nature communications

Article

https://doi.org/10.1038/s41467-025-60504-y

Distinct neural population code and causal roles of primate caudate nucleus in multimodal decision-making

Received: 7 October 2024

Accepted: 23 May 2025

Published online: 06 June 2025

Check for updates

Zhao Zeng[1,2], Ce Zhang[1,2], Yue Xu[1,2], Hua He[3] ✉ & Yong Gu[1,2] ✉

Perceptual decision-making involves distributed networks spanning both association cortices and subcortical areas. A fundamental question is whether such a network is highly redundant, or each node is distinct with unique function. Using a visuo-vestibular decision-making task, here we show the subcortical caudate nucleus (CN) of male primates displays distinct population code compared to association cortices along the modality dimension. Specifically, in a low-dimensional state subspace, neural trajectory in the frontal and posterior-parietal association cortical activity during multimodal-stimulus condition evolves along the visual trajectory, whereas along the vestibular trajectory in the CN. We then show CN population activity is consistent with the animal's behavioral strategy employed within a generalized drift-diffusion framework. Importantly, causal-link experiments, including application of GABAa-receptor agonist, D1-receptor antagonist, and electrical microstimulation, further confirmed CN's critical contributions to perceptual behavior. Our results confirm CN's vital importance to decision making in complex environments with multimodal information.

Cognitive functions are mediated by neural networks in the brain, and it is frequent to see many nodes in the network share similar properties, leading to the so-called "redundant" information processing hypothesis[1–3]. Perceptual decision-making, for example, is one of the most important cognitive functions by which the brain transforms sensory inputs into perception, decision, and motor output[4]. During this process, ramping-like activity, proposed to reflect evidence accumulation leading to ultimate choice, has been observed widely distributed in both cortical and subcortical regions[3,5,6] (Fig. 1a), including lateral intraparietal area[7] (LIP), frontal eye field[8] (FEF), superior colliculus[9] (SC), caudate nucleus[10] (CN). It remains unclear whether each node may carry a unique role that could be easily masked by representational redundancy for the purpose of robust control of the system[11].

Among these areas, the dorsal medial striatum (DMS), named as CN in primates, is well known for mediating a variety of brain functions, including motor execution[12,13], value evaluation[14,15], time perception[16,17], and more. What is the role of CN in the distributed sensory-motor association network? One hypothesis is that the choice-related signals in CN largely stem from cortical regions such as LIP and FEF[18,19]. This hypothesis is consistent with works showing that the caudate receives heavy projections from the neocortex[20], and that CN activity largely reflects cortical activity[21,22]. However, recent studies have also indicated a number of differences between CN and the sensory-motor association cortex including: (1) In perceptual decision-making, ramping activity is typically sustained in cortex while declined in CN at the late stage of a trial[23]; (2) The bias in the accumulation process in the cortex but not CN, is more related with the animals'

[1]Key Laboratory of Brain Cognition and Brain-inspired Intelligence Technology, CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Neuroscience, International Center for Primate Brain Research, Chinese Academy of Sciences, Shanghai, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]Department of Neurosurgery, Shanghai Eastern Hepatobiliary Surgery Hospital, Shanghai, China.
✉e-mail: hehua1624@smmu.edu.cn; guyong@ion.ac.cn

https://doi.org/10.1038/s41467-025-60504-y

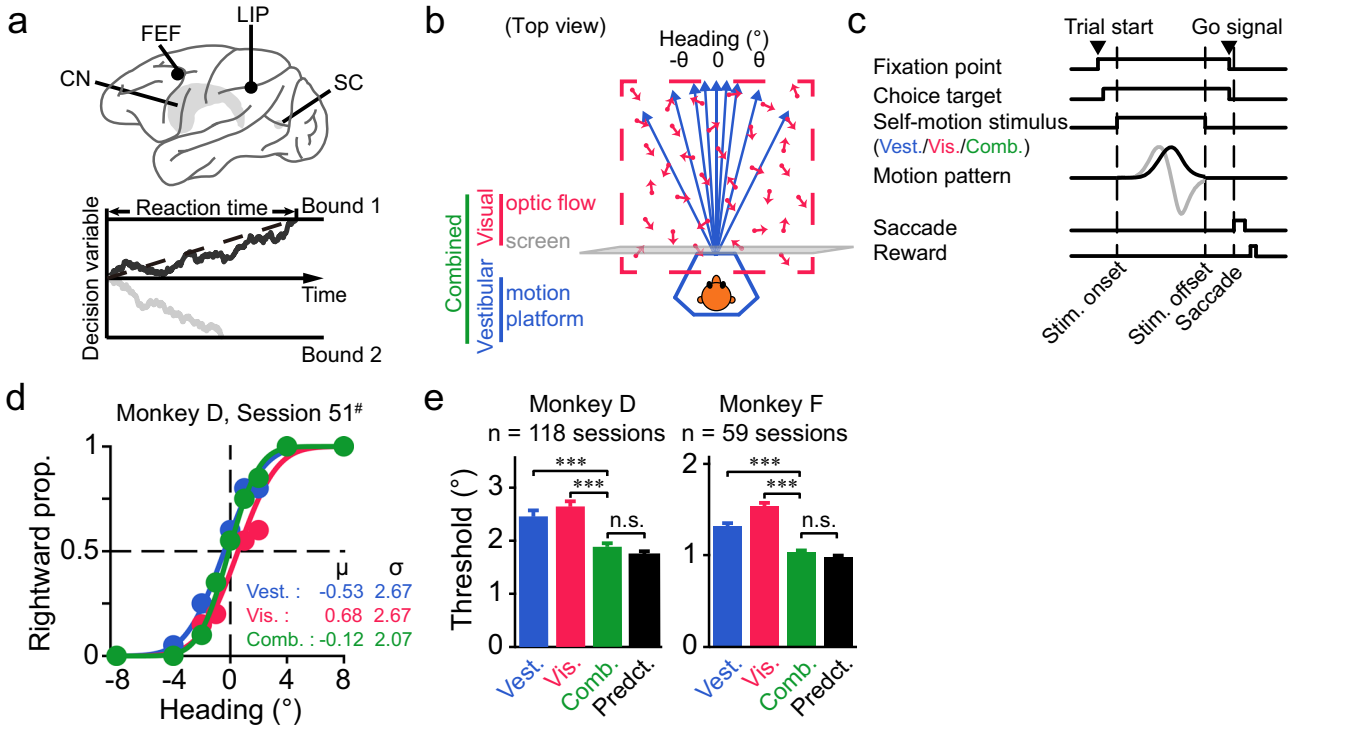


Fig. 1 | Comparing distributed perceptual decision-making signals using a multisensory paradigm. a Distributed brain regions (upper) representing evidence-accumulation signals (lower) in a perceptual decision-making task. CN caudate nucleus, FEF frontal eye field, LIP lateral intraparietal area, SC superior colliculus. **b** Schematic of a visuo-vestibular heading discrimination task using a virtual reality system. **c** Timeline of a trial. The motion stimulus profile includes a varied Gaussian velocity (black curve) and biphasic acceleration profile (gray curve). **d** Psychometric curves of an example session. The proportion of rightward choice is plotted as a function of heading for each modality condition (blue, vestibular-only; red, visual-only; green, combined). Dots are real data, and curves are fitted cumulative Gaussian functions. The μ and σ, respectively represent the mean and standard deviation of the best-fitted Gaussian function. **e** Averaged psychometric thresholds of two monkeys for three modality conditions and for optimal prediction from Bayesian cue integration theory (black bar). Error bars indicate SEM; *** and n.s., respectively, denote $p < 1e\text{-}03$ and no significance; two-tailed paired t-test.

reward-biased perceptual decision[24]; (3) In rodent, DMS but not association cortices is more causally related with decision performance[25–27].

In summary, although there is evidence suggesting that CN may play distinctive functions from the cortex, the so far discovered difference is restricted within the choice dimension, and its effective size is modest. To further explore how CN may be different from the cortical counterparts, in the current study, we introduced an additional dimension of modality by using a multisensory decision-making paradigm. Specifically, macaques were trained to discriminate heading directions based on reliability-matching visual (optic flow) and vestibular (inertial motion) cues (Fig. 1b, c), and the combination of both cues[28]. During the task, neural activity of single-units from CN was recorded and compared with that in FEF, and LIP recorded previously[29,30] (see "Methods"). We are interested in searching for similarity or heterogeneity across different areas, especially along the modality dimension.

Results

CN encodes multiple variables in a multimodal decision-making task

As briefly aforementioned, two macaque monkeys (monkey D and F) were trained to perform a visuo-vestibular heading discrimination task under a fixed-duration (FD) two-alternative forced choice paradigm (Fig. 1b). In each trial, a linear forward heading stimulus with a small deviation from straight ahead was offered, lasting 1500 ms. After a random delay (300–600 ms, uniform sampling), the monkeys were required to report their perceived direction, either left or right, by making a saccade to the corresponding target to earn reward (Fig. 1c). There were three sensory modality contexts: (1) vestibular-only condition, in which heading stimulus was provided solely by physical movement of a motion platform; (2) visual-only condition, in which heading was only simulated by optic flow on a display; (3) cue-combined condition, in which congruent vestibular and visual stimuli were offered at the same time. Note that cue reliability was carefully manipulated so that the behavioral performance would be similar between the two single cue conditions[31]. All stimulus conditions, including three sensory modalities and nine heading angles, were interleaved in one experimental session. After training, both monkeys performed well, illustrated by sigmoid heading-modulated psychometric curves (Fig. 1d). To quantify the performance, we fitted the psychometric curves via cumulative Gaussian function and took the parameters of mean (μ) and standard deviation (σ) to represent point of subjective equality (PSE) and psychophysical threshold, respectively. It turned out that in both animals, psychophysical threshold was comparable in the visual-only and vestibular-only condition, and became significantly smaller in the cue-combined condition (Fig. 1e). Importantly, performance under multimodal condition was close to that predicted from Bayesian optimal integration theory (Fig. 1e, green vs. black bar), which has been frequently observed in many previous studies[29–32].

While the animals performed the task, 200 well-isolated neurons were recorded from CN. Among them, 169 (monkey D, 117; monkey F, 52) putative medium spiny neurons (MSN) were identified according to their electrophysiological properties (Supplementary Fig. 1a–c)[33]. Peristimulus time histograms (PSTHs) of a few example neurons were shown to demonstrate that firing patterns of single CN neurons contained large heterogeneity (Supplementary Fig. 1d). To quantify the strength of task variables, we computed choice divergence (CDiv, Supplementary Fig. 1e) and modality divergence (MDiv) for each

neuron, using receiver operator characteristic (ROC) analysis (see "Methods"). Absolute values of CDiv and MDiv close to 1 respectively indicate stronger choice and modality selectivity, while they close to 0 represent small selectivity. For CDiv, many CN neurons significantly possessed choice signals in at least one modality condition (any choice cell, AnyC; permutation test, $p < 0.01$; monkey D, 60/117 = 51.3%; monkey F, 33/52 = 63.5%, Supplementary Fig. 3e). The mean CDiv was significantly larger than 0 (vestibular-only: 0.123, $p = 9.75e\text{-}06$; visual-only: 0.099, $p = 2.83e\text{-}04$; combined: 0.139, $p = 5.28e\text{-}06$; two-tailed t-test; Fig. 2a), indicating that CN overall prefers contralateral choice. For MDiv, many CN neurons also carried significant sensory modality signals (permutation test, $p < 0.01$; monkey D: 88/117 = 75.2%; monkey F: 36/52 = 69.2%), with roughly equal visual and vestibular preference (mean MDiv: −0.012, $p = 0.717$; two-tailed t-test; Fig. 2a). These cells with choice or modality preference were distributed along the longitudinal axes in a mixed manner (Supplementary Fig. 1f, g).

Previous studies have shown that sensory-motor association areas such as the posterior parietal cortex encode multiple task variables using a manner of category-free coding manner in multisensory decision tasks[29,34]. This was a similar case in CN. Firstly, numerous CN cells simultaneously had significant choice and modality preferences (permutation test, $p < 0.01$; vestibular-only: 47/169 = 27.8%; visual-only: 43/169 = 25.4%; combined: 54/169 = 31.9%). Secondly, there was neither a well-clustered distribution nor significant correlation between choice and modality signals (Fig. 2a). However, either task variable could be successfully extracted from CN population activity by using linear demixing algorithms such as demixed principal component analysis (dPCA, Supplementary Fig. 1h)[35]. We found the interaction term of choice and modality also captured response variances to some extent (Supplementary Fig. 1h, the third column). This nonlinear mixed selectivity has been previously shown to be computationally more flexible for dealing with complex cognitive tasks[36].

