Distinct neural manifolds and critical roles of primate caudate nucleus in multimodal decision-making

Zhao Zeng[1,2], Ce Zhang[1,2], Yue Xu[1,2], Hua He[3], Yong Gu[1,2,4,*]

[1]CAS Center for Excellence in Brain Science and Intelligence Technology, Key laboratory of Brain Cognition and Brain-inspired Intelligence Technology, Institute of Neuroscience, Chinese Academy of Sciences, 200031 Shanghai, China

[2]University of Chinese Academy of Sciences, 100049 Beijing, China

[3]Department of Neurosurgery, Shanghai Eastern Hepatobiliary Surgery Hospital, Shanghai，200438, China

[4]Lead contact

*Correspondence: guyong@ion.ac.cn

Summary

Perceptual decision-making is an important cognitive function involving distributed networks spanning cortico-subcortical levels, yet most studies only include single modality stimulus and reveal largely similar signals related with animal choice across areas, making it puzzling about individual's unique roles. Using a multimodal paradigm, here we showed dorsal medial striatum, namely caudate nucleus (CN) in primate, dramatically differs from association cortex in a reduced low-dimensional subspace. Specifically, bimodal neural state evolved towards nonvisual (vestibular) in CN, rather than towards visual as in frontal/parietal cortices. The distinct CN trajectories pattern may readily explain behaviors including close-to-vestibular but not visual reaction time under bimodal condition, or vestibular-overweighting during cue-conflict context. Further causal-link experiments including applying GABA-A-receptor agonist, and D1-receptor antagonist confirmed CN's essential role of dopaminergic input. Electrical-microstimulation also verified CN's sufficient contributions. Our results indicate beyond relay-station in cortico-striatal circuitry, CN plays distinct and critical roles in complex tasks with multimodal inputs.

Keywords

Caudate nucleus, Perceptual decision making, neural manifold, non-human primate, microstimulation, reversible chemical inactivation, causal link

Introduction

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. During this process, ramping-like activity, reflecting evidence accumulation that leads to ultimate choice, has been observed widely distributed in both cortical and subcortical regions (Figure 1A),[4,5] including lateral intraparietal area (LIP),[6] frontal eye field (FEF),[7] superior colliculus (SC),[8] caudate nucleus (CN).[9] A key but unclear question is whether there is any inter-area difference in the choice representation.

Among these areas, the dorsal medial striatum (DMS), named as caudate nucleus (CN) in primates, is well known for mediating a variety of brain functions, including motor execution,[10,11] value evaluation,[12,13] time perception[14,15] and more. What is the role of CN in the distributed sensory-motor association or integration network? A general hypothesis is that choice signals in CN largely stem from cortical regions such as LIP and FEF.[16,17] Recent works even suggest that caudate mainly reflects cortical activity,[18,19] serving as a relay station for the cortico-striatal circuitry. This hypothesis makes sense in term of that CN receives heavy projections from the neocortex,[20], yet it remains unknown what unique role CN may play in brain's sensory-motor integration network.

In the current study, we addressed this issue by using a multisensory decision-making paradigm, in which an additional dimension of sensory modality was provided. Specifically, macaques were trained to discriminate heading directions based on reliability-matching visual (optic flow) and vestibular (inertial motion) cue (Figure 1B and 1C). After training, the animals showed roughly equal performance in single cue conditions, as reflected by similar psychophysical threshold (Figure 1D and 1E). The threshold was significantly reduced in bimodal condition when congruent visuo-vestibular cues were provided, and the performance improvement was close to Bayesian prediction, a pattern typically seen in previous studies.[21-23] During the task, neural activity of single-units from CN, FEF, and LIP were recorded. In population, we expect manifold subspace reduced from high-dimension neural state space (Figure 1F), would exhibit two properties about choice and modality. First, choice-related signals should evolve, reflecting momentary sensory evidence accumulation e.g. visual velocity and vestibular acceleration,[22,23] a process which ultimately leads to different decisions (Figure 1G). Second, modality-related signals should evolve and diverge for each single cue conditions (e.g. visual and vestibular). As to bi-modal stimulus condition, neural trajectory is speculated to be in the middle of the two single cues who share similar reliabilities (Figure 1F), which can be quantified by a Euclid distance index among modalities (Figure 1H, middle panel, model #2). Alternatively, the bimodal neural trajectory may be biased towards either cue (Figure 1H, first and third panel, model #1 and #3, respectively). We then examine whether cortico-subcortical areas would present similar or distinctive patterns along these two main dimensions (choice and modality).

