natural and complex behavior may not be understood as stacked building blocks of simpler functions. It also opens up exciting avenues for translational therapies that address deficits in flexible decision-making associated with neuropsychiatric disorders. For instance, our results imply that cognitive flexibility is mediated by interactions between neural populations responsible for perception and belief. As such, therapies that affect communication between brain areas (e.g. by affecting neurotransmitters like dopamine(Botvinick and Braver, 2015; Mueller *et al.*, 2017)) have the potential to improve cognitive flexibility in health and disease. Indeed, stimulants that affect the dopamine system like methylphenidate or amphetamines can change focus and flexibility(Bagot and Kaminer, 2014; Mueller *et al.*, 2017). Going forward, studying the highly integrated belief-based decision-making system will open up

doors to potential treatments of conditions that affect cognitive flexibility and even solutions for healthy individuals to become better decision-makers in volatile environments.

Acknowledgments: We are grateful to Karen McKracken for providing technical assistance and to Josh Gold, Lori Holt, Douglas Ruff, Amy Ni, Ramanujan Srinath, and Morteza Sarafyazd for comments on an earlier version of this manuscript. This work was supported by the Simons Foundation (Simons Collaboration on the Global Brain award 542961SPI to MRC), the National Eye Institute of the National Institutes of Health (awards R01EY022930 and R01NS121913 to MRC), and the McKnight Foundation (McKnight Scholar award to MRC).

Author contributions: C.X., and M.R.C. designed research; C.X., and L.E.K. performed research; C.X. analyzed data; M.R.C. supervised the findings of this work; C.X., and M.R.C. wrote the paper.

Competing interests: The authors declare no competing interests.

Data and Code availability: Data and code are available upon reasonable request.

Additional information: Correspondence and requests for materials should be addressed to Cheng Xue (chx31@pitt.edu) or Marlene R. Cohen (cohenm@pitt.edu)

Materials and Methods

Experimental subjects

The subjects in our study were two adult male rhesus monkeys (*Macaca mulatta*, monkey F weighed 12 kilograms, monkey G weighed 9 kilograms), who were both experimentally naïve prior to the current experiments. All animal procedures were approved by the Institutional Animal Care and Use Committees of the University of Pittsburgh and Carnegie Mellon University. After we implanted each animal with a titanium head post, they were trained to perform two-interval, two-feature discrimination with stochastic rule switching (Figure 1A) (monkey F was trained for 11 months, monkey G for 9 months). We made sure the animals understood the essential requirements of the task based on their behavior (Figure 1C-D), before implanting each animal with 6×8 microelectrode arrays (Blackrock Microsystems) in both parietal cortical area 7a and visual cortical area V1. Each array was connected to a percutaneous connector that allowed daily electrophysiological recordings. The distance between adjacent electrodes was 400 µm, and each electrode was 1 mm long. We implanted 7a arrays on the crown of gyrus between intraparietal sulcus

and superior temporal sulcus at approximately 11mm lateral to the midline; and V1 arrays posterior to the lunate sulcus at approximately -13 mm to the intra-aural line and 7 mm lateral to the midline (Figure 1B).

Behavioral task

To study perceptual decision making under evolving task-beliefs in dynamic environment, we trained the animals to perform a two-interval, two-feature discrimination task with stochastic task switching. A trial began when the subjects fixated their gaze on a central dot on the screen and they were required to maintain fixation as long as the dot remained on the screen, or the trial would be aborted and unrewarded. A Gabor stimulus was then displayed for 200 ms. After a random delay (300ms to 500ms), a second Gabor stimulus was displayed for 200 ms with a slightly different spatial location (shifted left or right) and a slightly different spatial frequency (higher or lower), with independently randomized change amounts in the two features. The ranges of change amounts are titrated at the beginning of each session so that the overall perceptual performances of the spatial location task and spatial location task are both approximately 75%. Following a subsequent delay of 150ms, the fixation dot disappeared, and the animals were required to make a saccade to one out of four peripheral targets to indicate both the inferred relevant feature and the direction of change in that feature. The two cyan targets correspond to the increase and decrease of spatial frequency when it was believed to be the relevant feature, while the two magenta targets correspond to the left-shift and right-shift of spatial location. The location of the array of targets varied across experimental sessions but remained the same within each session. The monkeys were rewarded only if they correctly reported the direction of change in the relevant feature. The visual stimuli throughout a trial contain no information about the behavioral relevance of features. The relevant feature switches on a randomly chosen 2.5% of trials. The monkeys therefore needed to infer the relevant feature based on their choice and reward history.