At the population level, we sorted trials in each modality condition according to animal choices and averaged firing rates of all neurons to construct population PSTH, finding significant choice modulation under all three contexts (Fig. 2b). Similar to LIP/FEF[29,30], choice-related signals in visual-only context were delayed compared to the other two stimulus conditions (Supplementary Fig. 2a), implying that CN also integrates different dynamic signals. Specifically, choice signals in CN diverged around the peak time of the acceleration profile in the vestibular condition, and around the velocity peak time in the visual condition, suggesting that the vestibular acceleration but visual velocity information was collected for evidence accumulation (Supplementary Fig. 1i). Majority of CN neurons are tuned to subjects' internal choices but not external physical heading stimuli, which was supported by similar firing pattern in correct and error trials (Supplementary Fig. 3a, b) and markedly larger partial correlation between choice and CN neuronal responses (Supplementary Fig. 3c). More directly, we tested heading tuning of CN single neurons in the passive task where decision reports were not required, and found that very limited proportion of CN neurons were significantly selective to sensory headings (Supplementary Fig. 3d). Besides, choice signals for each neuron were highly correlated across modalities (Supplementary Fig. 3f).

In summary, these basic neural properties indicate that CN is within the sensory-motor association network, by showing similar patterns at first glance as seen in other association cortices such as LIP and FEF.

CN shows a distinct neural manifold from the frontal and parietal association cortex

Sensory-motor association areas usually contain dynamic and mixed signals that are not easily identified on single neurons, thus, we constructed a high-dimensional neural state space[37] for population activity recorded in CN in the current study, as well as in FEF and LIP from previous studies[29,30]. The high-dimensional space was then reduced to a low-dimensional subspace to capture principal traits. Specifically, population activity in each region was denoised by using principal component analysis (PCA), and the first three PCs (capturing more than 65% variance) were adopted to form a 3-D neural subspace (Fig. 2c). Two features could be readily extracted from this manifold subspace. First, neural trajectories sorted by choice evolved from a common starting point, and then gradually diverged in all three stimulus conditions as a function of time. Such a pattern is qualitatively similar in all areas. Indeed, Euclidean distance between each pair of trajectories (solid versus dashed curves), defining choice distance (Fig. 2d), illustrated a ramp-like trait with an obvious visual delay compared to vestibular in all three areas. Second, in contrast to the choice dimension, we found that neural trajectories sorted by modality dramatically differed across areas. Specifically, neural trajectory in the multimodal condition was fairly close to vestibular in CN, while was on contrary close to visual in LIP and FEF, as clearly seen in the manifold subspace (Fig. 2c). To quantify this, the Euclidean distance between each pair of modality comparison in the neural state subspace was computed for each area (Fig. 2e). Such a cortical-subcortical distinction was further confirmed by normalizing multimodal-unimodal (area under cyan or magenta curve in Fig. 2e) distance to visuo-vestibular (area under black curve in Fig. 2e) distance (Fig. 2f). All these results were also supported by a targeted dimensionality reduction (TDR) algorithm[38], by which the high-dimensional neural population activities were directly projected onto a meaningful subspace spanned by choice and modality axes (Fig. 2g). These results held across animals (Supplementary Fig. 4).

Recurrent neural network reproduces inter-area neural state heterogeneity

To help understand possible sources of inter-area heterogeneity in neural state trajectory along modality dimension, we turned to a recurrent neural network (RNN) constituting fully connected, nonlinear neurons that performed a visuo-vestibular heading discrimination task, in a way similar to the animals. Specifically, RNN neurons received noisy heading evidence from vestibular acceleration and visual velocity inputs (Fig. 3a), and were connected to a single linear read-out decoder, which made decisions based on all hidden units' activities via weighted summation. In each simulation, the RNN's only goal was to categorize heading directions either leftward or rightward as accurately as possible, based on vestibular-only, visual-only, or multimodal sensory inputs, without the need to separate the cue modality. During training, the network initially showed relatively high thresholds, yet they decreased soon after tens of trials before reaching a plateau (Fig. 3b, e, middle panels). Interestingly, the network automatically showed improved performance under multimodal stimulus conditions, close to the Bayesian prediction as seen in the animals' behavior (Fig. 3b, e, right panels).

We then explored what factors might mediate the neurophysiological findings, in particular, the heterogeneity in the CN and cortical areas. We first examined the factor of the time-lag, suspecting that the leading vestibular choice signal (aligned with acceleration) might cause the vestibular bias in the CN. Thus, we temporally-reversed the two stimulus inputs such that the visual input was ahead of vestibular (Supplementary Fig. 5a). We found in such a case, the patterns were largely reversed in both the choice and modality dimension (Supplementary Fig. 5b, c), compared to the CN manifolds. Thus, the leading vestibular signal can already explain the data in the CN in both of the choices (vestibular choice leading) and modality dimension (vestibular bias, see the CN Model in Fig. 3b–d). However, simply temporally-reversing the two sensory inputs cannot account for the FEF/LIP data, because while the operation successfully achieved visual bias in the modality dimension, the visual signal became leading the vestibular along the choice dimension (Supplementary Fig. 5a–c), which was

https://doi.org/10.1038/s41467-025-60504-y

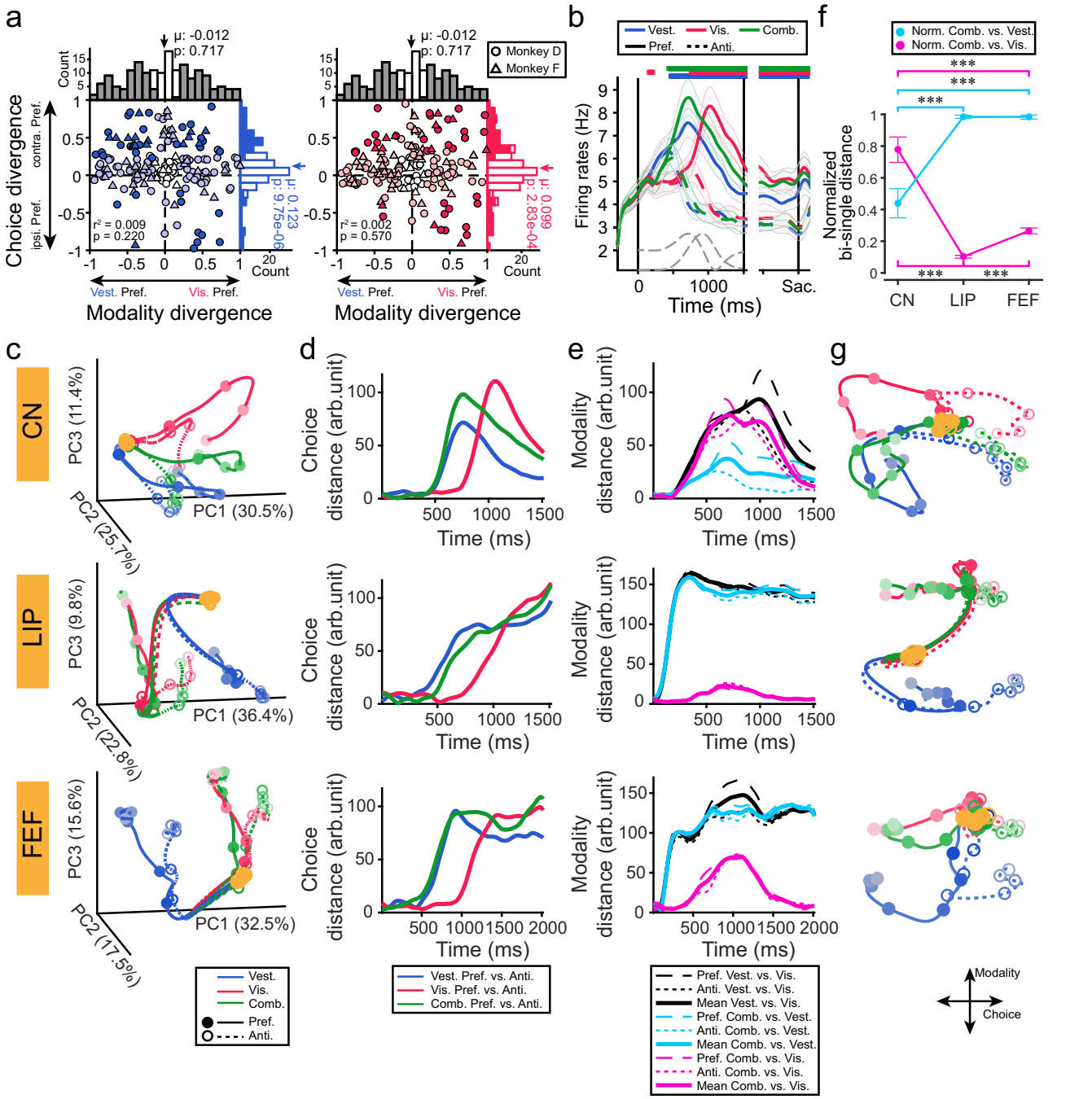


Fig. 2 | CN manifold and comparison with frontal and parietal association cortices. a Choice and modality signals of each CN neuron in vestibular-only (left) and visual-only (right) conditions. Dark dot: both choice and modality signals are significantly different from 0; Light dot: either variable is significant; Empty dot: neither variable is significant ($p < 0.01$; permutation test). Shape of dot: different monkeys. The p values in the bar and scatter subplots are from a two-tailed t-test and Pearson correlation, respectively. Note that the left and right subpanels share the same modality divergence distribution. **b** Averaged population PSTHs of CN neurons. Firing rates were aligned to stimulus onset (1st vertical line) or saccade onset (3rd vertical line). The 2nd vertical line indicates stimulus offset. Shaded error bars are SEMs. Horizontal colored bars indicate significant choice divergence ($p < 0.01$; two-tailed t-test). Solid and dashed lines represent animals choosing preferred and anti-preferred targets, respectively. Line color indicates sensory modality: blue, vestibular-only; red, visual-only; green, combined. **c** High-dimensional neural state spaces in CN (upper), LIP (middle), and FEF (lower) were projected onto a 3-dimensional subspace constructed by the first three principal components (PCs). The percentage in each axis represents the explained total variance of this PC. Curve color and style are the same as (**b**). **d** Choice distance evolves within trials in CN (upper), LIP (middle), and FEF (lower). Color indicates stimulus condition: blue, vestibular-only; red, visual-only; green, cue-combined. Time is from stimulus onset to offset. The arb. unit indicates an arbitrary unit. **e** Modality distance evolves within trials in CN (upper), LIP (middle), and FEF (lower). Color indicates each comparison of different modalities: black, vestibular-only versus visual-only; cyan, cue-combined versus vestibular-only; magenta, cue-combined versus visual-only. Thin dashed and dotted lines, respectively, represent modality distance of PREF and anti-PREF trials, while thick solid lines are the mean values of them. **f** Normalized single- versus multimodal Euclidean distance comparison across brain areas. Cyan and magenta represent normalized combined-vestibular and combined-visual distance, respectively. Dot and error bar represent mean and STD of 100-time samplings; *** indicates $p < 1.0e\text{-}10$, from two-tailed t-test. **g** High-dimensional neural activities in CN (left), LIP (middle), and FEF (right) are projected onto the subspace spanned by the two orthogonal axes of choice and modality using TDR analysis. Curve color and style are the same as (**b**).