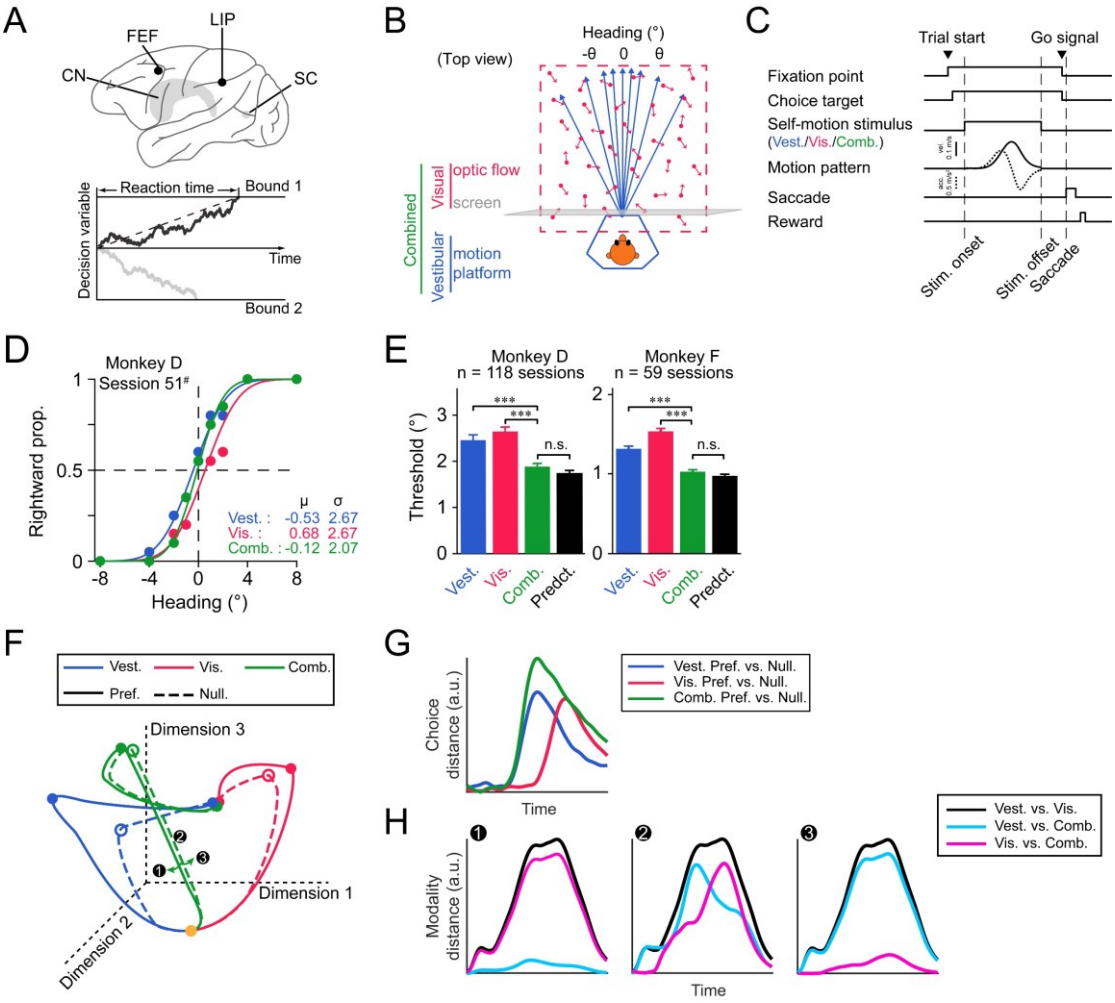

Figure 1. Comparing distributed perceptual decision-making signals using multisensory paradigm

(A) Distributed brain regions (upper) representing evidence-accumulation signals (lower) in perceptual decision-making task.