Electrophysiological recording

All visual stimuli were displayed on a linearized CRT monitor (1,024 × 768 pixels, 120-Hz refresh rate) placed 57 cm from the animal. We monitored eye position using an infrared eye tracker (Eyelink 1000, SR Research) and used custom software (written in Matlab using the Psychophysics Toolbox(Brainard, 1997) to present stimuli and monitor behavior. We recorded eye position and pupil diameter (1,000 samples per s), neuronal responses (30,000 samples per s) and the signal from a photodiode to align neuronal responses to stimulus presentation times (30,000 samples per s) using hardware from Ripple.

We recorded neuronal activity from Utah arrays during daily experimental sessions for several months in each animal (89 sessions from monkey F and 68 sessions from monkey G). We set the threshold for each channel at three times the standard deviation and used threshold crossings as the activity on that unit. We positioned the stimuli to maximize the overlap between potential stimulus locations and the joint receptive fields of V1 units, as determined using separate data collected while the monkeys fixated and Gabor stimuli were flashed across a range of retinal positions. The stimulus locations did not overlap with the receptive field of any 7a unit, and we confirmed that our 7a units did not have stimulus feature selective responses (Figure S1B).

We included experimental sessions if they contained at least 480 completed trials (where monkeys successfully maintained fixation until they indicated their choice). We analyzed the activity of area 7a units

during the first 300ms after the offset of the first stimulus, when there is no Gabor stimulus on the screen; and the activity of area V1 units during stimulus display periods, shifted with 34 ms visual latency. Units from area 7a were included if their average activity during the delay period was at least 5 sp/s. Units from V1 were included in the analyses if their average stimulus response was 1) at least 25% larger than baseline activity, measured 100ms before stimulus onset, and 2) larger than 5 sp/s. These procedures resulted in 89 sessions from Monkey F and 68 sessions from Monkey G; average 53 7a units, 46 V1 units, and 1053 completed trials per session.

Recurrent Neural Network

We trained recurrent neural networks to perform the two-interval, two-feature discrimination task with stochastic task switching. Similar to the monkeys, the network model was trained to infer the relevant feature and make the correct perceptual discrimination based on the implicit task rule. The training and testing of the neural network were implemented with custom code based on the open source package PsychRNN (Ehrlich *et al.*, 2021) and Tensorflow (Abadi *et al.*, 2016). The model consists of 100 units with all-to-all recurrent connections, without any a priori constraints on the connection weights. All results are qualitatively consistent with network sizes of 50 or 200 units.

To produce a choice on each iteration, the recurrent neural network receives input from four channels (Figure S3A): the spatial location change, the spatial frequency change, the recent choice history of the model, and the reward feedback associated with those choices. The recurrent neural network is connected to four output units, which correspond to the four saccade targets presented to the monkey. We take the output unit with the highest activity during choice period as the behavioral choice of the model. Gaussian noises were added to the stimulus changes so that the overall perceptual performances of the model would match those of the monkey (Figure S3B). We trained the model using custom code modified from PsychRNN that allowed the model's previous choice output to be part of the trial history input for the following trials. The past trial history information was fed into the network in chronological order: each trial starts with the stimulus changes followed by (in order) the choice the model made on the previous iteration, the reward feedback, and finally an intertrial interval before processing information from the next trial. For the model results shown in Figure 2 and Figure S3, the neural network considered a trial history of 7 trials before the current trial, which is longer than most of the monkeys' task-switch delays. The network behaviors shown in Figure 2 and Supplementary Figure S3 are qualitatively similar if the history input contained 4 trials or 10 trials before the current trial.

The loss function is defined as the squared error between the network output and the target output during the choice period. The target output is 1 on the output unit of the correct choice, and 0 for the other three output units. We initialized the weight matrix with random connections and used a learning rate of 0.001 during training.

Population analyses

To obtain a continuous neuronal measure of the animals' belief state, we analyzed the activity of the population of 7a neurons during the delay period in a high dimensional space in which the activity of each unit was one dimension. We used linear discriminant analysis to identify the best hyperplane to discriminate between 7a population activity on trials where monkeys chose spatial location targets from trials when they chose spatial frequency targets. We defined the belief strength on each trial as the Euclidean distance from the 7a population response to the discriminant hyperplane. Similarly, we obtained a continuous neuronal measure of the discriminability of stimulus change using V1 activity.

Statistical tests

All p-values reported in this study are from Wilcoxon signed rank test unless otherwise specified.