https://doi.org/10.1038/s41467-025-60504-y

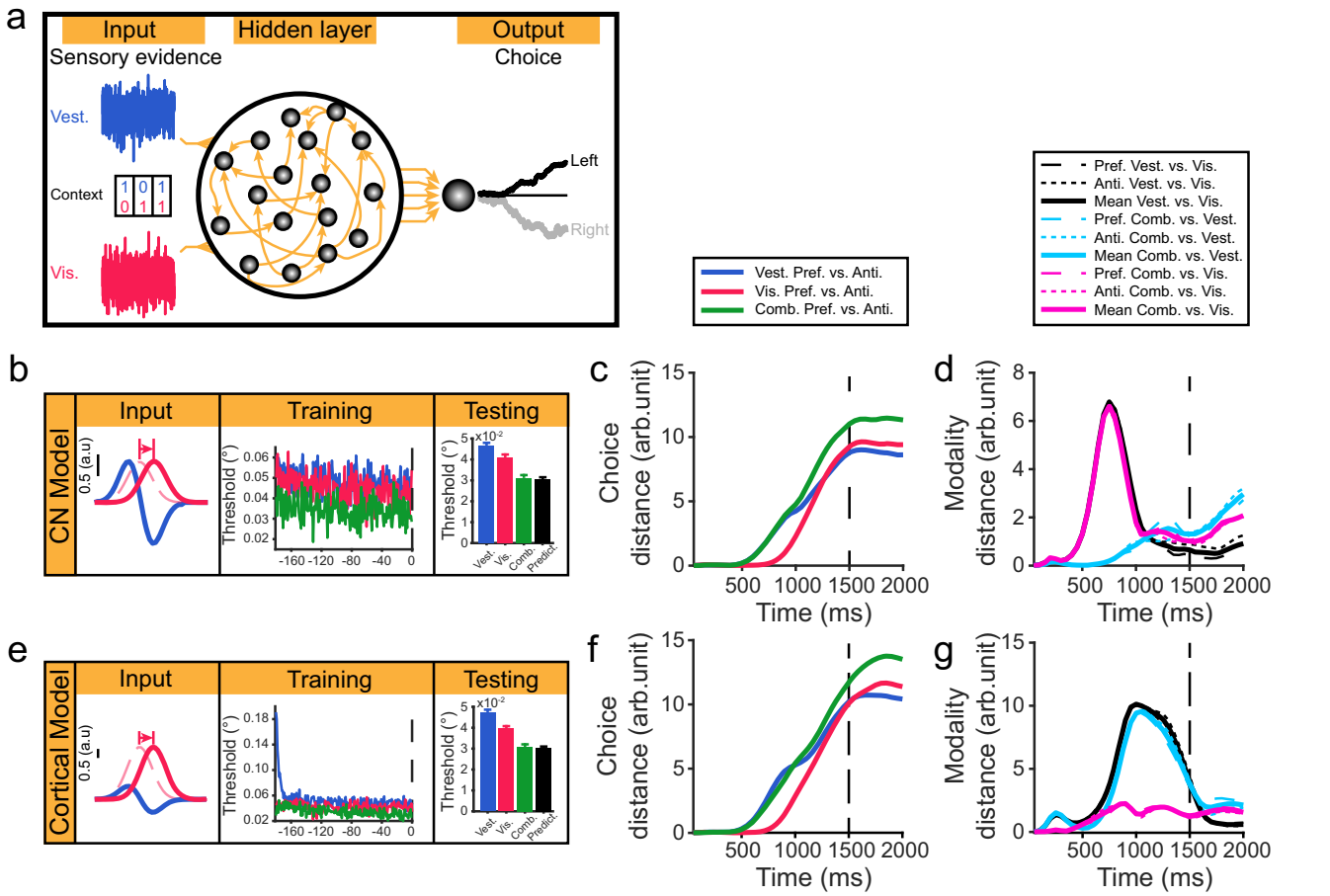


Fig. 3 | RNN simulations reproduced inter-area heterogeneity in modality distance. a Structure of a recurrently connected artificial network receiving noisy vestibular and visual heading inputs. Comparable vestibular and visual input magnitude and signal-to-noise ratio (**b**, left panel), RNN output behavioral performance (**b**, mid and right panels), choice (**c**), and modality (**d**) patterns. The red arrow indicates that we artificially delayed the visual input (solid versus dashed Gaussian profile in the left panel of **b**) to reproduce the visual-choice delay in (**c**). The solid profiles are real inputs. The vertical dashed line in (**c**, **d**) represents stimulus offset. Colors represent sensory modalities in (**b**, **c**). Meaning of curve color and style in (**d**) is the same as Fig. 2e, just as shown. Higher visual input magnitude with comparable signal-to-noise ratio (**e**, left panel), RNN output behavioral performance (**e**, mid and right panels), choice (**f**), and modality (**g**) patterns.

against the neurophysiological data. Therefore, in a second step, we explored the factor of the magnitude ratio of the two unisensory inputs under conditions when the acceleration and velocity peaks are temporally aligned to largely remove the time-lag effect (Supplementary Fig. 5d–f). Importantly, the input noise level was always controlled so that the signal-to-noise ratio of either input was comparable, to guarantee output of analogous visual and vestibular performance for each single cue. Under this operation, we found that implementing a relatively larger gain in the input magnitude, for example, the vestibular, could achieve vestibular bias in the modality dimension (Supplementary Fig. 5d–f), supporting the relative input magnitude as another critical factor. Thus, we then factored in the magnitude together with the time-lag to capture the cortical manifolds: a temporal delay and a higher magnitude in the visual input (Fig. 3e). We found that this achieved patterns in both choice (visual delay) and modality (visual bias) dimensions as in the LIP/FEF (the Cortical Model in Fig. 3e–g). In conclusion, the RNN simulations indicate that a visual delay is enough to generate the population activity in the CN, while the cortical population pattern requires an additional stronger visual input. The latter may be biologically plausible because it is consistent with the well-known roles of FEF/LIP involved in the vision-based decision networks.

Note, however, compared to the neurophysiology, RNN still produced a few dissimilarities in some details. First, the modality signals declined at the end of the trial, which was dissimilar to the FEF/LIP data. This is probably due to that the modality information is not a must for the task. Indeed, when the Cortical Model RNN was also asked to discriminate modality in addition to heading direction, the modality signals remained to the end of the trial (Supplementary Fig. 6a–c). Second, for the non-sustained choice signals in CN, this could be simulated if we trained the CN Model to perform an early-termination version task (Supplementary Fig. 6d–f), in which we artificially instructed the RNN to stop decision formation in advance, although the sensory evidence was still present. In this case, the choice-related activity would decline soon, consistent with the hypothesis that primate CN may contribute more to early stage of decision formation[18], particularly in a long fixed-duration task.

CN dynamics reflect perceptual performance in a multisensory context

In the previous sections, we have shown that the CN exhibits distinct population representation from the FEF/LIP, suggesting that the subcortical area is not simply reflecting everything in the cortical partners. In the following, we then examined whether CN activity is critically involved in perceptual decision-making. We first addressed how CN neuronal dynamics were functionally coupled with the animals' perceptual performance using a number of analytical methods, and then examined whether such correlations were "causal" using perturbation methods as would be illustrated in the next sections.

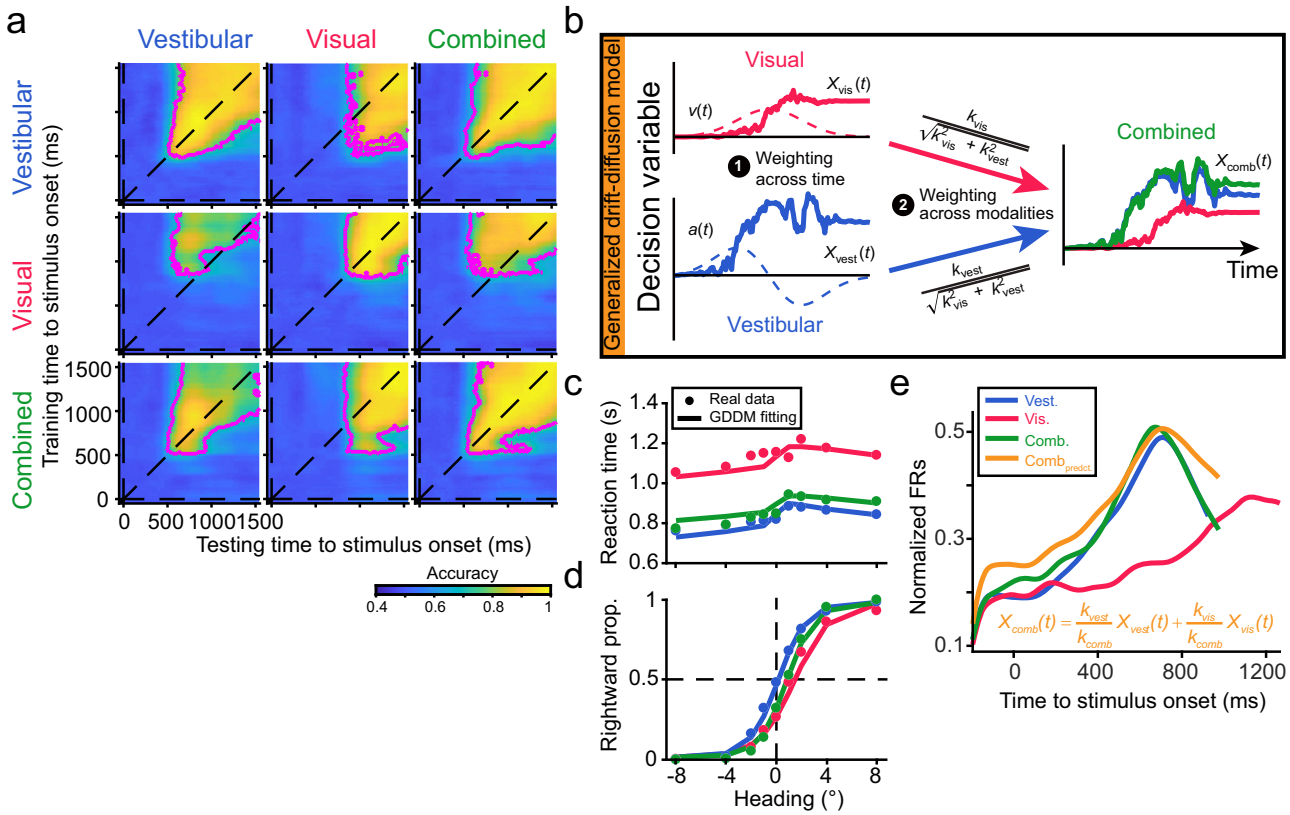


Fig. 4 | CN activity reflects animal multisensory decision behavior. a Decoding accuracy of cross-temporal Lasso classifiers presented as heat maps. Subpanels on the principal diagonal are within-modality decoding, while those off the principal diagonal represent across-modality predictions. For each subpanel, the ordinate represents the center of the time window used to train decoders, while the abscissa marks the testing time. Magenta lines outline the areas with accuracies significantly higher than chance level (50%, $p < 0.01$, two-tailed permutation test). **b** Schematic of the GDDM. Plot modified with permission from Jan Drugowitsch, Gregory C DeAngelis, Eliana M Klier, Dora E Angelaki, Alexandre Pouget (2014) Optimal multisensory decision-making in a reaction-time task eLife 3:e03005[39]. https://doi.org/10.7554/eLife.03005, under a CC BY license (https://creativecommons.org/licenses/by/4.0/). **c** Chronometric curves. Dots are measured data from monkey M. Lines are fitting curves from GDDM. Colors represent sensory modalities. **d** Psychometric curves. Labels are the same as (**c**). **e** The population firing rates averaged from 34 AnyC CN neurons (normalized). The orange line represents the predicted multimodal response from GDDM. Blue, red, and green lines are measured CN responses in the vestibular-only, visual-only, and multimodal conditions, respectively.

The first analysis was to train a series of Lasso decoders over time to predict an animal's choice on a single-trial basis from CN's population activity (see "Methods"). We found monkeys' upcoming choice could be reliably decoded about 500 ms after stimulus onset in the vestibular-only and multimodal conditions, and about 850 ms in the visual-only condition (Fig. 4a, subpanels along the principal diagonal). Importantly, classifiers trained using activity during the early period could be used to predict final choice when fed with population responses of the later stage, which is demonstrated by the light-yellow pixels away from the unity line (Fig. 4a, subpanels along the principal diagonal). In addition, classifiers trained in a specific sensory modality could predict decisions in other stimulus conditions (Fig. 4a, subpanels off the principal diagonal). Such generalization implies that CN encodes choice signals by using a similar strategy across time and modalities. That is, neuronal weights are consistent when the CN population activity is read out by downstream areas.