(B) Schematic of a visuo-vestibular heading discrimination task using a virtual reality system.

(C) Timeline of a trial. Motion stimulus profile includes a varied Gaussian velocity (solid curve) and biphasic acceleration profile (dashed curve).

(D) Psychometric curves of an example session. The proportion of rightward choice is plotted as function of heading for each modality condition. Dots are real data, and curves are fitted cumulative Gaussian functions.

(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.

(F) Conceptual illustration of possible population activity projected onto a three-dimensional subspace during multisensory decision-making task. Color: modality (blue, vestibular-only; red, visual-only; green, cue-combined); Solid and dashed lines: animal choice in the preferred target (IN) and anti-preferred target (OUT). Orange dot: start point of a trial.

(G) Hypothetical choice-related signals quantified by Euclidean distance between IN and OUT trajectories in the subspace of (F).

(H) Hypothetical modality-related signals quantified by Euclidean distance between neural trajectories across modalities in the subspace of (F). Three possible models were proposed for the bimodal neural trajectory: biased towards vestibular-only condition (left panel, model #1), balanced in the middle of the two single cues (middle panel, model #2), or biased towards visual-only condition (right panel, model #3). Color represents comparison between different modalities: (black, vestibular-only versus visual-only; cyan, bimodal versus vestibular-only; magenta, bimodal versus visual-only).

Results

CN encodes multiple variables in 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 (Figure 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 (300-500 ms), the monkeys were required to report their perceived direction, either left or right, by making a saccade to the corresponding target to earn reward (Figure 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.[21] All stimuli conditions, including three sensory modalities and nine heading angles, were interleaved in one experimental session. After training, both monkeys performed well, illustrated by nice psychometric curves (Figure 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 in both animals, psychophysical threshold was comparable in the visual-only and vestibular-only condition, and became significantly smaller in the cue-combined condition (Figure 1E). Importantly, performance under bimodal condition was close to that predicted from Bayesian optimal integration theory (Figure 1E, green vs. black bar), which has been frequently observed in many previous studies.[21-24]

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 (Figure S1A-S1C).[25] Peristimulus time histograms (PSTHs) of a few example neurons were shown to demonstrate that firing patterns of single CN neurons contained large heterogeneity (Figure S1D). To quantify the strength of task variables, we computed

134 choice divergence (CDiv, Figure S1E) and modality divergence (MDiv) for each
135 neuron, using receiver operator characteristic (ROC) analysis. For CDiv, many CN
136 neurons significantly possessed choice signals in at least one modality condition
137 (monkey D, 60/117 = 51.3%; monkey F, 33/52 = 63.5%, Figure S3E). The mean
138 CDiv was significantly larger than 0 (vestibular-only: 0.123, p = 9.75e-06; visual-
139 only: 0.099, p = 2.83e-04; combined: 0.139, p = 5.28e-06; two-tailed t-test; Figure
140 2A), indicating that CN overall prefers contralateral choice. For MDiv, many CN
141 neurons also carried significant sensory modality signals (monkey D: 88/117 =
142 75.2%; monkey F: 36/52 = 69.2%), with roughly equal visual and vestibular
143 preference (mean MDiv: -0.012, p = 0.717; two-tailed t-test; Figure 2A). These cells
144 with choice or modality preference were distributed along the longitudinal axes in a
145 mixed manner (Figure S1F and S1G).

146 Previous studies have shown that sensory-motor association areas such as the
147 posterior parietal cortex encodes multiple task variables using a manner of category-
148 free coding in multisensory decision tasks.[22,26] This was the similar case in CN.
149 Firstly, numerous CN cells simultaneously had choice and modality preferences
150 (vestibular-only: 47/169 = 27.8%; visual-only: 43/169 = 25.4%; combined: 54/169 =
151 31.9%). Secondly, there were neither well-clustered distribution nor significant
152 correlation between choice and modality signals (Figure 2A). However, either task
153 variable could be successfully extracted from CN population activity by using linear
154 demixing algorithms such as demixed principal component analysis (dPCA, Figure
155 S1H).[27] We found the interaction term of choice and modality also captured response
156 variances to some extent (Figure S1H, the third column). This nonlinear mixed
157 selectivity has been previously shown to be more computationally flexible for dealing
158 with complex cognitive tasks.[28]