Normative behavioral model

We use a normative model to characterize belief updating of an ideal observer, given the trial history and the perceptual ability of the monkey (Sarafyazd and Jazayeri, 2019; Purcell and Kiani, 2016a; Glaze *et al.*, 2018). Based on the monkeys' psychometric curve in an experiment session and the change amount of the chosen feature in each trial, we estimated the trial-by-trial probability that their perceptual choice was incorrect. For a non-reward trialed, the odds of likelihoods that the actual task is different from the monkeys' subjective belief is given by

$$\frac{p(diff|nr,\theta,c)}{p(same|nr,\theta,c)} = \frac{p(nr,\theta,c|diff)\cdot p(diff)}{p(nr,\theta,c)} \cdot \frac{p(nr,\theta,c)}{p(nr,\theta,c|same)\cdot p(same)} = \frac{p(nr,\theta,c|diff)}{p(nr,\theta,c|same)}$$

where θ and c refer to the stimulus change amount and perceptual choice in the feature the monkeys believed to be relevant; and nr refers to a non-reward trial outcome. We assumed that overall, the monkeys experienced an equal number of trials where the feature was the same or different from their current belief (i.e., $p(diff) = p(same)$). The monkeys were never rewarded on trials when their subjective task-belief was different from the actual task rule, so $p(nr,\theta,c|diff) = 1$. Meanwhile when subjective task-belief is consistent with the actual rule, the probability of perceptual error can be simply derived from the psychometric function associated with that choice:

$$\frac{p(diff|nr,\theta,c)}{p(same|nr,\theta,c)} = \frac{1}{1-p(\theta|c)}$$

where $p(\theta|c)$ is the psychometric function associated with the perceptual choice c (Figure 1C). In Figure 4, the psychometric function is replaced with the neurometric function (Figure 4A, also see *Population analyses* section).

For n consecutive non-reward trials, the likelihood ratio grows larger as perceptual evidence for a task switch grows as

$$\frac{1-\mathcal{L}_n^{same}}{\mathcal{L}_n^{same}} = \prod_{i=1}^{n} \frac{1}{1-p(\theta_i|c_i)}$$

where $\mathcal{L}_n$ is the likelihood that the task has not changed after n consecutive non-rewarded trials (examples in Figure S4A left panel). Aside from perceptual evidence, the observer presumably also has prior knowledge about the volatility of the task environment. After n consecutive non-reward trials, the prior probability that the task stays the same with the last rewarded trial is

$$Pr_n^{same} = Pr_{n-1}^{same}\cdot(1-h) + (1-Pr_{n-1}^{same})\cdot h$$

where h represents the hazard rate of task change at each trial (for an ideal observer, $h = 0.025$, see *behavioral task* section), with

$$Pr_0^{same} = 1-h$$

Examples of Pr_n^{same} under different environment volatility are shown in Figure S4A middle panel. Taking both perceptual evidence and prior knowledge of environment volatility into account, the model shows

that an ideal observer should switch tasks if the posterior probability is higher for actual task rule being different from the subjective task-belief than when they are the same (Figure S4A right panel), i.e.:

$$switch\ odds = \log((1 - Pr_n^{same}) \cdot (1 - \mathcal{L}_n^{same})) - \log(Pr_n^{same} \cdot \mathcal{L}_n^{same}).$$

$when\ positive,\ switch\ task$

References:

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J. and Devin, M. (2016) 'Tensorflow: Large-scale machine learning on heterogeneous distributed systems', *arXiv preprint arXiv:1603.04467*.

Alport, A., Styles, E. A. and Hsieh, S. (1994) '17 Shifting Intentional Set: Exploring the Dynamic Control of Tasks'.

Bagot, K. S. and Kaminer, Y. (2014) 'Efficacy of stimulants for cognitive enhancement in non-attention deficit hyperactivity disorder youth: a systematic review', *Addiction,* 109(4), pp. 547-57.

Bartolo, R. and Averbeck, B. B. (2020) 'Prefrontal Cortex Predicts State Switches during Reversal Learning', *Neuron,* 106(6), pp. 1044-1054.e4.

Bondy, A. G., Haefner, R. M. and Cumming, B. G. (2018) 'Feedback determines the structure of correlated variability in primary visual cortex', *Nat Neurosci,* 21(4), pp. 598-606.

Botvinick, M. and Braver, T. (2015) 'Motivation and cognitive control: from behavior to neural mechanism', *Annu Rev Psychol,* 66, pp. 83-113.