The second analysis was to more directly assess whether CN dynamics could show optimality as seen in the behavior (Fig. 1e). Previous research has proposed a remarkable framework named as generalized drift-diffusion model (GDDM, Fig. 4b, see "Methods"), which can successfully explain human subjects' heading performance under reaction-time (RT) context when both factors of choice correctness and response time were taken into account to estimate multisensory-integration optimality[39]. To do so, we trained a third monkey (M) to perform an RT version of the same multisensory heading discrimination task, in which the only difference was that the animal was allowed to make saccadic choice at any time after a minimum requirement of 700 ms from the stimulus onset (see "Methods"), without having to wait for the disappearance of the fixation point indicating the go-signal. In such a situation, we found that the animal showed different response times across modality conditions. In particular, response time was significantly longer in the visual condition (~1150 ms on average), compared to that in the vestibular (~820 ms on average) and combined (~850 ms on average) conditions (Fig. 4c, dot symbols). These patterns were similar to the human study[39]. GDDM well-fitted behaviors, including both chronometric (Fig. 4c, solid curves) and psychometric curves (Fig. 4d, solid curves). Importantly, the fitted drift rates, k-values (see "Methods"), representing subjective stimulus sensitivity, were close between the multimodal condition and prediction from optimal integration theory (Table 1, $k_{predict}$: 166.79, calculated by formula #9), suggesting that monkey M integrated visual and vestibular cues near optimally under RT context. After acquiring k-values, the predicted moment-to-moment multimodal response in CN was consequently calculated, which turned out to be very similar with the real neurophysiological data (Fig. 4e, dark yellow versus green curves), supporting that CN dynamics could reflect the animals' behavioral performance within the GDDM framework under more complex context when multiple factors need to be considered including response time, multimodal cues, and performance accuracy.

Table 1 | Fitted mean value of parameters in GDDM

k_{vest}	131.88	b_{vest}	−0.16
k_{vis}	102.11	b_{vis}	−1.47
k_{comb}	179.07	b_{comb}	−0.86
θ_{vest}	0.22	$t_{(nd,vest)}$	0.23
θ_{vis}	0.19	$t_{(nd,vis)}$	0.40
θ_{comb}	0.14	$t_{(nd,comb)}$	0.33
p_{lapse}	0.03		

Unilateral chemical inactivation in CN biases perceptual choice

We then further identified causal-link roles of CN activity in the animal's perceptual performance. Specifically, CN's activities were manipulated through a number of methods, including chemical inactivation to test essential contributions and electrical microstimulation to test sufficient contributions.

We first applied muscimol, agonist of $GABA_A$ receptor to non-selectively suppress CN's overall excitatory activity (Fig. 5a). We found that unilateral inactivation (monkey D, 5 cases; monkey F, 11 cases; Fig. 5b, c), but not saline control (monkey D, 6 cases; monkey F, 9 cases; Fig. 5f) induced significant ipsilateral bias in both animals' choice. This ipsilateral bias was concordant with our recording data showing that majority of CN neurons preferred contralateral choice (Fig. 5a). In contrast, unilateral muscimol inactivation in LIP (monkey D, 3 cases; monkey F, 7 cases) and FEF (monkey D, 3 cases; monkey F, 7 cases; Supplementary Fig. 7a, b) surprisingly produced negligible effect on the animals' perceptual choice (Fig. 5d, e), although neurons in these two areas also mainly preferred contralateral choice (Supplementary Fig. 2b). The inactivation in FEF, however, obviously impaired saccadic precision (Supplementary Fig. 7c), an effect which was absent in LIP (Supplementary Fig. 7d). This different influence from perturbation on eye movements in the two areas, may be accordant with previous findings that inactivating LIP, especially the dorsal subdivision (Supplementary Fig. 7b)[40], only produces minor saccade deficits compared to FEF[41]. Other than PSE, muscimol application in CN did not significantly affect perceptual sensitivity (Supplementary Fig. 8a), or history information usage strategy (Supplementary Fig. 8b, c).

Because CN is strongly innervated by midbrain dopaminergic inputs[42], we wondered whether selectively modulating dopamine receptors of CN neurons would affect the animals' task performance. Thus, we then applied the antagonist of D1 receptor, SCH23390, unilaterally into CN. We found that this manipulation induced significantly more contralateral choice (monkey D, 5 cases; monkey F, 10 cases; Fig. 5g), opposite to muscimol-mediated bias direction, implying that dopaminergic inputs into CN could modulate the evidence-accumulation process. Finally, similar to muscimol, blocking D1 receptors neither impacted heading sensitivity (Supplementary Fig. 8d) nor trial-history usage strategy significantly (Supplementary Fig. 8e, f).

These chemical-inactivation experiments suggest that CN plays a critical role in the sensory-motor decision task, yet one possibility is that the observed causal effects may be mainly due to motor execution instead of cognitive perception, particularly when considering that CN is well-known for modulating movement executing[12,43]. We think this is less likely to be the case in our study based on a few observations. First, chemical inactivation mainly affected difficult trials when heading directions were around the reference, which required more cognitive load compared to easy trials with large heading directions (Fig. 5b and Supplementary Fig. 8d, h). Second, we further trained the animals to perform a simple color-selection task that required much reduced cognitive load than the fine heading discrimination task (Fig. 5h). In this controlled task, muscimol and SCH23390 produced only minor influence on the animals' saccadic behavior (Fig. 5i, j), supporting that perturbation of CN activity mainly affected the animals' cognitive process.

Unilateral electrical microstimulation in CN biases perceptual choice

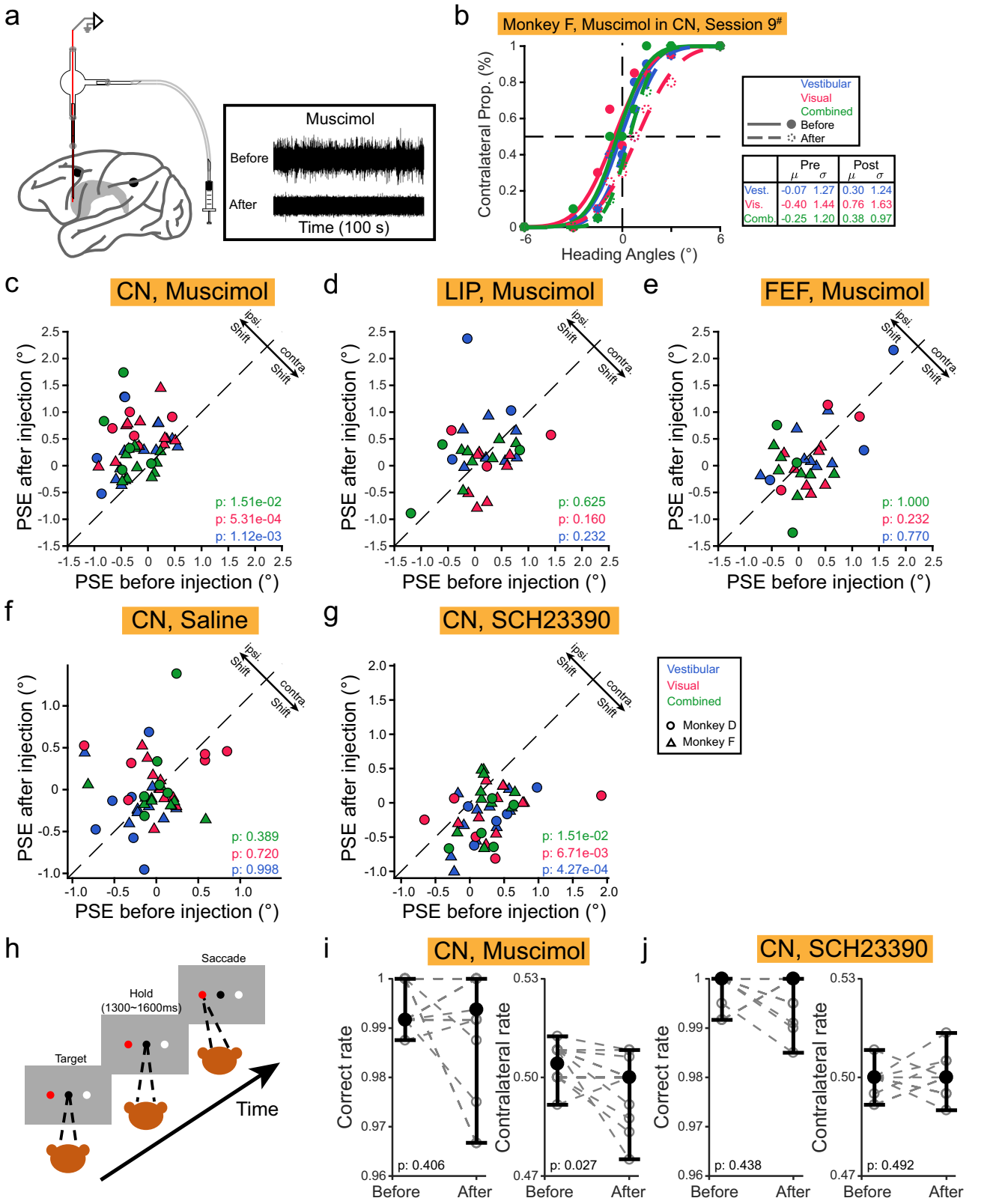
The third method for causal-link test was to use electrical microstimulation (50–80 μA, 300 Hz, cathode-leading biphasic, Fig. 6a). Microstimulation is typically thought to identify sufficient contributions by artificially activating neural activity[44,45], although it still remains a debate about its exact impact on evoked neural dynamics[46,47]. Specifically, in each experimental session, microstimulation was applied unilaterally in CN during the stimulus duration period, randomly for half of the total trials. Significances of PSE and threshold changes in the psychometric functions were assessed by a bootstrap test (Fig. 6b, c). Overall, microstimulation in CN was able to significantly bias PSE in a large proportion of cases (monkey D: 38/70 = 54.3%; monkey F: 52/76 = 68.4%, Fig. 6d). This proportion could be underestimated because of adaption of the currents[48] (Supplementary Fig. 9f). The effect was fairly consistent across sensory modality (Supplementary Fig. 9b). In contrast to PSE, microstimulation only modestly affected psychophysical threshold (Supplementary Fig. 9a), indicating that microstimulation in CN mainly introduced signals instead of noise[49], supporting that CN sufficiently contributes to the animals' perceptual choice.

We noticed, however, some inconsistency in the PSE shift direction between the two animals. Specifically, microstimulation tended to induce more contralateral choice in monkey D (Fig. 6d, top row), but more ipsilateral choice in monkey F (Fig. 6d, bottom row). This individual difference could not be explained by factors such as anatomical location (Supplementary Fig. 9c, d), or coding properties of the stimulated sites (Supplementary Fig. 9e). To test other possibilities, we applied microstimulation during the delay period after stimulus offset and before saccadic choice (Fig. 6e). We found in such a case, microstimulation effects were also different between the two monkeys. In particular, while microstimulation effect was absent for monkey D, the effect remained for monkey F (ipsilateral bias, Fig. 6f). Thus, unlike the consistence of bias directions induced by chemical inactivation, microstimulation appeared to induce more confound results that might be due to different strategy of the animals, or limitations in the technique per se (see more in Discussion).

Discussion

It is common to see that many brain functions are mediated by a distributed brain network involving multiple regions that share seemingly similar properties, leading to the "redundant" hypothesis[1–3]. One obvious benefit of this redundancy is that it helps to deal with dysfunction when some of the nodes are hampered. Alternatively, the seemingly redundant nodes may carry distinct functions for tasks under more complex contexts. Here we provide clear evidence, showing that for perceptual decision-making, the subcortical area CN is markedly different from the frontal and parietal association cortices along the modality dimension in a low-dimensional manifold space. In particular, neural trajectory under multimodal conditions is largely biased towards vestibular in CN, in contrast to a clear visual bias in LIP and FEF (Fig. 2). An RNN simulation provides a possible explanation that this cortical-subcortical distinction may result from inter-area differences in relative input strength of different sensory modalities (Fig. 3). We further establish the correlated (Fig. 4) and causal (Figs. 5 and 6) links between CN population activity and animals' behavioral performance in the multisensory-decision task. In conclusion, these results support the idea that CN is not a simple relay station for the cortico-striatal circuitry[23], and it is critically involved in perceptual decision-making.

https://doi.org/10.1038/s41467-025-60504-y



CN representational patterns may reflect animal decision strategy

Considering that CN shows a non-vision (vestibular) biased neural state during multimodal conditions when both single cues overall share similar cue-reliabilities, neither similar to the frontal (FEF) and parietal (LIP) sensory-motor cortex, nor to the polysensory area (MSTd, Supplementary Fig. 10), a question is, what does this imply for CN in multisensory perception and decision-making? Although in the current study, we have not directly tested this function (the animals were not required to discriminate modality information), we think this finding is consistent with a few behavioral observations in previous studies.