159 At the population level, we sorted trials in each modality condition according to
160 animal choices and averaged firing rates of all neurons to construct population PSTH,
161 finding significant choice modulation under all three contexts (Figure 2B). Similar to
162 LIP/FEF,[22,23] choice-related signals in visual-only context were delayed compared to
163 the other two stimulus conditions, implying that CN also integrates different dynamic
164 signals, that is, vestibular acceleration but visual velocity (Figure S1I and S2A).
165 Majority of CN neurons are tuned to subjects' internal choices but not external
166 physical heading stimuli, which was supported by similar firing pattern in correct and
167 error trials (Figure S3A and S3B) and markedly larger partial correlation between
168 choice and CN neuronal responses (Figure S3C). More directly, we tested heading
169 tuning of CN single neurons in the passive task where decision reports were not
170 required, and found that very limited proportion of CN neurons were significantly
171 selective to sensory headings (Figure S3D). Besides, choice signals for each neuron
172 were highly correlated across modalities (Figure S3F).

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

CN shows distinct neural manifold from 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 high-dimensional neural state space[29] for population activity in each area including CN, FEF and LIP, and then performed dimension reduction to capture principal traits. Specifically, population activity in each region was denoised by using principal component analysis (PCA), and the first three PCs (captured more than 65% variances) were adopted to form a 3-D neural subspace (Figure 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, defining choice distance (Figure 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 bimodal 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 (Figure 2C), as well as the quantified Euclid distance between each pair of modality comparison (Figure 2E, F). All these results were further confirmed by a targeted dimensionality reduction (TDR) algorithm,[30] by which the high-dimensional neural population activities were directly projected onto a meaningful subspace spanned by choice and modality axes (Figure 2G). These results also held across animals (Figure S4).

Thus, along modality dimension, CN manifold exhibits a distinct pattern from the frontal and posterior-parietal cortices. In fact, none of the association areas shows the expected symmetric pattern (model #2 in Figure 1H), which is surprising because the two single cue reliabilities have been carefully matched in our study, as evident from two observations. First, in behavior, all animals showed analogous psychophysical threshold in each single cue condition (Figure 1D and 1E). Second, in neurophysiology, one earlier stage polysensory area—the dorsal medial superior temporal sulcus (MSTd) in fact exhibited the expected model #2 pattern: bimodal neural trajectory was in the middle of the two single cues (Figure S5). Thus, CN may have received vestibular signals elsewhere in addition to FEF/LIP/MSTd, leading to an internally higher vestibular "gain" that attracts the neural trajectory towards vestibular during bimodal stimulus condition.