Brady, K. T., Gray, K. M. and Tolliver, B. K. (2011) 'Cognitive enhancers in the treatment of substance use disorders: clinical evidence', *Pharmacol Biochem Behav,* 99(2), pp. 285-94.

Brainard, D. H. (1997) 'The Psychophysics Toolbox.', *Spatial vision,* 10(4), pp. 433-6.

Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D. and Miller, E. K. (2012) 'Synchronous oscillatory neural ensembles for rules in the prefrontal cortex', *Neuron,* 76(4), pp. 838-846.

Cohen, M. R. and Maunsell, J. H. (2011a) 'Using neuronal populations to study the mechanisms underlying spatial and feature attention', *Neuron,* 70(6), pp. 1192-204.

Cohen, M. R. and Maunsell, J. H. (2011b) 'When attention wanders: how uncontrolled fluctuations in attention affect performance', *J Neurosci,* 31(44), pp. 15802-6.

Dickstein, D. P., Nelson, E. E., McCLURE, E. B., Grimley, M. E., Knopf, L., Brotman, M. A., Rich, B. A., Pine, D. S. and Leibenluft, E. (2007) 'Cognitive flexibility in phenotypes of pediatric bipolar disorder', *J Am Acad Child Adolesc Psychiatry,* 46(3), pp. 341-355.

Ebitz, R. B., Tu, J. C. and Hayden, B. Y. (2020) 'Rules warp feature encoding in decision-making circuits', *PLoS Biol,* 18(11), pp. e3000951.

Ehrlich, D. B., Stone, J. T., Brandfonbrener, D., Atanasov, A. and Murray, J. D. (2021) 'PsychRNN: An Accessible and Flexible Python Package for Training Recurrent Neural Network Models on Cognitive Tasks', *eNeuro,* 8(1).

Glaze, C. M., Filipowicz, A. L., Kable, J. W., Balasubramanian, V. and Gold, J. I. (2018) 'A bias–variance trade-off governs individual differences in on-line learning in an unpredictable environment', *Nature Human Behaviour,* 2(3), pp. 213-224.

Gold, J. I. and Shadlen, M. N. (2007) 'The neural basis of decision making', *Annu Rev Neurosci,* 30, pp. 535-74.

Hangya, B., Sanders, J. I. and Kepecs, A. (2016) 'A Mathematical Framework for Statistical Decision Confidence', *Neural Comput,* 28(9), pp. 1840-58.

Kable, J. W. and Glimcher, P. W. (2009) 'The neurobiology of decision: consensus and controversy', *Neuron,* 63(6), pp. 733-45.
Kamigaki, T., Fukushima, T. and Miyashita, Y. (2009) 'Cognitive set reconfiguration signaled by macaque posterior parietal neurons', *Neuron,* 61(6), pp. 941-51.
Ma, W. J. and Jazayeri, M. (2014) 'Neural coding of uncertainty and probability', *Annu Rev Neurosci,* 37, pp. 205-20.
Mante, V., Sussillo, D., Shenoy, K. V. and Newsome, W. T. (2013) 'Context-dependent computation by recurrent dynamics in prefrontal cortex', *Nature,* 503(7474), pp. 78-84.
Markov, N. T., Ercsey-Ravasz, M. M., Ribeiro Gomes, A. R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M. A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D. C. and Kennedy, H. (2014) 'A weighted and directed interareal connectivity matrix for macaque cerebral cortex', *Cereb Cortex,* 24(1), pp. 17-36.
Monsell, S. (2003) 'Task switching', *Trends Cogn Sci,* 7(3), pp. 134-140.
Mueller, A., Hong, D. S., Shepard, S. and Moore, T. (2017) 'Linking ADHD to the Neural Circuitry of Attention', *Trends Cogn Sci,* 21(6), pp. 474-488.
Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A. and Wilson, R. C. (2015) 'Reinforcement learning in multidimensional environments relies on attention mechanisms', *J Neurosci,* 35(21), pp. 8145-57.
Peixoto, D., Verhein, J. R., Kiani, R., Kao, J. C., Nuyujukian, P., Chandrasekaran, C., Brown, J., Fong, S., Ryu, S. I., Shenoy, K. V. and Newsome, W. T. (2021) 'Decoding and perturbing decision states in real time', *Nature,* 591(7851), pp. 604-609.
Purcell, B. A. and Kiani, R. (2016a) 'Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy', *Proc Natl Acad Sci U S A,* 113(31), pp. E4531-40.
Purcell, B. A. and Kiani, R. (2016b) 'Neural Mechanisms of Post-error Adjustments of Decision Policy in Parietal Cortex', *Neuron,* 89(3), pp. 658-71.
Sarafyazd, M. and Jazayeri, M. (2019) 'Hierarchical reasoning by neural circuits in the frontal cortex', *Science,* 364(6441).
Stoet, G. and Snyder, L. H. (2004) 'Single neurons in posterior parietal cortex of monkeys encode cognitive set', *Neuron,* 42(6), pp. 1003-12.
Stoet, G. and Snyder, L. H. (2009) 'Neural correlates of executive control functions in the monkey', *Trends Cogn Sci,* 13(5), pp. 228-34.
Thapar, A., Martin, J., Mick, E., Arias Vásquez, A., Langley, K., Scherer, S. W., Schachar, R., Crosbie, J., Williams, N., Franke, B., Elia, J., Glessner, J., Hakonarson, H., Owen, M. J., Faraone, S. V., O'Donovan, M. C. and Holmans, P. (2016) 'Psychiatric gene discoveries shape evidence on ADHD's biology', *Mol Psychiatry,* 21(9), pp. 1202-7.
Uchida, N., Kepecs, A. and Mainen, Z. F. (2006) 'Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making', *Nat Rev Neurosci,* 7(6), pp. 485-91.