First, we find that CN population dynamics are consistent with predictions using parameters regressed from the GDDM fitting of

Fig. 5 | Unilateral inactivation of CN biases animal choice. a Chemical injection into targeted areas, including CN, LIP, and FEF. The inset shows an example case of diminished neural activity after muscimol delivery. **b** Example of one inactivation experiment, showing performance before (solid lines and dots) and after (dashed lines and open dots) muscimol unilateral injection in CN. Comparison of PSE before and after unilateral muscimol injections in CN (**c**), LIP (**d**), FEF (**e**), saline injections in CN (**f**), and SCH23390 injections in CN (**g**). Shapes represent data from different monkeys. Colors are different modality conditions. All *p* values are from a two-tailed Wilcoxon signed-rank test; significance level, 0.05. Dots above or below the dashed diagonal, respectively, represent that animal choices are biased toward the ipsilateral or contralateral side compared to the manipulated hemisphere. **h** Schematic of the color-guided saccade task. **i** Comparison of saccadic accuracy (left) and bias (right) before and after unilateral muscimol injection in CN. Solid dots and error bars indicate median and 95% confidence interval, respectively. Open dots with dashed lines are from the same experiment (14 pairs). All *p* values are from a two-tailed Wilcoxon signed-rank test; significance level, 0.05. **j** Same as (**i**) but for unilateral SCH23390 injection in CN (10 pairs).

animal behavior. It's noticeable that the fitted drift rate, reflecting speed of evidence accumulation, is higher for vestibular-only than visual-only conditions (Table 1), indicating an intrinsically higher sensitivity to vestibular stimulus[39]. Second, the vestibular bias may explain "vestibular-overweighting", as reported previously in both human and non-human primate studies by using spatially conflicting visuo-vestibular heading stimuli. Specifically, subjects' PSE during multimodal conditions with conflict unimodal cues is biased towards vestibular, even when the two unimodal cues carry matching reliability and analogous perceptual sensitivity (psychophysical threshold)[32,50]. Third, the vestibular bias is consistent with the difference in response time under different modality conditions. Specifically, researchers have shown that under an RT version of the task, human subjects tend to make choices faster in vestibular-only and multimodal stimulus conditions compared to that in the visual-only condition[39]. This is also exactly what we see in our current study, by training one of the animals to perform the RT task.

From the anatomical perspective, CN could selectively gain-modulate inputs from the cortical partners, or receive sensory evidence from elsewhere, for example, thalamus[51], or other sensory cortices[52–54], leading to the vestibular-biased pattern (Supplementary Fig. 11). Contrary to CN, the multimodal neural trajectory is highly akin to the visual in LIP and FEF. This may happen due to the fact that LIP and FEF in the primates are within the classic visual pathway, in which both areas receive heavy projections from many extrastriate visual areas[55,56], and mediate visual attention[57,58].

In addition to the vestibular bias in the neural state space, another feature in CN is that the Euclidean distance between different conditions ultimately decreases at the late stage of trials, unlike the sustained plateau in frontal and parietal cortices (Fig. 2d, e). This descending Euclidean distance actually also happens in the choice dimension, which may reflect overweighting of early phase stimulus and early termination of evidence accumulation during fixed-duration task, as seen from previous psychophysical analysis[59–61]. Especially in our stimuli, visual sensory evidence becomes fairly weak at the late stage due to the Gaussian velocity motion profiles, and vestibular evidence is mainly concentrated at the early acceleration period. Thus, the later phase of sensory information may be largely neglected by the animals. Indeed, when switching to RT context, the animal made saccadic choice much earlier (vestibular: ~820 ms, visual: ~1150 ms, multimodal: ~850 ms) than the whole 1.5 s as required in the fixed-duration context. Nevertheless, the dynamic choice signals in CN in our current study are consistent with previous findings, supposing that compared to cortex, CN should be particularly important for decision formation at an early stage[18]. Moreover, a recent work suggests that suppressing striatum activity will significantly impair rodents' decision-making only when animals are in the evidence-accumulation state[62]. In addition to the choice signals, the descending modality information in CN may also be due to similar reasons: discrimination of modality (e.g., visual or vestibular) is not required in our task, therefore, modality-related signals are ultimately faded at the end of trials. It would be interesting to test in future experiments by requiring subjects to discriminate modality conditions, in addition to heading stimuli, and examine whether related CN activity may be sustained consequently. Finally, compared to CN, both choice-related activity and modality signals are largely sustained throughout the stimulus duration in FEF/LIP. The exact reason is unknown here, but a few possibilities may include strong recurrent connections[63], redundancy and compensation[11], and non-evidence accumulation related premotor effect[60].

Using multisensory paradigms to parse the perceptual decision networks

Most previous perceptual decision-making studies have used only one modality, such as the visual random-dots task[64]. If the ultimate perceptual decision about a latent variable, for example, motion direction is really "abstract", then choice-related signals are expected to be independent on sensory modality, motor effector[65], or spatial location of choice targets[66,67]. In the current study, we show that monkey's CN encodes decision-related signals similarly across time and also stimulus conditions (Fig. 4a), implying that the latent variable of heading direction is computed in an abstract way. This idea may be consistent with our causal-link experiments in which manipulating CN activity affects the animals' perceptual choice similarly in all three modality conditions (Figs. 5 and 6, Supplementary Fig. 9b), further advancing our understanding about the roles of CN in perceptual decision-making based on the vision[10,68]. Importantly, under a multimodal context, a higher dimensional subspace could be constructed and may reveal more different traits among network nodes, as shown in our current study (Fig. 2). One pitfall is that we have only included congruent visuo-vestibular cues in the current study, and it would be interesting and necessary to investigate how the vision dominant cortices and non-vision dominant CN may contribute differently to multisensory decision when cues from different sensory modalities conflict[69].

For the results of causal manipulations, we show that, unlike CN, unilateral inactivation of FEF or LIP does not significantly induce consistent choice bias, which is surprising based on the knowledge about neural coding in these areas, as well as some previous studies indicating their causal effects[70,71]. However, the overall causality effects in sensory-motor association cortices are controversial, especially in monkey LIP[72,73] and rodent PPC[34,74,75]. For example, a recent study shows that inactivating monkey LIP can impair behavioral performance only when a local random-dots stimulus[73], but not the choice target[72] is placed within the response filed of the multi-units at the inactivation site, implying that it is mainly the sensory evidence evaluation, instead of the motor deliberation in LIP that shows the causal effect[73]. In our current study, we used a global optic flow stimulus, which could involve a large area spanning both hemispheres that can easily compensate for each other[11]. Indeed, our current inactivation has not been simultaneously applied to both LIP and FEF, as well as both hemispheres at the same time in one experiment. Hence, the current negative inactivation results in FEF/LIP should be cautiously interpreted. Nevertheless, the positive inactivation results in CN, in the current study, together with previous works in primates and rodents, stress the great importance of subcortical regions, including striatum[26,27,62,68,76] and SC[77,78], for perceptual decision-making. This is plausible because the subcortical regions are typically more conserved and less redundant, making it likely that any perturbations would cause more significant consequences than those applied in the neocortex. It would be interesting to further explore other subcortical areas such as the SC in multisensory decision-making, as this region is

https://doi.org/10.1038/s41467-025-60504-y

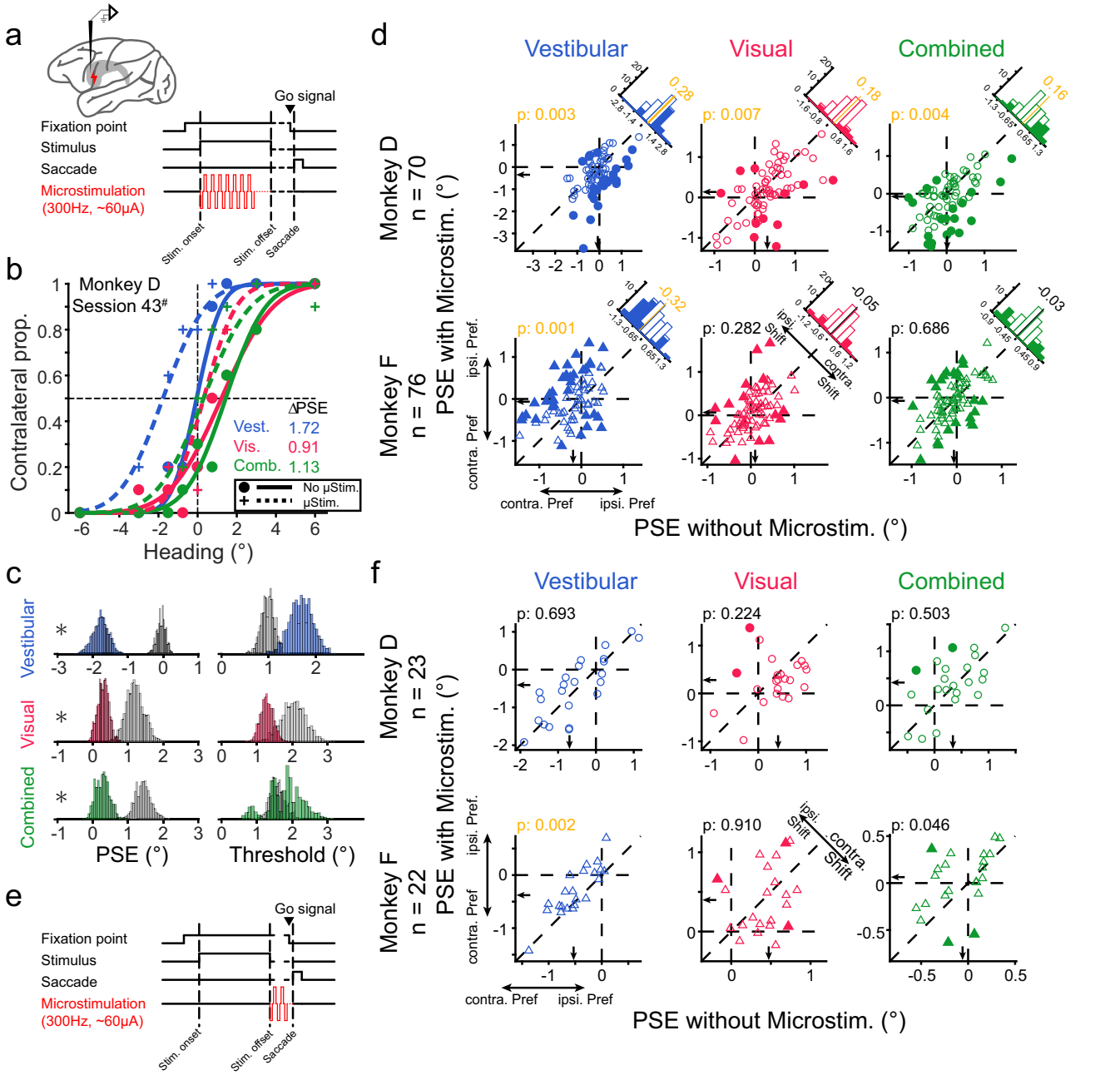


Fig. 6 | Unilateral microstimulation in CN affects animal choice. a Schematic of unilateral microstimulation applied in CN during the stimulus-presentation period (1.5 s). **b** Example of one experiment with microstimulation (dashed curves and cross symbols) and without microstimulation (solid curves and dots). Color represents sensory modality. **c** Distributions of PSE (left column) and threshold (right column) constructed by bootstrapping for microstimulation (color) and without microstimulation (gray) trials in the example session of (**b**). Asterisk indicates significant change of the PSE or threshold ($p < 0.01$; 1000-time bootstrap test, two-tailed). **d** Comparison of PSE with (ordinate) and without (abscissa) microstimulation during the stimulus-presentation period. Black arrows mark the median PSE. Marginal histogram counts the frequency of ΔPSE, with the median value indicated by a perpendicular solid line. Solid bars and dots indicate significant shifts; open bars and dots are insignificant changes. All p values are from a two-tailed Wilcoxon signed-rank test; significance level, 0.01. Dots above or below the dashed diagonal, respectively, represent that animal choices are biased toward ipsilateral or contralateral side compared to the manipulated hemisphere. **e** Microstimulation is applied during the delay period after the stimulus disappears. **f** Comparison of PSE with (ordinate) and without (abscissa) microstimulation during the delay period. Labels are the same as (**d**).

multisensory[79], and is also within the perceptual decision network[77,78], receiving heavy projections from both association cortices and basal ganglia circuits[18].

