

Neuronal stability in medial frontal cortex sets individual variability in decision-making

Tomoki Kurikawa¹, Tatsuya Haga^{ID 1}, Takashi Handa^{ID 2}, Rie Harukuni¹ and Tomoki Fukai^{ID 1*}

In the brain, decision making is instantiated in dedicated neural circuits. However, there is considerable individual variability in decision-making behavior, particularly under uncertainty. The origins of decision variability within these conserved neural circuits are not known. Here we demonstrate in the rat medial frontal cortex (MFC) that individual variability is a consequence of altered stability in neuronal populations. In a sensory-guided choice task, rats trained on familiar stimuli were exposed to unfamiliar stimuli, resulting in variable choice responses across individuals. We created a recurrent network model to examine the source of variability in MFC neurons, and found that the landscape of neural population trajectories explained choice variability across different unfamiliar stimuli. We experimentally confirmed model predictions showing that trial-by-trial variability in neuronal activity indexes the landscape and predicts individual variation. These results show that neural stability is a critical component of the MFC neural dynamics that underpins individual variation in decision-making.

Animals must respond to sensory input from the environment and make adaptive decisions. These responses can vary greatly depending on whether the sensory information is ambiguous or unfamiliar. When animals are familiar with a given sensory stimulus, their behavioral responses become stereotyped. However, when sensory input is unfamiliar, behavioral responses become more variable and develop a highly probabilistic nature. With ambiguous information, decision-making can show a wide spectrum of individual differences depending on a wide array of factors, including the intrinsic behavioral traits of each subject and extrinsic modifiers that can bias choice.

Many studies have explored the behavioral characteristics and neural mechanisms of probabilistic decision-making^{1–3}, including the role of choice-specific firing sequences within dedicated neural populations^{4–6}. In the rat medial frontal cortex (MFC), choice-specific neural trajectories have been recorded during the performance of an alternative choice task in response to auditory stimuli⁷. In this task, animals were trained with familiar tone stimuli, and then were required to respond to unfamiliar tone stimuli. After training, MFC neurons formed choice-specific trajectories for familiar stimuli. However, with unfamiliar stimuli, the probabilistic choice responses showed large individual differences in psychometric curves, presumably reflecting sensitivity to environmental changes in choice behavior.

These findings motivated us to investigate the neural mechanisms linking the formation of choice-specific sequences and the expression of individual differences. Cortical mechanisms for general tendencies in decision behavior have been extensively studied^{2,3,8}, but the neural origin of individual behavioral differences is poorly understood. More broadly, the origins of variability of neural dynamics in brain circuits and its relationship to behavioral variability across animals are a largely unexplored area of brain function. Here we examine choice-specific sequences in the MFC, which engages in goal-directed decision-making tasks^{9,10}, to explore experimentally and computationally the neural basis of individual variation and its underlying computational principles.

We first addressed how neural trajectories emerge for familiar stimuli and guide decision-making for unfamiliar stimuli. Then, we

asked whether and how the properties of neural trajectories influence choice responses and the resultant psychometric curves by constructing a recurrent neural network model^{11–14}. On the basis of experimental observations from the rat MFC, we trained the model with a reinforcement learning algorithm for an association task^{15–17} to generate a spectrum of decision behaviors that could be compatible with the observed behavioral variability across animals. In our analysis, we found that the stability of intrinsic neural trajectories to perturbation strongly predicted sensitivity in choice behavior in response to different unfamiliar stimuli. This stability is masked in neural trajectories driven by sensory stimuli, and hence is a covert property of internal neural dynamics. The model further predicted that trial-by-trial variability in trajectories reflected neural stability and is correlated with behavioral variability, and we confirmed this experimentally. Our results suggest that the covert stability of neural dynamics corroborates the trial-by-trial variability in the MFC that is linked to individual differences in choice responses under sensory uncertainty.

Results

Behavioral variability in decision-making with unfamiliar sensory cues. Our hypothesis is that individual differences in ambiguous choice responses emerge partly from neural dynamics formed in the MFC through previous experiences. We examined this hypothesis in a sensory-guided alternative choice task. Rats were trained to make either 'left' or 'right' licking in response to high-frequency (13 kHz) or low-frequency (10 kHz) auditory cues (called familiar cues), respectively (Fig. 1a). Each cue was presented for 200 ms, and correct choices after the cue presentation were rewarded. Among 36 trained rats, 21 rats reached a criterion correct rate (75%), and multineuron recordings were made from the MFC of these rats. Each rat underwent one or two days of subsequent recording sessions, during which the rats were exposed to unfamiliar cue tones (10.5 to 12.5 kHz) besides the familiar tones in 20% of trials. Because these tones were unfamiliar to the rats, their choice responses were ambiguous and probabilistic. Correct choices for familiar cues were always rewarded, but choices for unfamiliar

¹Laboratory for Neural Coding and Brain Computing, RIKEN Center for Brain Science, Wako, Japan. ²Department of Behavior and Brain Organization, Research Center Caesar, Bonn, Germany. *e-mail: tfukai@riken.jp

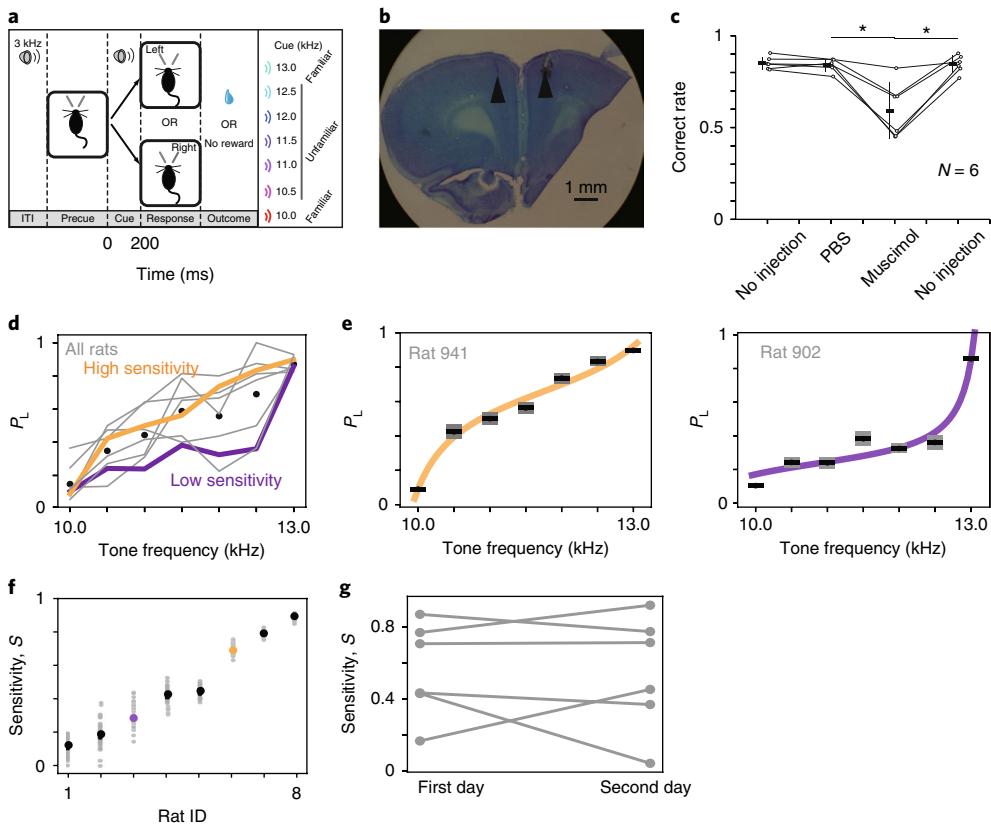


Fig. 1 | Individual differences in probabilistic choices across rats. Psychometric curves and task-related activity were studied in eight rats. **a**, Schematic illustration of the sensory-guided decision-making task. Colors indicate tone differences. Only familiar tones were used for training, and unfamiliar tones were inserted during recordings. **b**, Nissl-stained coronal section indicating injection sites (arrows) in bilateral MFC. The staining was performed in an additional six rats (around ten sections per rat). **c**, Results of the lesion study are shown for the six rats (circles). Center values and error bars show averages and standard deviation. Asterisks indicate statistical significance (PBS versus muscimol, $P=0.003$; muscimol versus no injection, $P=0.0030$; paired t -test, Bonferroni). **d**, Psychometric curves (gray) and average left-choice probabilities (filled circles) of the eight rats are shown. Typical examples of high (orange) and low (purple) sensitivity to auditory cues are also indicated. **e**, Psychometric curves of high (left) and low (right) sensitive rats were fitted with tangent functions. Center values and error bars show the average and standard deviation of resampled ensembles. **f**, Sensitivity values, S , of the resampled (gray) and original (purple and orange for the insensitive and sensitive rats, respectively, and black for others) curves were calculated. The rats are sorted in increasing order of their original sensitivity values. The average values of resampled S increase monotonically while preserving the original order. In **e** and **f**, 30 psychometric curves were generated by resampling 30 times for each rat. **g**, Sensitivity for the six rats that were recorded over two days is plotted.

cues were also rewarded in a probabilistic manner (Methods). Finally, eight rats yielded enough neurons for the present analysis. Unless otherwise stated, all results shown below are from the first day of recordings.

To examine whether the MFC is engaged in this sensory-guided decision-making, we also studied the decision behavior of six rats trained in the same procedure as described before⁷ with minor changes (Methods). Injection of the GABA_A receptor agonist muscimol into the bilateral MFC, but not injection of phosphate-buffered saline (PBS), significantly impaired the correct choice responses to familiar cues (Fig. 1b,c), indicating that the task requires the MFC.

An interesting feature was revealed in unfamiliar situations: the choice behavior showed large differences across the rats. The probability of choosing left licking generally increased with the tone frequency in all rats, but psychometric curves for unfamiliar cues exhibited large individual differences (Fig. 1d). Some rats varied choice probability sensitively to the frequency of unfamiliar cues (Fig. 1e, left), but other rats made biased or chance-level choices irrespective of the tone frequency (Fig. 1e, right). We quantified individual differences in the sensitivity (S) of psychometric curves by fitting the curves with a nonlinear function, P_L (Methods and Supplementary Fig. 1). We confirmed the stability of sensitivity

of each rat by examining whether the sensitivity stably characterizes the behavioral tendency of each rat against the resampling of the values of S (Fig. 1f). We further confirmed the stability of our fitting scheme by choosing another fitting function (Supplementary Fig. 2; see Methods). In addition, we examined whether the sensitivity was stable in the six rats that underwent the two subsequent days of recordings. Increases in the sensitivity could suggest that the rats had become familiar with the ‘unfamiliar cues’ through learning. Although two rats showed increased sensitivity on the second day, three rats were relatively sensitive and the other three rats remained relatively insensitive across the two days (Fig. 1g). The sensitivity across the population also showed no significant changes (Student’s t -test, $P=0.86$).

Reinforcement learning for a reservoir network model. To clarify how the behavioral difference emerges among the rats, we constructed a recurrent network based on the anatomical structure of the MFC (Fig. 2a). Our model extends the reservoir computing proposed for supervised motor learning^{11–14} to reward-driven sequence learning^{15–17}. The recurrent network (the reservoir) represents the MFC and consists of excitatory and inhibitory spiking neurons, a subset of which receive random projections from input terminal

neurons having different preferred tone ‘frequencies’. Anatomically, the input terminal may correspond to the auditory cortex^{18,19}, and random projections were assumed because tone-selective MFC neurons showed no systematic frequency dependence⁷. Excitatory neurons in the reservoir project to two rate-coding neurons (readout neurons), the L-neuron and R-neuron, which inhibit mutually and send feedback connections to the reservoir. The readout neurons may be representative of the motor cortex generating licking behavior²⁰. The network model makes a decision when the difference in output between L- and R-neurons reaches a threshold, θ . Owing to the noise applied to the reservoir and trial-to-trial variability in the initial states, decisions are stochastic even for an identical sensory input.

