Localized and global computation for integrating prior value and sensory evidence in the mouse cerebral cortex

Kotaro Ishizu[1], Shosuke Nishimoto[1,2], Akihiro Funamizu[1,2*]

1: Institute for Quantitative Biosciences, University of Tokyo, Laboratory of Neural Computation, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, JAPAN

2: Department of Life Sciences, Graduate School of Arts and Sciences, University of Tokyo, 3-8-2, Komaba, Meguro-ku, Tokyo 153-8902, JAPAN

Correspondence:

Akihiro Funamizu, Ph.D.
Laboratory of Neural Computation
Institute for Quantitative Biosciences (IQB), University of Tokyo
1-1-1 Yayoi, Bunkyo-ku, Tokyo, 113-0032, JAPAN
E-mail: funamizu@iqb.u-tokyo.ac.jp
Phone: +81-3-5841-7862

Abstract

Adaptive behavior requires integrating prior knowledge of action outcomes and sensory evidence for making decisions while maintaining prior knowledge for future actions. As outcome- and sensory-based decisions are often tested separately, it is unclear how these processes are integrated in the brain. In a tone frequency discrimination task with two sound durations and asymmetric reward blocks, we found that a temporal sequence of neurons in the medial prefrontal cortex of mice updated the prior reward expectations with incoming sensory inputs to compute sensory-confidence-dependent action values. The sensory inputs and choices were selectively decoded from the auditory cortex irrespective of reward priors and the secondary motor cortex, respectively, suggesting localized computations of task variables are required within single trials. In contrast, all the recorded regions represented prior values that needed to be maintained across trials. These results propose localized and global computations of task variables in different time scales in the cerebral cortex.

Introduction

Optimal decision-making requires estimating the current context from sensory inputs and making choices based on expected outcomes[1–5]. The neural basis of sensory-based context estimation is often investigated as perceptual decision-making in which humans and animals estimate contexts by accumulating sensory evidence[6,7]. In contrast, outcome-based action selection is investigated as value-based decisions in which animals update action values with recent choices and outcomes[8–10]. Although some theoretical and physiological studies integrate perceptual and value-based decision making[11,12], it is unclear how these processes are integrated into the brain to optimize behavior.

In theory, prior knowledge of the context probability and action values are used to compute the decision threshold of choices[1,13]. For simplicity, context probability was fixed but choice outcomes were varied in our task to investigate prior values[14]. Based on the decision threshold and sensory inputs, subjects made a choice. The choice outcome then determined the prior values in the next trial so that the values are maintained across trials. Our study investigates how the brain integrates the prior value and sensory input for a current choice while maintaining the values for future actions.

The cerebral cortex, including the medial prefrontal cortex (mPFC), the secondary motor cortex (M2) and the auditory cortex (AC), are involved in perceptual or value-based decisions. The AC modulates sensory-evoked activity with reward expectations, attention, and task engagement[15–18], although previous study finds that the non-sensory modulations in AC alone are not enough to drive behavior[5]. The mPFC neuron activity represents action values, rewards, or choices[19,20], and these neurons are causally involved in value-based decisions[21,22]. A recent study reported that the mPFC represents a confidence value[12] that combines prior values and sensory inputs, although it is unclear where and how prior values and sensory inputs are represented and integrated in the brain. M2 represents action values faster than other brain areas, including the striatum and mPFC, in rats[10]. Part of M2, the frontal orientation field (FOF), is necessary in auditory perceptual decision-making tasks[23,24]. These studies suggest a hypothesis that the cerebral cortex, including the sensory and frontal cortices, are orchestrated to integrate prior knowledge and sensory stimuli, although the role of each cortical region is still unclear.

Here, we used a tone-frequency discrimination task for head-fixed mice based on our previous studies[5,14,25], in which mice discriminated a tone frequency to choose between a spout on the left and another on the right. Our task presented either a long or short auditory stimulus in each trial and biased the choice outcome according to a block of trials. We found that mice made accurate choices in response to long-duration tones compared to their choices in response to

short-duration tones and made biased the choices in the asymmetric reward blocks, indicating that mice utilized both sensory uncertainties and reward expectations to guide behavior. Our brainwide electrophysiological study showed that a temporal sequence of neurons in the mPFC modulated the prior-value representations with incoming sensory inputs to compute sensory-confidence-dependent action values, i.e., confidence values. The value computation requires sensory inputs that were decoded from the AC irrespective of asymmetric reward blocks. The choices of the mice are represented by M2 activity. These results suggest the localized computations of confidence values, sensory inputs, and choices that were needed within single trials. In contrast, the prior values were equally represented by the activity across the three cortical regions, suggesting global maintenance. These results propose a localized and global computation of task variables that were needed within and across trials, respectively, in the cerebral cortex.

Results

We first performed the tone frequency discrimination task and showed that mice exploit not only the tone frequencies but also tone durations to make more accurate choices in long-sound trials than in short-sound trials. The outcome experiences after short-duration or difficult-frequency sounds have large effects on the next trial compared to the effects of long-duration or easy-frequency sounds. Based on these results and a previous study[12], we modeled how the prior value is sequentially modulated with sounds within a single trial by reinforcement learning (RL). An electrophysiological study showed that such value modulation is represented by a population sequence of neural activity in the mPFC. Decoding analysis showed that AC neurons have stable sound-frequency readouts orthogonal to prior values, while the M2 neurons had choice readouts. The prior values were decoded from all 3 recorded regions before and during the sound presentation.

Performance of mice depends on stimulus durations and asymmetric reward blocks in a tone frequency discrimination task.

To investigate how neurons in the cerebral cortex integrate the prior value and sensory evidence, we updated our previous tone-frequency discrimination task (**Figure 1A**)[14]. Mice were head-fixed and placed on a treadmill. Each trial started by moving the spout away from the mice. 1.5 s after moving the spout, we presented either a long (1.0 s) or a short sound. The duration of the short sound was either 0.2 or 0.4 s for 5 or 2 mice, respectively. In the other mouse, the duration of the short sound was 0.4 s for 1 session and 0.2 s for the remaining 4 sessions (8

mice in total). The sound consisted of short-duration pure tones (0.03 s), in which the frequency of each tone was either low (5–10 kHz) or high (20–40 kHz), named the tone cloud[5,25–27]. The proportion of high-frequency tones in the tone cloud was selected from 6 settings (0, 0.25, 0.45, 0.55, 0.75, and 1), and the dominant frequency determined the correct choice. The stimulus settings were named easy (0 or 1), moderate (0.25, 0.45), and difficult (0.45, 0.55) depending on the proportions. 0.5 s after the end of the sound, the spouts were moved forward, and the mice then chose the left or right spout. The correct or error choice provided 10% sucrose water (2.4 μl) or a noise burst (0.2 s), respectively.

After the first 40 trials in each session, we biased the reward size of the left and right spouts for correct choices in each block of 90–130 trials. The size of reward biases varied from 2.8–2.0 μl to 3.8–1.0 μl (large – small reward pair) for aligning the reward-dependent choice biases in each session. We defined the asymmetric reward blocks as left or right blocks according to the large-reward side. Left and right blocks were alternated until the end of the session.

We analyzed the choice behaviors of 8 mice in 122 sessions. The high-frequency tones were associated with the correct choice in the left or right spout in 4 mice for each side. In the example session (**Figure 1B and C**), the correct-response rates of the mouse were biased by the asymmetric reward blocks. In short-sound trials, the choice behavior was less accurate and more biased toward the large-reward side than in long-sound trials. The block-dependent choice biases were larger in short-sound trials than in long-sound trials (8 mice in 122 sessions, **Figure 1D**). Psychophysical kernels tested which tone-cloud timings and asymmetric reward blocks correlated with choices (**Figure 1E, Methods**)[14,28]. In long-sound trials, the choices of mice were significantly correlated with almost the entire duration of sounds, suggesting that long sounds provided longer sound evidence than the short sounds, thus resulting in accurate choices. The choices were also biased by the asymmetric reward blocks (linear mixed-effects model, $p = 1.3e-9$ and $5.5e-8$ for long- and short-sound trials, respectively; 8 mice in 122 sessions). These results were consistent with our previous study without neural recording[14].

A previous study showed that compared to high-confidence stimuli, low-confidence visual stimuli affected mouse choices in the next trial[12]. This matched our results showing that the previous rewarded trials with difficult tone clouds affected the choice in the next trial more than those with easy stimuli (**Figure 1F left**), suggesting that sensory confidence affected choice behavior. This was further supported by our results; the outcomes based on short sounds affected the choices more than those based on long sounds (**Figure 1F right**).

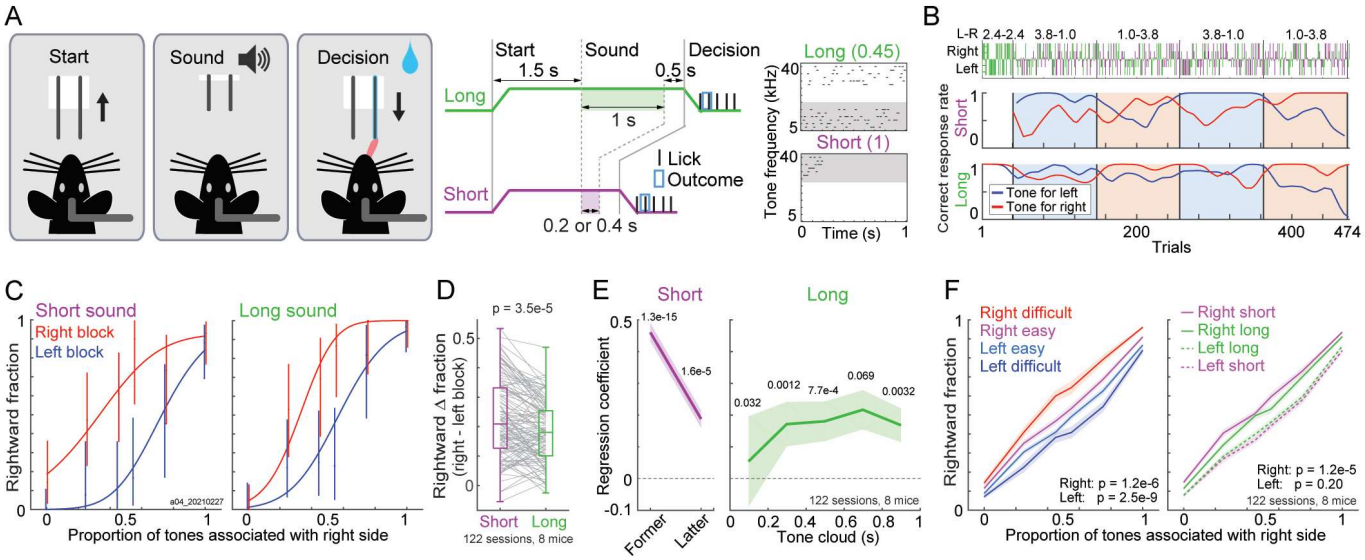


Figure 1. Tone-frequency discrimination task with two sound durations and asymmetric reward blocks.

A. Task structure. The task was based on our previous study[14]. Trial started by moving the spouts away from the mouse (Start). After a delay of 1.5 s, a sound stimulus was presented (Sound). The spouts were moved 0.5 s after the end of the sound, and mice licked the left or right spout to receive a water reward (Decision). Each trial had either a long or short sound duration. The right panels show example tone clouds. The parentheses show the proportion of high frequency tones. **B.** Choice behavior in an example session. The correct-response rate during tones associated with a left or right choice is a moving average with a Gaussian function ($\sigma = 10$ trials). The top panel shows the stimulus length and proportion of tones assciated with a left or right choice. Blue and red boxes show the asymmetric-reward blocks (left and right blocks). **C.** Psychometric function of choice behavior in **B**. Means and 95% confidence intervals. **D.** Stimulus duration affected reward-dependent choice biases. Rightward Δ fraction is the difference in the average fraction of rightward choices between the right and left blocks (linear mixed-effects model, 8 mice in 122 sessions) (central mark in the box: median; edges of the box: 25th and 75th percentiles; whiskers: most extreme data points not considered outliers (beyond 1.5 times the interquartile range), here and throughout). **E.** Psychophysical kernels. Logistic regression analyzed how tones at each time point and the asymmetric rewards correlated to the choice behavior. (Means and robust standard errors, linear mixed-effects model, 8 mice in 122 sessions). **F.** Sensory confidence affected choices in the next trial. Left: we categorized the rightward fraction of mice based on the tone difficulties in the previous correct trials. The choices were compared between the previous easy and diffcult stimuli. Right: we categorized the rightward fraction of mice based on the tone durations in previous correct trials. The choices were compared between the previous short and long durations. Means and robust standard errors (linear mixed-effects model, 8 mice in 122 sessions).

Reinforcement learning showed gradual modulation of prior values with sensory inputs to compute confidence values within single trials.

To understand the computation behind the behavior, we used the RL model to analyze the choices the mice made (**Figure 2A**). Based on the behavioral results (**Figure 1F**) and previous studies[12,14,29], we hypothesized that the updating of reward expectations for left and right choices (i.e., prior values) depended on the tone frequencies and durations, i.e., sensory confidence. The RL model had independent perceptual uncertainties for the long and short sound durations. First, the prior values were used to compute the decision threshold. Based on the threshold and perceived sensory inputs, the model estimated the choice probability in each trial (**Methods**). The confidence values were analyzed by multiplying the prior values with perceived sensory inputs and used to update the prior values. Thus, the RL model made choices by combining the prior values and sensory inputs, while it updated the prior values using confidence values.

In the example session, our model captured the block-dependent changes in the correct response rate of the mouse by updating the prior values (**Figure 2B left**). The simulated choices captured the differences in choice accuracy and block-dependent choice biases between sound durations (**Figure 2B right**). Our RL model matched the mouse choices better than the conventional psychometric functions or an RL model using only prior values (**Figure 2C and S1**)[14]. Parameters in the RL model suggested that the stimulus sensitivity (i.e., perceptual uncertainty) for short sounds was significantly larger than that for long sounds (**Figure 2D**). As our model parameterized the stimulus sensitivities for the duration of short- and long-sounds, we could simulate how the prior values were gradually modulated by sensory inputs to form the confidence values before, during (short-sound duration), and at the end of long-sound presentations by showing the proportion of right-action values relative to the left (**Figure 2E**). We found that the difference in the relative right action values (Δ right-action value) between the right- and left-correct trials gradually increased during the sound presentations, while the difference between the left and right blocks decreased (**Figure 2F**).

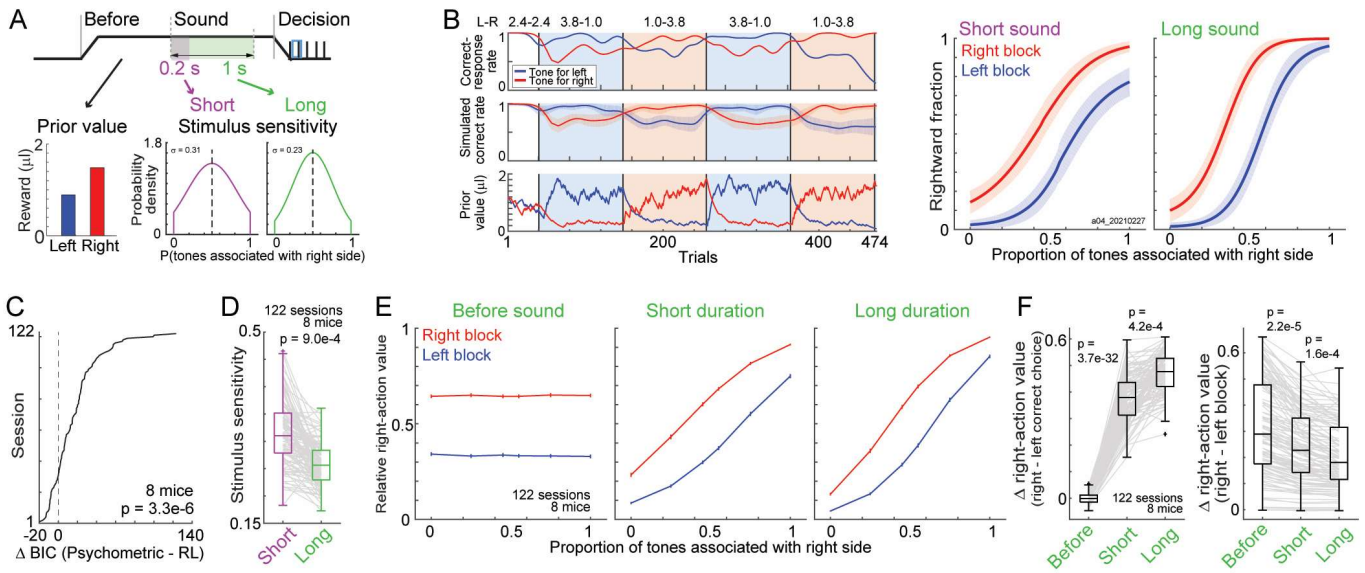


Figure 2. Reinforcement learning model with confidence values fit mouse behavior.
A. Scheme of the reinforcement learning (RL) model. The RL model had reward expectations for left and right choice (prior values), which were shared between the long and short-sound trials. Stimulus sensitivities were independently set for the short and long sounds. **B.** Choice simulation with the RL model in example session. Data are presented as in **Figure 1B**. Based on the fitted parameters in RL model, we simulated the choices of a mouse 100 times. Means and standard deviations. **C.** Comparison between the RL model and psychometric function. 122 sessions in 8 mice were sorted based on the difference of fitting between the models (Δ Bayesian information criterion: Δ BIC) (p value in linear mixed-effects model). **D.** Comparison of stimulus sensitivities for long- and short-sound durations in the RL model (linear mixed-effects model, 8 mice in 122 sessions). **E.** Modulation of prior values with sensory inputs to compute confidence values. Relative right-action value was the ratio between the right and left action values which were plotted for before sounds, short-duration sounds, and long-duration sounds. Means and starndard errors (8 mice in 122 sessions). **F.** Summary of value modulation before and during sounds. Left: value differences between right- and left-correct trials. Right: value differences between right and left blocks (linear mixed-effects model, 8 mice in 122 sessions).

The temporal sequence of mPFC neurons represents the modulation of prior values into confidence values.