Dopaminergic modulation in CN is essential for perceptual decision

It has been recognized that dopaminergic regulation in the neocortex is essential for many cognitive functions, such as working memory and associative learning, as indicated by selectively injecting chemical ligands of dopamine receptors into the prefrontal cortex[57,80–82]. Here, we showed that suppressing CN D1 receptors using an antagonist consistently biased animals' choice toward the contralateral side, opposite to muscimol inactivation. Our result is thus concordant with the finding that a low dose of D1 receptor antagonist typically elevates neuronal activity[83,84].

It is well known that the basal ganglia circuit contains two pathways of direct and indirect that are associated with D1 and D2

receptors, respectively. The two pathways are likely responsible for complimentary and opposite control of decision behavior[62]. It would be interesting to further test the D1 and D2 receptor pathways in future experiments, as has been performed in rodents using state-of-the-art techniques[13,62,85]. For example, a recent work developed a recombinase-free retrograde adeno-associated virus-dependent strategy to precisely modulate CN direct MSNs[86]. It would be promising to identify primate CN circuit-specific contributions to cognitive functions in future studies.

Implications of the microstimulation experiment

Electrical microstimulation has been widely used in primates to identify causal roles of neural activity in perceptual choice[68,70,87]. The technique is quite successful when applied in sensory areas, as evident by induced PSE shift in the direction consistent with the labeled-line of the stimulated neurons (encoded feature)[49]. The other advantage of applying microstimulation in these sensory areas is that in primates, neuronal signals sharing similar functions are typically well clustered in a local spatial domain (e.g., a few microns)[49]. Microstimulation in higher-level areas, or less sensory-dominant areas, could be trickier, because different neuronal signals could be more mixed together, and they are less spatially clustered[68]. Indeed, in the current study, we found that although microstimulation in CN frequently induced significant PSE shift, confirming its sufficient causal role, the bias direction was not consistent between the monkeys. In particular, the ipsilateral bias is counterintuitive, based on the fact that the majority of CN neurons prefer contralateral choice. This result, however, may be consistent with previous studies: CN may contain two components that show opposite preference for perceptual and motor process[68,88]. Other factors may also confound the microstimulation results. For example, the electrical currents are easily spread and subsequently activate passing fibers, leading to nonlocal effects[46,47].

Different stimulus parameters may also account for heterogeneous results across different studies. For example, previous studies have shown significant dependency of CN microstimulation-induced PSE shift and activity properties of the stimulated neurons[68,88]. In the current study, however, we found pretty weak correlations. One possible reason is that our optic flow occupies a full screen with a very large visual field (90° × 90°), which may be more easily compensated after local perturbations of neurons with restricted receptive fields. Thus, an activity-dependent perturbation effect may be corrupted from being seen clearly. Last but not least, electrical currents lack specific targeting and may impact direct and indirect pathways in the striatum differently, consequently generating heterogeneous results. Optogenetic or chemogenetic tools, that are typically much more relieved from these issues, may be applied in future experiments to further test the sufficient roles of certain CN populations in perceptual decision tasks.

Methods

Animals and surgeries

Three adult male monkeys (*Macaca mulatta*), monkey D, F and M, weighting 7–10 kg, were trained in this work. Monkeys were chronically implanted with a 6-cm lightweight plastic ring, which served as a recording chamber as well as for head restraint. Then, monkeys sat comfortably in a customized primate chair fixed within a visuo-vestibular virtual reality system. All animal procedures were approved and supervised by the Animal Care Committee of the Center for Excellence Brain Science and Intelligent Technology, Chinese Academy of Sciences.

Apparatus

The visuo-vestibular virtual reality system consists of a motion platform (MOOG MB-E-6DOF/12/1000KG) and an LED display (~30 cm of view distance and ~90° × 90° of visual angle; Samsung ED55C) vertically mounted on the platform, to offer vestibular and visual self-motion stimuli, respectively. The stimuli were controlled by a customized C++ software and synchronized with electrophysiological recording (AlphaOmega SnR, Israel) and eye tracking (SR Research, EyeLink 1000 Plus, Canada) systems by TEMPO (Reflective Computing, U.S.A).

Behavioral tasks

Memory-guided saccade task (MGS). Similar to previous works[89], monkeys were trained to perform a memory-guided saccade task to identify the response field (RF) of the recorded neuron. However, note that we didn't use this task to screen neurons in the later heading discrimination task, considering that there is only a weak correlation between the strength of choice signal and spatial selectivity in CN[10]. In brief, monkeys were required to fixate on a central fixation point for 100 ms to initiate a trial, after which a peripheral target flashed for 500 ms in one of the eight positions with 10° away from the display center. Monkeys were required to keep central fixation for another 1000 ms, before the fixation point disappeared, indicating a go-signal for the memorized peripheral target.

Multisensory heading discrimination tasks. Monkeys initiated a trial by fixating on the central fixation point, after which two choice targets appeared symmetrically on each side of the display with an eccentricity of 10°. After a random delay (100–300 ms), a 1.5-s, linear forward self-motion stimulus, defined by either visual, vestibular, or multimodal cues, was provided in the horizontal plane with a small deviation from the reference of straight forward (0° heading). After the stimulus offset, the animals waited for another random delay (300–600 ms) before the fixation point disappeared as a go-signal. The animals were then allowed to saccade toward one of the two choice targets to report whether their experienced self-motion direction (heading) was either in the leftward or rightward category. The correct answer led to reward of a drop of juice. For dead ahead trials (0° heading), the reward was randomly delivered.

In each experimental block, there were three modality conditions: (1) vestibular-only, where self-motion stimuli were solely offered by physical motion of the motion platform; (2) visual-only, where self-motion stimuli were simulated via optic flow on the display, while the motion platform was stationary; (3) cue-combined, where congruent vestibular and visual cues were provided synchronously. Similar to previous works[29,90], all stimuli followed a Gaussian velocity with a biphasic acceleration profile to simulate transient self-motion in the environment. In the heading discrimination task, the Gaussian profile had a displacement of 0.12 m, a peak velocity of 0.285 m/s appearing at 920 ms after stimulus onset, and a peak acceleration of 0.826 m/s^2 appearing at 735 ms after stimulus onset. Note that to best see a cue-combined effect, we have downgraded the visual cue reliability through coherence[91] (10%–30%) to match the vestibular, so that psychophysical threshold under the two single cue conditions would be roughly similar[31]. At last, we generated a star field with a volume of 100 × 100 × 40 cm^3, a density of 0.01 stars/cm^3 on the display. Task difficulty was determined by heading angles (monkey D: ±8°, ±4°, ±2°, ±1°, 0°; monkey F: ±6°, ±3°, ±1.5°, ±0.75°, 0°; monkey M: ±8°, ±4°, ±2°, ±1°, 0°). Thus, each repetition contained 27 interleaved trials (3 stimulus modality × 9 heading angle), and each repetition was typically repeated more than 20 times, leading to more than 500 trials.

Other than the fixed-duration task (1.5 s), an RT version of the task was also introduced for monkey M. Specifically, the animal was allowed to make choice before the go-signal appeared, without having to maintain central fixation across the whole stimulus duration period (1.5 s). To discourage over-pursuing speed for guessing, a minimum of

700 ms from stimulus onset to the saccadic initiation was imposed on all stimulus conditions.

Color-selection task. The aim of this task is to test motor execution and bias. The animals were only required to saccade toward the red target to earn a liquid reward. The structure of the trial was identical to the heading discrimination task, but with no self-motion stimulus appearing.

Electrophysiology

Caudate mapping. Area mapping was made through cross-validation between structural MRI and physiological properties. Along each electrode penetration, baseline changes of electrophysiological signals were carefully monitored to determine transition patterns of gray/white matter. As illustrated by Supplementary Fig. 1g, the recording sites mainly cover head and body parts of CN (AC: −12 to 8 mm), similar to previous works[10,92]. Electrophysiological recordings in LIP, FEF, and MSTd were conducted in previous experiments[29–31]. We noted that screening CN neurons with the MGS task here was not as strict as the previous LIP and FEF studies[29,30], because the correlation between neuronal choice encoding and memory spatial selectivity was not well-established in CN[10]. To confirm whether this difference would affect the main finding of neural state distinction in the current study. We selected CN neurons with significant memory spatial selectivity and LIP neurons without this in the MGS task, and then performed the neural state analysis for them separately (Supplementary Fig. 12). We found that the influence of this recording difference is modest.

Online recording and offline sorting. Once encountered a single neuron in CN, we identified its response field using the MGS task. Three epochs were focused on: (1) visual response period, 75–400 ms after the appearance of the peripheral visual target; (2) memory period, 25–900 ms after the peripheral target disappeared; (3) pre-saccade period, 50–300 ms before saccade initiation. If this neuron showed significant spatial selectivity during any of the three windows, we then used it to guide the placement of the two choice targets during the following heading discrimination task. If a neuron did not show significant modulation during the MGS task, we also ran a heading discrimination task, albeit with the choice targets placed in the horizontal meridian. All online recorded data were performed offline sorting, using Spike2 for single-channel data and Kilosort-2.0 for linear-array data.

Reversible chemical inactivation

Two drugs, Muscimol·HBr and SCH23390·HCl (from Sigma-Aldrich), were first dissolved in saline to the concentration of 10 mg/mL and 0.25 mg/mL, respectively. They were accurately injected (~4 μL) into the targeted areas at a speed of 0.15 μL/min using the hand-made injectrode with an electrode embedded (Fig. 5a). Because the half-life eliminations of the two drugs are very different, their effects on the animal behavior can last differently, which is about 12 h for muscimol, and less than 1 h for SCH23390. Thus, we applied two consecutive injections with an interval of 48 h for muscimol experiments, and 24 h for SCH23390, respectively.

Electrical microstimulation

For a traditional microstimulation experiment, particularly applied in sensory cortices, it's necessary to find the neuronal cluster with similar electrophysiological properties near the electrode tip[70,87,93]. However, it has been reported that there is no such cluster in CN[10]. So we did not systematically measure clustering to guide microstimulation, similar to previous work[68]. For the sake of comparison with previous work[68,88], the current was set to be negative-leading bipolar pulses, 300 Hz, 50–80 μA. These parameters would not evoke saccades[68,94]. In a microstimulation block, 50% of trials were randomly chosen to deliver the electric current, making it unlikely for the animals to quickly adapt their strategy.

Data analysis

Psychophysics. To quantitatively measure the animals' behavioral performance in the perceptual heading discrimination task, we constructed psychometric curves by plotting the proportion of "rightward" choices as a function of heading angles. These curves were fit with cumulative Gaussian functions, whose mean (μ) and standard deviation (σ) were defined as PSE and psychometric threshold, respectively. The Bayesian optimal prediction of psychometric threshold in the multimodal condition for each session was calculated by

$$\frac{1}{\sigma_{predct.}} = \sqrt{\frac{1}{\sigma_{vest.}^2} + \frac{1}{\sigma_{vis.}^2}} \quad (1)$$

where $\sigma_{\text{vest.}}$ and $\sigma_{\text{vis.}}$ indicate psychometric thresholds measured in vestibular-only and visual-only conditions, respectively[31].

Choice and modality preference. To construct PSTH, raw spike trains were aligned to events of stimulus onset, saccade onset, and feedback onset. Firing rates were computed in every 10-ms time window and smoothed with a Gaussian kernel ($\sigma = 50$ ms). Next, all correct trials were grouped according to the animals' choice. When averaging, PSTHs were normalized for each cell across time and modalities by

$$r_{norm} = \frac{r - r_{min}}{r_{max} - r_{min}} \quad (2)$$

where $r_{\max}$ and $r_{\min}$ are respectively maximum and minimum firing rates for each neuron. To quantify the strength of choice signals, we used ROC analysis to define an index of CDiv[29]. CDiv ranged from −1 to 1, with 1 meaning always stronger firing when the PREF target is chosen, and vice versa for −1. CDiv = 0 indicates no choice preference. Grand CDiv was also computed by using firing rates across the whole stimulus duration to identify choice preference for each neuron in each modality condition.