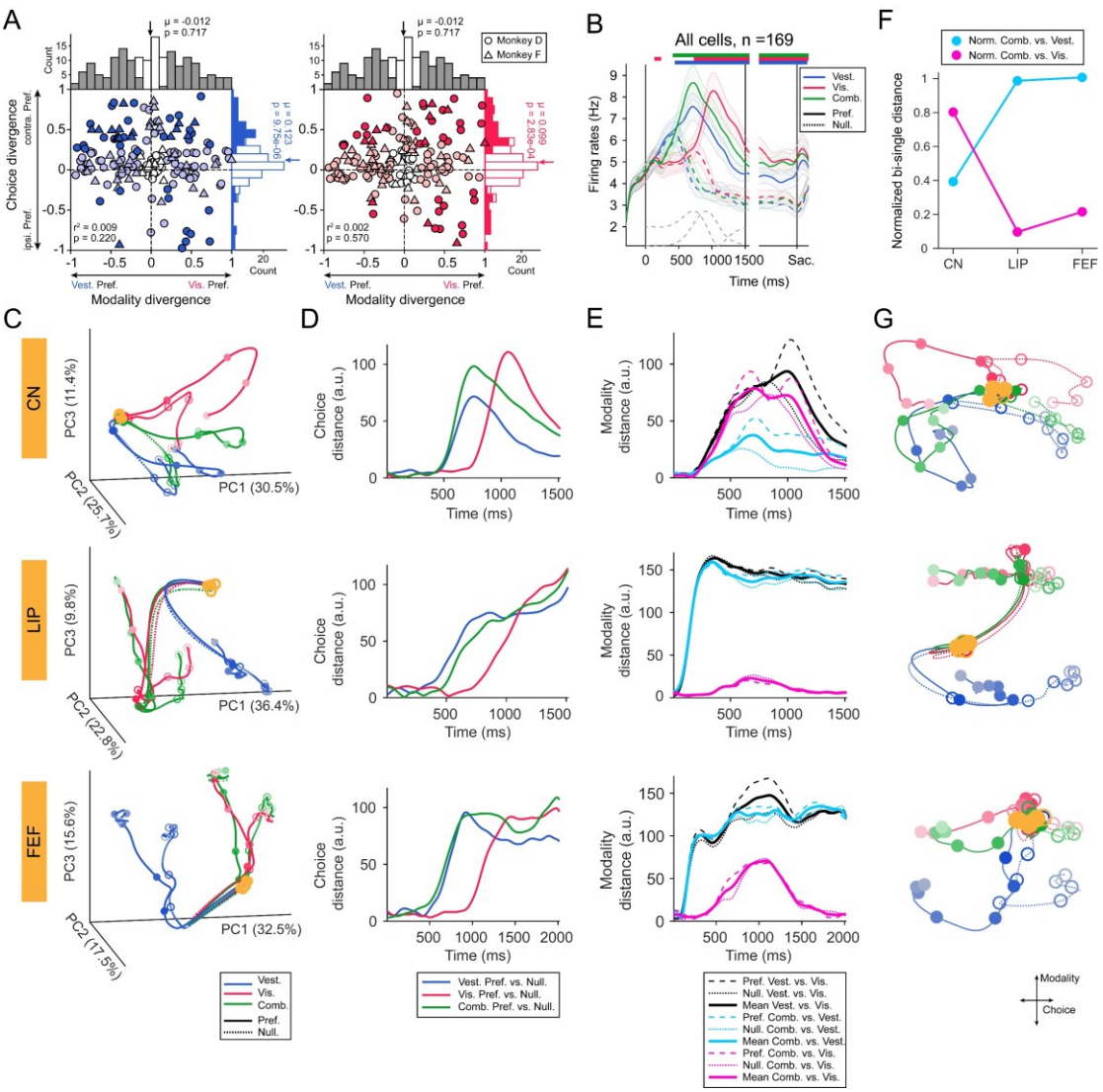

Figure 2. CN manifold and comparison with frontal and parietal association cortices

(A) Choice and modality signals in each CN neuron in vestibular-only (left) and visual-only (right) conditions. Dark dot: both choice and modality signals are significant; Light dot: either variable is significant; Empty dot: neither variable is significant ($p<0.01$; permutation test). Shape of dot: different monkeys.

(B) Averaged population PSTHs of CN neurons. Firing rates are aligned to stimulus onset (1st vertical line) and offset (2nd vertical line), or saccade onset (3rd vertical line). Shaded error bars are SEM. Horizontal colored bars indicate significant choice divergence ($p < 0.01$, two-tailed t-test).

(C) High-dimensional neural state space in CN (upper), LIP (middle), and FEF (lower) are projected onto 3-dimensional subspace constructed by the first three principal components. All labels are same as Figure 1F.

(D) Choice distance evolved within trials in CN (upper), LIP (middle), and FEF (lower). All labels are same as Figure 1G.

(E) Modality distance evolved within trials in CN (upper), LIP (middle), and FEF (lower). Thin dashed and dot lines respectively represent modality distance of IN and OUT trials, while thick solid lines are mean value of both IN and OUT trials.

(F) Normalized modality Euclid distance computed by averaged modality distance across time for the bimodal versus unimodal conditions (cyan or magenta curves in E), normalized by modality distance of two uni-modal conditions (black curve in E).
(G) High-dimensional neural activities in CN (left), LIP (middle), and FEF (right) are projected onto subspace spanned by the two orthogonal axes of choice and modality using TDR analysis.