Supplemental information:

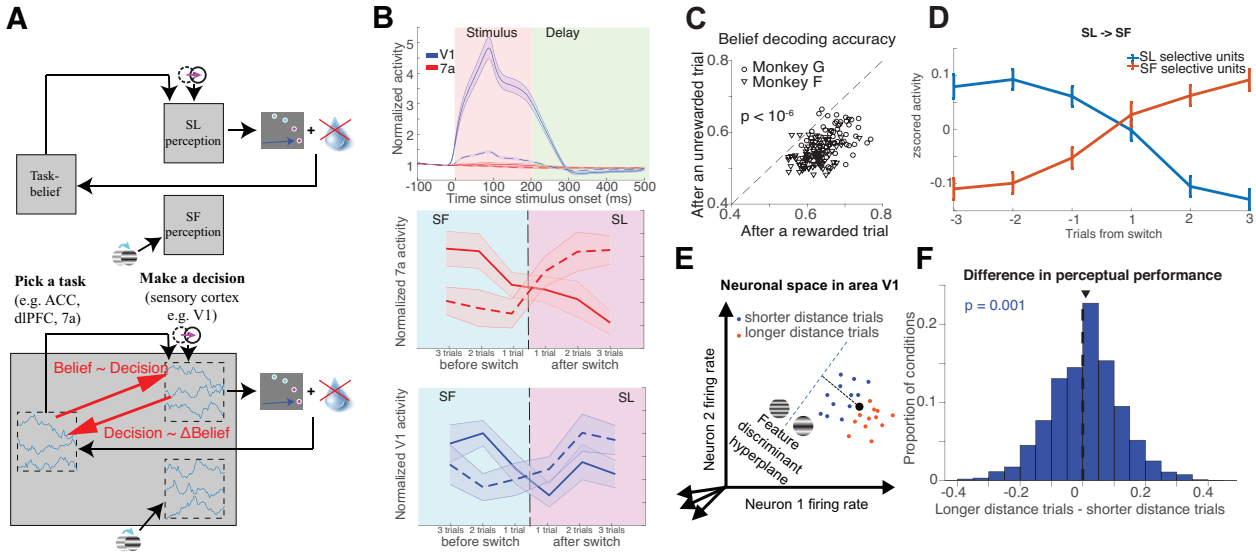

Figure S1. Hypotheses and neuronal measures of interaction between task-belief and decision-making. (**A**) Illustrative summary of the hypotheses on belief-based decision-making. Take the task in Figure 1A as example, intuitive view of natural decision-making consists of two separately studied steps: 1) flexibly shift among tasks in response to evolving environment contingencies, or task-switching; 2) based on the believed task rule and sensory input, choose the correct behavioral response (perceptual decision-making). These two hierarchical steps are assumed to have independent biological mechanisms mediated by different areas in the brain. However, we found that 1) the supposedly internal fluctuations of task-belief strength covaries with the fluctuations in the perception of the believed relevant feature (upper red arrow), and 2) that for apparently identical experiences, the uncontrolled fluctuations in perception influence how task-belief is adjusted next (lower red arrow). These results show that task-switching and perceptual decision-making are two cognitive processes inextricably linked in the brain. (B) Single unit selectivity for stimuli and task. Upper panel, normalized peri-stimulus time histograms (PSTHs) of the most spatially selective V1 unit (blue) and 7a unit (red) from each session. Solid lines show responses to the preferred spatial location, and dashed lines show responses to the non-preferred spatial location. Shaded area indicate s.e.m. Lower two panels show PSTHs of the most task selective unit from area 7a (**b**) and V1 (**c**) for each session around the monkey's decision to switch tasks. The solid lines represent SF-selective units, and the dashed lines represent SL-selective units. (C) Task-belief decoded from a neuronal population. Task-belief can be better classified from area 7a population activity following rewarded than unrewarded trials. The abscissa and ordinate of each point show the performance of a linear classifier following rewarded and unrewarded trials, respectively. Each point represents one experimental session. (**b**) For trials leading up to the animals' decision to switch from spatial location task to spatial frequency task, the normalized activity of spatial location task-selective neurons decreases while those of spatial frequency task-selective neurons increases. Error bars indicate standard error. (D) Fluctuations in V1 reflect fluctuations in perceptual accuracy. Left panel: we divide trials from the same stimulus condition according to the neuronal distance between the V1 population response to the feature discriminant hyperplane. Right panel: distribution of differences in perceptual performance between longer and shorter distance trials across conditions. The overall positive mean of the distribution shows that for the same stimulus difficulty, trials with better relevant-feature discriminability have higher perceptual performance.