To take advantage of sequence generation^{21,22}, we made recurrent excitatory connections obey a lognormal weight distribution, as found in cerebral cortices^{22,23}. We assumed a minimal model in which only readout connections were modifiable. Nevertheless, the reservoir dynamics is modulated through feedback connections from readout neurons. We trained the network model through a reward-driven reinforcement learning rule such that it correctly generated a ‘left’ or ‘right’ choice for familiar ‘high-frequency’ (FH) or ‘low-frequency’ (FL) inputs, respectively. We used eligibility traces^{15–17} to selectively train the readout connections that influence decision-making at each learning step (that is, within each trial during the learning process). Such connections were potentiated or depressed after every success or failure trial, respectively. All mathematical details of the model are described in the Methods. Although our learning rule is somewhat heuristic, the minimal model accounted for most of the essential experimental findings.

Figure 2b shows a typical example of a rewarded trial during learning. We trained 50 networks with different realizations of recurrent connections by randomly applying either FH or FL inputs (see Methods). More than 60% of model networks (31 networks) achieved the criterion correct rate of 80% (Fig. 2c). Successful learners occupied similar fractions in both models and rats, indicating that the task difficulty was adequately modeled. Below, we analyze the neural dynamics and behavioral variabilities of these 31 successful learners and 8 rats.

Next, we analyzed the choice responses of our model to unfamiliar cues (U1 to U5 in Fig. 2a). The psychometric curves of the successful learners displayed large individual differences (Fig. 2d,e), with their sensitivity values robust against resampling of psychometric curves (Fig. 2f). The width of tuning curves (the parameter L in equation (1); see Methods) affects the initial states set by external stimuli, and hence influences the behavioral characteristics of each network. The networks tend to exhibit a broader S distribution for a narrower width or a higher input selectivity (Fig. 2g). Both models and rats exhibited similar ranges of sensitivity for the widths of 0.3 to 0.4. Unless otherwise stated, we use the width of 0.4, which gives a smaller average fitting error than the other value.

Neural trajectories formed in the MFC and reservoir. The network models produced the divergent psychometric curves that are consistent with experimental observations. However, whether neural activities in the reservoir and in the MFC are also similar has to be clarified. To this end, we compared neural activities averaged over familiar trials between a rat (Fig. 3a) and a reservoir (Fig. 3b) under given cue and choice conditions. Both experimental and modeled activities exhibited distinct trajectories when neurons were sorted in the order of their peak firing times in correct trials for given stimuli. However, no clear trajectories appeared when neurons were sorted in the order observed in different stimulus conditions, demonstrating that neural trajectories are choice specific in both the MFC and the reservoir.

The principal components (PCs) of population neural responses in the MFC demonstrated that decision-making with familiar cues

formed neural trajectories selective to stimulus and choice (Fig. 3c). Note that the first principal component (PC1) was a frequency-nonselective component of auditory responses (Supplementary Fig. 3) and hence is not shown here. In the reservoir, similar condition-selective trajectories are formed, although their PC1 and PC2 corresponded to PC2 and PC3, respectively, in the MFC (Fig. 3d). Frequency-nonselective stimulus-evoked responses (PC1 of MFC) were almost missing in our models. The average variances explained by the first three PCs were about 78% (MFC) and 84% (models).

We then investigated how these familiar trajectories lead to correct choice responses. We applied a linear regression method²⁴ (Methods) to trial-averaged neural activities for given combinations of stimulus and choice, and identified two task-relevant axes: the stimulus and choice axes optimally explain changes in trajectories between FH and FL or between left and right choices, respectively. In both rat (Fig. 3e) and model (Fig. 3f), familiar trajectories projected onto the stimulus axis or the choice axis started to separate with cue onset. Cohen’s d also started to deviate from zero on both axes, indicating the beginning of substantial separation (Methods). The results indicate that both MFC and reservoir integrate sensory evidence during cue presentation. Supplementary Fig. 4 shows the separation of neural trajectories in the other five rats that yielded sufficient (30 or more) neurons. Note that the small separation of pre-cue trajectories observed in Fig. 3e along the choice axis is not a common tendency of all rats.

Familiar trajectories revealed interesting differences between the models and rats. First, trajectory separation along the stimulus axis was generally smaller in the rats compared with the models, as indicated by Cohen’s d in Fig. 3e,f. Although the cause of this discrepancy remains unclear, this was partly due to the small sample of tone-selective MFC neurons and small differences in their responses to different tones (on average, several spikes per second). Second, trajectories were separated almost maximally along the choice axis just before cue termination in the rats⁷, whereas the separation continued to grow in the models. This continuous separation was caused by the continuation of positive feedback from readout neurons. The separation was terminated if we cut off the feedback after cue termination (Supplementary Fig. 5). Thus, our computational model well reproduced neural dynamics in the rat MFC.

Neural dynamics in MFC and reservoir predicts individual choice behaviors. We further examined whether neural trajectories have sufficient information for predicting choice responses in the MFC and our model. Fisher discriminant analysis (FDA) determined a hyperplane (equivalently, its normal vector \mathbf{W}_{opt}) that optimally divides the population activity patterns corresponding to left and right choices at the time of familiar choice responses (Methods). The normalized weight vector of post-learning readout connections should be close to \mathbf{W}_{opt} (however, the two vectors are not identical because of the nonlinearity of readout neurons and mutual inhibition between them).

The averaged familiar trajectories (projected onto \mathbf{W}_{opt}) that link given sensory cues to different choices were distinctive in both rats (Fig. 4a, left) and models (Fig. 4a, right). Notably, \mathbf{W}_{opt} obtained from familiar trials was also valid for trajectory discrimination in unfamiliar trials, meaning that familiar and unfamiliar trajectories converge at similar patterns. We computed the distributions of projected activities in a rat (Fig. 4b) and a model (Fig. 4e) at the times of choice. Because the left- and right-choice distributions have an overlap in the rats, we adopted the mid-point of their means as a naive criterion boundary for inferring trial-by-trial choice responses. On average about 75% of trajectories in the MFC could be discriminated by this criterion (Fig. 4c). In the models, correct inference was obtained in about 90% of trials (Fig. 4f). Thus, neural trajectories in the MFC and reservoir determine choice responses on a single-trial basis. Choice-related single-unit activity has been reported in the MFC of rodents^{25,26}.

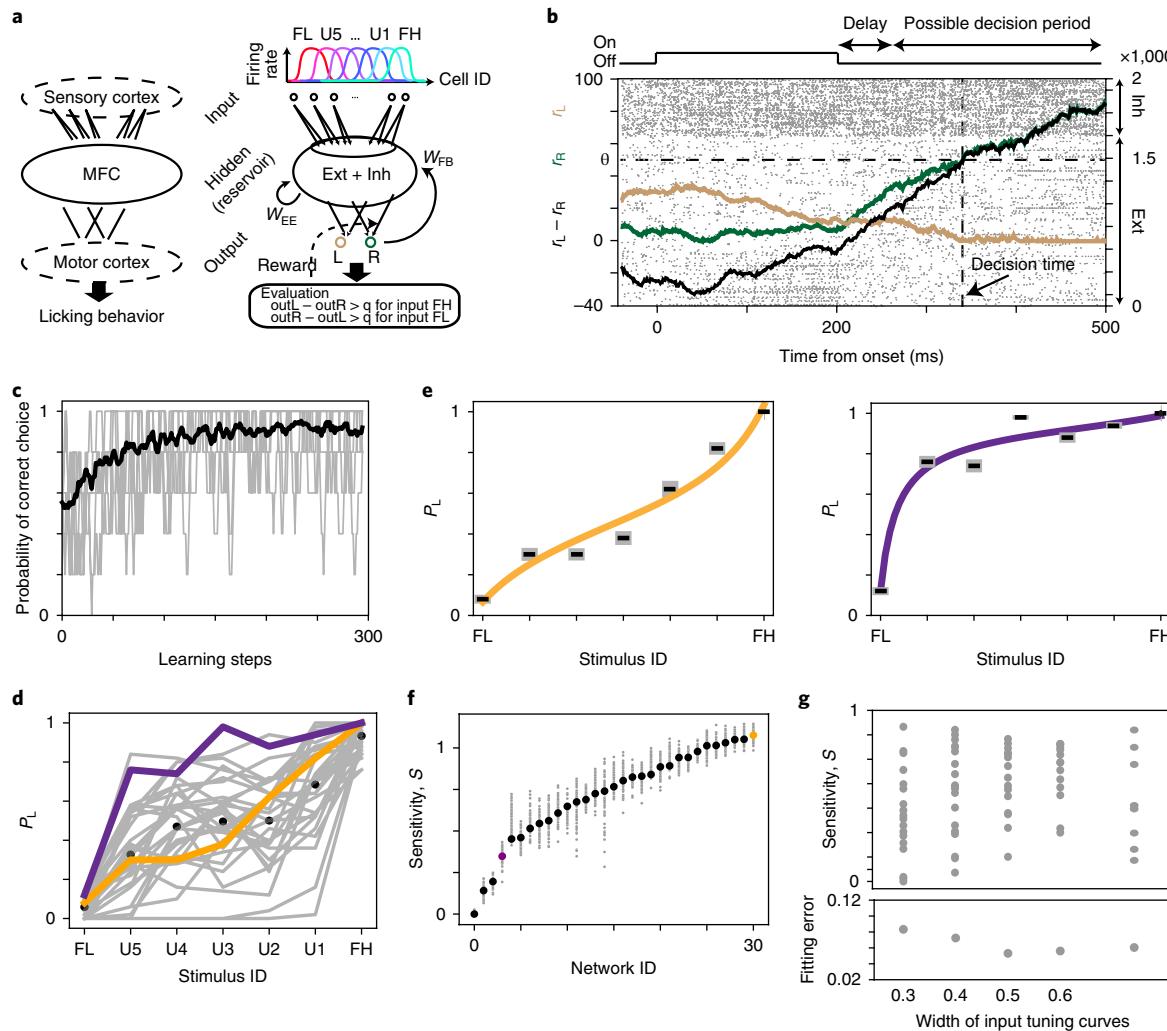


Fig. 2 | Responses of network models reproduce individual differences. **a**, The presumed anatomical circuits (left) and model network structure (right) are shown schematically. Tone selectivity of input neurons is indicated by the color code used in Fig. 1a. **b**, Spike raster showing the reservoir (dots) and the firing rates of L (r_L , brown) and R (r_R , green) neurons in a trial. Here we show 500 excitatory neurons receiving input projections, 1,000 excitatory neurons receiving no input projections, and 500 inhibitory neurons that were randomly sampled. Inh, inhibition; Ext, excitation. **c**, Averaged correct-choice probability over 50 different realizations of the network (black) and five examples (gray). Correct-choice probability was evaluated every five learning steps. **d**, Psychometric curves of 31 successful learners (gray) and two typical curves with high (orange) or low (purple) sensitivity. **e**, The typical sensitive (left) and insensitive (right) psychometric curves are fitted with the same tangent functions as used in Fig. 1e for the rats. The center values and error bars indicate the average and standard deviation. **f**, Sensitivity S of the original (black) and resampled (gray) psychometric curves for 31 successful learners. The values of S are marked for the original sensitive (orange) and insensitive (purple) psychometric curves. **g**, Distributions of S are plotted for the rats and the models at the top. The S values of 18, 20, 16 and 14 networks were plotted for the widths of 0.3, 0.4, 0.5 and 0.6, respectively. Averaged fitting errors are shown at the bottom.

However, FDA was insufficient for predicting individual differences in the sensitivity, S . We found no significant correlations between the sensitivity and the discriminability of trajectories in both rats (Fig. 4d; $P=0.87$, two-sided Student's t -test) and models (Fig. 4g; $P=0.71$, two-sided t -test). Although this could be partly due to the small size of the data in the case of rats, the different behavioral characteristics may not merely reflect neural states at the choice point. Therefore, we next investigated how neural dynamics preceding the choice point determine the behavioral characteristics.