Recurrent neural network reproduced 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 who 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 (Figure 3A), and were connected to a single linear read-out decoder, who 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 bimodal sensory inputs. During training, the network initially showed relatively high thresholds, yet they decreased soon after tens of trials before reaching a plateau (Figure 3B, 3E and 3H, middle panels). Interestingly, the network automatically showed improved performance under bimodal stimulus condition, close to the Bayesian prediction as seen in the animals' behavior (Figure 3B, 3E and 3H, right panels).

We then intuitively varied the magnitude ratio of the two unisensory inputs fed into the network (Figure 3B, 3E and 3H, left panels). Importantly, noise level was always controlled so that signal-to-noise ratio of either input remained unchanged, to guarantee output of analogous visual and vestibular performance for each single cue (Figure 3B, 3E and, right panels). However, we hoped that the different magnitude factor, might represent some different impacts (e.g. over- or under-weighting) on particular cues during bimodal stimulus condition. This was indeed the case for our RNN simulation. In particular, RNN's bimodal state trajectory evolved towards vestibular when vestibular has an equal (Figure 3D) or relatively larger magnitude (Figure 3G) compared to visual, generating a CN-like pattern. On contrary, only when visual has a much larger magnitude (Figure 3J), would RNN produce a frontal/parietal-like visual bias during bimodal stimulus condition.

In contrast to modality dimension, RNN produced similar choice patterns that were independent on input magnitude-ratio (Figure 3C, 3F and 3I). Therefore, RNN results could largely reproduce neurophysiological findings across cortico-subcortical areas. The heterogeneity in modality dimension across areas though, may suggest that CN receives different vestibular inputs other than simply inheriting everything from frontal and posterior-parietal association cortices.