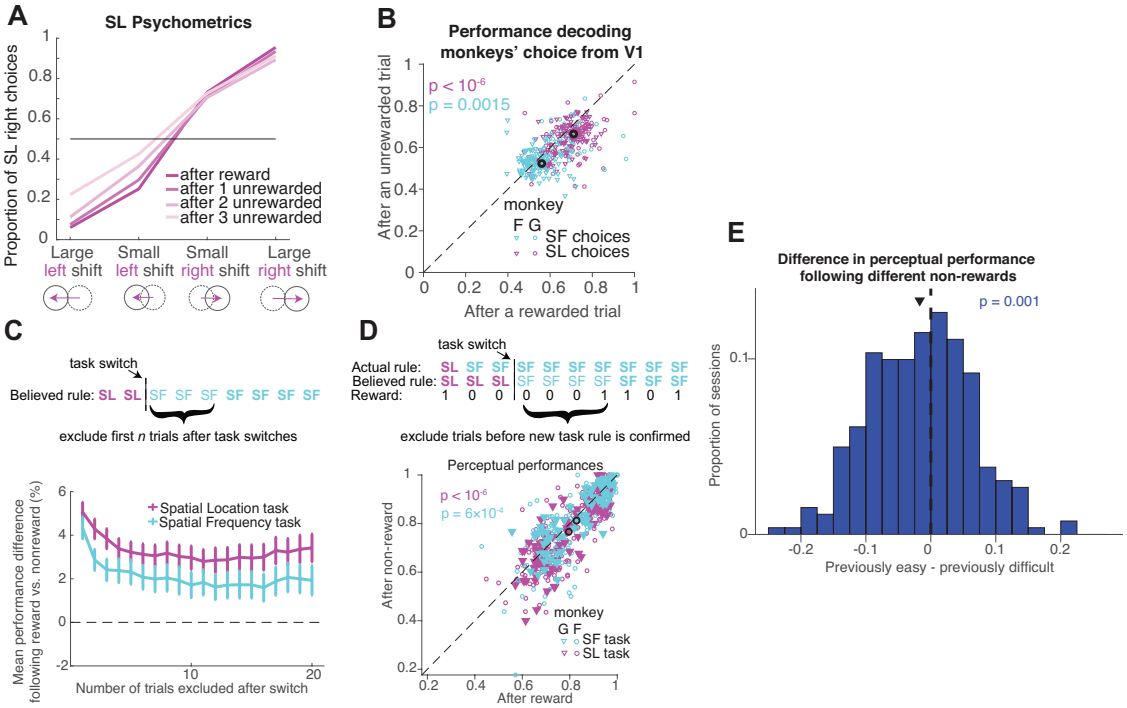

Figure S2. Task uncertainty is associated with worse perceptual decisions, independent of task-set inertia or task switching costs. (**A**) Psychometric curves for trials in which the monkey chose to perform a spatial location discrimination in one example session as a function of reward history. The psychometric curves (proportion rightward choices as a function of size and sign of the shift) are flatter (less accurate) with increasing consecutively-unrewarded trials, which correspond to more task uncertainty. (**B**) The abscissa and ordinate of each point represents performance decoding the animals' perceptual choices following rewarded and unrewarded trials in each experimental session. The animals' perceptual choices are significantly better decodable from V1 activity when task belief is certain (after reward) than when it is uncertain (after missing a reward). (**C**) If the perceptual performance decrement following non-rewards (as in Figure 2A) were caused by task-set inertia or task switching costs, the effect would disappear when subjects were fully adapted or committed to a task. Cyan and magenta curves show the difference between perceptual performance following rewards and non-rewards as a function of the number of trials after each task switch that were excluded. Error bars indicate s.e.m. As the subject adapts to a chosen task, the performance difference asymptotes at a significant non-zero value, instead of disappearing. (**D**) The perceptual performance decrement persists even when we exclude trials following a task switch but before the first reward after the task switch (which confirms the subjects' choice of task), leaving trials where the subjects are committed to a task. Perceptual performance following a reward (abscissae) remains significantly higher than following a non-reward (ordinates). (E) Perceptual