Stability of neural dynamics influences behavioral characteristics. Because decision-making depends on an interplay between internal dynamics and sensory input in our model, we explored whether different characteristics of internal dynamics generate different behavioral characteristics. To characterize the pure effect

of internal dynamics, we applied no sensory stimuli. Instead, we applied a perturbative input to each network 30 times in different directions when the network evolved into a state (perturbed state) to which sensory stimuli were originally applied. We measured the average Euclidian distance (neural responsiveness, χ , in equation (22); see Methods) of perturbative trajectories from their geometric center. Note that χ is perturbed state dependent. Perturbed trajectories evolving from a state with high χ may diverge broadly, while those from a state with low χ may hardly diverge (Supplementary Fig. 1).

We first show typical network responses to sensory stimuli. In a sensitive network, unfamiliar stimuli U1, U2 and U3 (close to FH) evoked trajectories evolving into left choice, whereas those evoked by U4 and U5 (close to FL) evolved into right choice, showing that the responses are sensitive to similarity between unfamiliar

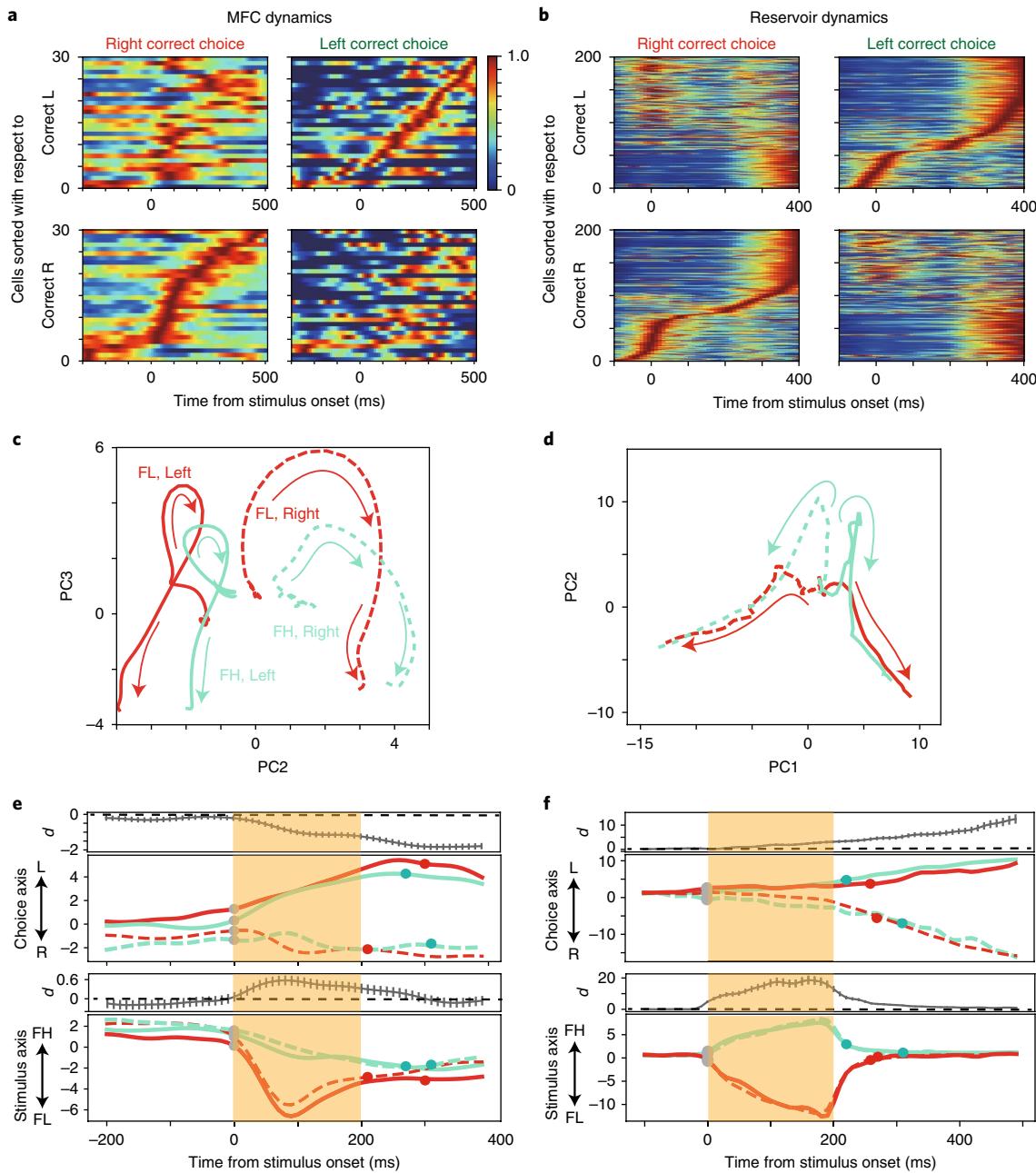


Fig. 3 | Neural dynamics in MFC and in the reservoir. **a**, Activities of the 30 most active neurons out of 44 in a rat are averaged and normalized over successful familiar trials. Activities are sorted in the temporal order of peak responses in correct left (top) or correct right (bottom) trials. **b**, Activities of the 200 most active excitatory neurons out of 5,000 in a model. **c,d**, Trajectories in the MFC (**c**) and the reservoir (**d**) are projected onto a two-dimensional PC-space from -300 ms to $+300$ ms (MFC) or -100 ms to $+900$ ms (models) after stimulus onset. We averaged activities over 30 MFC neurons with the largest rate differences between pre- and post-familiar-stimulus periods and the 100 most active model neurons under each stimulus (FH, green; FL, red) and choice (left, solid; right, dotted) condition. In the MFC, PC1 is not shown as it merely represented the presence of cues. **e,f**, Population neural activity in the rat (**e**) and model (**f**) is projected onto the choice (top) and stimulus (bottom) axes in the same colors and lines as in panel **c**. Gray and color (green, red) circles indicate stimulus onset and mean reaction time, and the y-axis d shows Cohen's d with 95% confidence levels (vertical bars). Trajectory separation is significant when $d=0$ lies outside the range of the vertical bars. All of the above results were obtained from 379 (44) FH→left and 309 (44) FL→right successful trials in rats (and model), and 101 (6) FH→right and 67 (6) FL→left unsuccessful trials.

and familiar stimuli (Fig. 5a, top). By contrast, all unfamiliar trajectories resulted in left choice in an insensitive network, showing insensitivity to stimuli and a strong bias in the responses (Fig. 5a, bottom). Next, perturbed trajectories diverged broadly in the sensitive network, giving high χ values (0.22; Fig. 5b, top), whereas the insensitive network generally yielded more localized perturbed trajectories and lower χ values (0.18; Fig. 5b, bottom). Taken together,

these results suggest that networks that respond differentially to different unfamiliar stimuli likely show high neural responsiveness, whereas those responding similarly to different stimuli likely give low responsiveness.

To quantify the above relationship, we measured the response bias defined as $|left - 7/2|$ with the number of left choices obtained in response to the seven stimuli (FL, U1, ..., U5, FH). The response

bias is 0.5 if three or four of the seven choices are left (weak biases), while it is 3.5 if all choices are only left or right (strong biases). The response bias was significantly and negatively correlated with the value of χ in both sensitive (Fig. 5c, top; Pearson's correlation coefficient $r = -0.274$; $P = 7.56 \times 10^{-5}$; two-sided Student's t -test) and insensitive (Fig. 5c, bottom; $r = -0.562$; $P = 1.34 \times 10^{-21}$) successful learners. In the sensitive network, the distribution of the average χ values was biased towards higher χ values (Fig. 5c, top), whereas the insensitive network yielded a bias towards lower values (Fig. 5c, bottom). The sensitivity values, S , of 31 successful learners displayed a significant positive correlation with the average χ values (Fig. 5d; $r = 0.61$; $P = 0.003$). We further confirmed this correlation by dividing the networks into a subgroup of 15 networks with the upper half of χ values and a subgroup of the remaining 16 networks. These groups had significantly different average S values (0.237 in the higher- χ group and 0.158 in the lower- χ group; t -test, $P = 0.0041$). These results give the following picture of neural mechanisms of individual differences: higher neural responsiveness implies a shallower landscape of internal dynamics around the trajectory and produces a sensitive choice behavior, while lower sensitivity implies a deeper landscape and generates an insensitive choice behavior (Fig. 5e).

By contrast, perturbed and unperturbed trajectories showed no large quantitative and qualitative differences in the presence of FH in both network types (Fig. 5f). Similar results were obtained for FL (data not shown). Accordingly, no significant correlations exist between behavioral sensitivity and the average neural responsiveness in the presence of learned stimuli (Fig. 5g). Thus, our model suggests that internal network dynamics influence the behavioral variability, but that such influences are easily masked by external inputs (Fig. 5e). This may explain why no convincing relationship was found between neural population responses and behavioral variability in the rats.

Neural responsiveness generates trial-by-trial variance. Because neural responsiveness cannot be computed from our experimental data, we need another metric to characterize neural dynamics in the rats. We examined whether trial-by-trial variances reflect the neural dynamics of decision making^{27–29} (Methods). In fact, the normalized trial-by-trial variability calculated before and after stimulus onset was significantly correlated with neural responsiveness in the models (Fig. 6a). Here to reduce artifacts from neuron sampling, we normalized post-stimulus-onset variability by pre-stimulus-onset variability. The results imply that trial-by-trial variance can be a proxy of neural responsiveness. Furthermore, the variability was correlated with the sensitivity in the models (Fig. 6b, left).

Similarly, the normalized variability was correlated with the sensitivity in the eight rats (Fig. 6b, right). Because of the small number of rats, we resampled the experimental data 1,000 times and obtained 1,000 resampled correlation coefficients across the rats (Methods). All of the coefficients were positive (Fig. 6c), indicating that neural population dynamics also reflects the behavioral differences of individual rats, and that the normalized variability probes the characteristics of neural dynamics (presumably, neural responsiveness).

Relationship between behavioral sensitivity and learning. The above relationship between behavioral sensitivity and neural responsiveness could reflect a difference between 'better' and 'worse' learners. In Supplementary Fig. 6a, we addressed this question by comparing successful learners and unsuccessful learners ($n = 13$). From the total of 19, we excluded 6 unsuccessful learners that responded correctly if the time limit for decision making was relaxed. The remaining unsuccessful learners tended to exhibit less sensitivity compared with successful learners, but this tendency was statistically insignificant (t -test, $P = 0.07$). The average percentage of correct responses calculated from all (rats) or 50 (models)

test trials was also not significantly correlated with the sensitivity (Supplementary Fig. 6b,c).

Theoretically, we might anticipate that the sensitivity, or the degree of bias in decision-making, could be correlated with the speed of learning. In fact, sensitive learners learned slowly while insensitive learners learned quickly in our model (Supplementary Fig. 6d). Because behavioral sensitivity is correlated with neural responsiveness in the model, the latter is also correlated with the learning time as measured by the number of learning steps required to reach the criterion correct rate (Supplementary Fig. 6e). By contrast, correlation between the sensitivity and learning time was not convincing in the rats, primarily because of day-by-day fluctuations in decision performance (Supplementary Fig. 6f).

Reaction time is uncorrelated with behavioral characteristics. Reaction time is an important measure of the behavioral characteristics of individual animals. Interestingly, the median reaction times of the eight rats were not significantly correlated with sensitivity (Supplementary Fig. 7a–c), showing that the reaction time does not strongly reflect individual differences in choice behavior. We asked similar questions in our model and found that reaction times were also uncorrelated with sensitivity (Supplementary Fig. 7e–g). However, we also noticed an interesting difference between the model and experiment. In each model, reaction times to unfamiliar cues were significantly longer than those to familiar cues (Supplementary Fig. 7h; $P < 0.001$), which is consistent with our intuition. Unexpectedly, however, the reaction times of the rats were not significantly different between familiar and unfamiliar cues (Supplementary Fig. 7d; $P = 0.8$). Only four rats exhibited somewhat longer (15–30 ms) median reaction times for unfamiliar cues (Supplementary Fig. 7c). This discrepancy suggests that additional mechanisms not modeled here work in the MFC under unfamiliar sensory conditions.