During the tone-frequency discrimination task, we recorded the electrophysiological activity of neurons in the cerebral cortex with a Neuropixels 1.0 probe[30]. In each session, i.e., each day, we targeted one or two cortical areas and inserted Neuropixels probes. We repeated the procedure on multiple days to target the mPFC, M2, or AC. We also targeted the posterior parietal cortex (PPC); however, automated spike sorting with Kilosort 3 detected a relatively small number of neurons in the PPC compared to the number detected in the other 3 regions. We decided not to analyze the data from PPC.

In the mPFC, we analyzed the neural activity of both hemispheres in 8 mice in 34 sessions (**Figure 3A and S2A**). After spike sorting with Kilosort 3, we identified 9503 neurons in total

based on the approximate depth of electrodes from the brain surface (**Methods**). We then identified 3077 task-relevant neurons that significantly increased the activity at one of the 132 time-windows during the task compared to the baseline activity ($p < 1e-10$ in one-sided Wilcoxon signed rank test) (**Figure 3B left**). Each time window was 0.1 s between -1.5 and 2.5 s from sound onset in long-sound trials (40 windows) and between -1.5 and 1.7 s in short-sound trials (32 windows). The time windows were also set between -0.5 and 2.5 s from the choice timing both in long- and short-sound trials (30 x 2 = 60 windows; 40 + 32 + 60 = 132 windows in total). The baseline for sound-onset activity was defined as the activity at -0.2–0 s from the trial start, i.e., spout removal. The baseline for choice activity was defined as the activity at -0.2–0 s from the spout approaching, i.e., between the sound end and the choice.

We first found that the task-relevant neurons were sequentially recruited during the long-sound trials, rather than mainly observing the ramping-activity of neurons toward the end of sounds. In the mPFC, 5.6% of neurons on average were identified as task relevant in each time window of 0.1 s (**Figure 3B right**). 18.3% of neurons on average were newly recruited every 0.1 s during the task.

We analyzed the activity of each neuron in detail. Example neuron increased the activity when the mouse selected the right spout irrespective of tone categories, suggesting choice encoding (**Figure 3C**). We investigated whether the mPFC neurons represented the choices or sounds by comparing the tone indices between the correct and error trials. The tone index was defined as the activity difference between the right- and left-category tones ranging between -1 and 1 (**Methods**). If a neuron represented a tone category, the signs of tone indices between the correct and error trials became identical. However, the tone indices flipped for the choice representation. We found that the tone indices of mPFC neurons were flipped between the correct and error trials (**Figure 3D**), suggesting choice encoding. The absolute values of tone indices were increased from the initial to the end of long sounds in correct trials (Wilcoxon signed rank test, $p = 6.2e-6$, 425 neurons), suggesting that the choice representations became precise during sounds. The tone indices at the end of short sounds were correlated to those at the corresponding duration of long sounds, rather than those at the end of long sounds (**Figure S3A**), suggesting that the activity was gradually shaped from the sound onsets.

To analyze the left- and right-choice encoding neurons at the same time, we aligned the activity of neurons to the preferred side. The preferred side (left or right choice) was defined once based on the activity at the time window with the largest difference from the baseline activity in long-sound correct trials (one-sided Wilcoxon signed rank test). For example, when the activity was higher in left-choice trials than in right-choice trials, the preferred side was defined as left. Then, the preferred block and tone were defined based on the preferred side: the preferred block

had a large reward amount on the preferred side, and the preferred tone was associated with the reward on the preferred side. For example, the preferred block and tone for a left-choice preferred neuron were defined as the left block and left-rewarded tone frequency, respectively. The aligned population activity of mPFC neurons gradually separated between tone categories in correct long-sound trials (**Figure 3E**).

We analyzed the neural activity before sound onset. Example neuron showed increased activity in the left block irrespective of choice (**Figure 3F**). Before sound, the activity of mPFC neurons was modulated by blocks rather than by choices (**Figure S3B**). We investigated how the task-relevant neurons in the mPFC changed their activity before and during the correct long-sound trials (**Figure 3G and S4A**). Before sound onset, the activity was significantly mainly modulated by blocks (ANOVA p value). At the initial phase of sounds, activity was modulated both by the choices and blocks, while at the end of long sounds, activity was only modulated by choices. These gradual shifts in activity from block modulations to choice modulations were consistent with the group of neurons activated before or at the end of long sounds (**Figure S5**). We summarized how the block- and choice-dependent activity changed before and during long sounds (**Figure 3H**). The choice and block modulations were increased and decreased, respectively, consistent with the gradual modulation of prior values into confidence values in the RL model (**Figure 2E and 2F**).

We analyzed the neural activity of both hemispheres of M2 from 8 mice in 34 sessions. Similar to the results in the mPFC, M2 neurons represented the choices or asymmetric reward blocks (**Figure S6**). During correct long-sound trials, the choice modulation of activity gradually increased (**Figure S6H**). In contrast, the block modulation of activity was weak in M2 and did not show a decrease in block modulation during sounds. These results suggest the choice representations in M2.

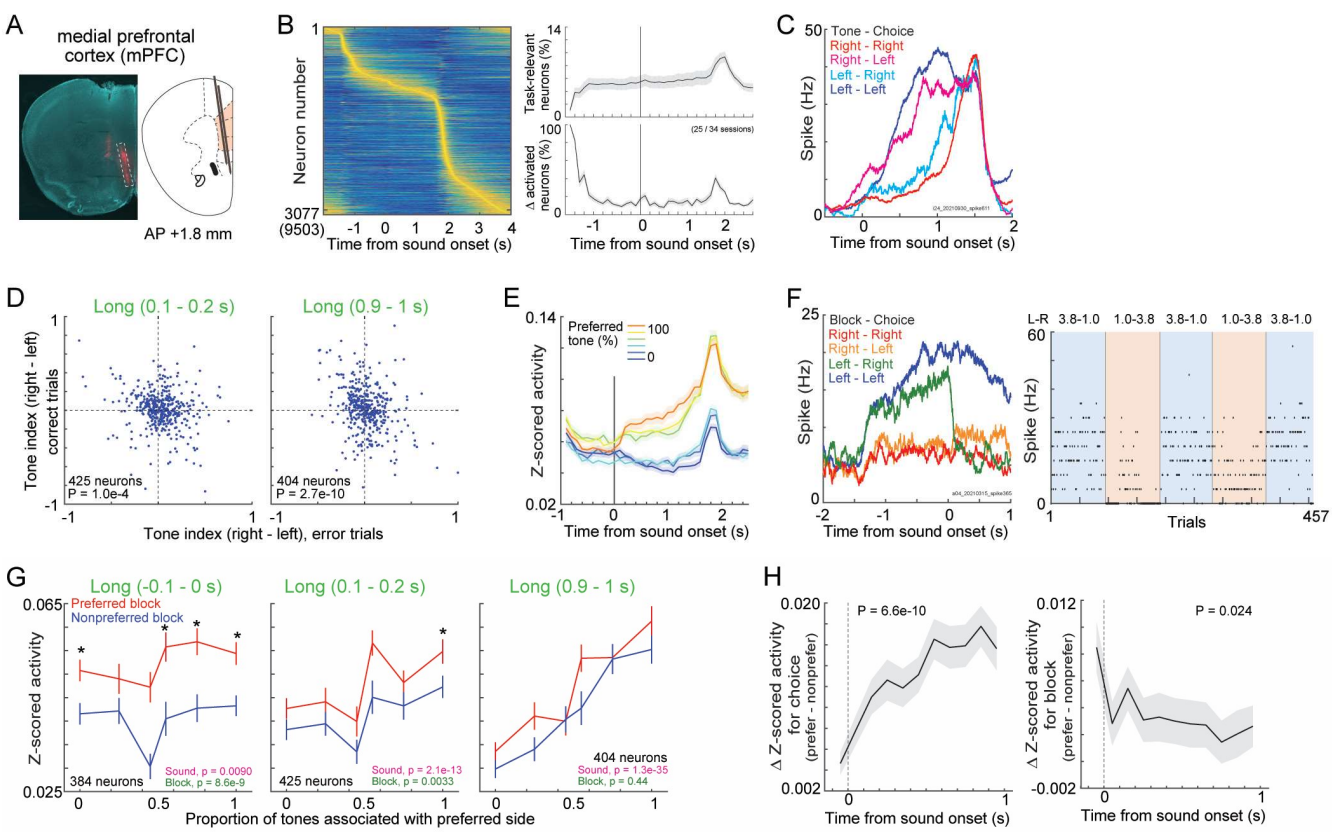


Figure 3. Temporal sequence of neurons in the medial prefrontal cortex represents the modulation of prior values into confidence values during sounds.
A. Location of Neuropixels 1.0 probe for the medial prefrontal cortex (mPFC) in an example mouse. The weak red color at the relative center of the brain slice was the location of the M2 probe. **B.** Average activity of task-relevant neurons in the mPFC (3077 out of 9503 neurons, $p < 1e\text{-}10$ in one-sided Wilcoxon signed-rank test). Left: average activity of each neuron was normalized between 0 and 1, and sorted by the peak activity timings. Right: proportion of task-relevant neurons in each time window of 0.1 s (up). Proportion of task-relevant neurons newly recruited from the previous time window (Δ activated neurons) (bottom). Means and standard errors. **C.** Average activity of one neuron after sound onset. The neuron showed increased activity when the mouse selected the left spout irrespective of sounds. **D.** Tone indices at the beginning and end of long sounds. Tone indices in correct and error trials were flipped, suggesting the choice representation. Task relevant neurons with a significant increase in activity at the analyzed time window were shown. **E.** Means and starndard errors of activity of task-relevant neurons in each time window. The activity of neurons were aligned to their preferred tones. **F.** Average activity of one neuron before sound onset. The neuron showed increased activity in the left block irrespective of choices. Right: neural activity between -0.2 and 0 s from sound onset in each trial. **G.** Activity of task-relevant neurons before, at the beginning, and at the end of long sounds in correct trials. The activity was compared between sound categories and blocks with ANOVA (p values in the bottom right) (* $p < 0.01$ in Wilcoxon signed rank test). Medians and robust standard errors. **H.** Modulation of choice- and block-dependent activity before and during sounds. The left panel shows the comparison of the activity between the preferred and nonpreferred sides. The right panel shows the comparison of the activity between the preferred and nonpreferred blocks. Medians and robust starndard errors. Robust linear regression showed that the differences in activity between choices or blocks were significantly increased or decreased, respectively, consistent with the confidence-value computation in the RL model (**Figure 2F**).

Neuronal activity in the AC represents the prior values and sensory inputs.

We analyzed the neuronal activity in both hemispheres of the AC from 7 mice in 29 sessions (**Figure 4A**). We identified 2722 task-relevant neurons from the recorded 5994 neurons (**Figure 4B**). At the sound onset, 70.7% of neurons on average were newly recruited (**Figure 4B right**): the proportion of neurons was significantly larger than that in the mPFC of 17.7% (Mann–Whitney U test, p = 3.4e-7, 16 and 25 sessions in the AC and mPFC). Example neuron in the AC showed increased activity during the right-rewarded tones irrespective of choices (**Figure 4C**). In the population of neurons, the tone indices for correct and error trials had the same signs, suggesting sound encoding (**Figure 4D**). Consistent with the results in the mPFC, the tone indices at the end of short sounds were correlated with the tone indices at the corresponding duration of long sounds, rather than those at the end of long sounds (**Figure S3A**). In contrast with the mPFC, the activity in the AC did not ramp but gradually decreased during sounds (**Figure 4E**).

We investigated the activity of the AC before sound onset. Example neuron showed increased activity during the left block irrespective of choices, suggesting block encoding. We summarized the activity before and during the long-sound correct trials (**Figure 4G**). 12.8% of neurons showed block-modulated activity before sound onset (**Figure S4A**), consistent with a previous finding of reward modulations in the AC[5,15]. The block modulations were observed in both directions: activity in the preferred block was increased or decreased with respect to the preferred tone frequencies. The population activity was significantly increased at the preferred blocks than nonpreferred blocks in one tone frequency in correct choice trials (**Figure 4G left**), and some frequencies in all trials (**Figure S4B**). Different from the mPFC and M2 (**Figure 3 and S6**), the sound-dependent activity of AC neurons rapidly changed at sound onset, while block-dependent activity was stable (**Figure 4H**).

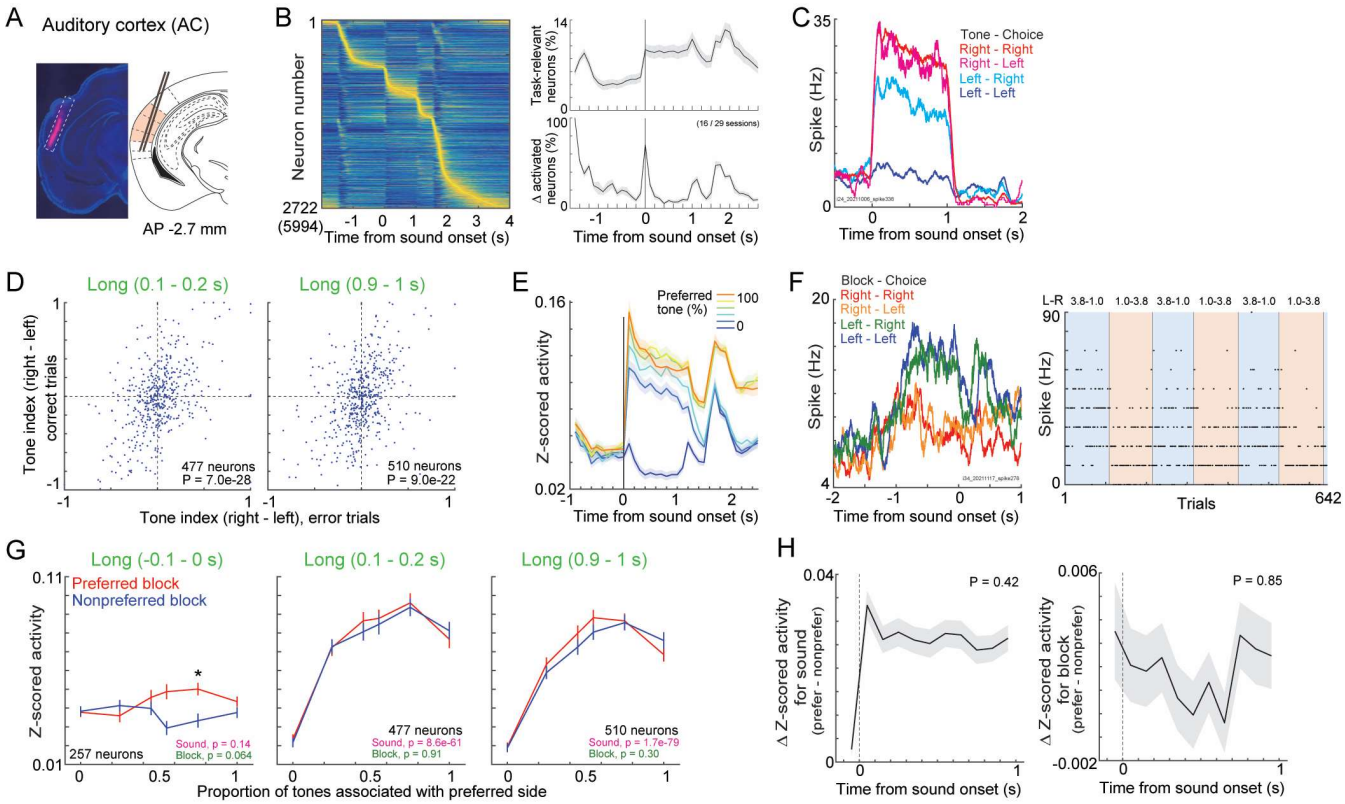


Figure 4. Block and sound representations in the auditory cortex.
A. Probe location for the auditory cortex (AC) in an example mouse. **B.** Average activity of task-relevant neurons in the AC. Data plots are same as in **Figure 3**. **C.** Average activity of an example neuron during long sounds. The neuron showed increased activity when the tones associated with right reward were presented irrespective of choices. **D.** Tone indices in task-relevant neurons. The tone indices of correct and error trials had same signs, suggesting the sound representation. **E.** Means and standard errors of activity in task-relevant neurons. **F.** Average activity of an example neuron before sound onset. The neuron showed increased activity in left block irrespective of choices. **G.** Activity of task-relevant neurons before, at the beginning, and at the end of long sounds in correct trials (ANOVA p values; * $p < 0.01$ in Wilcoxon signed rank test). Population activity of AC neurons represented the sound frequencies. Medians and robust starndard errors. **H.** Sound- and block-dependent activity before and during sounds. Difference in activity between preferred and nonpreferred tones (left), and between preferred and nonpreferred blocks (right) are shown. Medians and robust standard errors (p values in robust linear regression).

Localized representation of sound and choice and global representation of prior values in the cerebral cortex.

To further investigate the neural encoding of each cortical region and to compare the encoding across regions, we modeled the neural activity with a generalized linear model (GLM) (**Figure 5A, Methods**). The neural activity was averaged every 0.1 s and smoothed with a Gaussian filter ($\sigma = 0.25$ s). Our GLM had an equal number of kernels for sound categories, choices, and prior values, as the prior values were binarized depending on the dominant side. The prior value was estimated from the RL model, which achieved the maximum likelihood in each session (**Figure 2**). We plotted how the GLM fit the example neural activity in the AC (single trials: **Figure 5A**; average across trials, **Figure 5B**). The full GLM succeeded in modeling that the neuron increased activity in response to left-rewarded tones irrespective of choices, while the GLM without sound kernels failed to model the sound representation.

By comparing the log likelihood between the full GLM and GLM without sound, choice, or prior kernels, we analyzed the proportion of neurons encoding each task variable ($p < 0.01$ in the likelihood ratio test) (**Figure 5C**). In this analysis, we analyzed the subpopulation of task-relevant neurons that significantly increased the activity at least once from the start of the trial to before the choice was made ($p < 1e\text{-}10$ in the one-sided Wilcoxon signed rank test). We found that a larger proportion of AC neurons represented sounds than the M2 and mPFC (chi-square test, $p = 1.1e\text{-}15$ and $3.8e\text{-}15$). M2 neurons represented choices more than the mPFC ($p = 5.4e\text{-}4$). The proportion of neurons representing prior values did not have significant differences across the three cortical regions ($p = 0.023$–0.33 before Bonferroni correction). We also compared the GLM fit across regions (**Figure 5D**). The sound and choice were encoded in the AC and M2, respectively. The prior value was encoded in the mPFC and M2 rather than in the AC. The differences in prior encoding may come from the nonsignificant encoding neurons, as the proportion of prior-encoding neurons was similar across the three regions (**Figure 5C**). Thus, the sounds and choices were encoded in the AC and M2, respectively, while the prior values were encoded in all 3 recorded regions.