Similarly, comparing firing rates between vestibular-only and visual-only conditions, we defined modality preference (MDiv), with 1 indicating strong visual preference, and −1 indicating strong vestibular preference.

Dimensional reduction of neural activities. To better visualize neural trajectories in a low-dimensional subspace and demix task variables, we used three dimensionality reduction methods: PCA, TDR, and dPCA (see Supplementary Information).

For PCA, we first constructed a matrix **r** of size $N_{\text{unit}} \times (N_{T \times 3\,\text{modalities} \times 2\,\text{choices}})$, whose columns contained firing rates of all single neurons during a specific time bin in each stimulus condition when one of the targets was chosen. Then, standard singular value decomposition was conducted to get left singular vectors, **U**. The first three columns of **U** (3 PCs) were chosen to span a denoising subspace, onto which the population activities, **r**, were projected. In this three-dimensional space, six neural trajectories (3 modalities and 2 choices) were presented as a function of elapsed time. To quantify the strength of choice or modality selectivity in this subspace, we calculated the Euclidean distance between corresponding pair of trajectories. For directly comparing combined-single distance (cyan, combined-vestibular; magenta, combined-visual) within and across brain regions (Fig. 2f), we quantified it by calculating its area under the curve (AUC) and then normalized it by dividing the AUC of the visuo-vestibular distance (black curve). To have statistics, for each region, we randomly selected 100 neurons from the pool and computed PCA and the

related results. This process was repeated 100 times to extract variance for statistical assessment.

For TDR, a subspace of **D** was constructed using the first 15 PCs of PCA. Linear regression was then used to capture task-related variables (heading, choice, modality) in the neural state space. Specifically, z-scored firing rates of neuron i at time t were predicted as a linear combination of multiple variables:

$$\mathbf{r}_{i,t} = \beta_{i,t,1} \cdot \mathbf{heading} + \beta_{i,t,2} \cdot \mathbf{choice} + \beta_{i,t,3} \cdot \mathbf{modality} + \beta_{i,t,4} \quad (3)$$

where $\mathbf{r}_{i,t}$ is the z-scored firing rates of neuron i at time t in all trials, **heading** represents heading directions of all trials, **choice** includes animal decisions (−1: anti-PREF choice; 1: PREF choice), and **modality** contains sensory modalities (1: vestibular-only; −1: visual-only). The regression coefficients $\beta_{i,t}$ measured how much the trial-by-trial firing rates of neuron i at time t were dependent on the corresponding variables. After that, we denoised $\boldsymbol{\beta}_t$ (includes $\beta_{i,t}$ from all neurons) by projecting it onto the denoising subspace **D**, and acquired $\boldsymbol{\beta}_t^{pca}$. For each task variable, we then determined the time t_v when $\boldsymbol{\beta}_t^{pca}$ had the largest norm, so that the denoised regression subspace $\boldsymbol{\beta}_v$ was captured by that variable. Next, we orthogonalized each $\boldsymbol{\beta}_v$ to get the task-related subspace through QR-decomposition. Last, the averaged population activities were projected onto this targeted subspace. More detailed descriptions can be found in the supplementary part of a previous work[38].

Cross-temporal Lasso decoders. To decode monkeys' choice from CN population activity, a series of linear classifiers were trained and used as a function of time over the trial:

$$p_{T1}(t) = \beta_0(t) + \sum_{i=1}^{N} \beta_i(t) r_i(t) \quad (4)$$

where $p_{T1}(t)$ represents the probability of choosing T1 target (PREF choice) at time t, r_i indicates firing rates of neuron i at time t, N is number of neurons. Assuming sparse encoding, LASSO regularization was applied to prevent over-fitting when fitting each set of parameters $\boldsymbol{\beta}$.

Trials of each neuron were first categorized into 6 groups according to choice and modality, and firing rates in each trial in a sliding 100-ms bin (step size was 20 ms) were computed. 80 trials within each group were randomly selected for training, and the rest trials (>10) were used for testing. During training, the best regularization parameter of the model was determined by a 10-fold cross-validation. During testing, the model predicted T1 (PREF) choice if the $p_{T1}(t)$>0.5 and T2 (anti-PREF) choice if $p_{T1}(t)$<0.5 at any moment. This step was repeated 1000 times to assess significance over the chance level.

Generalized drift-diffusion model (GDDM). To predict firing dynamics in multimodal condition in RT task, we used GDDM proposed previously[39]. In brief, the model assumes that both sensory and evidence-accumulation processes are modulated by cue reliability:

$$x(t) = \mathrm{r}(t) k \sin(h) + \eta(t) \quad (5)$$

$$X(t) = \int r(t) x(t) \mathrm{d}t \quad (6)$$

where $x(t)$ and $X(t)$ are respectively sensory evidence and decision variable at time t; $r(t)$ indicates moment-to-moment reliability of sensory stimuli; k denotes constant drift rate, measuring the subjective sensitivity to stimuli; $\sin(h)$ is the horizontal projection of heading; $\eta(t)$ represents noise. Based on previous findings, the models assumes that subjects rely on vestibular acceleration but visual velocity to perform heading discrimination task[39]. Cue reliability during multimodal condition was formulated as:

$$r_{comb.}^2(t) = \frac{K_{vest.}^2}{k_{comb.}^2} r_{vest.}^2(t) + \frac{k_{vis.}^2}{k_{comb.}^2} r_{vis.}^2 \quad (7)$$

After being well-trained, the optimal subjective sensitivity was predicted by:

$$k_{predct.}^2 = k_{vest.}^2 + k_{vis.}^2 \quad (8)$$

The accumulated evidence under multimodal condition could also be predicted from:

$$X_{comb.}(t) = \frac{\mathrm{k}_{vest.}}{\mathrm{k}_{comb.}} X_{vest.}(t) + \frac{k_{vis.}}{k_{comb.}} X_{vis.}(t) \quad (9)$$

The model encompassed 13 parameters $[k_{vis.}, k_{vest.}, k_{comb.}, \theta_{vis.}, \theta_{vest.}, \theta_{comb.}, b_{vis.}, b_{vest.}, b_{comb.}, t_{(nd,vis.)}, t_{(nd,vest.)} t_{(nd,comb.)}, p_{lapse}]$ that were used to fit chronometric and psychometric curves of all three modality conditions simultaneously. Among the parameters, k denotes the drift rate; θ indicates the decision threshold; b is the heading bias; t_{nd} represents the non-decision time; p_{lapse} means lapse rate.

Recurrent neural network simulations

We trained randomly initiated RNNs, comprising 256 nonlinear hidden units, a multisensory heading discrimination task analogous to that performed by monkeys. For all units, a tanh activation function is used:

$$\mathbf{x}(t) = \tanh\left(\mathbf{W}_r \mathbf{x}(t-1) + \mathbf{W}_i \mathbf{I} + \mathbf{B} + \mathbf{Z}\right) \quad (10)$$

where matrices $\mathbf{W}_r$ and $\mathbf{W}_i$ denote recurrent and input weights, respectively; **B** represents activity offset; **Z** indicates that all units are affected by independently and identically distributed (i.i.d.) white noise, with a standard deviation of 0.1.

Network input and output. The network receives visual velocity but vestibular acceleration input, simulating self-motion in 9 heading directions ([±8, ±6, ±4, ±2, 0]). The stimuli last 1.5 s with a temporal resolution of 0.05s. The sensory evidence is influenced by i.i.d. noise, with variance proportional to the input magnitude, leading to analogous signal-to-noise ratio between the two single cues, that is, visual and vestibular conditions. Indeed, after training, the network produced analogous performance (psychophysical threshold) in visual and vestibular condition (Fig. 3). Noted that the visual input was artificially delayed by 250 ms (solid versus dashed red Gaussian profile in Fig. 3b, e, h) to capture the visual-choice delay found in the three decision-related regions.

An output neuron reads out RNN's population activity using a linear weighted sum, and generates a binary choice at the end of each trial:

$$x_{\mathrm{o}}(t) = \sum_{i=1}^{256} w_i x_i(t) \quad (11)$$

$$\mathrm{Choice} = \begin{cases} 1, & x_o(t=30)>0 \\ 0, & x_o(t=30)<0 \end{cases} \quad (12)$$

where x_o represents activity of the output neuron, and x_i indicates activity of the ith neuron in the RNN.

Network training. The weights of inputs-RNN, RNN-units, and RNN output are initialized through He initialization and updated simultaneously using Adam Optimizer and the back-propagation through

time. Loss in each time step is calculated using BCELoss combined with a Sigmoid layer (BCEWithLogitsLoss function):

$$\mathrm{Loss}(t) = \sum_{n=1}^{N} \left(y_n \log \sigma(x_o(t)) + (1 - y_n) \log(1 - \sigma(x_o(t)))\right) \quad (13)$$

where y_n is the label of trial n; N is batch size; σ is sigmoid function.

Reporting summary

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

Data availability

Data supporting this study are available in the Zenodo repository at https://zenodo.org/records/13923317. Source data are provided with this paper.

Code availability

The analysis codes used in this study are available in the GitHub repository at https://github.com/ZacZeng/CN-causally-contributes-to-MSDM as well as Zenodo at https://doi.org/10.5281/zenodo.15260721.

References

1. Lee, D., Seo, H. & Jung, M. W. Neural basis of reinforcement learning and decision making. *Annu. Rev. Neurosci.* **35**, 287–308 (2012).
2. Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R. & Haynes, J. D. The distributed nature of working memory. *Trends Cogn. Sci.* **21**, 111–124 (2017).
3. Steinmetz, N. A., Zatka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**, 266–273 (2019).
4. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
5. Siegel, M., Buschman, T. J. & Miller, E. K. Cortical information flow during flexible sensorimotor decisions. *Science* **348**, 1352–1355 (2015).
6. Khilkevich, A. et al. Brain-wide dynamics linking sensation to action during decision-making. *Nature* **634**, 890–900 (2024).
7. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).
8. Kim, J. N. & Shadlen, M. N. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* **2**, 176–185 (1999).
9. Horwitz, G. D., Batista, A. P. & Newsome, W. T. Representation of an abstract perceptual decision in macaque superior colliculus. *J. Neurophysiol.* **91**, 2281–2296 (2004).
10. Ding, L. & Gold, J. I. Caudate encodes multiple computations for perceptual decisions. *J. Neurosci.* **30**, 15747–15759 (2010).
11. Li, N., Daie, K., Svoboda, K. & Druckmann, S. Robust neuronal dynamics in premotor cortex during motor planning. *Nature* **532**, 459–464 (2016).
12. Hikosaka, O., Sakamoto, M. & Usui, S. Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. *J. Neurophysiol.* **61**, 780–798 (1989).
13. Kravitz, A. V. et al. Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* **466**, 622–626 (2010).
14. Cai, X., Kim, S. & Lee, D. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* **69**, 170–182 (2011).
15. Kim, H. youngF. & Hikosaka, O. Distinct basal ganglia circuits controlling behaviors guided by flexible and stable values. *Neuron* **79**, 1001–1010 (2013).
16. Mello, G. B., Soares, S. & Paton, J. J. A scalable population code for time in the striatum. *Curr. Biol.* **25**, 1113–1122 (2015).
17. Monteiro, T. et al. Using temperature to analyze the neural basis of a time-based decision. *Nat. Neurosci.* **26**, 1407–1416 (2023).
18. Ding, L. & Gold, J. I. The basal ganglia's contributions to perceptual decision making. *Neuron* **79**, 640–649 (2013).
19. Lo, C. C. & Wang, X. J. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* **9**, 956–963 (2006).
20. Van Hoesen, G. W., Yeterian, E. H. & Lavizzo-Mourey, R. Widespread corticostriate projections from temporal cortex of the rhesus monkey. *J. Comp. Neurol.* **199**, 205–219 (1981).
21. Peters, A. J., Fabre, J. M. J., Steinmetz, N. A., Harris, K. D. & Carandini, M. Striatal activity topographically reflects cortical activity. *Nature* **591**, 420–425 (2021).
22. Chen, S. et al. Brain-wide neural activity underlying memory-guided movement. *Cell* **187**, 676–691 e616 (2024).
23. Ding, L. Distinct dynamics of ramping activity in the frontal cortex and caudate nucleus in monkeys. *J. Neurophysiol.* **114**, 1850–1861 (2015).
24. Fan, Y., Gold, J. I. & Ding L. Frontal eye field and caudate neurons make different contributions to reward-biased perceptual decisions. *eLife* **9**, e60535 (2020).
25. Hanks, T. D. et al. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**, 220–223 (2015).
26. Yartsev, M. M., Hanks, T. D., Yoon, A. M. & Brody C. D. Causal contribution and dynamical encoding in the striatum during evidence accumulation. *eLife* **7**, e34929 (2018).
27. Cui, L. et al. Causal contributions of cell-type-specific circuits in the posterior dorsal striatum to auditory decision-making. *Cell Rep.* **44**, 115084 (2024).
28. Zeng, Z., Zhang, C. & Gu, Y. Visuo-vestibular heading perception: a model system to study multi-sensory decision making. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **378**, 20220334 (2023).
29. Hou, H., Zheng, Q., Zhao, Y., Pouget, A. & Gu, Y. Neural correlates of optimal multisensory decision making under time-varying reliabilities with an invariant linear probabilistic population code. *Neuron* **104**, 1010–1021 e1010 (2019).
30. Zheng, Q., Zhou, L. & Gu, Y. Temporal synchrony effects of optic flow and vestibular inputs on multisensory heading perception. *Cell Rep.* **37**, 109999 (2021).
31. Gu, Y., Angelaki, D. E. & Deangelis, G. C. Neural correlates of multisensory cue integration in macaque MSTd. *Nat. Neurosci.* **11**, 1201–1210 (2008).
32. Butler, J. S., Smith, S. T., Campos, J. L. & Bulthoff, H. H. Bayesian integration of visual and vestibular signals for heading. *J. Vis.* **10**, 23 (2010).
33. Friedman, A. et al. A corticostriatal path targeting striosomes controls decision-making under conflict. *Cell* **161**, 1320–1333 (2015).
34. Raposo, D., Kaufman, M. T. & Churchland, A. K. A category-free neural population supports evolving demands during decision-making. *Nat. Neurosci.* **17**, 1784–1792 (2014).
35. Kobak, D. et al. Demixed principal component analysis of neural population data. *eLife* **5**, e10989 (2016).
36. Rigotti, M. et al. The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013).
37. Churchland, M. M. et al. Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
38. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
39. Drugowitsch, J, DeAngelis, G. C., Klier, E. M., Angelaki, D. E. & Pouget, A. Optimal multisensory decision-making in a reaction-time task. *eLife* **3**, e03005 (2014).