For instance, the exposure to unfamiliar tones could slow down the overall responses of the rats. To examine this, we compared the reaction times of the eight rats between the last day of training (familiar only) and the mixed familiar/unfamiliar environment, and found similar reaction-time distributions in the two environments (Supplementary Fig. 8a,b). However, five rats exhibited significantly longer reaction times and one rat showed a significantly shorter reaction time in the mixed environment (Supplementary Fig. 8c; $P < 0.05$, U-test). Therefore, although the overall effect was small, unfamiliar cues prolonged the reaction times in the majority of rats. Such effects were also not modeled.

Discussion

In this study, we have shown how neural populations in the MFC process sensory-guided alternative decision-making when rats are exposed to unfamiliar sensory stimuli. The choice responses to unfamiliar stimuli were variable across different rats, and the psychometric curves of probabilistic choices ranged from highly to poorly sensitive types across individual rats. By building a network model, we have revealed that the covert stability of internal neural dynamics predicts individual differences in choice. The model predicted that highly sensitive rats should demonstrate large trial-by-trial variability in neural trajectories, and our experimental results validated this prediction.

Choice-selective neural dynamics and its relevance to behavioral variance. Neural mechanisms of decision-making under uncertainty have been explored in various species and behavioral tasks^{1,30,31}, and it is now widely accepted that the recruitment of neural ensemble sequences is essential for organizing decision-making behavior^{4–6}. While decision-making is immediate under familiar sensory conditions, the decision response becomes ambiguous under uncertainty and exhibits large variance across individuals.

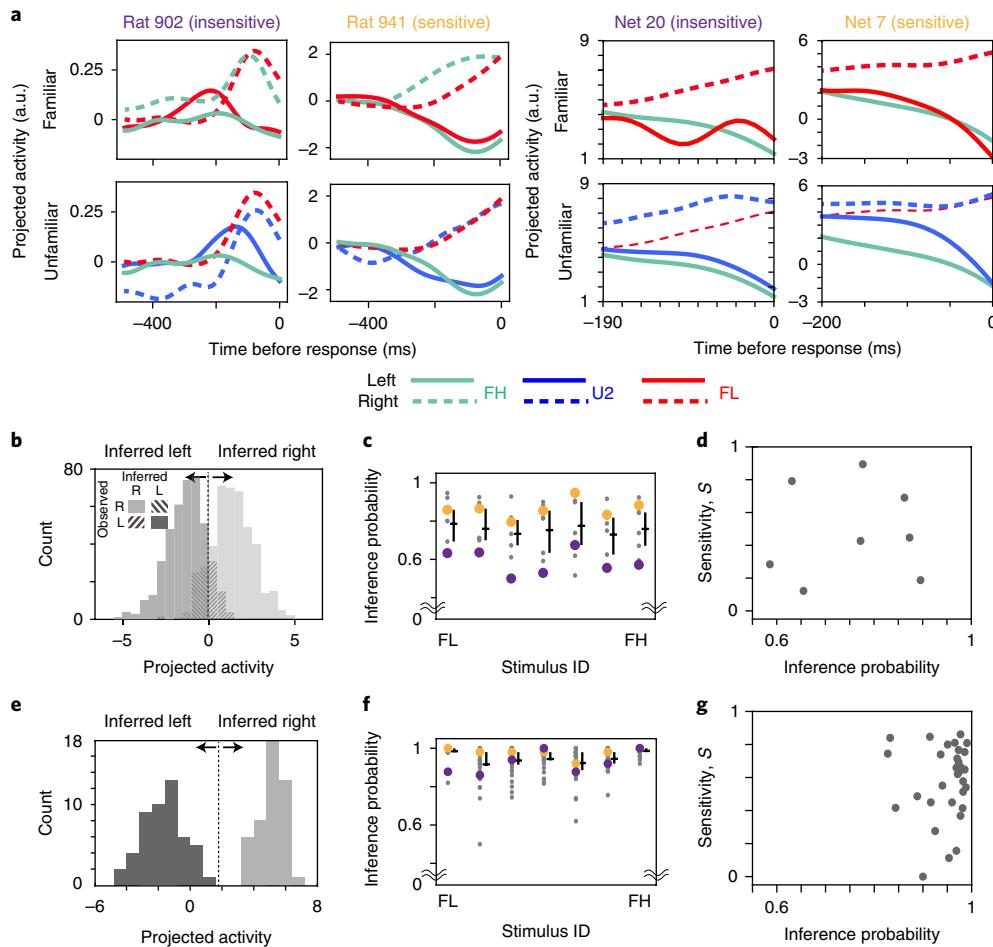


Fig. 4 | Predictions of choice behavior by MFC and model. **a**, Trajectories were optimally separated by FDA in rats (left) and models (right) with different sensitivity. Projected trajectories are shown in left-choice (solid) and right-choice (dotted) trials for familiar (top) and unfamiliar (bottom) stimuli. Trial numbers in rats 902, 941 and networks 7 and 20: FH→left 352, 422, 50, 50; FH→right 354, 433, 46, 43; FH→right 59, 6, 0, 0; FL→left 41, 43, 4, 6; U2→left 13, 42, 31, 43; U2→right 27, 21, 19, 6. **b**, Distributions of projected population activity are shown at the times of left choices for FH (dark gray) and right choices for FL (light gray) in the sensitive rat shown in panel **a**. The dotted line shows the midpoint between the means of the two distributions and its left-side area predicts left-choice probability. The shaded area indicates the trials incorrectly inferred. **c**, The probabilities that left and right choices are correctly inferred from projected trajectories are shown for all eight rats (gray circles) and for the sensitive (orange) and insensitive (purple) rats. Vertical and horizontal bars show the medians and first and third quartiles, respectively. Trial numbers for FL, U5 to U1, FH are 3120, 390, 348, 330, 304, 289, 3198 (total for all rats), respectively. **d**, Relationships between the probability of correct inference and sensitivity are shown for the eight rats. **e–g**, As for panels **b–d**, but for the 31 successful model networks. The number of trials per stimulus is 1,550 (50 × 31).

Here we have explored whether and how choice-selective trajectories determine behavioral variance in the rodent MFC, which plays a pivotal role in processing sensory-guided and value-based goal-directed decisions^{7,9,10}. Our results suggest that the covert stability of MFC neural dynamics is a determinant of individual behavioral variation in perceptual decision-making. This stability is ‘covert’ because the neural trajectories exhibited no significant differences between network models with different behavioral sensitivity under the influences of sensory stimuli (Fig. 5f,g). This prediction could be tested by optogenetically stimulating MFC neurons³². More specifically, future experiments in which the MFC network is perturbed with optogenetic activation of a set of pyramidal cells at the timing of sensory cues while the variability of the resulting network outcomes is measured would allow one to estimate the ‘depth of the landscape’ and extend these results further.

Behavioral variability between individuals has been a focus of psychological and behavioral studies in humans³³. Individual preferences in the trade-off between exploration and exploitation under uncertain conditions have been studied in a large-scale

brain network including cortical and subcortical structures³⁴. Furthermore, recent studies using functional magnetic resonance imaging have revealed that macroscopic brain activity influences the behavioral variability of individual animals³⁵. Our results further suggest that not only macroscopic states of the brain circuitry, but also local cortical dynamics, influence individual differences in a perceptual choice behavior that does not require strategic exploratory decisions. In addition, our modeling study demonstrates that perturbation of neural circuits is an effective way to uncover a link between a covert property of neural dynamics and individual behavioral variance. Such a perturbative method recently revealed hidden transient dynamics in working-memory-guided human behavior³⁶. These results also suggest the potential of perturbative methods for clarifying the underlying mechanisms of individual variance.

Covert stability determines responses to different inputs. Neural trajectories are thought to transform information on sensory stimuli into decision behavior in sensory-guided decision-making^{4–6}. Some recent studies investigated the underlying dynamical structure

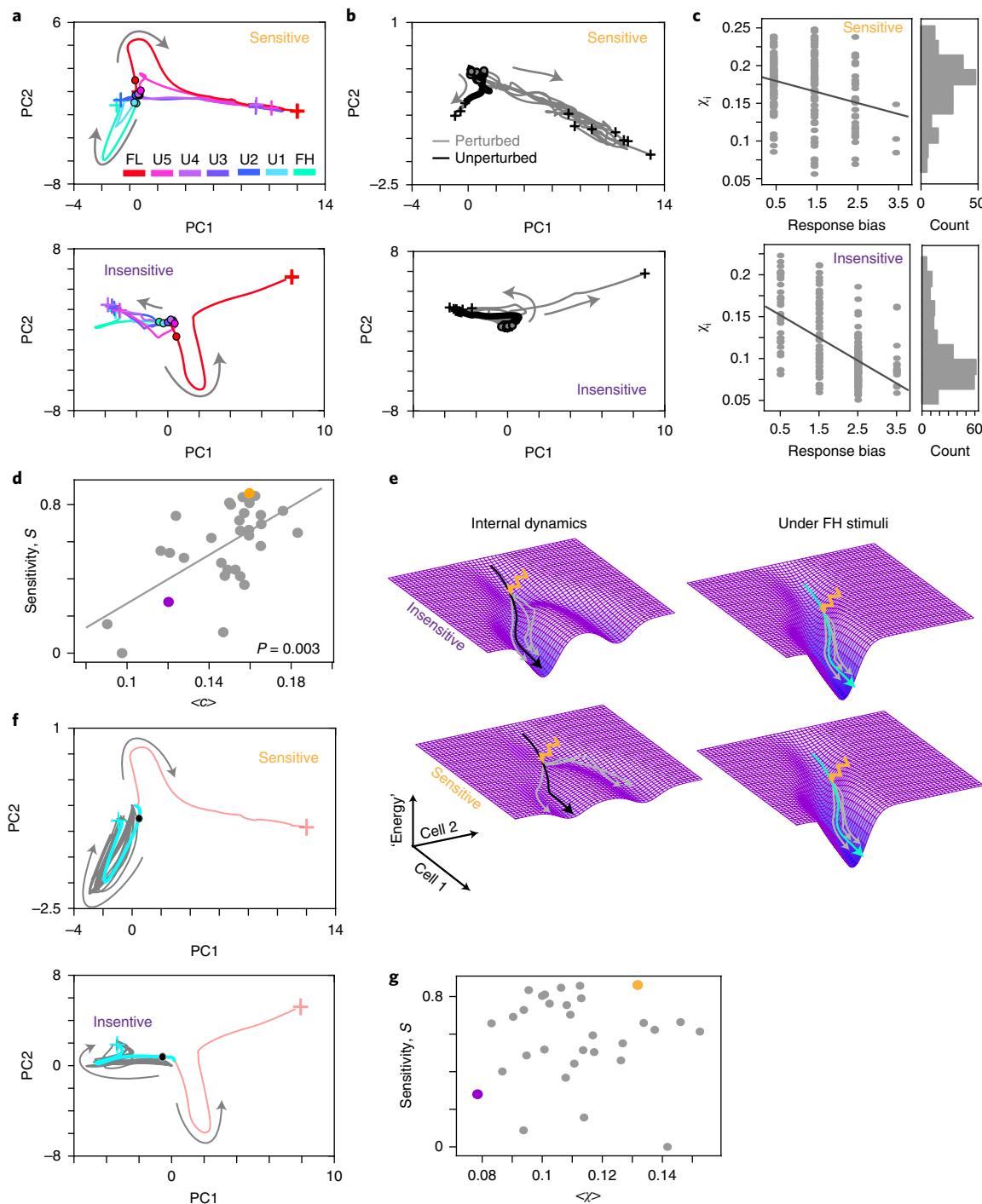


Fig. 5 | Characterization of neural dynamics for different choice behaviors. **a**, Trajectories evoked by familiar and unfamiliar stimuli are shown (one trial per stimulus) in terms of PCs for sensitive (top) and insensitive (bottom) networks, respectively. **b**, Ten perturbed trajectories (gray) chosen randomly from 30 trials (per perturbed state) and an unperturbed trajectory (black) in a reference trial are shown in the same networks and PC spaces as in panel **a**. Perturbed states (gray circles) and decision points (crosses) are indicated. All perturbed states are identical, although they look different owing to Gaussian smoothing (s.d. 30 ms). **c**, Left, neural responsiveness for 50 perturbed states is shown against the response bias $|left-3.5|$ in the five most sensitive (top) and five most insensitive (bottom) networks. Right, histograms of neural responsiveness. **d**, Neural responsiveness averaged over randomly chosen 50 perturbed states (30 perturbations per perturbed state) is plotted against the sensitivity for all successful learners. Orange and purple circles indicate the two networks shown in panel **a**. **e**, Schematic images of the landscape of neural dynamics with (right) and without (left) familiar stimuli. **f**, Ten perturbed trajectories (gray) and an unperturbed trajectory (cyan) are shown in the presence of FH (gray) in the same two networks and perturbed states (circles) as in panel **a**. The unperturbed trajectory in the presence of FL is plotted for reference (red). **g**, Average neural responsiveness over the perturbed trajectories shown in panel **f** and sensitivity are plotted for all successful learners.