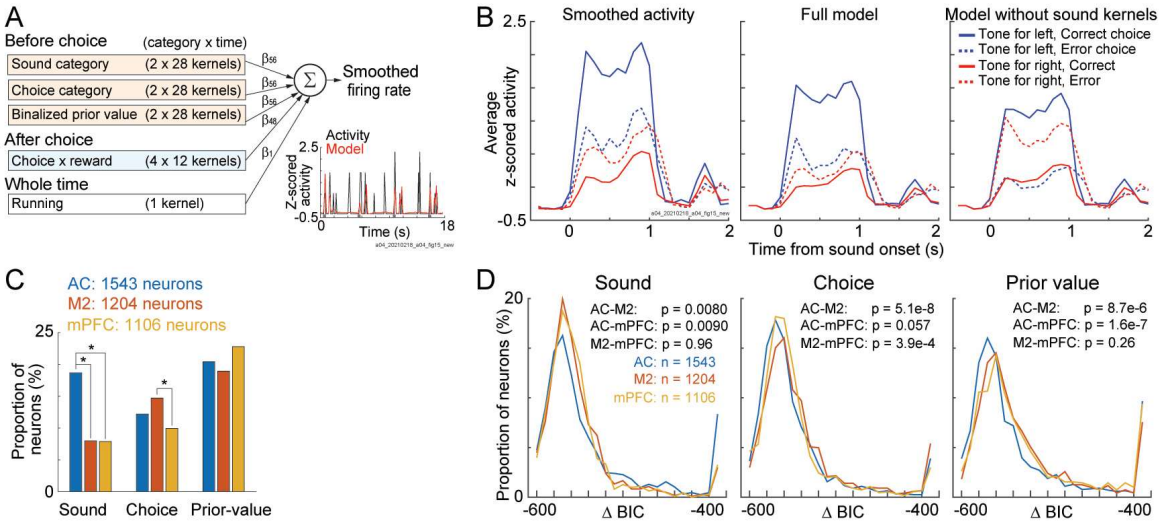


Figure 5. Sound, choice, and prior value encoding in the AC, mPFC, and M2.
A. Schematic of the genelarized linear model (GLM) to test the neural encoding of task variables. Our GLM estimated the log-scaled smoothed activity of each neuron from the 217 kernels and constant term. The right bottom plot shows the smoothed activity and the output of the GLM in example trials in one auditory cortical neuron. **B.** GLM in the auditory cortical neuron in **A**. Left: average neural activity. The neuron showed increased activity during left-reward associated tones. Middle: average trace of the full GLM. Right: average trace of the GLM without sound kernels. The GLM without sound kernels failed to model the sound encoding of the neuron. **C.** Propotion of neurons significantly represented sound, choice, and prior value in each cortical region. We compared the log likelihood between the full GLM and the model without corresponding kernels ($p < 0.01$ in the likelihood ratio test). The proportion of neurons were compared across regions with the chi-square test (* $p < 0.01$). **D.** Comparison of GLM fitting across cortical regions. Δ Bayesian information criterion (BIC) shows the difference of model fitting between the full GLM and model without each task kernel (Wilcoxon signed rank test).

Localized readout of sound and choice and global readout of prior value from the AC, M2, and mPFC.

It is important to investigate not only what neurons represent (encoding) but also what information other brain areas can extract from the brain region (i.e., decoding)[5,31]. We used sparse logistic regression (SLR) and analyzed the decoding performance of sound categories, choices, and binarized prior values (**Figure 6A**). The three task variables were independently decoded in single trials with 10-fold cross validation (CV). The CVs were repeated 100 times to reduce the noise from randomization. We first compared the decoding performance of sounds, choices, and prior values within each cortical region (correct rate: **Figure 6B**; d prime: **Figure S7A**). In the decoding analyses, we used the sessions with simultaneously recorded activity of more than 20 neurons. Before sound onset, the prior readout was better than the sound and choice in all 3 regions (Wilcoxon signed rank test, p = 5.4e-7–3.8e-6). During long sounds, the AC had a better readout of sound than choice and prior values (p = 8.1e-6–0.0017). In M2, choice readout gradually improved during sounds (comparison between the choice and prior decoding at initial and late sounds: p = 0.13 and 7.9e-6), while the mPFC did not have significant differences between the choice and prior decoding even at the late phase of sounds (p = 0.35). In all 3 regions, the decoding performance of the prior value was above the chance level before and during long sounds (Wilcoxon signed-rank test, p = 5.4e-7–4.9e-4), suggesting a stable prior readout. The prior value decoding at the block changes showed that the updating speed of value decoding did not have significant differences across the 3 regions (**Figure S8A and B**).

We then directly compared the decoding performance of 3 cortical regions by subsampling the number of neurons and trials in long-sound trials (**Figure 6C and S7B**). In each session, we randomly picked 100 neurons and 180 trials for 10-fold CV. This subsampled CV was repeated 1000 times. We found that the prior readout did not have significant differences among the AC, M2, and mPFC (Wilcoxon signed rank test, p = 0.080–0.55). In contrast, the sound and choice readouts were better in the AC and M2, respectively, than in the other two regions (sound: p = 8.7e-9 and 1.1e-6; choice: p = 2.7e-5 and 2.8e-4). These results were consistent when we subsampled the 50 neurons and 180 trials in each session (**Figure S7B and S8C**). These results suggest the global and localized readout of prior value and sound/choice, respectively.

We further analyzed how sound and choice decoding changed during sounds. In the AC, the sound readouts were improved by long sounds (**Figure 6D left and S7C**), suggesting a temporal integration of sound at the stage of AC. The sound readout from the AC was stable across blocks[5], while the readouts from the M2 and mPFC were biased by the blocks (**Figure 6D middle and right**). The unbiased sound readout from the AC was important to compute the confidence value in our RL model (**Figure 2**). The choice decoding of the M2 and mPFC was improved during long sounds (**Figure 6E, S7C, and S8D**).

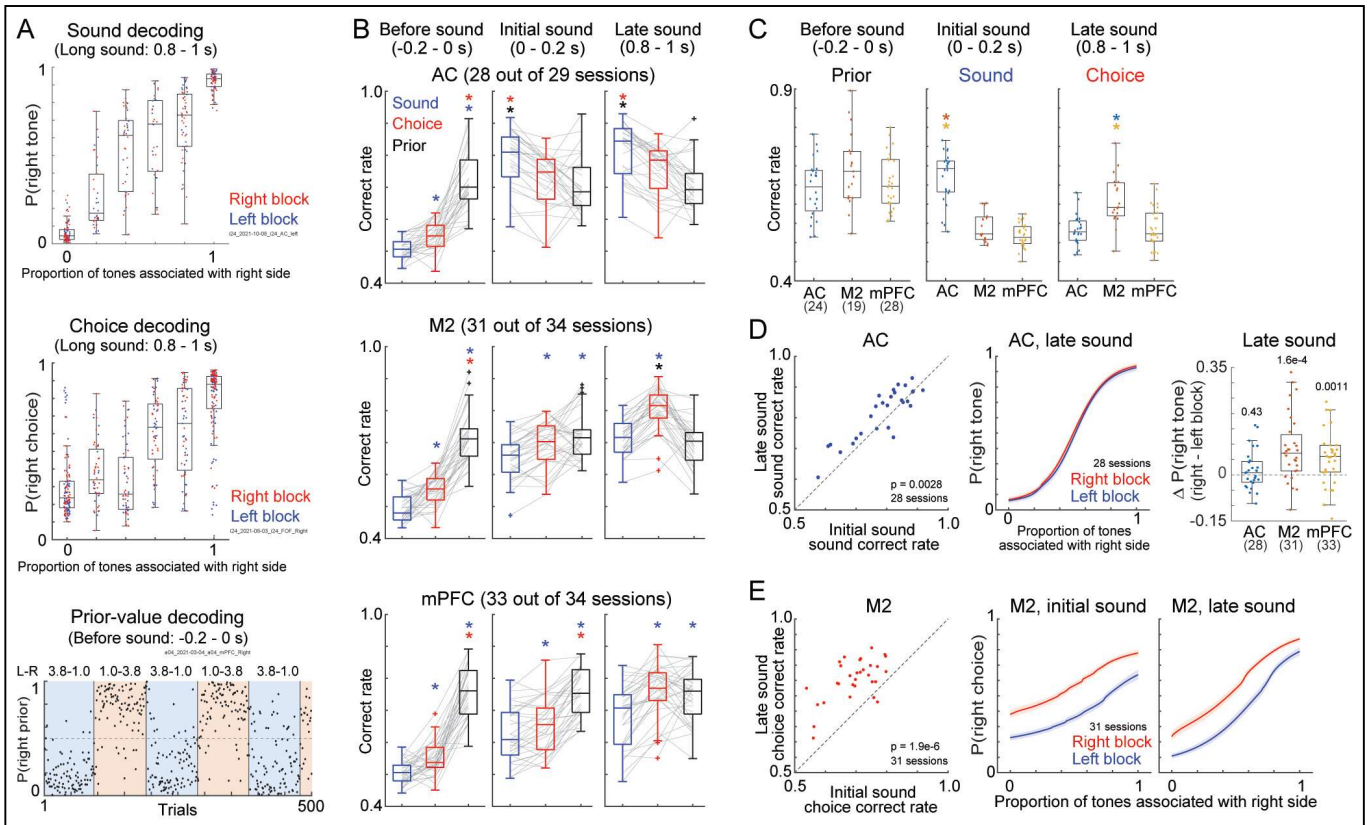


Figure 6. Localized readout of sound and choice and global readout of the prior value from the AC, M2, and mPFC.

A. Sound, choice, and prior value decoding in example sessions. Top: sound decoding in an AC session. Dots show the probability estimate as a right-rewarded tone in each trial. Middle: choice decoding in an M2 session. Dots show the decoded right-choice probability. Bottom: Prior-value decoding in an mPFC session. Dots show the decoded binarized prior value. **B.** Comparison of decoding performance within each cortical region. Sessions with more than 20 neurons were analyzed. Astarisks show the significant differences in decoding performance (* $p < 0.01$ in the Wilcoxon signed rank test). **C.** Comparison of decoding performance across cortical regions. We subsampled the number of neurons and long-sound trials in each session. Sessions with more than 100 neurons were used (* $p < 0.01$ in the Mann–Whitney U test). Parentheses show the number of sessions. **D.** Stable sound decoding in the AC. Left: Comparison of sound-decoding performance between initial (0–0.2 s) and end (0.8–1 s) of long sounds (Wilcoxon signed rank test). Middle: neurometric function of sound decoding. Means and standard errors. Right: stability of sound decoding. Δ p(right tone) was the difference of average neurometric functions between the right and left blocks. (p value in the Wilcoxon signed rank test). **E.** Choice decoding in the M2. Data presentations are same as in **D** but for choice decoding.

Orthogonal representations of sounds and blocks in the AC and the global separation of blocks in the AC, M2, and mPFC.

Our decoding analyses proposed that (i) the AC had a stable sound readout and (ii) the AC, M2, and mPFC had a prior-value readout before and during sounds. To elucidate these two points, we analyzed the state dynamics of blocks, sounds, and choices by introducing "Coding Direction" as was done in previous studies (**Figure 7A**)[32,33]. This analysis utilized our task structure to compute the population activity vectors for separating the asymmetric reward blocks (block mode) and for separating the sound category (sound/choice mode). As our tone-index analyses showed that the activity of mPFC and M2 neurons mainly represented choices while the activity of AC neurons represented sounds (**Figure 3, 4, and S6**), we used only the correct trials and made the separation of sound and choice categories identical. Thus, the sound/choice modes in the mPFC and M2 were mainly for choice, while it was for sound in the AC. Similar to the decoding analyses, we analyzed the sessions that simultaneously recorded the activity of more than 20 neurons. Before applying QR decomposition in the block and sound/choice modes for orthogonalization, we analyzed the cosines between the two modes. We found that the AC and mPFC had a better separation of the two modes than the M2 (**Figure 7B**). The orthogonalized average state dynamics in the AC showed that the block and sound modes were separated throughout the trials (**Figure 7C**). The explained variances of the two modes were better than chance in all the recorded regions (Wilcoxon signed rank test, $p = 5.4e\text{-}7–3.8e\text{-}6$) (**Figure 7D**).

We further tested the orthogonality of block and sound/choice modes (**Figure 7E**)[34]. In block mode, we compared the trajectories between choices; in the sound/choice mode, we compared the trajectories between blocks. In the AC, the neural trajectories did not have significant differences between sounds and blocks. On the other hand, the neural trajectories of one mode were affected by the other in the M2 and mPFC. These results suggest the orthogonal state space of blocks and sounds in the AC. Block modes were separated throughout the trials, especially in the AC and mPFC (**Figure 7E**), suggesting stable block separations.

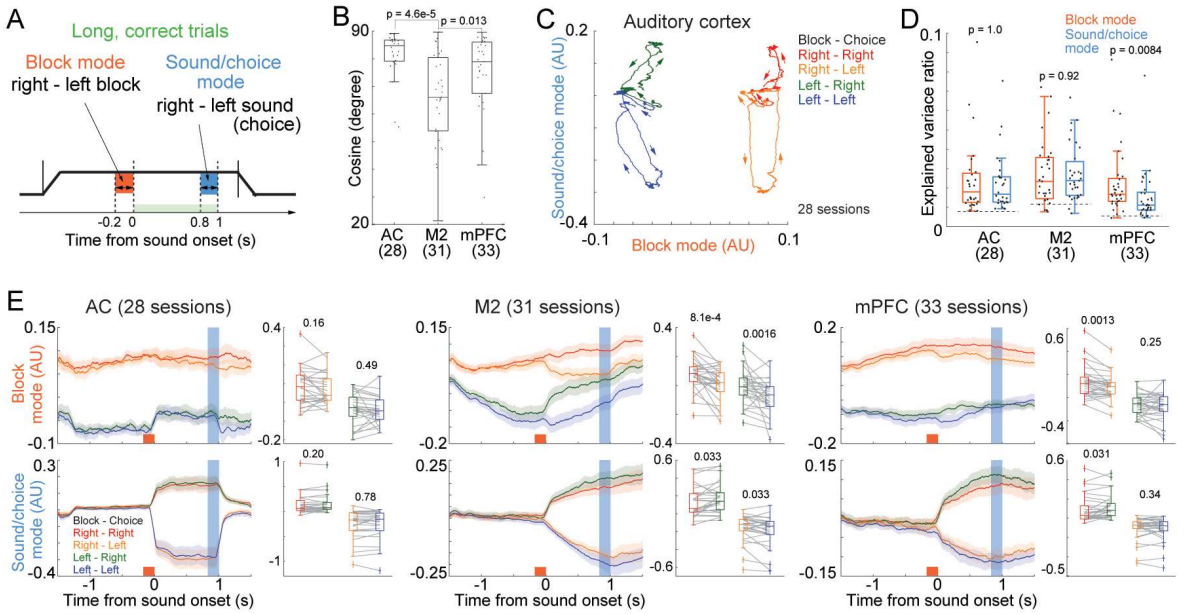


Figure 7. State space analysis explains orthogonalized separations of block and sound in the AC and stable block separations in the AC, M2, and mPFC.
A. Schematic of state-dynamics analysis. We defined the block mode and sound/choice mode based on the simultaneously recorded activity in each session. As we defined the modes from the activity in correct trials, the sound- and choice-separating modes were identical. **B.** Separation of block and sound/choice modes. Parentheses show the number of sessions (Mann–Whitney U test). **C.** Population trajectories in the AC averaged across sessions. **D.** Explained variance ratio. Dotted lines show the baseline. Wilcoxon signed rank test compared the ratio between the block and sound/choice modes. **E.** Block and sound/choice modes in each block-choice combination. Left: means and starndard errors. Right top: block modes were compared between choices at the sound end. Right bottom: sound/choice modes were compared between blocks (Wilcoxon signed rank test).

Discussion

Our results showed that mouse behavior in a tone-frequency discrimination task depended not only on the tone frequencies but also on the tone durations (**Figure 1**): the mouse choices were less accurate and more biased toward the large-reward side in short-sound trials than in long-sound trials. In addition, the outcome experiences based on less accurate stimuli in terms of both the tone frequencies and durations affected the choices in the next trial compared to those based on accurate stimuli, consistent with a previous finding of confidence-dependent value updating[12,29]. Our RL model suggested that the reward expectation of each choice (i.e., prior value) was modulated by sound inputs to compute confidence values (**Figure 2**). This gradual modulation was represented in a sequential population activity of mPFC neurons (**Figure 3**).

The sound inputs and choices were selectively represented in the AC and M2, respectively (**Figure 4 and S6**). In contrast, the prior values were decoded from all 3 recorded regions before and during sounds (**Figure 6 and 7**). As the representations of confidence values, sounds, and choices were required within trials, while the prior values needed to be maintained across trials for future actions, our results suggest a localized and global computation of task variables for short- and long-time scales, respectively, in the cerebral cortex (**Figure S9**).

It has been reported that an RL framework based on a partially observed Markov decision process (POMDP) successfully models the ramping activity of neurons in the posterior parietal cortex (PPC) in a perceptual decision-making task[7,11], supporting that the RL model can apply to both perceptual and value-based decisions[9,12,35]. By utilizing the short-sound trials, our RL model estimated how the prior values were gradually modulated by sensory inputs to form the confidence values during long-sound presentations (**Figure 2**). We used the model parameters of short-sound trials as the parameters at the initial phase of long sounds because the long- and short-sound trials had identical settings until the end of short sounds (**Figure 1**). This analysis was supported by our neuron data in which the neural activity at the end of short sounds was correlated with the activity of the corresponding duration of long-sound stimuli, rather than the end of long sounds (**Figure S3**). Mice temporally integrated the short sound pulses in the tone cloud and made accurate choices in long-sound trials compared to those in short-sound trials (**Figure 1**), showing that the confidence values were computed even at the late phase of long sounds.