performance is worse following obvious but unrewarded changes in a visual feature. Unrewarded trials with larger believed-relevant feature changes indicate more strongly that the current task-belief is wrong, thus predicting an overall weaker task-belief in the next trial. Correspondingly, we see significantly worse perceptual performance following perceptually-easy yet unrewarded trials (overall negative mean of the distribution).

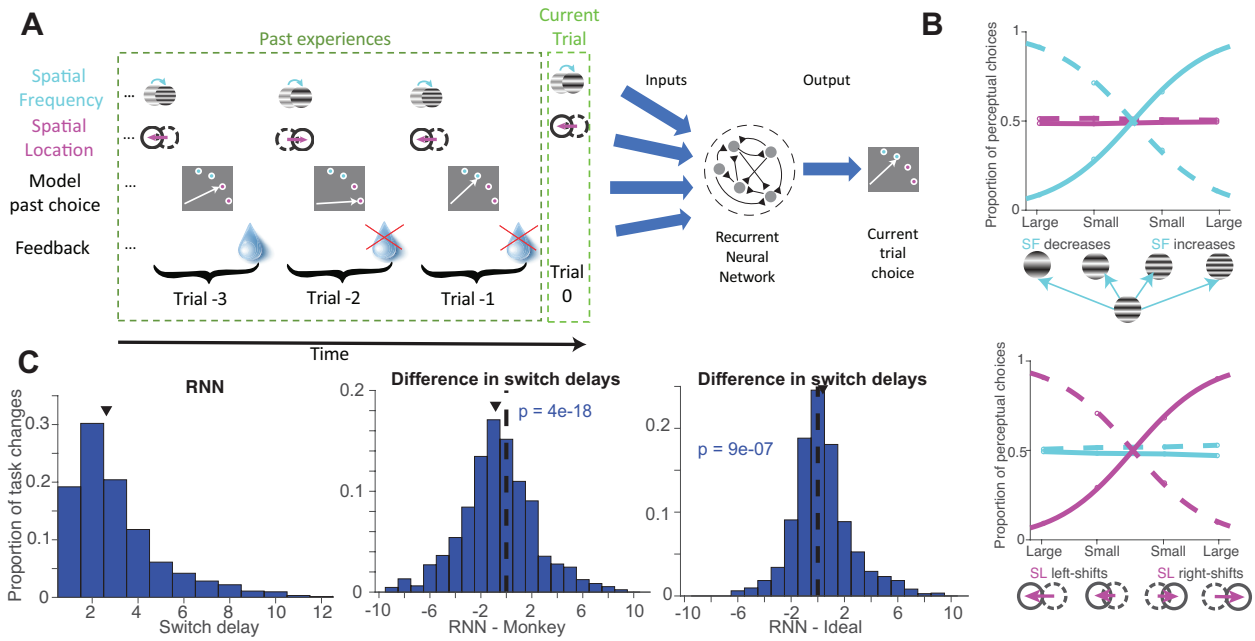

Figure S3. Recurrent neural network trained to perform two-feature discrimination task with stochastic rule-switching. (A) Schematic of the inputs and outputs of the network. The network receives sequential inputs containing information about stimulus changes, past model choices, and choice feedback (reward history) over the course of the last several trials. Like the monkey, the network model is trained to infer the implicit task rule from recent history and make corresponding decisions. **(B)** Similar to the monkeys' behavior in Figure 1C, the choices of a trained network are informed by the believed relevant feature information, while independent of the believed irrelevant feature information. Conventions as in Figure 1C. **(C)** Like the monkeys, the network learned to adapt to changes in task demands. Left panel shows the distribution of number of trials it took the network to adapt to task changes across experimental sessions. Middle panel shows the distribution of number of trials the network took to adapt to the task change relative to the monkey. Overall, the network is more efficient than the monkey at task-switching. Right panel shows the distribution of number of trials the network took to adapt to the task change relative to the ideal observer model (see Star Methods). Overall, the network's task-switching behavior is close to but slightly trails behind the ideal observer model's prediction.