of neural-state space to clarify how such trajectories process sensory stimuli in order to organize an adequate behavioral output according to task demands^{24,37–39}. For instance, line attractors underlie the

context-dependent selection and integration of noisy visual input²⁴. In a delayed tactile discrimination task, the separatrix of neural trajectories has to be adequately located to make decisions in a required

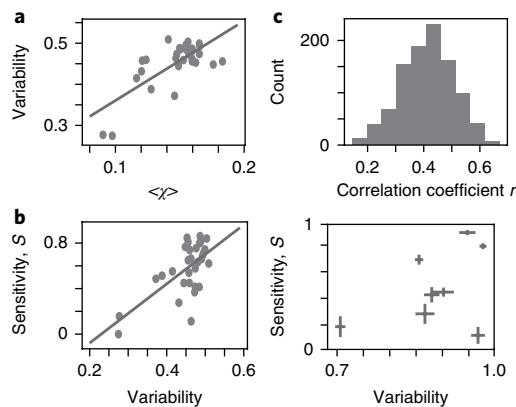


Fig. 6 | Trial-by-trial variability predicts individual difference. **a**, The neural responsiveness, χ , and the normalized trial-by-trial variability are significantly correlated in all 31 successful learners (Pearson's $r=0.65$; $P=3.36 \times 10^{-6}$; two-sided t-test). **b**, Correlations between normalized variability and sensitivity across 31 successful learners (left; Pearson's $r=0.73$; $P=7.13 \times 10^{-5}$; two-sided t-test) and eight rats (right; $P=0.29$; no significance) (see Methods). For the rats, error bars in the ordinate show standard deviations, calculated as in Fig. 1f, and error bars in the abscissa show standard deviations over neuron ensembles by resampling 90% of trials 50 times for each rat. Center values represent averages. **c**, Histogram of Pearson's correlation coefficients between normalized variability and sensitivity in the rats. We resampled the experimental data 1,000 times per rate and calculated 1,000 correlation coefficients.

time window³⁷. In an arm-reaching task, the neural-state space of the motor cortex has a ‘null space’, in which neural trajectories do not correlate with current movements but do influence the selection of subsequent movements⁴⁰.

Our model reveals crucial influences of the covert landscape of neural-state space around the choice-specific trajectories in an unfamiliar sensory environment. A shallow landscape around a perturbed state generates decision behavior that is highly sensitive to unfamiliar stimuli. Conversely, a deep landscape predicts poorly sensitive choice behavior. The behavioral importance of robust trajectories was recently pointed out in a classification task³⁸. Computational models^{41,42} and experimental studies^{43,44} also demonstrated the impact of internal neural dynamics on processing sensory information and behavioral parameters in cognitive tasks. We further suggest a link between the covert property (that is, the degree of robustness) of internal neural dynamics and behavioral variance across individuals.

Another consequence of the covert landscape is that a weak stability of trajectories predicts large trial-by-trial variability in neural activity, and that this variation, in turn, is correlated with behavioral variances. We confirmed the latter in MFC neural activities (Fig. 6). Although trial-by-trial variability reflected decision-making previously^{27–29}, the underlying mechanisms are not clear. We suggest that the covert stability of neural dynamics is a possible mechanism. In humans, larger trial-by-trial variability in movements predicted faster rates of motor learning, although whether high variability in movements implies high variability in neural activity remains unclear^{45,46}. In the present task, large variability in neural responses may increase the flexibility in controlling neural trajectory evolution, thus enabling stimulus-sensitive rats to flexibly adjust their probabilistic behavioral responses according to cue tones. This may explain why trial-to-trial neural variability is correlated with psychometric curves. However, because of the small size of samples, these correlations were too weak in the rats to derive a convincing conclusion on this point (Fig. 6).

Limitations of our models. Our model largely replicated the choice behavior and neural dynamics observed in the rats, but some discrepancy exists. First, the reaction times for unfamiliar cues are comparable to those for familiar cues in the rats (Supplementary Fig. 7c: the mean reaction times differed by at most 15 to 30 ms between the two cue types), whereas in our model the mean reaction time was much shorter for familiar stimuli than for unfamiliar ones. Although the exposure to unfamiliar cues increased the reaction time to familiar cues in some rats (Supplementary Fig. 8c), the actual causes of this discrepancy remain unclear. A likely explanation is that learning of a sensory environment increases the threshold for decision-making to improve the accuracy of behavioral responses (that is, the threshold may be different between familiar and unfamiliar cues). Yet another explanation is that bottom-up inputs from the primary auditory cortices are less effective for unfamiliar cues in activating MFC neurons, as they have not learned these inputs. Indeed, modeling studies suggest various plasticity mechanisms to effectively gate the learned inputs⁴⁷.

Second, trajectory separation somewhat slows down in the rats about 100 ms after cue onset (Fig. 3e), whereas the separation continues in our model without slow down (Fig. 3f) if feedback input still exists (Supplementary Fig. 5). Around this timing, a sharp increase in the gamma power of the local field potential occurs in the MFC⁷. As suggested for rodents^{48,49} and primates⁵⁰, the increased gamma power may indicate the onset of cross-area communication in the MFC, possibly for preparing behavioral outputs. This communication across brain areas, which is not implemented in our model, might suppress trajectory separation (possibly by terminating feedback from the motor cortex to the MFC).

The divergent behavior of the rats may result from different decision strategies without invoking internal neural dynamics. For instance, some animals may adopt an internal decision threshold, above which it licks left, while other animals may detect a particular familiar tone and would lick left for any other tone. Although we cannot exclude the possibility that the MFC recruits additional mechanisms not modeled here, overall behavioral and neural data, including the broad spectrum of behavioral responses across the rats, are consistent with the predictions of our neural dynamics model. Evidence for the role of neural dynamics in decision-making is also accumulating for various decision-making tasks and brain regions^{4,24,37}.

In sum, we have shown that the covert local landscape of internal neural dynamics in the MFC influences global behavioral trends in decision-making under unfamiliar sensory conditions. Our work provides evidence that differences in local circuit dynamics determine behavioral variance across individuals.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41593-018-0263-5>.

Received: 23 July 2018; Accepted: 2 October 2018;
Published online: 12 November 2018

References

- Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
- Machens, C. K., Romo, R. & Brody, C. D. Functional, but not anatomical, separation of ‘what’ and ‘when’ in prefrontal cortex. *J. Neurosci.* **30**, 350–360 (2010).
- Wang, X. Decision making in recurrent neuronal circuits. *Neuron* **60**, 215–234 (2008).
- Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484**, 62–68 (2012).
- Fujisawa, S., Amarasingham, A., Harrison, M. T. & Buzsáki, G. Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nat. Neurosci.* **11**, 823–833 (2008).

6. Mazor, O. & Laurent, G. Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons. *Neuron* **48**, 661–673 (2005).
7. Handa, T., Takekawa, T., Harukuni, R., Isomura, Y. & Fukai, T. Medial frontal circuit dynamics represents probabilistic choices for unfamiliar sensory experience. *Cereb. Cortex* **27**, 3818–3831 (2017).
8. Deco, G., Rolls, E. T. & Romo, R. Synaptic dynamics and decision making. *Proc. Natl Acad. Sci. USA* **107**, 7545–7549 (2010).
9. Erlich, J. C., Bialek, M. & Brody, C. D. A cortical substrate for memory-guided orienting in the rat. *Neuron* **72**, 330–343 (2011).
10. Narayanan, N. S., Cavanagh, J. F., Frank, M. J. & Laubach, M. Common medial frontal mechanisms of adaptive control in humans and rodents. *Nat. Neurosci.* **16**, 1888–1895 (2013).
11. Maass, W., Natschläger, T. & Markram, H. Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput.* **14**, 2531–2560 (2002).
12. Jaeger, H. & Haas, H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* **304**, 78–80 (2004).
13. Sussillo, D. & Abbott, L. F. Generating coherent patterns of activity from chaotic neural networks. *Neuron* **63**, 544–557 (2009).
14. Laje, R. & Buonomano, D. V. Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nat. Neurosci.* **16**, 925–933 (2013).
15. Izhikevich, E. M. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* **17**, 2443–2452 (2007).
16. Williams, R. J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* **8**, 229–256 (1992).
17. Seung, H. S. Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron* **40**, 1063–1073 (2003).
18. Condé, F., Maire-lepoivre, E., Audinat, E. & Crépel, F. Afferent connections of the medial frontal cortex of the rat. II. Cortical and subcortical afferents. *J. Comp. Neurol.* **352**, 567–593 (1995).
19. Hoover, W. B. & Vertes, R. P. Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat. *Brain. Struct. Funct.* **212**, 149–179 (2007).
20. Reep, R. L., Corwin, J. V., Hashimoto, A. & Watson, R. T. Efferent connections of the rostral portion of medial agranular cortex in rats. *Brain Res. Bull.* **19**, 203–221 (1987).
21. Teramae, J., Tsubo, Y. & Fukai, T. Optimal spike-based communication in excitatory networks with strong-sparse and weak-dense links. *Sci. Rep.* **2**, 485 (2012).
22. Ikegaya, Y. et al. Interpyramidal spike transmission stabilizes the sparseness of recurrent network activity. *Cereb. Cortex* **23**, 293–304 (2013).
23. Buzsáki, G. & Mizuseki, K. The log-dynamic brain: how skewed distributions affect network operations. *Nat. Rev. Neurosci.* **15**, 264–278 (2014).
24. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
25. Sul, J. H., Jo, S., Lee, D. & Jung, M. W. Role of rodent secondary motor cortex in value-based action selection. *Nat. Neurosci.* **14**, 1202–1210 (2011).
26. Siniscalchi, M. J., Phoumthipphavong, V., Ali, F., Lozano, M. & Kwan, A. C. Fast and slow transitions in frontal ensemble activity during flexible sensorimotor behavior. *Nat. Neurosci.* **19**, 1234–1242 (2016).
27. Arieli, A., Sterkin, A., Grinvald, A. & Aertsen, A. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science* **273**, 1868–1871 (1996).
28. Churchland, M. M. et al. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* **13**, 369–378 (2010).
29. Churchland, A. K. et al. Variance as a signature of neural computations during decision making. *Neuron* **69**, 818–831 (2011).
30. Kepcs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231 (2008).
31. Znamenskiy, P. & Zador, A. M. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* **497**, 482–485 (2013).
32. Kopeć, C. D., Erlich, J. C., Brunton, B. W., Deisseroth, K. & Brody, C. D. Cortical and subcortical contributions to short-term memory for orienting movements. *Neuron* **88**, 367–377 (2015).
33. Kusev, P. et al. Understanding risky behavior: the influence of cognitive, emotional and hormonal factors on decision-making under risk. *Front. Psychol.* **8**, 1–10 (2017).
34. Frank, M. J., Doll, B. B., Oas-Terpstra, J. & Moreno, F. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat. Neurosci.* **12**, 1062–1068 (2009).
35. Finn, E. S. et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* **18**, 1664–1671 (2015).
36. Wolff, M. J., Jochim, J., Akyürek, E. G. & Stokes, M. G. Dynamic hidden states underlying working-memory-guided behavior. *Nat. Neurosci.* **20**, 864–871 (2017).
37. Carnevale, F., de Lafuente, V., Romo, R., Barak, O. & Parga, N. Dynamic control of response criterion in premotor cortex during perceptual detection under temporal uncertainty. *Neuron* **86**, 1067–1077 (2015).
38. Chaisangmongkon, W., Swaminathan, S. K., Freedman, D. J. & Wang, X.-J. Computing by robust transience: how the fronto-parietal network performs sequential, category-based decisions. *Neuron* **93**, 1504–1517 (2017).
39. Shenoy, K. V., Kaufman, M. T., Sahani, M. & Churchland, M. M. A dynamical systems view of motor preparation: implications for neural prosthetic system design. *Prog. Brain. Res.* **192**, 33–58 (2011).
40. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null space: permitting preparation without movement. *Nat. Neurosci.* **17**, 440–448 (2014).
41. Rajan, K., Abbott, L. F. & Sompolinsky, H. Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys. Rev. E* **82**, 1–5 (2010).
42. Curto, C., Sakata, S., Marguet, S., Itskov, V. & Harris, K. D. A simple model of cortical dynamics explains variability and state dependence of sensory responses in urethane-anesthetized auditory cortex. *J. Neurosci.* **29**, 10600–10612 (2009).
43. Tavor, I. et al. Task-free fMRI predicts individual differences in brain activity during task performance. *Science* **352**, 216–220 (2016).
44. Luczak, A., Bartho, P. & Harris, K. D. Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron* **62**, 413–425 (2009).
45. Wu, H. G., Miyamoto, Y. R., Gonzalez Castro, L. N., Ölveczky, B. P. & Smith, M. A. Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nat. Neurosci.* **17**, 312–321 (2014).
46. Singh, P., Jana, S., Ghosal, A. & Murthy, A. Exploration of joint redundancy but not task space variability facilitates supervised motor learning. *Proc. Natl Acad. Sci. USA* **113**, 14414–14419 (2016).
47. Shim, Y., Philippides, A., Staras, K. & Husbands, P. Unsupervised learning in an ensemble of spiking neural networks mediated by ITDP. *PLoS Comput. Biol.* **12**, 1–41 (2016).
48. Colgin, L. L. Rhythms of the hippocampal network. *Nat. Rev. Neurosci.* **17**, 239–249 (2016).
49. Yamamoto, J., Suh, J., Takeuchi, D. & Tonegawa, S. Successful execution of working memory linked to synchronized high-frequency gamma oscillations. *Cell* **157**, 845–857 (2014).
50. Fries, P. Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu. Rev. Neurosci.* **32**, 209–224 (2009).