The confidence-value representations in the mPFC were reported in a previous study[12]. We advanced the finding by that a temporal sequence of population neuronal activity in the mPFC modulated the prior-value representations with sensory inputs to compute the confidence values (**Figure 3**). As the confidence values were essential for value updating, these results suggest that the mPFC is involved in both the value computation and updating.

Consistent with previous findings[5,15], the activity of individual AC neurons was modulated by reward blocks by increasing or decreasing the activity at the preferred tone frequency (**Figure S4A**). By setting proper high-dimensional population vectors (**Figure 6**), the AC generated block and sound readouts, consistent with a recent finding of the population readout from the visual and motor cortices[36]. We also found that the block- and sound-population vectors were orthogonal (**Figure 7**), suggesting that other brain areas can receive sound frequency inputs from the AC irrespective of reward priors (**Figure 6D**)[5]. The choices of mice are represented in M2 (**Figure 5, 6, and S6**). These results suggest that the confidence values, sensory inputs, and choices were locally represented in the mPFC, AC, and M2, respectively. These

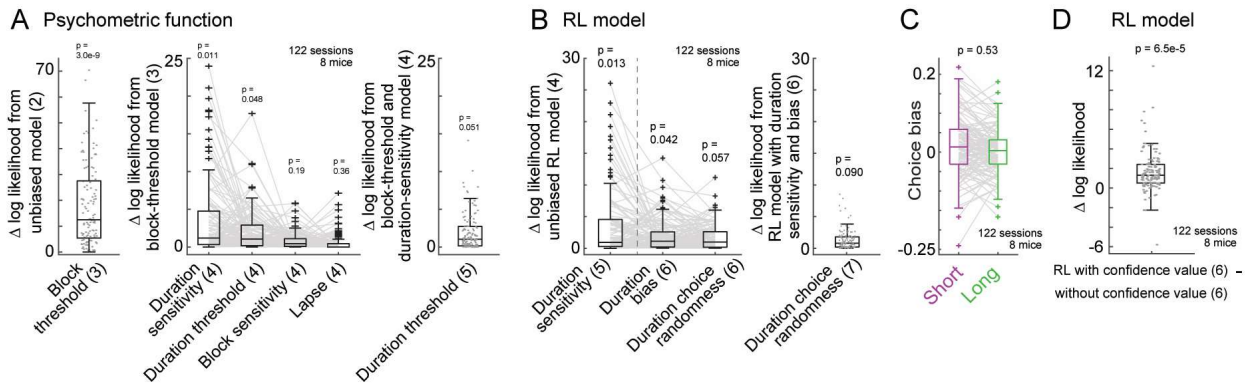
representations were required rapid changes in every single trials. A recent study utilizing Neuropixels 1.0 found that pre-action events, such as visual stimuli and upcoming choices, are represented in restricted regions in the neocortex[37], consistent with the localized computation in our study. In contrast, the prior values were maintained in all three recorded regions (**Figure 6 and 7**), consistent with a global representation of choices and task engagements in a previous study[37]. As the prior values need to be maintained across trials for future actions, such long maintenance may need global coding.

To further investigate the localized and global computations in the cerebral cortex, it is important to test the role of the orbitofrontal cortex (OFC) and PPC. The OFC is known to represent sensory confidence[38], implying that the sensory inputs are once stored in the OFC and integrated into the mPFC for value computations. The PPC temporally integrates visual evidence for choices in a perceptual decision task[7]. In addition, the PPC is causally involved in the representation of prior knowledge of sensory or choice[39–41], implying that the PPC is involved in the integration of prior knowledge and sensory inputs[6].

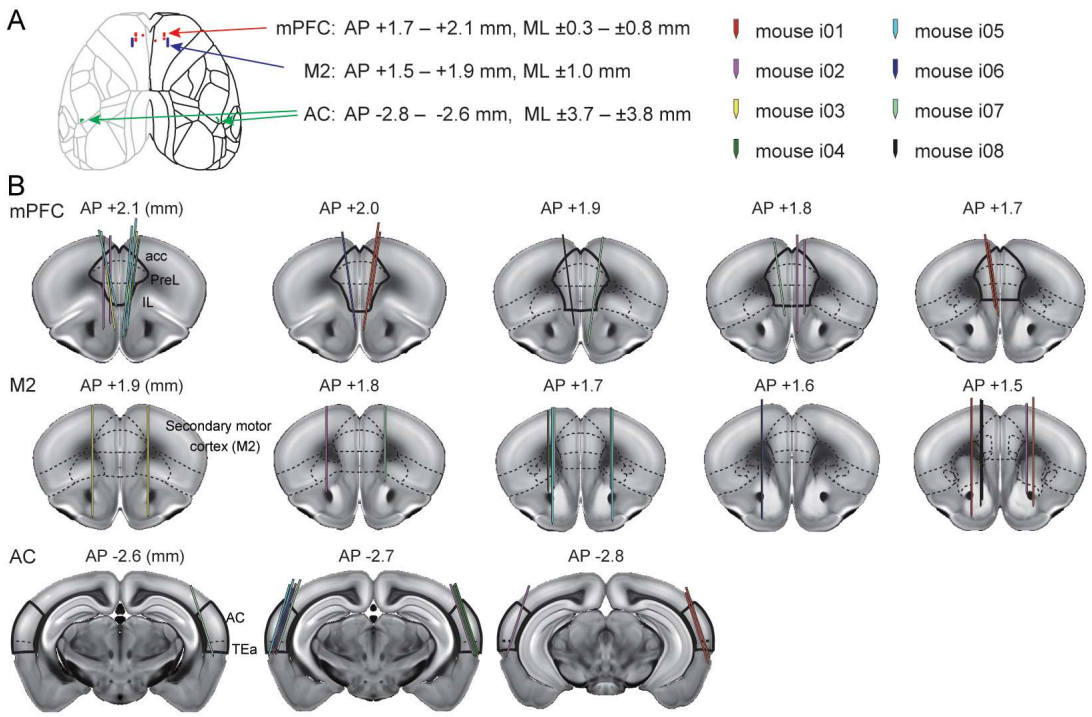
Finally, our decoding analyses found that the AC improved the tone-frequency readout during long sounds (**Figure 6D**), suggesting that sound evidence was temporally integrated at the stage of AC. This is surprising because in a visual perceptual decision-making task in monkeys, the temporal integration of visual stimuli occurs in the PPC, while the visual cortex mainly passes the visual information to the PPC[7,42]. This difference between the auditory and visual processes may come from different contributions of brain regions for decision-making. For example, the M2 but not PPC is necessary for choices in auditory perceptual decisions in rodents[23]. A feedback circuit from the PPC to the AC may enhance the frequency discrimination in the AC, as the AC-PPC circuit is involved in integrating sound evidence in gerbils[43].

In summary, we found that the cerebral cortex locally represented the confidence values, sensory inputs, and choices, which changed rapidly in single trials. In particular, a population sequence of activity in the mPFC integrated the prior values and sensory inputs to compute the confidence values. In contrast, the mPFC, M2, and AC represented the prior values that were slowly updated and maintained across trials. These results suggest a localized and global computation of task variables used in short- and long-time scales, respectively.

Supplementary figures



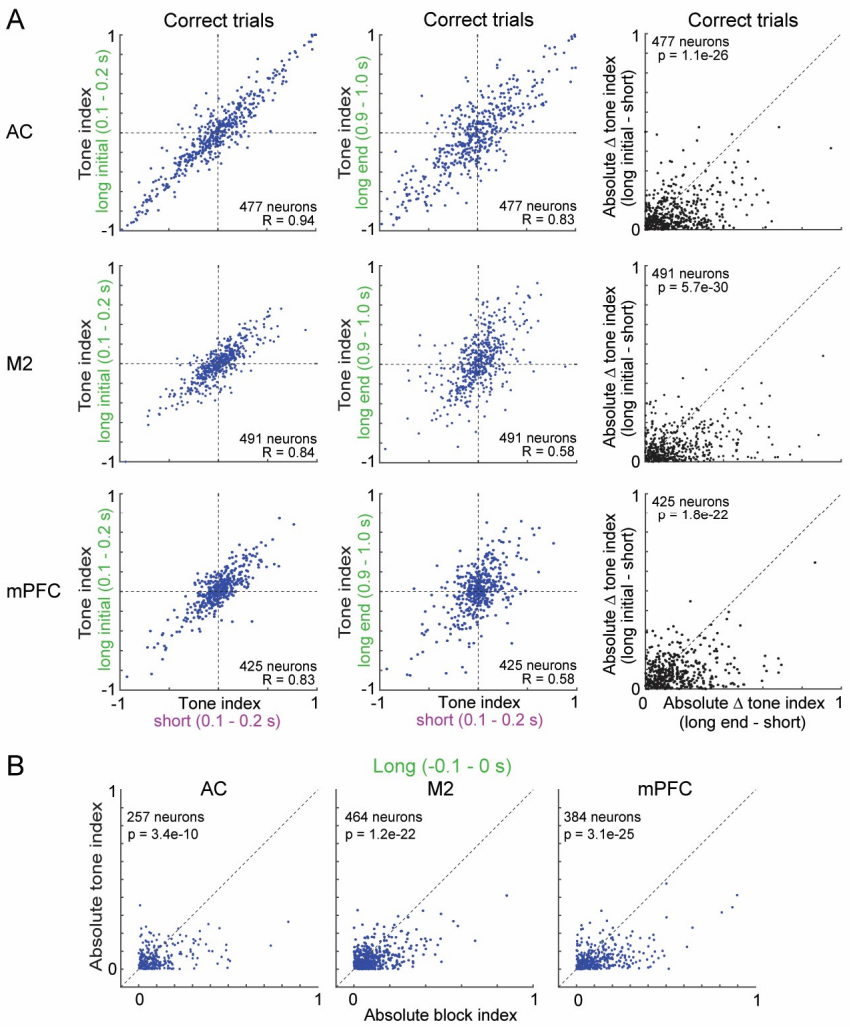
Supplementary Figure 1. Behavioral model comparison. A. Psychometric function. We first analyzed the average log likelihood per session in each mouse, and then averaged aross mice for the likelihood ratio test (**Methods**). Parentheses show the number of parameters in the psychometric-function models. The unbiased model had the parameters for stimulus sensitivity and choice threshold irrespective of blocks and sound durations (2 parameters). The likelihood ratio test was used to investigate whether the model with additional block-dependent or duration-dependent parameters fit mice choices ($p < 0.05$). The model with block-dependent threshold and duration-dependent sensitivity fit the mice choices (4 parameters) (central mark in box: median; edges of the box: 25th and 75th percentiles; whiskers: most extreme data points not considered outliers (beyond 1.5 times the interquartile range), here and throughout).
B. Reinforcement learning (RL) model using the confidence value. Data are presented as in **A**. The unbiased RL model had the parameters of learing rate, choice randomness (inverse temperature), stimulus sensitivity, and choice-threshold irrespective of blocks and sound durations (4 parameters). The likelihood ratio test was used to investigate whether the duration-dependent sensitivity, threshold, and choice randomness improved the model fitting ($p < 0.05$). The RL model with duration-dependent sensitivity and choice bias fit the mice choices (6 parameters).
C. Duration-dependent choice bias in the RL model. The choice bias did not have significant differences between the short and long sounds (linear mixed-effects model).
D. Comparison of RL models with and without confidence value. The RL model with confidence value fit the mice choices better than our previous model which only used prior value (linear mixed-effects model)[14].



Supplementary Figure 2. Locations of Neuropixels probes

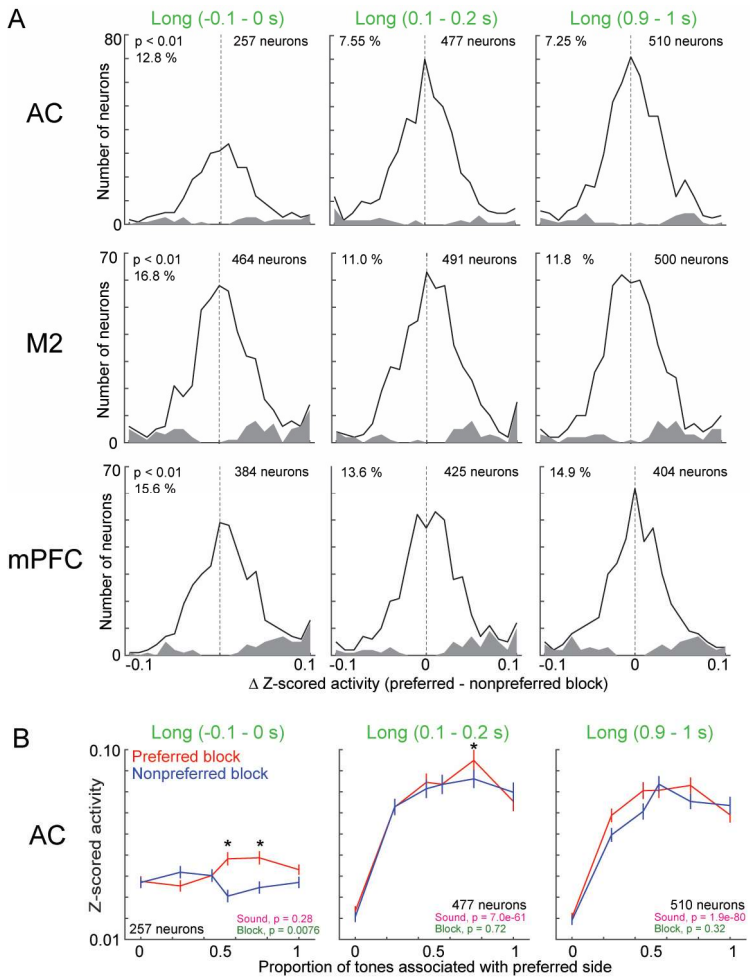
A. Probe entry locations from the surface of dorsal cortex based on the Allen Brain Atlas coordinates. Red dots: the medial prefrontal cortex (mPFC). Blue dots: the secondary motor cortex (M2). Green dots: the auditory cortex (AC). AP and ML indicate the anterior-posterior and medial-lateral axes from bregma (mm), respectively. In each brain region and hemishere, we prepared one tiny hole in the skull for probe entry. In 34 out of 42 holes (80.9%), we identified the probe locations in the post hoc fixed brain slices (**Methods**). All the identified probe locations were successfully located in the target regions.

B. Probe tracks in 8 mice on coronal brain slices registered on Neuropixels trajectory explorer (https://github.com/petersaj/neuropixels_trajectory_explorer).



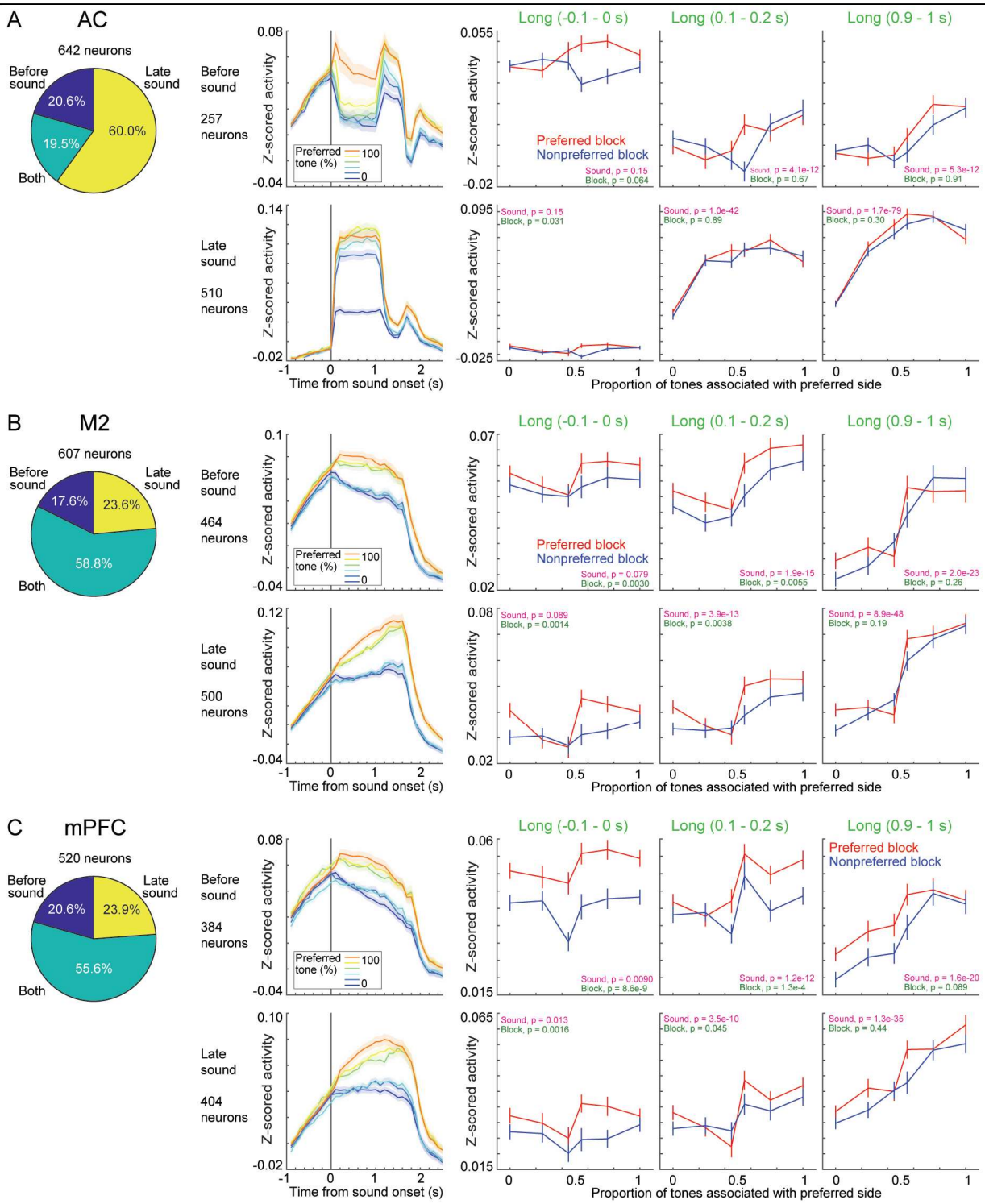
Supplementary Figure 3. Tone and block index during short- and long-sound trials.
A. Tone index. We compared the tone indices between correct short- and long-sound trials. The tone indices of short-sound trials (0.1–0.2 s) were correlated to those of the initial long sounds (0.1–0.2 s) rather than to the end of long sounds (0.9–1.0 s) (left and middle: Pearson correlation; right: p value in the Wilcoxon signed rank test).
B. Block index. The block index was defined as the activity difference between the left and right blocks. The absolute block indices were larger than the tone indices before sound onset (Wilcoxon signed rank test), suggesting the block representations in all 3 regions.