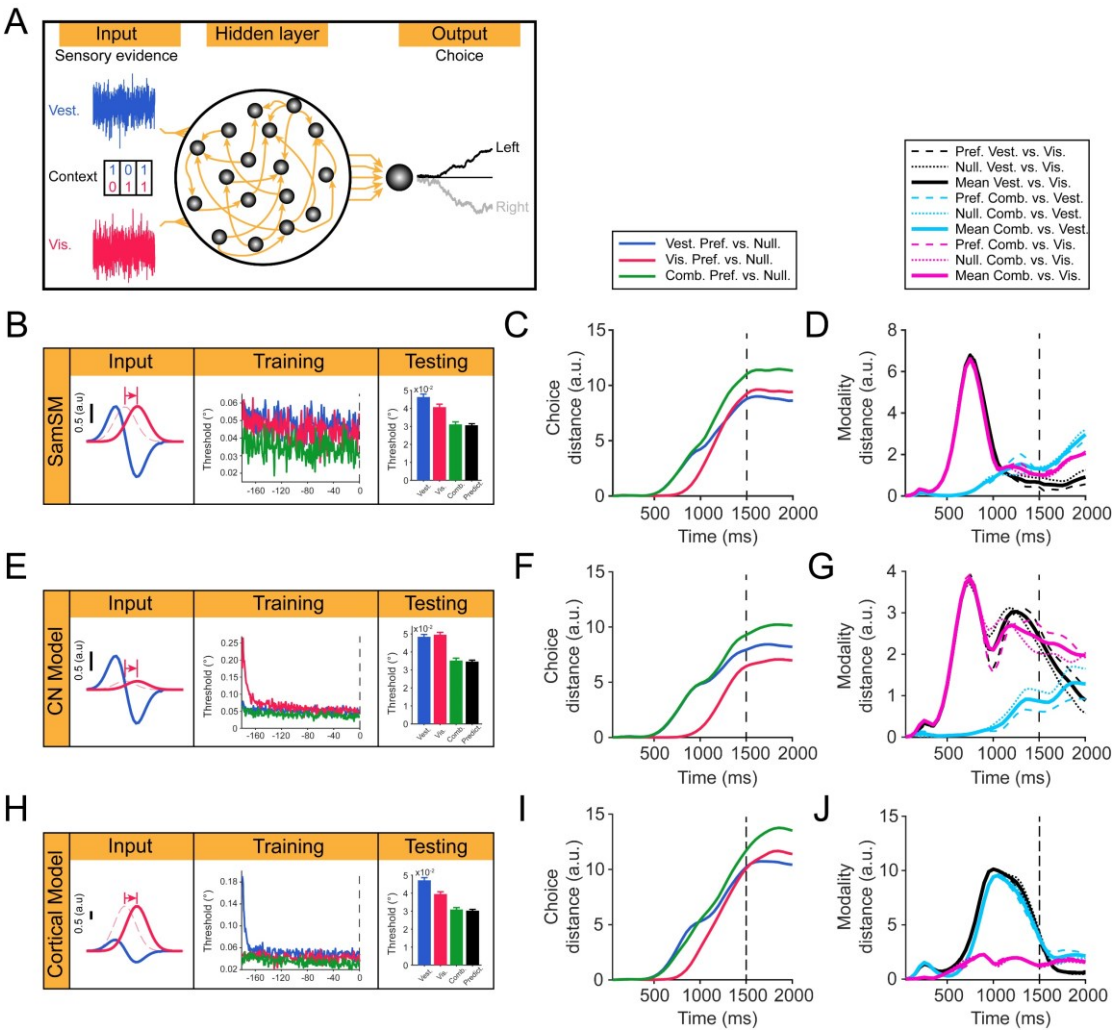

Figure 3. RNN simulations reproduced inter-area heterogeneity

(A) Structure of a recurrently connected artificial network constituting 256 nonlinear units receiving noisy vestibular and visual heading inputs.

(B-D) 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. Vertical dashed line in (C) and (D) represents stimulus offset.

(E-G) Higher vestibular 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.

(H-J) Higher visual input magnitude with comparable signal-to-noise ratio (H, left panel). RNN output behavioral performance (H, mid and right panels), choice (I) and modality (J) patterns.

CN dynamics reflect perceptual performance in multisensory context

We then assessed whether the distinct CN in perceptual decision-making network would play critical roles in our multimodal heading discrimination task. We first addressed how CN neuronal dynamics were functionally coupled with the animals' perceptual performance on a single-trial basis 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.

The first analysis was to train a series of Lasso decoders over time to predict animal's choice on a single-trial basis from CN's population activity (see Method). We found monkeys' upcoming choice could be reliably decoded about 500 ms after stimulus onset in the vestibular-only and bimodal conditions, and about 850 ms in visual-only condition (Figure 4A, principal diagonal). Importantly, classifiers trained using activity during early period could be generalized to later stage both within, and across stimulus conditions (Figure 4A, off the principal diagonal), implying that choice signals in CN share similar strategy across time and modalities (e.g. abstract decision).

The second analysis was to more directly assess whether CN dynamics could show optimality as seen in the behavior. Previous research has proposed a remarkable framework of generalized drift-diffusion model (GDDM, Figure 4B, see Method), 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 assess optimality.[31] To do so, we trained a third monkey (M) to perform a RT version of the same heading discrimination task, in which the only difference was that the animal was allowed to make saccadic choice at any time right after the stimulus onset, without having to wait for the disappearance of the fixation point indicating the go-signal. In such a situation, we found the animal showed different response time across modality conditions. In particular, response time was significantly longer in the visual condition (~1150 ms on average), compared to that in vestibular (~820 ms on average) and combined (~850 ms on average) condition (Figure 4C, dot symbols). These patterns were similar to the human study[31]. GDDM well fitted behaviors including both of the chronometric (Figure 4C, solid curves) and psychometric curves (Figure 4E, solid curves). Importantly, the fitted drift rate, k-values (see Method), representing subjective stimulus sensitivity, were close between bimodal condition and prediction from optimal integration theory (Table 1), suggesting that monkey M integrated cues near optimally under reaction-time context. After acquiring k-values, the predicted moment-to-moment bimodal response in CN was consequently calculated, which turned out to be very similar with the real neurophysiological data (Figure 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.