Acknowledgements

We thank J. Johansen and C. Yakoyama for critical comments on the manuscript, and M. Tatsuno and K. Watanabe for discussions about the analysis of correlation. This work was partly supported by KAKENHI (grants 16H01289 and 17H06036 to T.F.) from MEXT.

Author contributions

T.F. and T.K. designed the work and constructed the model. T.K. conducted numerical simulations and experimental and numerical data analyses. T. Handa performed behavioral and electrophysiological experiments and analyzed experimental data. T. Haga performed behavioral and lesion experiments, and R.H. performed behavioral training, surgery and histological staining. T.F., T.K., T. Handa and T. Haga wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41593-018-0263-5>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.F.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2018

Methods

Experimental procedure. All experiments were carried out according to the Animal Experiment Plan approved by the Animal Experiment Committee of RIKEN. The details of experimental procedures are given in ref.⁷. Briefly, head-restrained adult Long-Evans rats (male, 210–240 g, strain SLC) were trained to associate the licking of spouts with the delivery of a reward. We presented either of two pure tones (familiar cue tones: 13.0 kHz and 10.0 kHz, for 0.2 s) in a pseudo-random order as a cue for licking a spout located at the left or right side of the rat, respectively (Fig. 1a). The rats were required to lick a correct spout within 5.0 s from the cue onset to obtain a reward (0.1% saccharin water). If the licking response was incorrect, we did not deliver the reward and we prolonged the duration of the immediately following post-response period (3.0 s) by 5.0 s as an aversive experience. We continued the training until the correct rate finally reached a criterion (75%) without error correction. During the recordings, we presented the two familiar cue tones (10.0 kHz and 13.0 kHz) and five unfamiliar cue tones (10.5 kHz, 11.0 kHz, 11.5 kHz, 12.0 kHz and 12.5 kHz) with an occurrence probability of 80% or 20% (4% for each unfamiliar tone), respectively. A typical rat underwent several tens of unfamiliar trials per recording day; these unfamiliar trials were intermingled with a four-times larger number of familiar trials. The correct familiar cue trials were always rewarded, whereas the reward probability was linearly varied for unfamiliar cue trials along its cue-tone frequency: (left/right) = 10.5 kHz (0.17/0.83), 11.0 kHz (0.33/0.67), 11.5 kHz (0.5/0.5), 12.0 kHz (0.67/0.33), and 12.5 kHz (0.83/0.17). We trained 36 rats with familiar tones and only 21 reached the criteria for successful learning. After surgery, 15 were available for multi-neuron recordings and 8 of them finally yielded qualitatively and quantitatively satisfactory data for the succeeding analysis.

We recorded multiunit activity mainly from the deep layers (depth from pia matter: 1.0–2.0 mm) of the MFC (+2.7 to +3.6 mm anterior, 0.6–2.0 mm lateral of Bregma) through a 32-channel silicon probe consisting of 4 shanks (Neuro Nexus Technologies, USA), each with 2 tetrode sites separated vertically by 0.5 mm. We primarily analyzed the behavioral and neuronal data obtained on the first day of the recording sessions, when the rats were still not habituated to unfamiliar cues, but also analyzed the behavioural data from the second day for six rats. Spikes were sorted with a custom-made semi-automatic spike-sorting program, EToS⁵¹, and the sorted spike clusters were further analyzed manually using Klusters and NeuroScope^{52,53}. The number of isolated units was $n=37$ in rat 941, $n=68$ in 897, $n=22$ in 902, $n=12$ in 807, $n=44$ in 880, $n=40$ in 879, $n=50$ in 940 and $n=51$ in 949. Neuronal activity and behavioral performance were analyzed using MATLAB (MathWorks).

Muscimol injection in the MFC. We trained another seven rats by the same procedure as described before⁷, with a few minor changes: first, we presented familiar tones for 0.5 s throughout the training and experiments; second, a post-response period was extended to 4 s, and a temporal penalty for incorrect choices was shortened to 3 s. Once rats reached a performance criterion (a correct rate of more than 75%), the rats were used for a series of consecutive experiments. After the completion of training sessions, the rats underwent a second surgery under anesthesia (2–3% isoflurane) in order to make a tiny cranial window above the MFC, to allow the insertion of an injection needle in each hemisphere. On the next day (first day) and fourth day, the task performance was tested without any injections. On the second and third days, 0.1 M PBS and the GABA_A-receptor agonist muscimol (dissolved in 0.1 M PBS; 1 mg ml⁻¹; Sigma-Aldrich)⁵⁴, respectively, were infused prior to performing the task. Briefly, a microinjection syringe (Hamilton) filled with PBS or muscimol solution was set on a microsyringe pump (Legato 130, KD Scientific), which was installed on a micromanipulator (1760–61; David Kopf Instruments). We vertically inserted the 31-gauge needle into the MFC (+3.0 mm to Bregma, 1.0 mm to midline, −1.5 mm from pia surface) and injected PBS or muscimol in both hemispheres (0.3 μL total; 0.2 μL min⁻¹). The animals were left for half an hour and then were required to perform the task. We analyzed data on six out of seven rats, and excluded data from a rat that did not show licking behavior at all after the muscimol injection. The difference in behavioral performance between conditions were statistically estimated by a paired Student's *t*-test with Bonferroni-corrected criteria.

Histological analysis. Under deep anesthesia, animals were perfused intracardially with ice-chilled 0.9% saline followed by 4% paraformaldehyde (PFA) in 0.1 M phosphate buffer. Brains were removed and placed in PFA for a few weeks. Postfixed brains were sliced coronally into 100-μm-thick serial sections. We stained the sections with a thionin Nissl staining procedure to identify the injection site of muscimol in the dorsal frontal cortex.

Neuron and network models. We constructed a network model comprising three parts: an input layer, a reservoir network and a readout layer (Fig. 2a). The input layer has 200 neurons, each of which responds to the stimulus k at the following firing rate r_i :

$$r_i = (R_{\max}/L) \exp(-((i/N_{in}) - 0.1(k+1))^2/2(\sigma L)^2) \quad (1)$$

where i , R_{\max} , N_{in} and σ are the neuron index, the maximum firing rate, the number of input neurons, and the standard deviation, respectively. The values of these

parameters were set as $R_{\max} = 110$, $N_{in} = 200$ and $\sigma = 0.15$; $k = 1, \dots, 7$ specifies the preferred stimulus of the neuron, with $k=1$ and 7 corresponding to familiar high (FH) and familiar low (FL) stimuli, respectively, and $k=2, \dots, 6$ corresponding to five unfamiliar inputs (U1–5), respectively. The parameter L is the width of frequency tuning curves, and most results were calculated for $L=0.4$ except in Fig. 3g.

The reservoir network is essentially the same as the recurrent network model studied in ref.²¹, except for the introduction of NMDA receptors and background noise as well as minor modifications of model parameters. The reservoir network has 5,000 excitatory (E) and 1,000 inhibitory (I) leaky integrate-and-fire neurons obeying the membrane dynamics:

$$du_i/dt = -(u_i - V_{\text{leak}})/\tau + I_i^E + I_i^I + I_i^{IN} + I_i^{FB} + I_i^N \quad (2)$$

where $V_{\text{leak}} = -70$ (mV), $\tau = 20$ or 10 (ms) for excitatory and inhibitory neurons, respectively, and the refractory period is 1 (ms). Other abbreviations are defined below. If u reaches threshold $V_{\text{th}} = -45$ (mV), the neuron fires and u is reset at $V_r = -60$ (mV). Excitatory and inhibitory recurrent synaptic inputs, I_i^E and I_i^I , are described by the following equations:

$$I_i^E = -g_i^{\text{AMPA}}(u_i - V_E) - 0.1g_i^{\text{NMDA}}(u_i - V_E) \quad (3)$$

$$I_i^I = -g_i^{\text{GABA}}(u_i - V_I) \quad (4)$$

$$dg_i^X/dt = -g_i^X/\tau^X + \sum_j G_{ij}^Y \sum_k \delta(t - t_j^k - d_j) \quad (5)$$

where G_{ij}^Y and d_j are the weight and delay of synaptic connection from neuron j to neuron i , respectively; t_j^k is the k th spike time of neuron j ; and X is AMPA, NMDA or GABA. The reversal potentials of synaptic inputs are $V^E = 0$ (mV) and $V^I = -80$ (mV); synaptic time constants are $\tau^{\text{AMPA}} = 8$ (ms), $\tau^{\text{NMDA}} = 100$ (ms) and $\tau^{\text{GABA}} = 8$ (ms); and δ is the Kronecker's delta function. All synaptic conductances, g , are normalized by the membrane capacitance to have the dimension of 1 ms⁻¹. At excitatory-to-excitatory (E-to-E) connections, synaptic delays are chosen randomly from 1 to 3 (ms), and at other connections they are from 0.5 to 1.5 (ms). Synaptic input from input neurons I_i^{IN} , feedback input I_i^{FB} from readout neurons (see below), and background noise I_i^N are all excitatory and defined as:

$$I_i^Y = -g_i^Y(u_i - V_E) \quad (6)$$

$$dg_i^Y/dt = -g_i^Y/\tau^{\text{AMPA}} + \sum_j G_{ij}^Y \sum_k \delta(t - t_j^k - d_j) \quad (7)$$

where G_{ij}^Y is the synaptic weight from neuron j to neuron i and t_j^k is the k th spike time of presynaptic neuron j of input type Y (IN, FB or N). Here $G^N = 0.1$, and other weights obey lognormal distributions, as described later. In equation (7), spikes in I_i^N and I_i^{FB} are generated by non-stationary Poisson processes with the instantaneous firing rates given by equation (1) or those of the readout neurons, respectively. The feedback spikes are generated independently for individual postsynaptic reservoir neurons to avoid strongly correlated activation of these neurons. Background noise is given by a Poisson spike train of 20 (Hz).