448



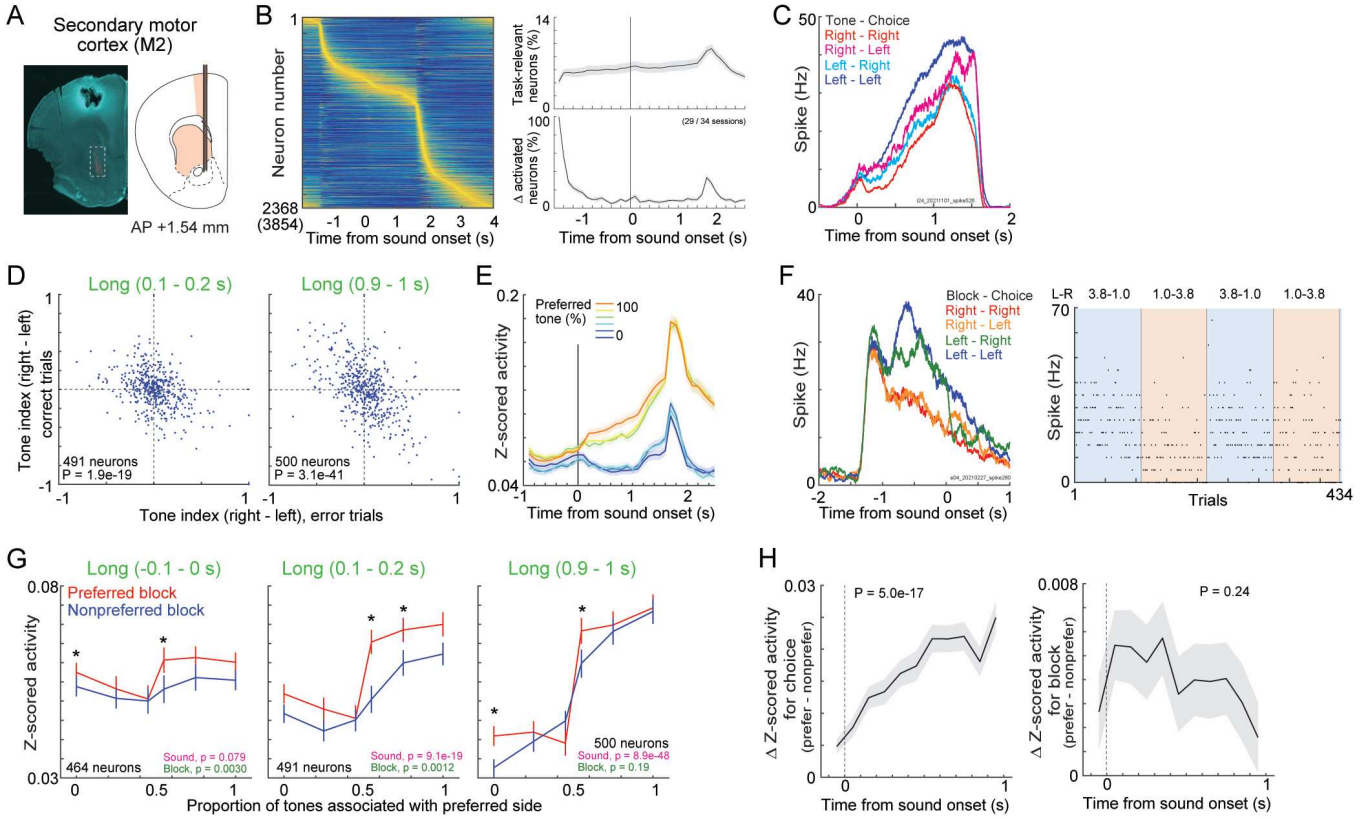
Supplementary Figure 4. Block modulation of single neurons.
A. Block modulations in correct long-sound trials. We analyzed the difference in activity between the preferred and nonpreferred blocks in each neuron. The gray areas show the neurons with significant change in activity between blocks ($p < 0.01$ in Wilcoxon signed rank test). All 3 cortical regions had block-modulated neurons before and during long sounds. **B.** Activity of task-relevant neurons before, at the beginning, and at the end of long sounds in all trials (ANOVA p values; * $p < 0.01$ in Wilcoxon signed rank test). Data are presented as in **Figure 4G** for all trials.

449

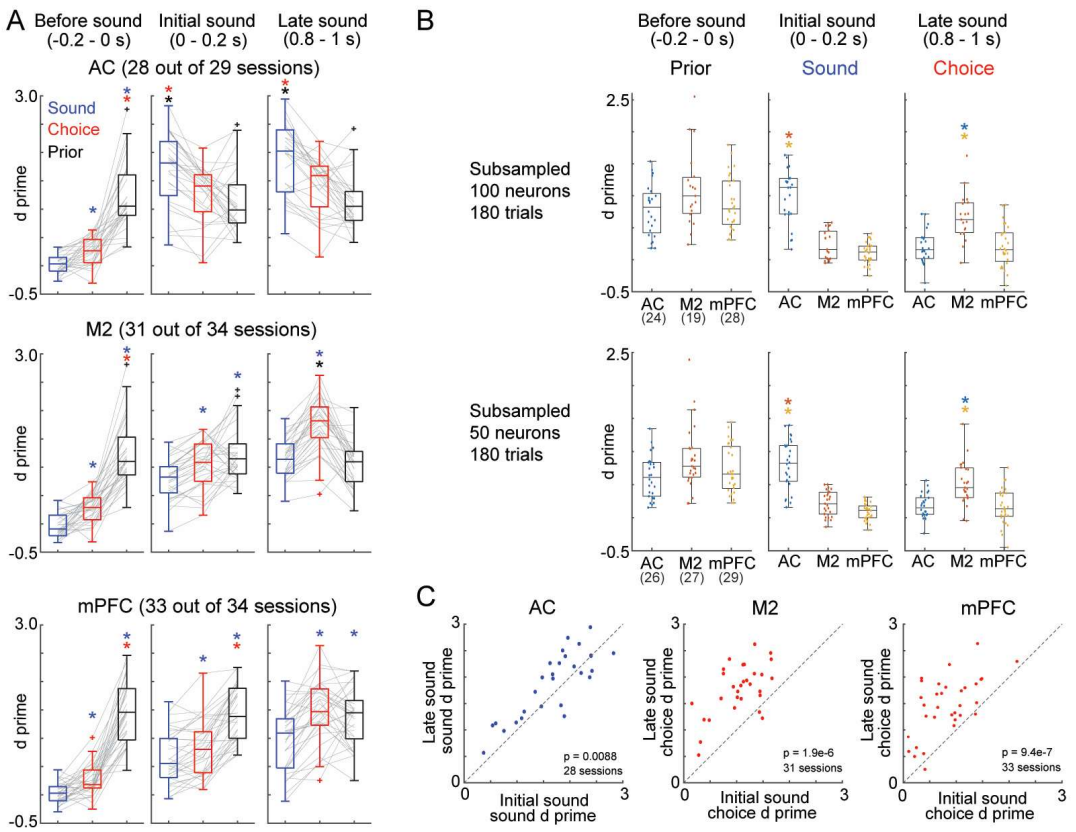


Supplementary Figure 5. Block and choice representations in before- and late-sound activated neurons in the M2 and mPFC and stable sound representation in the AC.

We extracted 2 neuron groups from the task relevant neurons. One group increased the activity before sound onset in long-sound trials (-0.1–0 s). The other group increased activity at the end of the sound (0.9–1.0 s). Pie charts show the distribution of neurons activated during before and late sounds. **A – C.** Top and bottom rows show the activity of before- and late-sound activated neurons, respectively. In the M2 and mPFC, both groups of neurons showed that the modulation of activity by block and choice gradually decreased and increased, respectively (ANOVA p value). Left: Means and starndard errors. Right three figures: medians and robust starndard errors.

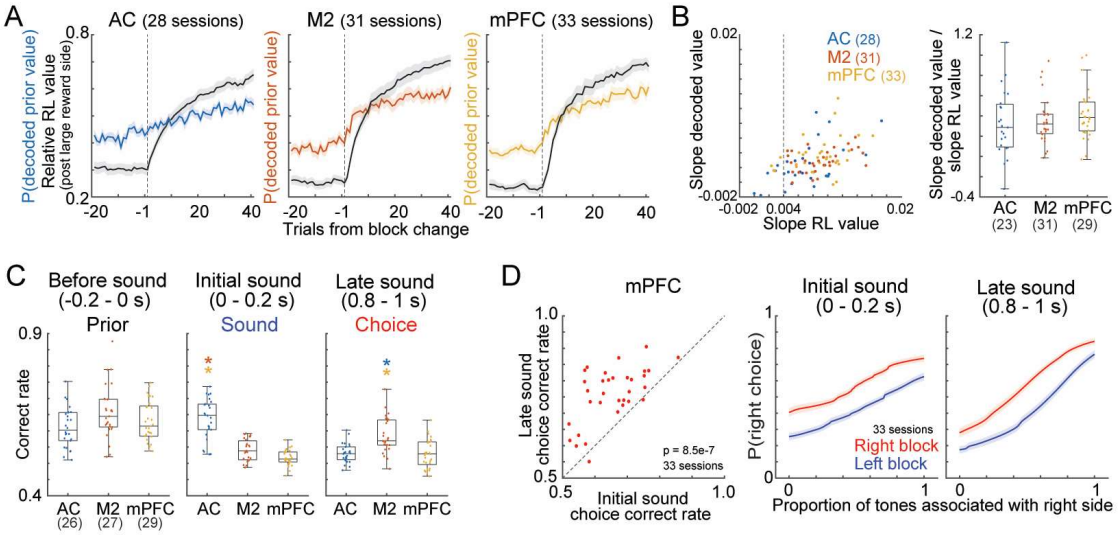


Supplementary Figure 6. Block and choice representations in the M2.
Data are presented as in **Figure 3 and 4** for the M2 neurons.



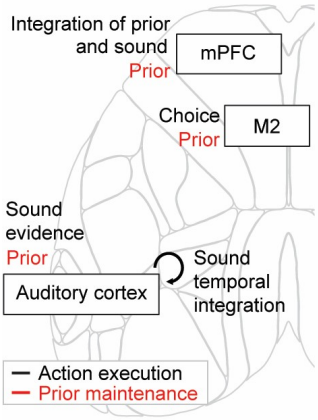
Supplementary Figure 7. Comparison of the decoding performance in the AC, M2, and mPFC.

Data are presented as in **Figure 6** but the decoding performance was plotted with d prime. **A.** Comparison of decoding performance within each cortical region (* $p < 0.01$ in the Wilcoxon signed rank test). **B.** Comparison across regions (* $p < 0.01$ in the Mann–Whitney U test). **C.** Comparison within each cortical regions between initial and late sound phases (p value in the Wilcoxon signed rank test).



Supplementary Figure 8. Correct rate of decoding in the AC, M2, and mPFC.
A. Prior value decoding after the block changes. Relative RL value and P(decoded prior value) show the prior value in our RL model and the decoded prior value of high rewarded side after the block change, respectively. Both values were first averaged across the trials within a session and then averaged across the sessions. Error bars show the standard errors.
B. Updating of prior value decoding. We used the first 40 trials after the block changes and investigated the slope of RL prior value and that of decoded prior value in each session (regress in Matlab) (left). We used the sessions when the slope of RL value beyond 0.004 and investigated the ratio between the slopes of decoding and RL values (right). Parentheses show the number of sessions. Comparison of the slope ratio showed that the AC, M2, and mPFC did not have significant differences in the updating speed of prior value decoding (Mann–Whitney U test, $p = 0.17–0.65$).
C. Comparison of the decoding performance in the AC, M2, and mPFC. Data are presented as in **Figure 6C**, but 50 neurons and 180 trials were subsampled in each session for 10-fold cross validation (* $p < 0.01$ in the Mann–Whitney U test). **D.** Choice decoding in the mPFC. Data are presented as in **Figure 6E** (Wilcoxon signed rank test).

460



Supplementary Figure 9. Proposed functions of the cerebral cortex for integrating the prior value and sound input for making a choice.
Prior values and sensory inputs were integrated to compute the confidence values in the mPFC. The sounds and choices were represented in the auditory cortex and M2, respectively. These results suggest a localized computation of task variables which are required within single trials. All the recorded regions represented the prior values which needed to be maintained across trials. Our results suggest a localized and global computation of task variables essential for short- and long-time scales, respectively, in the cerebral cortex.

Materials and Methods

We used 8 male CBA/J mice (Charles River, Japan) that were 8 to 13 weeks of age at the start of behavioral training. All animal procedures were approved by the Animal Care and Use Committee at the Institute for Quantitative Biosciences (IQB), University of Tokyo. Mice were housed in a temperature-controlled room with a 12 h/12 h light/dark cycle. All experiments were performed during the dark cycle.

Surgery

Surgical procedures and the original versions of the behavioral task were previously described[5,14,25]. Before surgery, the mice were restricted to 1.5 mL of water per day or had free access to citric acid water for at least one week. The weight of each mouse was checked daily to avoid dehydration. Surgery had two steps. The first step was to implant a custom-designed lightweight head bar for behavioral training. The second step was a craniotomy for electrophysiology. In both surgeries, mice were anesthetized with isoflurane (1.5% at induction, below 1% at maintenance) with an additional analgesic (meloxicam 2 mg/kg, subcutaneous) and eye ointment. The mice were placed in a stereotaxic apparatus. For head-bar implantation, the scalp was removed above the entire cortical area. The skull was cleaned with povidone iodine and hydrogen peroxide. The head bar was attached to the skull with superbond adhesive (Sun Medical or Parkell S380) and cyanoacrylate glue (Zap-A-Gap, PT03)[44].

Craniotomy was performed under isoflurane anesthesia. A tiny hole within 1 mm in diameter was made by drilling cyanoacrylate glue and the skull at a recording site to expose the brain surface. The recording site was chosen from the following candidates: the mPFC recording sites were centered at +1.8 mm anterior-posterior (AP), ±0.8 mm medio-lateral (ML) from bregma; the M2 sites were centered at +1.5 AP, ±1.0 ML; the AC sites were centered at -3.0 AP, ± 3.7 ML (**Figure S2**). The M2 sites corresponded to the location of the frontal orienting field (FOF) in rats[24]. The dura mater was removed, and the hole was filled with a mixture of agar and PBS. The agar was coated with silicone oil (30,000 cSt, Aldrich) and sealed with Kwik-Sil (World Precision Instruments) to prevent the brain surface from drying.

Tone-frequency discrimination task

After recovery from head-bar implantation, behavioral training started. The weight of each mouse was carefully monitored, and additional water was given after daily training to prevent dehydration. All the training was performed inside a sound-attenuating booth (O'hara Inc.) or a custom-made training box, while all the electrophysiological recordings were performed inside the sound-attenuating booth. Mice were head-fixed and positioned over a cylinder

treadmill. One speaker (#60108, Avisoft Bioacoustics) was placed diagonally to the right of the mice for auditory stimulation. The speaker was calibrated with a free-field microphone (Type 4939, Brüel and Kjaer). Precalibrated water was delivered through two spouts connected to an infrared lick sensor (O'hara Inc.). The spouts were moved back and forth in each trial. The position of the spouts was controlled by an Arduino with a servo motor (hsb-9475sh, Hitech) and measured with a rotary encoder (E6A2-CS3C, Omron electronics). The behavioral system was controlled by a custom MATLAB (MathWorks) program running on the Bpod r0.5 framework (https://sanworks.io) on Windows.

Similar to our previous behavioral task[14], mice selected the left or right spout to receive a water reward depending on the frequency of the sound stimulus (**Figure 1**). Each trial started by retracting the two spouts away from the mice. After a constant interval of 1.5 s, a sound stimulus of the tone cloud began. The tone cloud consisted of a series of 30 ms pure tones with rise/decay ramps of 3 ms at a rate of 100 tones per second[26,27]. The frequency of each tone was sampled from 18 logarithmically spaced slots (5 to 40 kHz). The tone cloud in each trial contained the low (5–10 kHz) and high frequency tones (20–40 kHz) and was categorized as low or high depending on the dominant frequency. The proportion of high tones in each tone cloud was selected from 6 settings (0, 0.25, 0.45, 0.55, 0.75, or 1) with a probability of 25%, 12.5%, 12.5%, 12.5%, 12.5%, or 25% (i.e., 2:1:1:1:1:2). Although the proportion of high tones had low variability from the 6 settings except for the 100% low or high tone during the training phase, we used the exact 6 settings for the electrophysiological recording phase. In each setting, we used 20 varieties of tone clouds. The settings for short sounds were identical to the corresponding time window of long sounds.

After a 0.5 s delay from the sound end, the two spouts were moved forward, and the mice could select the left or right spout. The high- or low-category tone in a correct trial was associated with a reward of 2.4 μl of 10% sucrose water in either the left or right spout. The association between the tone category and rewarded choice was determined for each mouse. The selection of the opposite spout triggered a noise burst (0.2 s), marking an error trial. The outcome was delivered immediately after the choice. When mice did not select a spout within 15 s from the start of the trial, a new trial started.