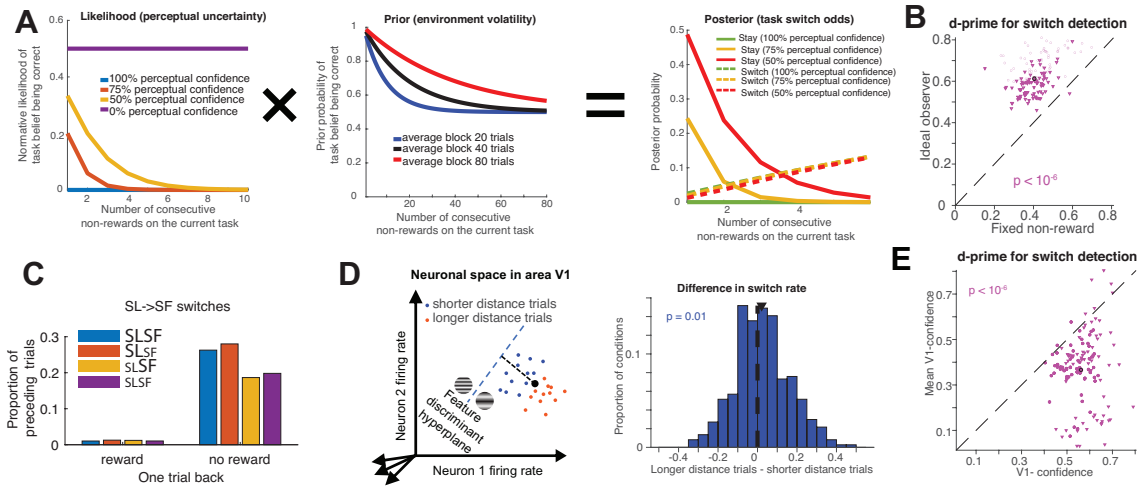

Figure S4. . **Fluctuations in perception informs task-belief update in the future.** (**A**) A normative model predicts task switching behavior**.** Left panel: the likelihood (y axis), which is defined as the probability that the task has remained the same, decreases with the number of consecutive unrewarded trials (x axis) for each perceptual certainty (colored lines). The rate of decay is slower for lower perceptual confidence. Middle Panel: the prior probability (y axis) describes, according to the monkey's prior knowledge of the environment volatility, the probability that two trials that are n trials apart (x axis) have the same task rule. Two trials are more likely to have the same task when the average block is longer. The average block length in the current dataset is 40. Right Panel: Combining the likelihood and prior, the model predicts a task switch when the posterior probability of switch becomes larger than that of stay. (**B**) An alternative to the ideal observer model (y-axis) is one in which the observer disregards all trial information and switches tasks after a fixed number of consecutive non-rewards (fixed non-reward model, x-axis). For almost all the sessions, the ideal observer model produce higher switch-predicting sensitivity (d') than the fixed non-reward model. (**C**) Proportion of each trial type that precedes a switch from the spatial location task to the spatial frequency task. Big fonts in the legend indicate large changes in the corresponding visual feature. Monkeys are more likely to switch from SL to SF tasks when there is a big SL change and yet they did not get a reward. Switches from SF to SL are qualitatively similar. **(D)** Noise fluctuations in V1 affects task-belief update. Left panel: we divide trials from the same condition (stimuli, choice, feedback) according to the neuronal distance between the V1 population response to the feature discriminant hyperplane. Right panel: Distribution of differences in switch rate on the trial following longer or shorter distance unrewarded trials across conditions, where trials in each condition have identical stimuli, chosen target, and reward (all unrewarded). The overall positive distribution shows that for the same stimuli and choice, unrewarded trials with better relevant-feature discriminability are more likely to be followed by a task switch. This is expected if fluctuations in V1 reflect fluctuations in perceptual confidence. (E) A model that incorporates V1 response variability can better explain switching behavior. The model using the original V1 population response each trial (V1-confidence, x-axis) consistently outperforms the model using the averaged V1 population response for all trial within the same condition (same stimulus, choice, and reward outcome; mean V1-confidence, y-axis) for monkey G (circles) and monkey F (triangles). The abscissa and ordinate represent the sensitivity (d') of switch predictions using the two models for each experimental session.