The readout layer has two rate-based neurons, L- and R- neurons, which integrate spike inputs from the reservoir and undergo mutual inhibition. Their firing rates $r_{L,R}$ obey

$$dr_{L,R}/dt = -(r_{L,R} - r_{R,L})/\tau + \sum_i W_{i \rightarrow L,R} \sum_k \delta(t - t_i^k) \quad (8)$$

where $\tau = 50$ (ms). The value of $r_{L,R}$ is set to zero if it takes a negative value owing to the mutual inhibition and $W_{i \rightarrow L,R}$ is a readout connection from the neuron i in the reservoir to the L- (R-) neuron (see the next section in detail).

Network architecture. Among 5,000 excitatory neurons in the reservoir network, 2,000 neurons receive excitatory external input from input neurons through sparse connections. The connection probability $P = 0.01$ and their weights are set at 0.02. The wiring probabilities of excitatory-to-excitatory (E-to-E), excitatory-to-inhibitory (E-to-I), inhibitory-to-excitatory (I-to-E) and inhibitory-to-inhibitory (I-to-I) connections are $P^{EE} = P^{EI} = 0.1$ and $P^{IE} = P^{II} = 0.5$. The weights of E-to-E connections are generated according to the following log-normal distribution:

$$p(x) = \exp(-(\log x - \mu)^2/(2\sigma^2)) / (\sqrt{2\pi}\sigma x) \quad (9)$$

where $\sigma^{EE} = 0.8$ and $\mu^{EE} = \log(0.01) + (\sigma^{EE})^2$. The weights of E-to-I, I-to-E and E-to-I connections are fixed at 0.01, 0.002 and 0.0025, respectively.

Each readout neuron is projected to by a subset of reservoir neurons not receiving sensory input through readout connections $W_{i \rightarrow L,R}$ and feed their outputs back to another subset of reservoir neurons. The former and latter subsets of neurons consist of randomly and independently chosen 30% and 50%, respectively,

of such reservoir neurons as are connected to none of the input neurons. An overlap can exist between the L-neuron-projecting and R-neuron-projecting subsets. The weights of $W_{i \rightarrow L,R}$ are randomly chosen between 0 and 60. The weights of feedback connections G^{FB} obey a lognormal distribution given in equation (9), with the variance and mean given as $\sigma^{\text{FB}} = 1.2$ and $\mu^{\text{FB}} = \log(0.012) + (\sigma^{\text{FB}})^2$.

Learning protocol. The present network model learns to correctly associate external stimuli with choice responses by modifying the readout connections $W_{i \rightarrow L,R}$ through reinforcement learning. Recurrent connections are not modifiable to keep the lognormal weight distribution stable against spiking dynamics. The learning procedure is as follows: first, the network undergoes a prestimulus run for 100 (ms), which induces a baseline activity in the initial network state; subsequently, one of the familiar stimuli (FH or FL) is applied from 0 to 200 (ms); after the termination of stimulus followed by a delay period of 50 (ms), the network is allowed to make a decision if the activity of a readout neuron exceeds that of the other readout neuron by a criterion difference θ (Fig. 3b). If the network does not reach a decision criterion by 500 (ms) from the end of delay period, the trial is reset (failed) and a novel trial is initiated. In some simulations, the time limit of decision was set to 900 ms. If the decision by the network is correct, the readout weights are modified with a positive reward, whereas the weights are punished with a negative reward if the decision is incorrect. If the network fails to give a decision in a trial, the network is also punished. If the network reaches a decision criterion before the end of delay period, we discard this trial and start a new trial.

Readout connections are modified in terms of eligibility trace $e_{i \rightarrow L,R}$ (refs 2–17) as:

$$de_{i \rightarrow L,R}/dt = (a_i(t)(r_{L,R}(t) - r_{L,R}^{\text{slow}}(t)) - e_{i \rightarrow L,R})/\tau_e \quad (10)$$

where $a_i(t)$ is the average activity of presynaptic reservoir neuron, $r_{L,R}(t)$ and $r_{L,R}^{\text{slow}}$ are the firing rates of readout neurons and their low-pass-filtered versions, respectively, and s is the time scale of the eligibility. To measure the extent to which a particular connection contributes to decision-making, an eligibility trace is assigned to each readout connection and is calculated from correlations between $a_i(t)$ and high-pass-filtered readout activity ($r_{L,R}(t) - r_{L,R}^{\text{slow}}(t)$). The variables $a_i(t)$ and $r_{L,R}^{\text{slow}}$ were computed in terms of presynaptic and postsynaptic activities as:

$$a_i(t) = \sum_k 5 \exp\left(-\frac{t-t_i^k}{\tau_a}\right) \quad (11)$$

$$dr_{L,R}^{\text{slow}}/dt = (r_{L,R}(t) - r_{L,R}^{\text{slow}})/\tau_{\text{slow}} \quad (12)$$

where $\tau_a = 10$ (ms) and $\tau_{\text{slow}} = 50$ (ms). The term $(r_{L,R}(t) - r_{L,R}^{\text{slow}})$ detects a rapid change in $r_{L,R}(t)$ faster than $r_{L,R}^{\text{slow}}$, so the eligibility trace is increased by coincidence between a high presynaptic firing rate and a rapid increase in the readout activity.

We modify readout connections by reward expectancy U and the limited eligibility at learning step T :

$$W_{i \rightarrow L,R}(T+1) \leftarrow W_{i \rightarrow L,R}(T) + \alpha \Delta W_{i \rightarrow L,R}(T) \quad (13)$$

$$\Delta W_{i \rightarrow L,R}(T) = U(T)(e'_{i \rightarrow L,R} - \langle e'_{i \rightarrow L,R} \rangle) \quad (14)$$

where α is a learning rate, which was set as $\alpha = 10$ throughout this study; $e'_{i \rightarrow L,R}$ is the normalized eligibility trace of the i th connection defined below; and $\langle e'_{i \rightarrow L,R} \rangle$ is the average over all presynaptic reservoir neurons i . The normalized eligibility trace is defined as:

$$e'_{i \rightarrow L,R} = 5 \langle e_{i \rightarrow L,R} \rangle_i \tanh(e_{i \rightarrow L,R}/5 \langle e_{i \rightarrow L,R} \rangle_i) \quad (15)$$

where $\langle \dots \rangle_i$ denotes averaging over presynaptic reservoir neurons. We manually bound the value of $W_{i \rightarrow L,R}(T)$ below 60.

The modifications of readout connections occur only after one learning step (trial) finishes. After a success trial, U is taken to be positive (that is, potentiation of influential connections); U is negative after a failure trial (depression of such connections). More specifically, reward expectancy and decision criterion are adaptively modified during learning as follows:

$$U(T+1) = \begin{cases} U^{\max} - U(T)^{\text{slow}} & \text{rewarded trial} \\ -U^{\max} & \text{failure trial} \end{cases} \quad (16)$$

$$\theta(T+1) = \begin{cases} (1-\gamma_\theta)\theta(T) + \gamma_\theta\theta^{\max} & \text{rewarded trial} \\ (1-\gamma_\theta)\theta(T) + \gamma_\theta\theta^{\min} & \text{failure trial} \end{cases} \quad (17)$$

$$U^{\text{slow}}(T+1) = \begin{cases} (1-\gamma_U)U^{\text{slow}}(T) + \gamma_U U^{\max} & \text{rewarded trial} \\ (1-\gamma_U)U^{\text{slow}}(T) & \text{failure trial} \end{cases} \quad (18)$$

At an arbitrary T , $-U^{\max} \leq U(T) \leq U^{\max}$ and $\theta^{\min} \leq \theta \leq \theta^{\max}$. Parameter values are $U^{\max} = 1$, $\theta^{\min} = 10$, $\theta^{\max} = 30$ (values in Hz), $\gamma_U = 0.05$ and $\gamma_\theta = 0.005$. The variable $U^{\text{slow}}(T)$ is rapidly decreased while the criterion $\theta(T)$ is increased when the network has been rewarded in successive trials. Therefore, we may regard $U^{\text{slow}}(T)$ as reward expectancy at the learning step.

Fitting psychometric curves. Psychometric curves were calculated from the probability of left choices. In the training of our model, the criterion of decision θ was initially kept low and then gradually increased until it was finally fixed at 50 (Hz) after learning. This manipulation made the separation of neural trajectories easier and clearer at the decision timing without changing the qualitative behavior of the model with constant θ . The left-choice probabilities for familiar and unfamiliar stimuli were calculated for each rat over a few hundreds or a few tens of trials, respectively. The probabilities for all stimuli were calculated for each network model over 50 trials, except in Fig. 5c in which we simulated 100 trials. In a small fraction of trials (less than 1%), the network model did not reach the decision criterion within the time limit for simulations (less than 900 ms). For such trials, we assigned a ‘relative distance’ to the decision criterion θ at the time limit to the left-choice probability: $P_L = \{(r_L - r_R) + \theta\}/20$.

Unless otherwise stated, we fitted psychometric curves (the probability of left choices) by a nonlinear function in the least-squares method: $P_L = c(\tan(f-a) + b)$ (Supplementary Fig. 1). The variable f presents a normalized tone frequency $\{-1, -2/3, -1/3, 0, 1/3, 2/3, 1\}$, which correspond to $\{10, 10.5, 11, 11.5, 12, 12.5, 13\}$ (kHz) respectively in the rats, and $\{FL, U5, U4, U3, U2, U1, FH\}$ respectively in the models. After fitting a psychometric curve, we normalized the parameter c to define the sensitivity S such that $S=1$ refers to a linearly increasing psychometric curve $P_L(f) = (f+1)/2$ and $S=0$ to a highly biased psychometric curve, $P_L(-1) = P_L(-2/3) = \dots = P_L(+2/3) = 0$ and $P_L(1) = 1$. The larger the value of S , the more sensitive a psychometric curve. A conventional sigmoid function was not suitable to characterize sensitivity in psychometric curves.

We examined whether the evaluation of sensitivity remains stable for each rat and model if a different fitting scheme was used to characterize their psychometric curves. In this scheme, we used $P'(f) = b'\tan(a'f) + (1-b')\tan(a'')$ as a new fitting function, and defined a parameter as $s' = a' \times b'$ (Supplementary Fig. 2). This definition is reasonable as a' and b' scale the x -axis and y -axis, respectively, to modify the slope of the fitting function. Then, a new sensitivity S' was defined from s' by the normalization process mentioned previously.

Stability of parameter fitting for psychometric curves. We resampled a different set of 30 trials out of the entire data set (comprising several tens of unfamiliar trials and several hundreds of familiar trials) for each tone and rat, and generated a psychometric curve per rat. Then, for each rat we repeated this procedure 30 times to obtain 30 samples of psychometric curves and the corresponding 30 values of sensitivity S . We then plotted the resampled values of S in an increasing order of the original sensitivity of the rats, and examined whether the average values of the resampled sensitivity values preserve the original order of S .

Linear discriminant analysis. We examined the neural dynamics underlying decision-making in the rats and models by FDA. We grouped neural states of familiar trajectories at the time of decision-making (in the rats, the time of licking responses) into two groups: a group of neural states in left-choice trials and a group of neural states in Right-choice trials. FDA identifies such an $(N-1)$ -dimensional hyperplane that maximizes the ratio of the mean distance between the two groups (intergroup distance) to the sum of standard deviation from this hyperplane within each group (intragroup distance). Here N is the dimension of neural state, that is, the number of neurons in the population. We defined a one-dimensional line, called \mathbf{W}_{opp} , that is orthogonal to the identified $(N-1)$ -dimensional hyperplane.