A duration of either a long (1.0 s) or short sound was randomly selected for each trial, except that the first 40 trials in each session were always long stimuli with 100% low- or high-tone clouds. The duration of the short sound was 0.2 s for 5 mice and 0.4 s for 2 mice. In the other one mouse, the short sound had a duration of 0.4 s at the initial session of electrophysiology, and it was 0.2 s for the remaining 4 sessions. The intensity of the tone cloud was constant in

each trial but was sampled from either 60, 65, or 70 dB SPL (sound pressure level in decibels with respect to 20 μPa). After the first 40 trials, the reward amounts of the left and right spouts were switched by a block of 90–130 trials. In each block, the reward amount for a correct choice of a large – small reward pair was selected from 2.8–2.0 μl to 3.8–1.0 μl in each session. We defined the left and right blocks depending on the large reward side. If mice did not complete 4 asymmetric-reward blocks, we did not use the session for analysis. After the behavioral task on each day (session), we passively presented the mice with sound stimuli to record the sound-responsive activity of neurons. Data for passive sound presentations are not shown in this manuscript.

Electrophysiological recording with Neuropixels 1.0 and spike sorting

Electrophysiology was performed with Neuropixels 1.0 (IMEC) based on the procedure in a previous study[45]. Every day, i.e., every session, of the behavioral task, we inserted one or two Neuropixels probes. For the mPFC recordings, the probe was tilted 8 degrees in the medial direction except in one mouse (0 degrees). For the AC recordings, the tilted angle was set to 20 degrees in the lateral direction. The angle was 0 degrees for the M2 recording. To identify the probe location in the post hoc fixed brain, the probe was soaked with a diluted solution of CM-DiI or DiA (ThermoFisher Product #V22888 or #D3883). The probe was manually lowered to the brain surface through a mixture of agar and PBS. The brain surface was defined at the depth when spikes were first observed at the recording electrodes around the tip during insertion. We then inserted the probe at a speed of 120 μm/min to prevent damage to the brain tissue[46]. Recording was started at least 5 minutes after the end of probe insertion.

The head bar and tip electrode of the Neuropixels probe were used as the ground and reference, respectively. The 384 electrodes from the tip were used for recording. Open-Ephys GUI acquired the neural data at a sampling rate of 30 kHz with a gain of 500 (PXIe acquisition module, IMEC)[47]. The task events, including treadmill rotations, were sampled at 2.5 kHz (BNC-2110, National Instruments). After recording, the probe was slowly extracted, and the probe-inserted hole was covered with Kwik-Sil. We inserted the probe 1 to 4 times (i.e., 1 to 4 days) in each hole in both the left and right hemispheres of the brain.

For spike sorting, the recording data during the behavioral task and passive sound presentations in each day from one mouse were concatenated. Spike sorting and manual curation were performed with Kilosort 3 in MATLAB (https://github.com/cortex-lab/KiloSort) and Phy in Python (https://github.com/cortex-lab/phy).

Kilosort-3 spike sorting tracked the approximate depth of spikes from each unit during a

session. We defined the depth of each unit from the approximate location of the probe and the electrode position, which measured the maximum amplitude of spikes on average. For the Neuropixels probe for M2 recording, we used the units for analysis when the estimated spike depth from the brain surface was above 1.5 mm. For the mPFC recording probe, we used the units for analyses when the spike depth was (i) above 3.2 mm for the 8-degree tilted probe (7 mice) and (ii) above 3.0 mm for the 0-degree probe (1 mouse). We analyzed all the recorded units for the AC probe.

Histology

After electrophysiological recording, the mice were perfused with 10% formalin solution. Mice were deeply anesthetized with an induction of 5% isoflurane and further with a mixture of 0.3 mg/kg medetomidine, 4.0 mg/kg midazolam, and 5.0 mg/kg butorphanol. The brains were preserved in 10% formalin solution for 24 hours and stored at 4 °C in PBS. The brains were sectioned by a vibratome (VT1000S, Leica Biosystems) along the coronal plane with a thickness of 100 μm. The slice tissues were mounted on glass slides with DAPI mounting medium (Vectashield). Images were acquired with a confocal laser scanning microscope (FV3000, Olympus) or a fluorescence microscope (BZ-X710, Keyence) under 4x magnification (**Figure S2**).

Data analysis

All analyses were performed in MATLAB (MathWorks), except the manual curation of spike-sorting data that was performed in Python. On each day of the session, trials in which mice succeeded in selecting the left or right spout were analyzed. We analyzed the behavioral data from 8 mice in 122 sessions (mouse i01 – i08: 23, 12, 11, 21, 20, 20, 10, and 5 sessions). We did not use the sessions for analyses (i) when mice did not select the correct sides at least 75% of the time for the 100% low- and high-tone clouds in long-sound trials (7 sessions) and (ii) when mice did not complete the 4 asymmetric reward blocks (2 sessions). Among the 122 sessions, the neural activity was recorded from the AC from 7 mice in 29 sessions (mouse i01 – i08: 5, 1, 0, 4, 7, 7, 4, and 1 sessions), the M2 from 8 mice in 34 sessions (6, 6, 6, 6, 3, 4, 1, and 2 sessions), and the mPFC from 8 mice in 34 sessions (8, 5, 5, 3, 3, 4, 4, and 2 sessions). The other 32 sessions (7 sessions were simultaneously recorded with AC) targeted the posterior parietal cortex (PPC). However, as the number of neurons from PPC was small compared to the other 3 regions, we did not use the PPC data for analyses.

Behavioral analysis

We used logistic regression (MATLAB: glmfit) to quantify which tone-cloud timings and asymmetric reward blocks were correlated with the choice in each trial (**Figure 1E**). The detailed method for the psychophysical kernel was reported in previous studies[14,28]. This analysis only used trials with moderate (75% low or high frequency) and difficult (55% low or high) tone clouds.

We modeled the choice behavior of mice with a psychometric function for which the equations were reported in our previous study[14]. In summary, the psychometric function modeled the stimulus sensitivity of mice with a truncated Gaussian ranging between 0 and 1[38]. Our model investigated whether the choice biases and stimulus sensitivities of mice depended on the asymmetric reward blocks and stimulus durations. We also tested whether a lapse rate was necessary to model the choice behavior by the likelihood ratio test[48] (**Figure S1A**) (**Methods, Model comparison**).

Reinforcement learning model

We analyzed the choices the mice made with a reinforcement learning (RL) model[9,11,12,14]. The RL model had the expected reward of left and right choice Q_a in each trial t, defined as the prior value. The model used the prior value to compute the decision threshold x_0 with a softmax equation and an inverse temperature parameter β:

$$x_0(t) = \frac{\exp(\beta Q_{left}(t))}{\exp\left(\beta Q_{left}(t)\right)+\exp\left(\beta Q_{right}(t)\right)}. \quad (1)$$

The softmax equation modeled a perceived reward size that might be different from the actual amount of water[49,50]. The choice probability was estimated from duration-dependent perceptual uncertainty σ_s and a bias parameter d:

$$P(right, t) = \int_{x_0(t)+d}^{1} ZN(x|E_{right}(t), \sigma_{s(t)}^2)dx. \quad (2)$$

$E_{right}(t)$ was the proportion of tone frequency associated with a rightward choice in a tone cloud. Z truncated the Gaussian distribution between 0 and 1 here and throughout. Independent from the choice probability, the perceived probability of left- and right-category tone was estimated as follows:

$$P(left\ category, t) = \int_{0}^{0.5} ZN\left(x\middle|E_{rig}\quad(t), \sigma_s^2\right)dx.$$

$$P(right\ category, t) = \int_{0.5}^{1} ZN\left(x\middle|E_{rig}\quad(t), \sigma_s^2\right)dx. \quad (3)$$

Because of σ_s, we were able to estimate the perceived tone category at the duration of short

632 sounds even in the long-sound trials. The confidence value $conf_Q_a$ was defined by multiplying
633 the prior value Q_a with the associated tone category $P(category, t)$[12], as the left- or right-
634 category tone was only associated with the reward on one side. We defined $conf_Q_a$ as follows:

635 $$conf_Q_{left} = P(left\ category, t)Q_{left}$$

636 $$conf_Q_{righ} = P(right\ category, t)Q_{right}. \quad (4)$$

637 We updated the prior value with forgetting Q-learning, which used the difference between the
638 actual outcome $r(t)$ and confidence value:[9]

639 $$Q_a(t+1) = \begin{cases} Q_a(t) + \alpha\big(r(t) - 2conf_Q_a(t)\big) & if\ a = a(t) \\ (1-\alpha)Q_a(t) & if\ a \neq a(t) \end{cases}, \quad (5)$$

640 where α is the learning rate. We used forgetting Q-learning, which fit better in the choice
641 behavior of rodents than a standard Q-learning model in previous studies[35,50,51]. We tested
642 whether the model using confidence values fit the choice behavior compared to that using only
643 the prior values in our previous study (**Figure S1**)[14]. The initial prior value for each choice was
644 the average amount of expected reward (i.e., 1.2).

645

646 ***Model comparison***

647 For model fitting of psychometric functions and RL models, we defined the likelihood $l(t)$ from
648 the estimated choice probability in each trial (e.g., Equation 2):

649 $$l(t) = \begin{cases} P(right, t) & if\ a(t) = right \\ 1 - P(right, t) & if\ a(t) = left \end{cases}. \quad (6)$$

650 We then analyzed the likelihood in each session L using the trials during the left and right
651 blocks:

652 $$L = \prod_{t=1}^{T} l(t), \quad (7)$$

653 where T is the number of trials. Model comparison within the psychometric functions or the
654 RL models was performed using the likelihood ratio test[48], which investigated whether an
655 additional parameter in the model significantly increased the averaged log likelihood per
656 session ($p < 0.05$) (**Figure S1**). The model parameters were fit to achieve the maximum
657 likelihood. We used the Bayesian information criterion (BIC) to compare the performance
658 between the psychometric function and the RL model (**Figure 2C**)[52]:

659 $$BIC = -2\log(L) + klog(T), \quad (8)$$

660 where k is the number of free parameters.

661

662 ***Spike data analysis***

663 We first identified task-relevant neurons that significantly increased the activity during the task
664 ($p < 1e-10$ in the one-sided Wilcoxon signed rank test). The activity of each neuron was aligned
665 at the sound onset or at the choice timing. For sound-aligned activity, the task-relevant neuron
666 increased the activity compared to the baseline activity in at least one time window (0.1 s)
667 between -1.5 and 2.5 s for the long-sound trials (40 windows) or between -1.5 and 1.7 s for the
668 short-sound trials (32 windows). For the choice-aligned activity, the task-relevant neuron
669 increased the activity in at least one time window between -0.5 and 2.5 s from the choice onset
670 either for the long- or short-sound trials (30 x 2 = 60 windows, in total 40 + 32 + 60 = 132
671 windows). The baseline for sound-aligned activity was the activity at -0.2–0 sec from spout
672 removal, i.e., before trial initiation. The baseline for choice-aligned activity was the activity at
673 -0.2–0 sec from the spout approaching, i.e., between the end of the sound and making a choice.

674

675 In the task-relevant neurons, we analyzed the tone index, which was the difference between the
676 average activity of left- and right-choice trials in a time window of 0.1 s:

677 $$Tone\ index = \frac{\text{mean(spikes (right choice trials))} - mea\ \ \text{(spikes (left choice trials))}}{\text{mean(spikes (right choice trials))} + mean\text{(spikes (left choice trials))}}. \quad (9)$$

678 The tone index ranged between -1 and 1 and was independently analyzed for correct and error
679 trials (**Figure 3D**). We also analyzed the block index, which was the difference in average
680 activity between the left- and right-block trials ranging between -1 and 1 (**Figure S3B**).

681

682 We defined the preferred sides (left or right choice) of task-relevant neurons. The preferred side
683 was defined once based on the activity during the long-sound correct trials at the time window
684 with the largest activity difference from the baseline (one-sided Wilcoxon signed rank test).
685 The average activity on the preferred side was higher than that on the nonpreferred side. For
686 example, when the average activity in correct left-choice trials was higher than that in correct
687 right-choice trials, the preferred side was defined as left. The preferred tone and preferred block
688 were defined based on the preferred side. The preferred tone was associated with the reward
689 on the preferred side. The preferred block had a large reward amount on the preferred side.

690

691 ***Encoding analysis with a generalized linear model (GLM)***

692 The neural activity was averaged every 0.1 s and smoothed with a Gaussian filter ($\sigma = 0.25$ s).
693 We modeled the smoothed log-scaled activity with a generalized linear model (GLM)[9,33,37]. We
694 used the MATLAB software package glmnet (https://glmnet.stanford.edu/index.html) for the

analysis[53]. Our GLM had an equal number of kernels for the sound, choice, and prior value in the RL model, as the sounds and prior values were categorized based on the dominant tones and sides, respectively. In each of the 3 task variables, we prepared 28 time-locked kernels: 18 time-locked kernels in the long-sound trials between -0.2 and 1.5 s from the sound onset, and 10 kernels in the short-sound trials between -0.2 and 0.7 s from the sound onset. The total number of task-variable kernels was 168 (3 task variables x 2 categories x 28 time windows). We also prepared the time-locked kernels for the choice-outcome combinations between -0.1 and 1.0 s from the choice onset (4 combinations x 12 time windows = 48 kernels) and for the running speed of the mouse analyzed from the treadmill rotation (1 kernel). In total, we prepared 218 kernels, including the constant term (168 + 48 + 1 + 1 = 218).

Our GLM fit the log-scaled neural activity with the elastic net with 99% L2 and 1% L1 using 10-fold cross validation (CV) (cvglmnet). By comparing the log likelihood between the full GLM and the model without sound, choice, or prior-value kernels, we analyzed the proportion of neurons encoding each task variable ($p < 0.01$ in likelihood ratio test) (**Figure 5C**). We also analyzed the difference in BIC between the full model and the model without each task-variable kernel (Δ BIC) (**Figure 5D**).

Decoding analysis

We decoded the sound category, choice, and prior category (i.e., dominant side of the prior value) from the simultaneously recorded neural activity in each session. This decoding analysis used sessions with more than 20 neurons. We used a sparse logistic regression in glmnet (https://glmnet.stanford.edu/index.html) for the analysis. We first compared the decoding performance of 3 task variables within a session by using all the recorded neurons and 10-fold CV (cvglmnet) with a time window of 0.2 s. In the decoding analysis, 10-fold CV was performed within the training data to separate all the parameter estimations, including the regularization parameter, from the test data (nested CV). The trained decoder then predicted the task variables in the test data. These CVs were performed 100 times to reduce the noise from randomization. We then analyzed the neurometric function by fitting the binarized decoding outputs with the psychometric function (**Figure 6D and 6E**) (**Methods, Behavioral analysis**)[14].

To compare the decoding performance across cortical regions, we aligned the number of neurons and trials by subsampling 100 neurons and 180 trials in each session (**Figure 6C**). This analysis used sessions with more than 100 neurons. We also subsampled 50 neurons and 180 trials in each session to validate the analysis (**Figure S7A**). We repeated the subsampled 10-fold CV 1000 times. To remove the crosstalk across decoders with 3 task variables, the performance of sound decoding was independently analyzed for the left- and right-choice trials

and averaged. Prior decoding was also independently performed for the left- and right-choice trials. Choice decoding was performed every 4 tone-cloud-setting trials with moderate or difficult tone clouds, and the performances were averaged.

State space analysis

We used "Coding Direction" analysis to investigate the population vector for separating neural activity between blocks (Block mode) and the population vector for separating activity between sound categories (Sound/choice mode) (**Figure 7**)[32,33]. As this analysis only used the long-sound correct trials, the separations of sound- and choice-categories were identical. We utilized our task structure to separate the activity between the block and sound/choice modes. The block mode used the activity between -0.2 and 0 s from the sound onset and analyzed the average difference in z scored activity between the left and right blocks: W_{block}. The sound/choice mode analyzed the average activity difference between the left- and right-choice trials at the end of a long sound (0.8–1.0 s from sound onset): $W_{sound/choic}$. To orthogonalize the block and sound/choice modes, we used QR decomposition[32,33]:

$$W = QR, \tag{10}$$

where W consists of the columns of W_{block} and $W_{sound/choic}$. The first two columns of Q were the orthogonal subspace of $W_{block}^{\perp}$ and $W_{sound/choic}^{\perp}$. Similar to the decoding analyses, we used sessions with more than 20 neurons.

Quantification and statistical analysis

All analyses were performed with MATLAB (MathWorks), except the manual curation of spike-sorting data with Python. Model fitting of psychometric functions and RL models was performed using the likelihood ratio test[14,48] (**Figure S1**) and Bayesian information criterion (BIC) (**Figure 2**) (**Methods**, **model comparison**)[52]. Behavioral data analyses were performed using the linear mixed-effects model (MATLAB: fitlme)[14], as we analyzed multiple sessions from each subject (mouse). To compare samples in each session, the fixed and random effects were the samples and subjects, respectively. We used the chi-square test to compare the proportion of neurons encoding task variables (**Figure 5C**). In other analyses, we mainly used two-sided nonparametric statistical tests. Descriptions were made for one-sided statistical tests. Error bars for means are standard deviations or standard errors. Error bars for medians were median absolute deviations (MADs) or robust standard errors (1.4826*MAD/sqrt(n); n = number of data points)[54,55]. All the statistical details are in the figure legends, figures, or results.

Acknowledgments

This work was funded by JSPS LEADER, Kakenhi (20K23339, 21H00187, and 21H05243), the Senri Life Science Foundation, and the Japanese Neural Network Society (JNNS) for A. F; JSPS Kakenhi (20K23317) for K.I.

Author contributions

K.I. collected and analyzed the data. S.N. analyzed the data. A.F. designed the experiment, analyzed the data, and wrote the paper.

Competing interests

The author declares no conflict of interest.

Materials & Correspondence

Should be addressed to A.F.