40. Chen, M. et al. Two subdivisions of macaque LIP process visual-oculomotor information differently. *Proc. Natl. Acad. Sci. USA* **113**, E6263–E6270 (2016).
41. Wardak, C., Olivier, E. & Duhamel, J. R. Saccadic target selection deficits after lateral intraparietal area inactivation in monkeys. *J. Neurosci.* **22**, 9877–9884 (2002).
42. Bjorklund, A. & Dunnett, S. B. Dopamine neuron systems in the brain: an update. *Trends Neurosci.* **30**, 194–202 (2007).
43. Kitama, T., Ohno, T., Tanaka, M., Tsubokawa, H. & Yoshida, K. Stimulation of the caudate nucleus induces contraversive saccadic eye movements as well as head turning in the cat. *Neurosci. Res.* **12**, 287–292 (1991).
44. Nichols, M. J. & Newsome, W. T. Middle temporal visual area microstimulation influences veridical judgments of motion direction. *J. Neurosci.* **22**, 9530–9540 (2002).
45. Yu, X., Hou, H., Spillmann, L. & Gu, Y. Causal evidence of motion signals in macaque middle temporal area weighted-pooled for global heading perception. *Cereb. Cortex* **28**, 612–624 (2018).
46. Histed, M. H., Bonin, V. & Reid, R. C. Direct activation of sparse, distributed populations of cortical neurons by electrical microstimulation. *Neuron* **63**, 508–522 (2009).
47. Tolias, A. S. et al. Mapping cortical activity elicited with electrical microstimulation using FMRI in the macaque. *Neuron* **48**, 901–911 (2005).
48. Jeurissen, D., Shushruth, S., El-Shamayleh, Y., Horwitz, G. D. & Shadlen, M. N. Deficits in decision-making induced by parietal cortex inactivation are compensated at two timescales. *Neuron* **110**, 1924–1931 e1925 (2022).
49. Gu, Y., Deangelis, G. C. & Angelaki, D. E. Causal links between dorsal medial superior temporal area neurons and multisensory heading perception. *J. Neurosci.* **32**, 2299–2313 (2012).
50. Fetsch, C. R., Turner, A. H., DeAngelis, G. C. & Angelaki, D. E. Dynamic reweighting of visual and vestibular cues during self-motion perception. *J. Neurosci.* **29**, 15601–15612 (2009).
51. Potegal, M., Copack, P., de Jong, J., Krauthamer, G. & Gilman, S. Vestibular input to the caudate nucleus. *Exp. Neurol.* **32**, 448–465 (1971).
52. Maunsell, J. H. R. & Vanessen, D. C. The connections of the middle temporal visual area (Mt) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* **3**, 2563–2586 (1983).
53. Reig, R. & Silberberg, G. Multisensory integration in the mouse striatum. *Neuron* **83**, 1200–1212 (2014).
54. Borra, E., Biancheri, D., Rizzo, M., Leonardi, F. & Luppino, G. Crossed corticostriatal projections in the macaque brain. *J. Neurosci.* **42**, 7060–7076 (2022).
55. Blatt, G. J., Andersen, R. A. & Stoner, G. R. Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *J. Comp. Neurol.* **299**, 421–445 (1990).
56. Schall, J. D., Morel, A., King, D. J. & Bullier, J. Topography of visual cortex connections with frontal eye field in macaque: convergence and segregation of processing streams. *J. Neurosci.* **15**, 4464–4487 (1995).
57. Noudoost, B. & Moore, T. Control of visual cortical signals by prefrontal dopamine. *Nature* **474**, 372–375 (2011).
58. Colby, C. L. & Goldberg, M. E. Space and attention in parietal cortex. *Annu. Rev. Neurosci.* **22**, 319–349 (1999).
59. Nienborg, H. & Cumming, B. G. Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* **459**, 89–92 (2009).
60. Yates, J. L., Park, I. M., Katz, L. N., Pillow, J. W. & Huk, A. C. Functional dissection of signal and noise in MT and LIP during decision-making. *Nat. Neurosci.* **20**, 1285–1292 (2017).
61. Okazawa, G., Sha, L., Purcell, B. A. & Kiani, R. Psychophysical reverse correlation reflects both sensory and decision-making processes. *Nat. Commun.* **9**, 3479 (2018).
62. Bolkan, S. S. et al. Opponent control of behavior by dorsomedial striatal pathways depends on task demands and internal state. *Nat. Neurosci.* **25**, 345–357 (2022).
63. Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. & Suarez, H. H. Recurrent excitation in neocortical circuits. *Science* **269**, 981–985 (1995).
64. Newsome, W. T., Britten, K. H. & Movshon, J. A. Neuronal correlates of a perceptual decision. *Nature* **341**, 52–54 (1989).
65. Wang, J., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018).
66. Bennur, S. & Gold, J. I. Distinct representations of a perceptual decision and the associated oculomotor plan in the monkey lateral intraparietal area. *J. Neurosci.* **31**, 913–921 (2011).
67. Charlton, J. A. & Goris, R. L. T. Abstract deliberation by visuomotor neurons in prefrontal cortex. *Nat. Neurosci.* **27**, 1167–1175 (2024).
68. Ding, L. & Gold, J. I. Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task. *Neuron* **75**, 865–874 (2012).
69. Fetsch, C. R., Pouget, A., DeAngelis, G. C. & Angelaki, D. E. Neural correlates of reliability-based cue weighting during multisensory integration. *Nat. Neurosci.* **15**, 146–154 (2011).
70. Hanks, T. D., Ditterich, J. & Shadlen, M. N. Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nat. Neurosci.* **9**, 682–689 (2006).
71. Murd, C., Moisa, M., Grueschow, M., Polania, R. & Ruff, C. C. Causal contributions of human frontal eye fields to distinct aspects of decision formation. *Sci. Rep.* **10**, 7317 (2020).
72. Katz, L. N., Yates, J. L., Pillow, J. W. & Huk, A. C. Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature* **535**, 285–288 (2016).
73. Zhou, Y. & Freedman, D. J. Posterior parietal cortex plays a causal role in perceptual and categorical decisions. *Science* **365**, 180–185 (2019).
74. 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* **4**, e05457 (2015).
75. Zhong, L. et al. Causal contributions of parietal cortex to perceptual decision-making during stimulus categorization. *Nat. Neurosci.* **22**, 963–973 (2019).
76. Wang, L., Rangarajan, K. V., Gerfen, C. R. & Krauzlis, R. J. Activation of striatal neurons causes a perceptual decision bias during visual change detection in mice. *Neuron* **98**, 669 (2018).
77. Jun, E. J. et al. Causal role for the primate superior colliculus in the computation of evidence for perceptual decisions. *Nat. Neurosci.* **24**, 1121–1131 (2021).
78. Peysakhovich, B. et al. Primate superior colliculus is causally engaged in abstract higher-order cognition. *Nat. Neurosci.* **27**, 1999–2008 (2024).
79. Stein, B. E. & Stanford, T. R. Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* **9**, 255–266 (2008).
80. Ott, T., Jacob, S. N. & Nieder, A. Dopamine receptors differentially enhance rule coding in primate prefrontal cortex neurons. *Neuron* **84**, 1317–1328 (2014).
81. Puig, M. V. & Miller, E. K. The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron* **74**, 874–886 (2012).
82. Williams, Z. M. & Eskandar, E. N. Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat. Neurosci.* **9**, 562–568 (2006).

83. Nicola, S. M., Surmeier, J. & Malenka, R. C. Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annu. Rev. Neurosci.* **23**, 185–215 (2000).
84. Vijayraghavan, S., Wang, M., Birnbaum, S. G., Williams, G. V. & Arnsten, A. F. Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat. Neurosci.* **10**, 376–384 (2007).
85. Tang, S., Cui, L., Pan, J. & Xu, N. L. Dynamic ensemble balance in direct- and indirect-pathway striatal projection neurons underlying decision-related action selection. *Cell Rep.* **43**, 114726 (2024).
86. Chen, Y. et al. Circuit-specific gene therapy reverses core symptoms in a primate Parkinson's disease model. *Cell* **186**, 5394–5410 e5318 (2023).
87. Salzman, C. D., Murasugi, C. M., Britten, K. H. & Newsome, W. T. Microstimulation in visual area MT: effects on direction discrimination performance. *J. Neurosci.* **12**, 2331–2355 (1992).
88. Doi, T., Fan, Y., Gold, J. I. & Ding, L. The caudate nucleus contributes causally to decisions that balance reward and uncertain visual information. *eLife* **9**, e56694 (2020).
89. Roitman, J. D. & Shadlen, M. N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* **22**, 9475–9489 (2002).
90. Gu, Y., Watkins, P. V., Angelaki, D. E. & DeAngelis, G. C. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J. Neurosci.* **26**, 73–85 (2006).
91. Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* **12**, 4745–4765 (1992).
92. Hikosaka, O., Sakamoto, M. & Usui, S. Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *J. Neurophysiol.* **61**, 814–832 (1989).
93. Li, W., Lu, J., Zhu, Z. & Gu, Y. Causal contribution of optic flow signal in Macaque extrastriate visual cortex for roll perception. *Nat. Commun.* **13**, 5479 (2022).
94. Watanabe, M. & Munoz, D. P. Saccade reaction times are influenced by caudate microstimulation following and prior to visual stimulus appearance. *J. Cogn. Neurosci.* **23**, 1794–1807 (2011).

Acknowledgements

We thank Wenyao Chen for monkey care and training, and Ying Liu for C++ software programming. We thank Jan Drugowitsch for kindly offering the GDDM code. This work was supported by grants from the National Science and Technology Innovation 2030 Major Program (2022ZD0205000), Shanghai Municipal Science and Technology Major Project (Grant No. 2019SHZDZX02) to Y.G.

Author contributions

Z.Z., H.H., and Y.G. conceived the project and designed the experiments. Z.Z. and Y.X. performed the experiments. C.Z. and Z.Z. constructed RNN simulations. Z.Z. analyzed the data. Z.Z. and Y.G. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-60504-y.

Correspondence and requests for materials should be addressed to Hua He or Yong Gu.

Peer review information *Nature Communications* thanks Robbe Goris, who co-reviewed with Jongmin Moon, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.