Linear regression analysis. In addition to FDA, we used a regression method to identify the two (choice and stimulus) axes explaining differences in trajectories between the left- and right-choice conditions in the rats and between FH and FL stimuli in the models. Our method is the same as used in refs 24,37. Briefly, we obtained the average firing rates of 100 neurons that showed the largest differences in firing rate between pre- and post-stimulus onset in models, and the average rate of 30 neurons that showed the largest differences in firing rate between FH and FL stimuli after taking difference between pre- and post-stimulus onset in the rats. We used neural activities recorded from 100 (100) ms before to 500 (300) ms after the stimulus onset in models (rats). After Gaussian filtering, we calculated the z-scores of these activities, where the standard deviation of the filter was 10 ms for the models and 30 ms for the rats. Then, we performed a linear regression analysis:

$$r_{i,k}(t) = \beta_{i,1}(t) \times \text{choice}(k) + \beta_{i,2}(t) \times \text{stimulus}(k) + \beta_{i,3}(t) \quad (19)$$

where $r_{i,k}(t)$ represents the z-scored responses of neurons ($i = 1$ –100) at time t on trial k ; and $\beta_{i,1}(t)$, $\beta_{i,2}(t)$, and $\beta_{i,3}(t)$ represents the choice coefficient, stimulus coefficient and residual component, respectively. Here these coefficients are projected onto the subspace spanned by the ten largest PCs, and choice (k) and stimulus (k) are binary variables.

To quantify difference between regression trajectories for different conditions, we calculated Cohen's d for FH-left and FL-right trajectories at every 10 ms as follows:

$$d = (m_1 - m_2) / (((n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2) / (n_1 + n_2 - 2)) \quad (20)$$

where $m_{1,2}$, $n_{1,2}$ and $\sigma_{1,2}$ are the means, degrees of freedom, and standard deviations of FH-left (indexed by 1) and FL-right (indexed by 2) trajectories. Cohen's d represents an effect size to indicate the standardized difference between two means. We calculated the standard deviation of Cohen's d^{35} as follows:

$$\sigma_d = \sqrt{(n_1 + n_2) / (n_1 n_2) + \frac{1}{2} d^2 / (n_1 + n_2 - 2)} \quad (21)$$

Neural responsiveness. Neural responsiveness, x_u , characterizes the property of dynamics in a neural network by measuring how it evolves in response to a perturbative input given to a neural state u (Supplementary Fig. 1). To measure x_u , we set the membrane potentials in all reservoir neurons at the initial value of -60 (mV) and simulated the time evolution of the neural population up to time t_0 (100 ms), at which point we applied a perturbative input to the trajectory 30 times: $u_i(t_0) \rightarrow u_i(t_0) + \delta u_i$ ($i=1, \dots, N$; $k=1, \dots, 30$). Choosing a different vectored value of perturbation δu^k randomly at every time, we obtained 30 different perturbed trajectories $u^{(k)}$ ($k=1, \dots, 30$), all of which were perturbed at the same state $\{u_i\}$ by δu^k . Each component of δu^k is chosen from a uniform probability distribution of $[-0.05, 0.05]$ (mV). To evaluate the effect of the perturbation on state evolution, we measured the average distance D between the perturbed trajectories and an unperturbed trajectory $u^{(0)}$ at time $t_1 = 500$ (ms) as follows. The time of measurement did not change the essential results. Let $\{x_{PCi}^{(k)}\}$ be the i th PC of neural state at time t_1 along the k th perturbed trajectory $u^{(k)}(t_1)$, where $k=0$ refers to the unperturbed trajectory for convenience. If the network model reaches the decision criterion before time t_1 , $x_{PCi}^k(t_1)$ is calculated at this time. We define x_u for the perturbed state u as:

$$\chi_u = \left(\frac{1}{31} \right) \sum_{k=0}^{30} D(x_{PCi}^{(k)}, \langle x_{PCi}^{(k)} \rangle_k) / D(L, R) \quad (22)$$

where $D(a, b) = \left(\sum_{i=1}^5 (a_{PCi} - b_{PCi})^2 \right)^{1/2}$ is the Euclidean distance between state a and state b on the five-dimensional PC subspace; $\langle x_{PCi}^{(k)} \rangle_k$ is the average of the perturbed and unperturbed trajectories; and the distance from the average trajectory is averaged over the trajectories. $D(L, R)$ provides a normalization factor, and is the distance between left-choice and right-choice points on the PC subspace, where the choice points refer to the averaged neural states at which the network model reaches left or right choice.

We emphasize that the neural responsiveness is assigned to each neural state rather than each trajectory. Therefore, a neural trajectory could show a broad range of neural responsiveness depending on the specific state at which a perturbative input was given. In Fig. 5a,b,f, we plotted perturbed neural trajectories starting from various neural states with low to high x_u values. In Fig. 5d–g, we defined the neural responsiveness of a model network by summing up the x_u values of 50 neural states. Python scipy library (pearsonr) was used to calculate Pearson's correlation coefficients.

Trial-by-trial variability in rats and models. We selected those neurons that satisfy the following conditions. First, the average firing rates should be greater than 1 Hz (2 Hz) during the whole task period [−300 ms, 500 ms] ([−50 ms, 500 ms]) in the rats (or the models), where the origin 0 ms is the stimulus onset. Second, differences in firing rates before and after the stimulus onset should be greater than 2.5 Hz in the rats and 30 Hz in the models. Then, we applied

PCA to the activity of this neural ensemble during a pre-stimulus epoch (a 300 (50)-ms-long interval prior to stimulus onset) and a post-stimulus epoch (a 300 (500)-ms-long interval following stimulus onset) in the rats (the models). We computed the logarithm of cumulated product of PC variances for each condition: $\text{var}_{\text{pre}, \text{post}}^{\text{pre}, \text{post}}(\text{tone}, \text{choice}) = \sum_{i=1}^N \log(\text{var}_i^{\text{pre}, \text{post}}(\text{tone}, \text{choice})) / N$ where N is the number of selected neurons (if N is larger than 50, we set as $N=50$) and $\text{var}_i^{\text{pre}, \text{post}}(\text{tone}, \text{choice})$ is the variance of the i th PC for the pre- and post-stimulus ensemble for a given tone and choice. Then, we averaged $\text{var}_{\text{post}}^{\text{post}}$ across all conditions for each rat and network and normalized it by dividing $\text{var}_{\text{pre}}^{\text{pre}}$ to define the normalized variability. We performed the analysis in eight rats and 31 successful learners. To estimate the fluctuation of sampling, we resampled 90% of trials 50 times for each condition and each rat, and calculated 50 values of trial-by-trial variance for each rat. In Fig. 6b (right), error bars in the abscissa show the standard deviations of the resampled trial-by-trial variances; and error bars in the ordinate show standard deviations of sensitivity S for the resampled trials shown in Fig. 1f. In Fig. 6c, we calculated correlation coefficients between the resampled trial-by-trial variances and resampled sensitivities in the rats by randomly picking up them.

Statistics. We report means and standard deviations throughout unless otherwise noted. Correlation was by Pearson's r unless otherwise noted. All tests were two-tailed Student's t -tests unless otherwise noted. The data distribution was assumed to be normal unless otherwise stated, but this was not formally tested. No data points were excluded from the analyses unless otherwise stated. No statistical methods were used to predetermine sample sizes in both experiment and modeling, but our sample sizes in the experiment are similar to those reported previously^{37,9,44} and all networks were able to be trained to full performance in simulations. Data collection and analysis were not performed blind to the conditions of the experiments, as this did not apply to our simulations. Data collection and assignment to experimental groups also did not apply, because all networks were equivalent before training.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability. The in-house program codes used for numerical simulations and data analysis are available at GitHub (https://github.com/labyrinth2shell/individual_differences).

Data availability

The experimental data that support the findings in this study are available from the corresponding author upon reasonable request.

References

51. Takekawa, T., Isomura, Y. & Fukai, T. Spike sorting of heterogeneous neuron types by multimodality-weighted PCA and explicit robust variational Bayes. *Front. Neuroinform.* **6**, 1–13 (2012).
52. Harris, K. D., Henze, D. A., Csicsvari, J., Hirase, H. & Buzsáki, G. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J. Neurophysiol.* **84**, 473–478 (2000).
53. Isomura, Y., Harukuni, R., Takekawa, T., Aizawa, H. & Fukai, T. Microcircuitry coordination of cortical motor information in self-initiation of voluntary movements. *Nat. Neurosci.* **12**, 1586–1593 (2009).
54. Narayanan, N. S., Horst, N. K. & Laubach, M. Reversible inactivations of rat medial prefrontal cortex impair the ability to wait for a stimulus. *Neuroscience* **139**, 865–876 (2006).
55. Nakagawa, S. & Cuthill, I. C. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev.* **82**, 591–605 (2007).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

All experimental data were collected by using an in-house LabVIEW code, which was developed for our previous recording study (Handa et al., Cereb Cortex 2017). The code will be available upon request. Neural network simulations were performed by in-house C++ codes. The simulation codes are available at GitHub (https://github.com/labyrinth2shell/individual_differences).

Data analysis

The primary softwares used in the present data analysis and neural network simulations were in-house made. The former was written in python and the latter was written in C++. In-house computer codes for the major data analysis will be made available through GitHub (https://github.com/labyrinth2shell/individual_differences). Spikes were sorted with custom-made EToS ver 3 (<http://etos.sourceforge.net/>) and further analyzed by publicized software tools Neuroscope and Klusters (<http://neurosuite.sourceforge.net/>). Neuronal activity and behavioral performance were analyzed using MATLAB (2007 ver., The MathWorks, Inc.).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The experimental data necessary for replicating the present results will be available upon request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	As the task was difficult for the rat, we trained as many rats as possible every week. Predetermination of sample size was practically difficult and not attempted. After preliminary studies, 36 rats were trained on the task, 21 reached the criteria for success rate (>75%), after surgery 15 were left for multi-neuron recordings, and 8 rats finally yielded qualitatively and quantitatively satisfactory data (the number of neurons n=12 ~ 68) for modeling studies. Our experimental results as well as computational models, in which we can obtain sufficiently large datasets, consistently suggested certain relationships between neural trajectory dynamics and behavioral variance. We therefore concluded that the data size is sufficient for addressing the behavioral and neural characteristics of individual animals. In the lesion experiment, we analyzed the behavior of 6 rats, which were different from any of the above rats. The data size was chosen such that it is similar to the sample size in the analysis of neural activity.
Data exclusions	To investigate the individual choice strategy, we only used data from the well performing animals after the behavioral training, and excluded data from the other poorly performing animals (criteria for success rates > 75%). The success rate should be well above the chance level and the values around 80% have been conventionally used. We adjusted the value according to this convention and the difficulty of the present task. In addition, we even excluded data from well-behaving animals if the data was too small (i.e., n<12) to study neural population responses with sufficient statistics. In the lesion experiment, we excluded data from a rat which did not show licking behavior at all after muscimol injection.
Replication	In this experiments, the overall choice behavior of animals and neural responses to familiar stimuli were well replicated across animals. In addition, our lesion experiment impaired the learned decision behavior in all the rats that underwent this experiment. However, the focus of this study was the individual differences exhibited by the rats for unfamiliar stimuli and the possible neural origin of this behavioral variance. First, the large variation of behavioral tendency across rats is considered to be meaningful because the tendency remained stable within each rat across the two days of subsequent recordings. Second, although the behavioral tendency varied across rats, significant correlation was found between the behavioral variances and trial-by-trial variability in neural responses across the rats. Further, this correlation could be repeatedly replicated in a large number of simulation trials in the model. Thus, our study qualitatively and quantitatively replicated the broad spectrum of trial-by-trial variabilities in behavioral and neural data together with the relationships between them.
Randomization	Subjects were chosen at random for the training of the task, but subjects for recordings were chosen on the basis of criterion for the task performance after the training sessions. Presentation of tone stimuli was randomized according to certain probabilities.
Blinding	In the current study, investigators were blinded to group allocations during data collections and data analysis. Data collection and analysis were not performed blind to the conditions of the experiments as this did not apply to our simulations.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Unique biological materials
<input checked="" type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology
<input type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Human research participants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals male, Long-Evans rats (210-240 gr, 6-7 weeks old)

Wild animals This study did not involve wild animals.

Field-collected samples This study did not involve samples collected from the field.