References

1. Dayan, P. & Daw, N. D. Decision theory, reinforcement learning, and the brain. *Cogn. Affect. Behav. Neurosci.* **8**, 429–453 (2008).
2. Gold, J. I. & Ding, L. How mechanisms of perceptual decision-making affect the psychometric function. *Prog. Neurobiol.* **103**, 98–114 (2013).
3. Körding, K. P., Ku, S. P. & Wolpert, D. M. Bayesian integration in force estimation. *J. Neurophysiol.* (2004) doi:10.1152/jn.00275.2004.
4. Doya, K. Modulators of decision making. *Nat. Neurosci.* **11**, 410–416 (2008).
5. Funamizu, A., Marbach, F. & Zador, A. M. Stable sound decoding despite modulated sound representation in the auditory cortex. *bioRxiv* 2023.01.31.526457 (2023) doi:10.1101/2023.01.31.526457.
6. Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E. & Shadlen, M. N. Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J. Neurosci.* **31**, 6339–6352 (2011).
7. Gold, J. & Shadlen, M. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
8. Samejima, K., Ueda, Y., Doya, K. & Kimura, M. Representation of Action-Specific Reward Values in the Striatum. *Science (80-.).* **310**, 1337–1340 (2005).
9. Ito, M. & Doya, K. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J. Neurosci.* **29**, 9861–9874 (2009).
10. Sul, J. H., Jo, S., Lee, D. & Jung, M. W. Role of rodent secondary motor cortex in value-based action selection. *Nat. Neurosci. 2011 149* **14**, 1202–1208 (2011).
11. Rao, R. P. N. Decision making under uncertainty: A neural model based on partially observable Markov decision processes. *Front. Comput. Neurosci.* (2010) doi:10.3389/fncom.2010.00146.
12. Lak, A. *et al.* Dopaminergic and Prefrontal Basis of Learning from Sensory Confidence and Reward Value. *Neuron* (2020) doi:10.1016/j.neuron.2019.11.018.
13. Pisupati, S., Chartarifsky-Lynn, L., Khanal, A. & Churchland, A. K. Lapses in perceptual decisions reflect exploration. *Elife* **10**, 1–27 (2021).
14. Funamizu, A. Integration of sensory evidence and reward expectation in mouse perceptual decision-making task with various sensory uncertainties. *iScience* **24**, 102826 (2021).
15. Guo, L., Weems, J. T., Walker, W. I., Levichev, A. & Jaramillo, S. Choice-selective

neurons in the auditory cortex and in its striatal target encode reward expectation. *J. Neurosci.* (2019) doi:10.1523/JNEUROSCI.2585-18.2019.

16. Fritz, J., Shamma, S., Elhilali, M. & Klein, D. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* **6**, 1216–1223 (2003).

17. Otazu, G. H., Tai, L.-H., Yang, Y. & Zador, A. M. Engaging in an auditory task suppresses responses in auditory cortex. *Nat. Neurosci.* **12**, 646–654 (2009).

18. Hubel, D. H., Henson, C. O., Rupert, A. & Galambos, R. 'Attention' units in the auditory cortex. *Science (80-.).* (1959) doi:10.1126/science.129.3358.1279.

19. Pinto, L. & Dan, Y. Cell-Type-Specific Activity in Prefrontal Cortex during Goal-Directed Behavior. *Neuron* **87**, 437–450 (2015).

20. Rushworth, M. F. S., Noonan, M. A. P., Boorman, E. D., Walton, M. E. & Behrens, T. E. Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**, 1054–1069 (2011).

21. Balleine, B. W. & Dickinson, A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**, 407–419 (1998).

22. Corbit, L. H. & Balleine, B. W. The role of prelimbic cortex in instrumental conditioning. *Behav. Brain Res.* **146**, 145–157 (2003).

23. Erlich, J. C., Brunton, B. W., Duan, C. A., Hanks, T. D. & Brody, C. D. Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *Elife* **2015**, (2015).

24. Hanks, T. D. *et al.* Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**, 220–223 (2015).

25. Marbach, F. & Zador, A. M. A self-initiated two-alternative forced choice paradigm for head-fixed mice. *bioRxiv* (2016) doi:10.1101/073783.

26. Znamenskiy, P. & Zador, A. M. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* **497**, 482–5 (2013).

27. Xiong, Q., Znamenskiy, P. & Zador, A. M. Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* **521**, 348–351 (2015).

28. Odoemene, O., Pisupati, S., Nguyen, H. & Churchland, A. K. Visual evidence accumulation guides decision-making in unrestrained mice. *J. Neurosci.* (2018) doi:10.1523/JNEUROSCI.3478-17.2018.

29. Lak, A. *et al.* Reinforcement biases subsequent perceptual decisions when confidence is low, a widespread behavioral phenomenon. *Elife* **9**, (2020).

30. Jun, J. J. *et al.* Fully integrated silicon probes for high-density recording of neural activity. *Nature* (2017) doi:10.1038/nature24636.

31. Bialek, W., Rieke, F., De Ruyter Van Steveninck, R. R. & Warland, D. Reading a neural code. *Science (80-.).* (1991) doi:10.1126/science.2063199.

32. Allen, W. E. *et al.* Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. *Science* **364**, (2019).

33. Osako, Y. *et al.* Contribution of non-sensory neurons in visual cortical areas to visually guided decisions in the rat. *Curr. Biol.* **31**, 2757-2769.e6 (2021).

34. Stringer, C. *et al.* Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, (2019).

35. Funamizu, A., Ito, M., Doya, K., Kanzaki, R. & Takahashi, H. Uncertainty in action-value estimation affects both action choice and learning rate of the choice behaviors of rats. *Eur. J. Neurosci.* (2012) doi:10.1111/j.1460-9568.2012.08025.x.

36. Voitov, I. & Mrsic-Flogel, T. D. Cortical feedback loops bind distributed representations of working memory. *Nat. 2022 6087922* **608**, 381–389 (2022).

37. Steinmetz, N. A., Zatka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nat. 2019 5767786* **576**, 266–273 (2019).

38. Masset, P., Ott, T., Lak, A., Hirokawa, J. & Kepecs, A. Behavior- and Modality-General Representation of Confidence in Orbitofrontal Cortex. *Cell* (2020) doi:10.1016/j.cell.2020.05.022.

39. Akrami, A., Kopec, C. D., Diamond, M. E. & Brody, C. D. Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature* (2018) doi:10.1038/nature25510.

40. Hwang, E. J., Dahlen, J. E., Mukundan, M. & Komiyama, T. History-based action selection bias in posterior parietal cortex. *Nat. Commun.* (2017) doi:10.1038/s41467-017-01356-z.

41. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Matching Behavior and the representaiton of value in the parietal cortex. *Science (80-.).* **304**, 457–461 (2004).

42. Bair, W., Zohary, E. & Newsome, W. T. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J. Neurosci.* **21**, 1676–97 (2001).

43. Yao, J. D., Gimoto, J., Constantinople, C. M. & Sanes, D. H. Parietal Cortex Is Required for the Integration of Acoustic Evidence. *Curr. Biol.* **30**, 3293-3303.e4 (2020).

44. Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial neural dynamics are dominated by richly varied movements. *Nat. Neurosci.* (2019) doi:10.1038/s41593-019-0502-4.

45. Siegle, J. H. *et al.* Survey of spiking in the mouse visual system reveals functional hierarchy. *Nat. 2021 5927852* **592**, 86–92 (2021).

46. Fiáth, R. *et al.* Slow insertion of silicon probes improves the quality of acute neuronal recordings. *Sci. Reports 2019 91* **9**, 1–17 (2019).

47. Siegle, J. H. *et al.* Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. *J. Neural Eng.* **14**, (2017).

48. Daw, N. D. Trial-by-trial data analysis using computational models. *Decis. making, Affect. Learn. Atten. Perform. XXIII* **23**, 3–38 (2011).

49. Constantinople, C. M., Piet, A. T. & Brody, C. D. An Analysis of Decision under Risk in Rats. *Curr. Biol.* (2019) doi:10.1016/j.cub.2019.05.013.

50. Ito, M. & Doya, K. Distinct Neural Representation in the Dorsolateral, Dorsomedial, and Ventral Parts of the Striatum during Fixed- and Free-Choice Tasks. *J. Neurosci.* **35**, 3499–3514 (2015).

51. Hattori, R., Danskin, B., Babic, Z., Mlynaryk, N. & Komiyama, T. Area-Specificity and Plasticity of History-Dependent Value Coding During Learning. *Cell* (2019) doi:10.1016/j.cell.2019.04.027.

52. Stoilova, V. V. *et al.* Auditory cortex reflects goal-directed movement but is not necessary for behavioral adaptation in sound-cued reward tracking. *J. Neurophysiol.* **124**, 1056–1071 (2020).

53. Simon, N., Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *J. Stat. Softw.* **39**, 1–13 (2011).

54. Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z. J. & Scanziani, M. A neural circuit for spatial summation in visual cortex. *Nat. 2012 4907419* **490**, 226–231 (2012).

55. Funamizu, A., Kuhn, B. & Doya, K. Neural substrate of dynamic Bayesian inference in the cerebral cortex. *Nat. Neurosci.* (2016) doi:10.1038/nn.4390.

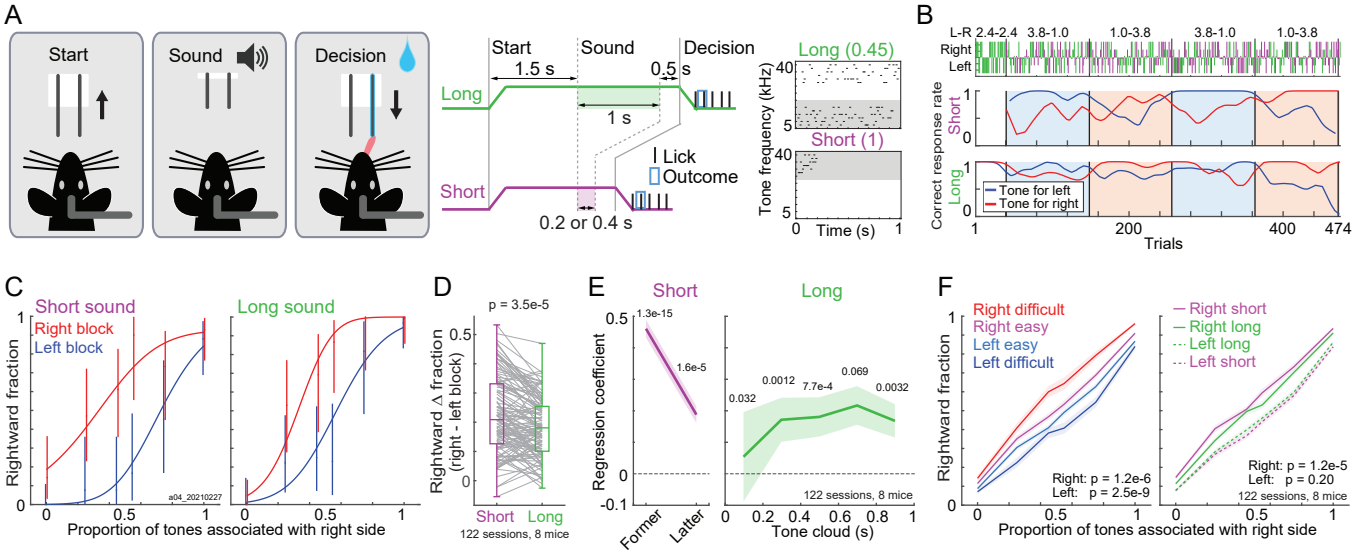

A
Start
Sound
Decision
Start
Sound
Decision
1.5 s
0.5 s
Long
1 s
Lick
Outcome
Short
0.2 or 0.4 s
Long (0.45)
Short (1)
Tone frequency (kHz)
Time (s)
B
L-R 2.4-2.4 3.8-1.0 1.0-3.8 3.8-1.0 1.0-3.8
Right
Left
Correct response rate
Short
Long
Tone for left
Tone for right
Trials
C
Short sound
Right block
Left block
Long sound
Rightward fraction
Proportion of tones associated with right side
D
p = 3.5e-5
Rightward Δ fraction (right - left block)
Short
Long
122 sessions, 8 mice
E
Short
Long
1.3e-15
1.6e-5
0.032
0.0012
7.7e-4
0.069
0.0032
Regression coefficient
Former
Latter
122 sessions, 8 mice
Tone cloud (s)
F
Right difficult
Right easy
Left easy
Left difficult
Right: p = 1.2e-6
Left: p = 2.5e-9
Right short
Right long
Left long
Left short
Right: p = 1.2e-5
Left: p = 0.20
122 sessions, 8 mice
Rightward fraction
Proportion of tones associated with right side

Figure 1

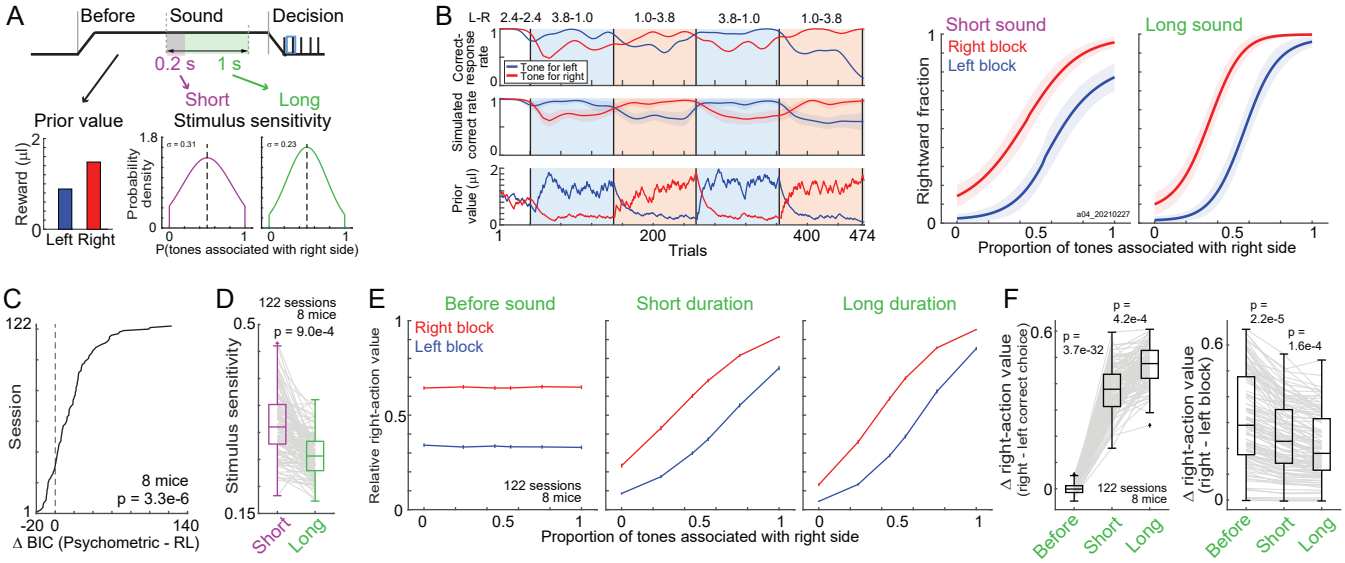

A
Before
Sound
Decision
0.2 s
1 s
Short
Long
Prior value
Stimulus sensitivity
Reward (μl)
Left Right
Probability density
σ = 0.31
σ = 0.23
P(tones associated with right side)
B
L-R
2.4-2.4
3.8-1.0
1.0-3.8
3.8-1.0
1.0-3.8
Correct-response rate
Tone for left
Tone for right
Simulated correct rate
Prior value (μl)
Trials
Short sound
Right block
Left block
Long sound
Rightward fraction
a04_20210227
Proportion of tones associated with right side
C
Session
8 mice
p = 3.3e-6
Δ BIC (Psychometric - RL)
D
122 sessions
8 mice
p = 9.0e-4
Stimulus sensitivity
Short
Long
E
Before sound
Short duration
Long duration
Right block
Left block
Relative right-action value
122 sessions
8 mice
Proportion of tones associated with right side
F
p = 4.2e-4
p = 3.7e-32
Δ right-action value (right - left correct choice)
122 sessions
8 mice
Before
Short
Long
p = 2.2e-5
p = 1.6e-4
Δ right-action value (right - left block)

Figure 2

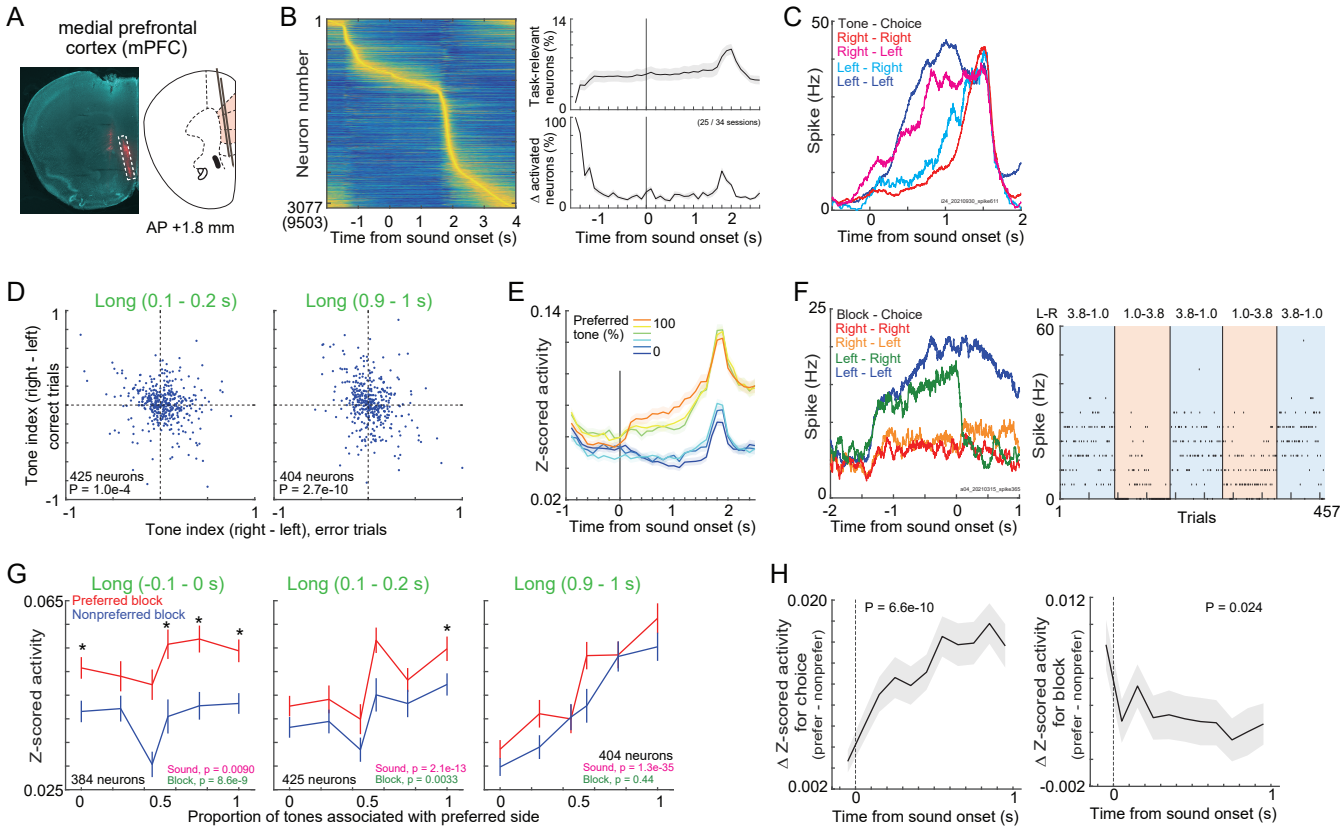

A
medial prefrontal cortex (mPFC)
AP +1.8 mm
B
Neuron number
3077 (9503)
Time from sound onset (s)
Task-relevant neurons (%)
Δ activated neurons (%)
(25 / 34 sessions)
C
Tone - Choice
Right - Right
Right - Left
Left - Right
Left - Left
Spike (Hz)
Time from sound onset (s)
D
Long (0.1 - 0.2 s)
Long (0.9 - 1 s)
Tone index (right - left) correct trials
Tone index (right - left), error trials
425 neurons
P = 1.0e-4
404 neurons
P = 2.7e-10
E
Preferred tone (%)
100
0
Z-scored activity
Time from sound onset (s)
F
Block - Choice
Right - Right
Right - Left
Left - Right
Left - Left
Spike (Hz)
Time from sound onset (s)
L-R
3.8-1.0
1.0-3.8
3.8-1.0
1.0-3.8
3.8-1.0
Trials
G
Long (-0.1 - 0 s)
Long (0.1 - 0.2 s)
Long (0.9 - 1 s)
Preferred block
Nonpreferred block
Z-scored activity
384 neurons
Sound, p = 0.0090
Block, p = 8.6e-9
425 neurons
Sound, p = 2.1e-13
Block, p = 0.0033
404 neurons
Sound, p = 1.3e-35
Block, p = 0.44
Proportion of tones associated with preferred side
H
P = 6.6e-10
Δ Z-scored activity for choice (prefer - nonprefer)
Time from sound onset (s)
P = 0.024
Δ Z-scored activity for block (prefer - nonprefer)
Time from sound onset (s)

Figure 3

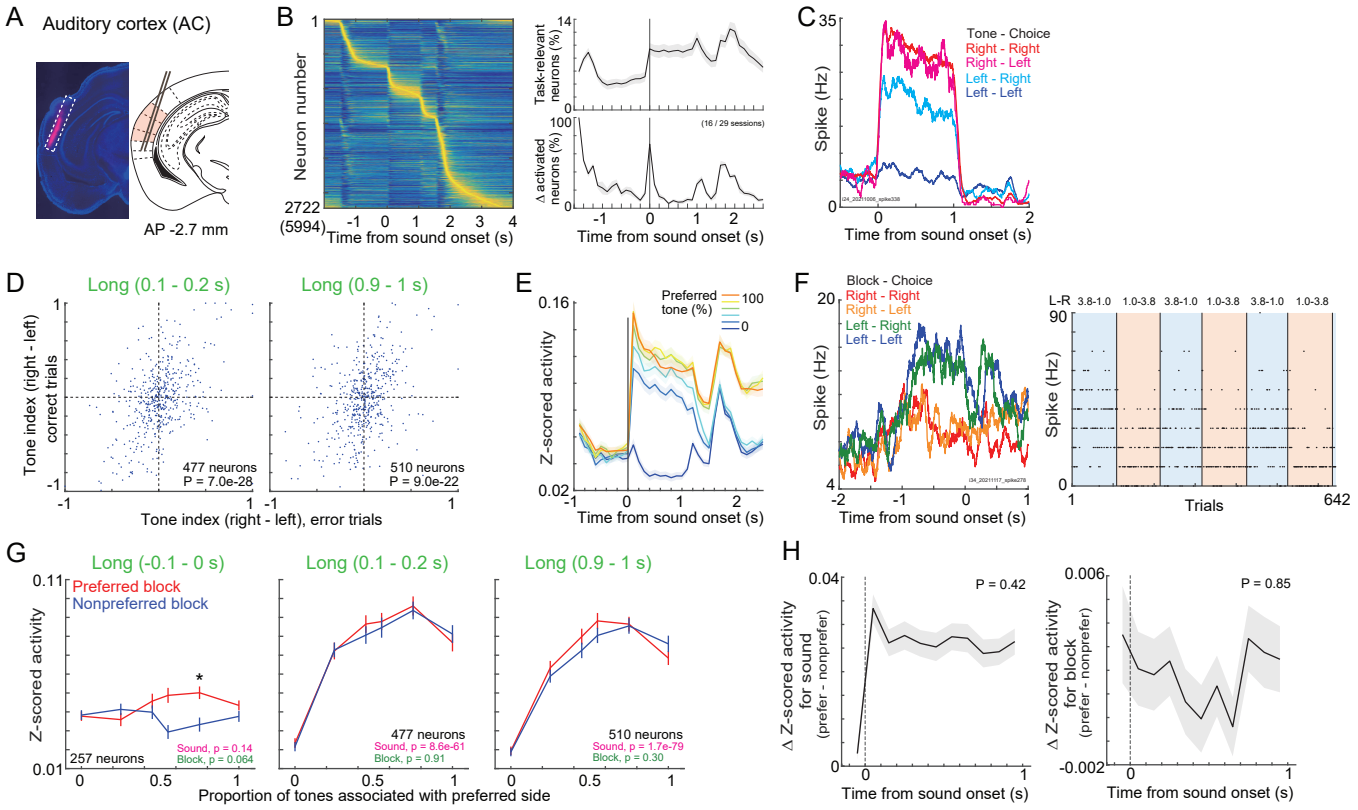

A
Auditory cortex (AC)
AP -2.7 mm
B
Neuron number
2722
(5994)
Time from sound onset (s)
Task-relevant neurons (%)
Δ activated neurons (%)
(16 / 29 sessions)
C
Spike (Hz)
Tone - Choice
Right - Right
Right - Left
Left - Right
Left - Left
Time from sound onset (s)
D
Long (0.1 - 0.2 s)
Long (0.9 - 1 s)
Tone index (right - left)
correct trials
477 neurons
P = 7.0e-28
510 neurons
P = 9.0e-22
Tone index (right - left), error trials
E
Preferred tone (%)
100
0
Z-scored activity
Time from sound onset (s)
F
Block - Choice
Right - Right
Right - Left
Left - Right
Left - Left
Spike (Hz)
Time from sound onset (s)
L-R
3.8-1.0
1.0-3.8
Trials
642
G
Long (-0.1 - 0 s)
Long (0.1 - 0.2 s)
Long (0.9 - 1 s)
Preferred block
Nonpreferred block
Z-scored activity
257 neurons
Sound, p = 0.14
Block, p = 0.064
477 neurons
Sound, p = 6.6e-61
Block, p = 0.91
510 neurons
Sound, p = 1.7e-79
Block, p = 0.30
Proportion of tones associated with preferred side
H
P = 0.42
Δ Z-scored activity for sound (prefer - nonprefer)
P = 0.85
Δ Z-scored activity for block (prefer - nonprefer)
Time from sound onset (s)

Figure 4

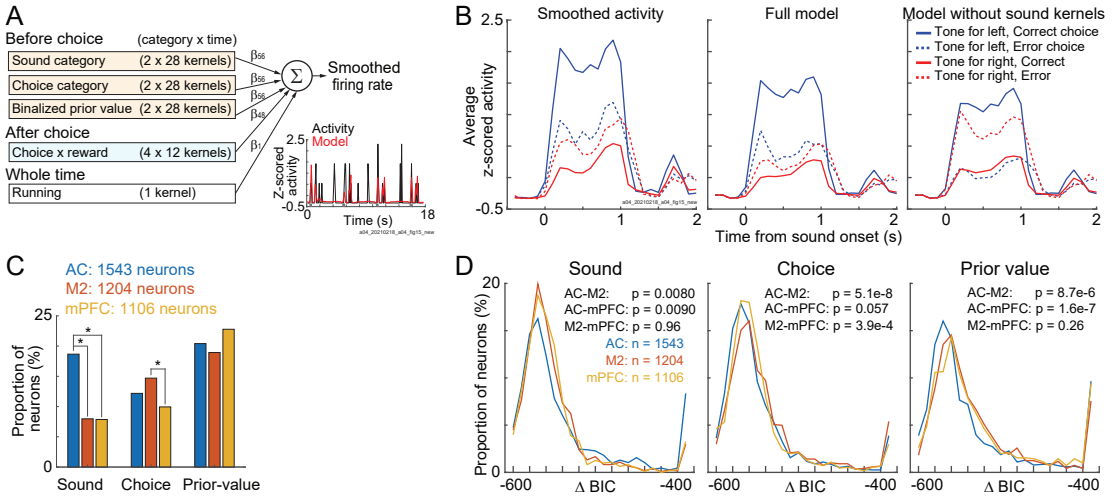

A
Before choice
(category x time)
Sound category
(2 x 28 kernels)
Choice category
(2 x 28 kernels)
Binalized prior value
(2 x 28 kernels)
After choice
Choice x reward
(4 x 12 kernels)
Whole time
Running
(1 kernel)
Smoothed firing rate
Activity
Model
Z-scored activity
Time (s)
B
Smoothed activity
Full model
Model without sound kernels
Tone for left, Correct choice
Tone for left, Error choice
Tone for right, Correct
Tone for right, Error
Average z-scored activity
Time from sound onset (s)
C
AC: 1543 neurons
M2: 1204 neurons
mPFC: 1106 neurons
Proportion of neurons (%)
Sound
Choice
Prior-value
D
Sound
AC-M2: p = 0.0080
AC-mPFC: p = 0.0090
M2-mPFC: p = 0.96
AC: n = 1543
M2: n = 1204
mPFC: n = 1106
Choice
AC-M2: p = 5.1e-8
AC-mPFC: p = 0.057
M2-mPFC: p = 3.9e-4
Prior value
AC-M2: p = 8.7e-6
AC-mPFC: p = 1.6e-7
M2-mPFC: p = 0.26
Proportion of neurons (%)
Δ BIC

Figure 5

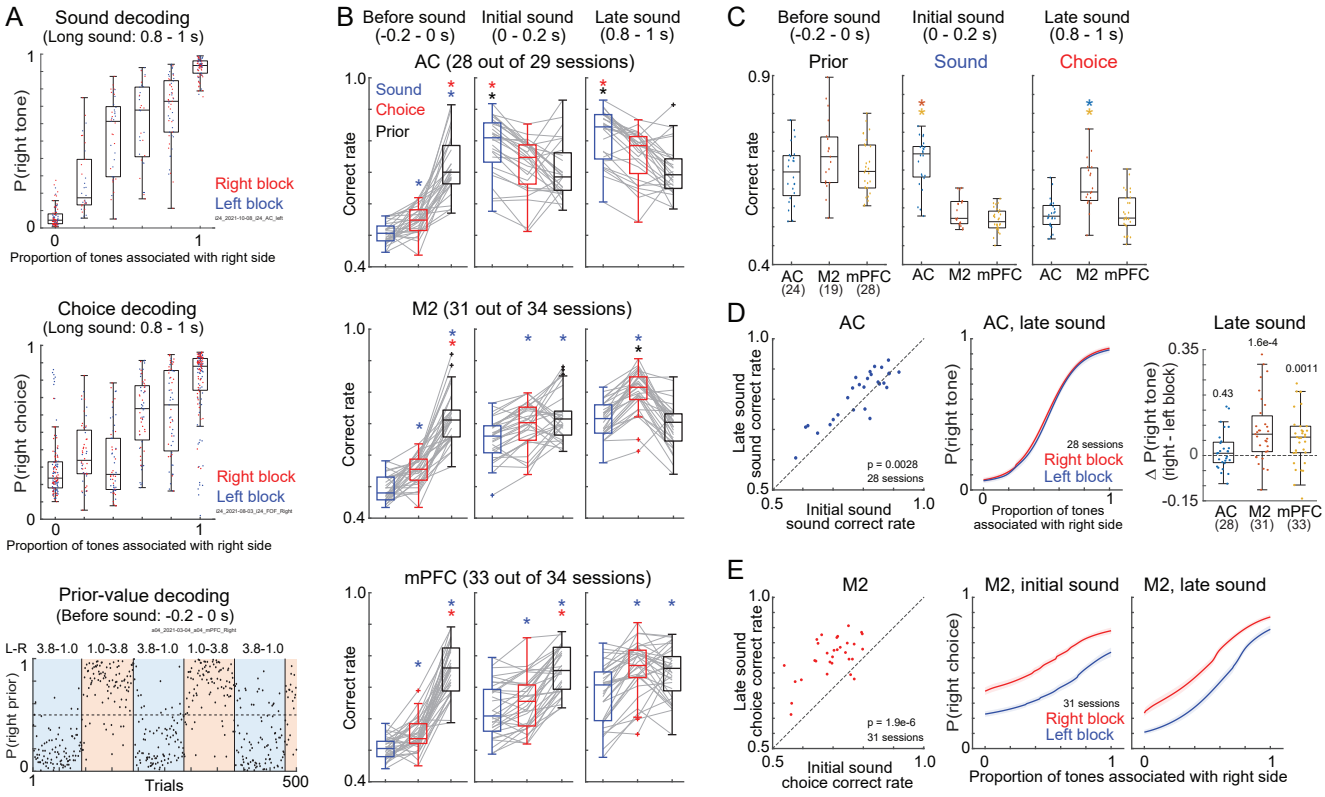

A
Sound decoding
(Long sound: 0.8 - 1 s)
P(right tone)
Right block
Left block
Proportion of tones associated with right side
Choice decoding
(Long sound: 0.8 - 1 s)
P(right choice)
Right block
Left block
Proportion of tones associated with right side
Prior-value decoding
(Before sound: -0.2 - 0 s)
L-R 3.8-1.0 1.0-3.8 3.8-1.0 1.0-3.8 3.8-1.0
P(right prior)
Trials
B
Before sound (-0.2 - 0 s)
Initial sound (0 - 0.2 s)
Late sound (0.8 - 1 s)
AC (28 out of 29 sessions)
Sound
Choice
Prior
Correct rate
M2 (31 out of 34 sessions)
mPFC (33 out of 34 sessions)
C
Before sound (-0.2 - 0 s)
Initial sound (0 - 0.2 s)
Late sound (0.8 - 1 s)
Prior
Sound
Choice
Correct rate
AC (24)
M2 (19)
mPFC (28)
D
AC
Late sound sound correct rate
Initial sound sound correct rate
p = 0.0028
28 sessions
AC, late sound
P(right tone)
28 sessions
Right block
Left block
Proportion of tones associated with right side
Late sound
Δ P(right tone) (right - left block)
1.6e-4
0.0011
0.43
AC (28)
M2 (31)
mPFC (33)
E
M2
Late sound choice correct rate
Initial sound choice correct rate
p = 1.9e-6
31 sessions
M2, initial sound
M2, late sound
P(right choice)
31 sessions
Right block
Left block
Proportion of tones associated with right side

Figure 6

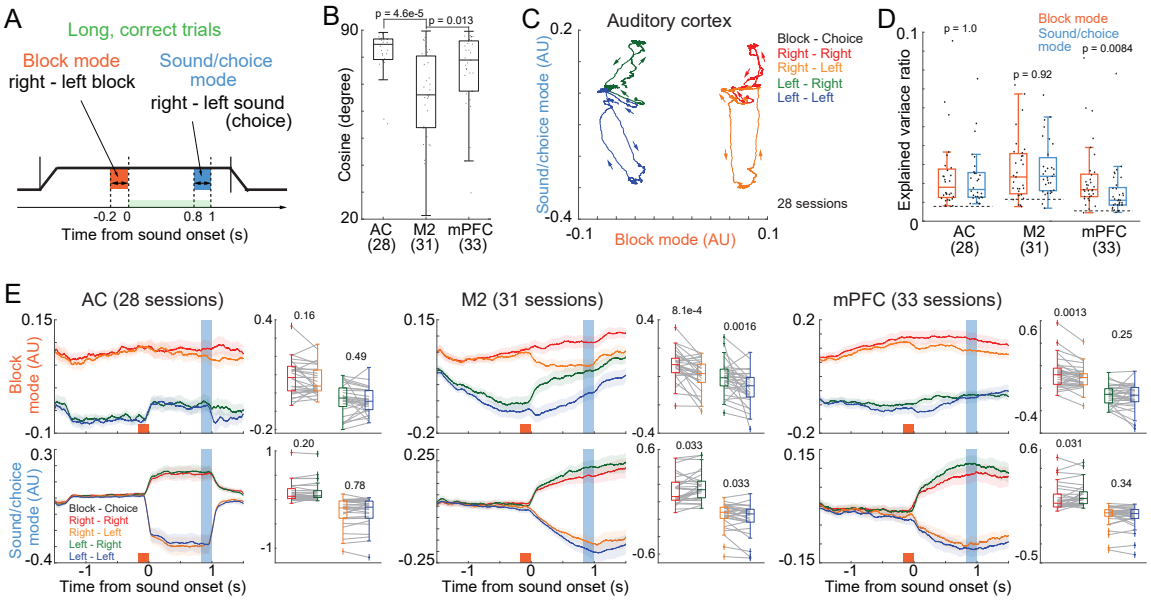

A
Long, correct trials
Block mode
right - left block
Sound/choice mode
right - left sound (choice)
-0.2 0
0.8 1
Time from sound onset (s)
B
p = 4.6e-5
p = 0.013
Cosine (degree)
AC (28)
M2 (31)
mPFC (33)
C
Auditory cortex
Sound/choice mode (AU)
Block mode (AU)
Block - Choice
Right - Right
Right - Left
Left - Right
Left - Left
28 sessions
D
Explained variace ratio
Block mode
Sound/choice mode
p = 1.0
p = 0.92
p = 0.0084
AC (28)
M2 (31)
mPFC (33)
E
AC (28 sessions)
M2 (31 sessions)
mPFC (33 sessions)
Block mode (AU)
Sound/choice mode (AU)
0.16
0.49
0.20
0.78
8.1e-4
0.0016
0.033
0.033
0.0013
0.25
0.031
0.34
Time from sound onset (s)

Figure 7