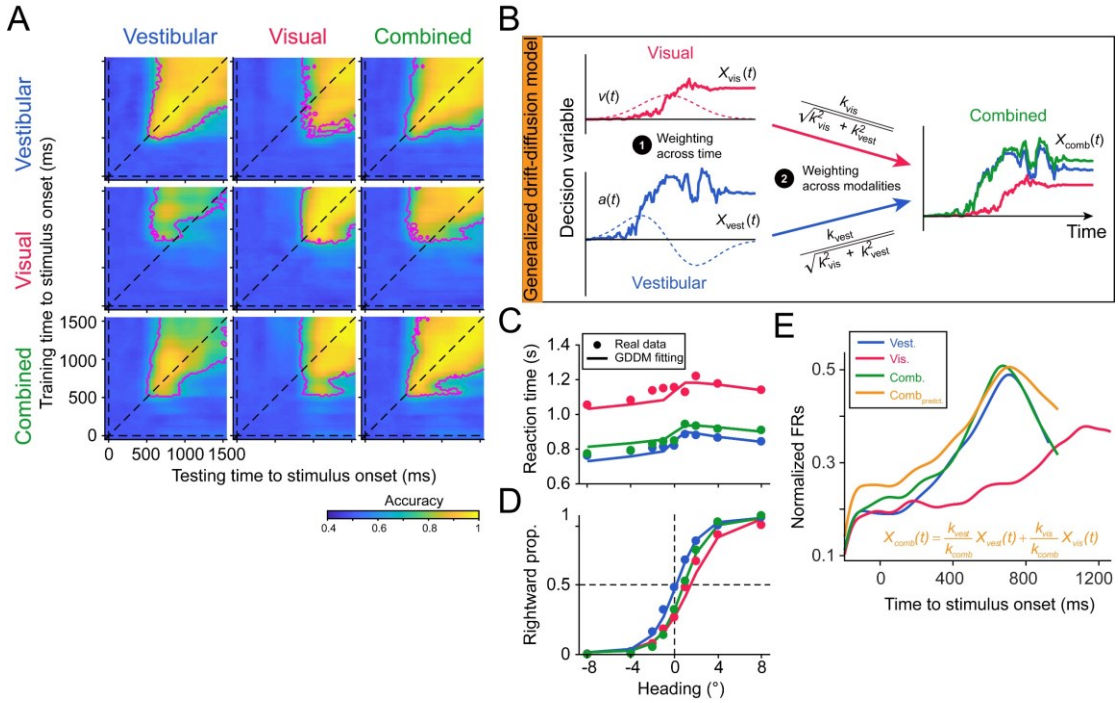

Figure 4. CN activity reflects animal multisensory decision behavior

(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 center of the time window used to train decoders, while the abscissa marks the testing time. Magenta lines outline the areas with accuracy significantly higher than chance level (50%, $p < 0.01$, permutation test).

(B) Schematic of the GDDM. Plot modified with permission from Drugowitsch et al.[31]

(C) Chronometric curves. Dots are measured data from monkey M. Lines are fitting curves from GDDM.

(D) Psychometric curves. Labels are same as (C).

(E) Population firing rates of CN (normalized). Orange line represents predicted bimodal response from GDDM. Blue, red, and green lines are measured CN responses in vestibular-only, visual-only, and bimodal condition, respectively.

Table 1. Fitting results 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		

337 **Unilateral chemical inactivation in CN biased perceptual choice**
338 We then examined causal-link roles of CN activity and the animal's perceptual
339 performance. Specifically, CN's activities were manipulated through a number of
340 methods, including chemical inactivation to test essential contributions, and electrical
341 microstimulation to test sufficient contributions.

342 We first applied muscimol, agonist of $GABA_A$ receptor to non-selectively suppress
343 CN's overall excitatory activity (Figure 5A). We found that unilateral inactivation
344 (Figure 5B and 5C), but not saline control (Figure 5F) induced significant ipsilateral
345 bias in both animals' choice. This ipsilateral bias was concordant with our recording
346 data showing that majority of CN neurons preferred contralateral choice (Figure 2A).
347 In contrast, unilateral muscimol inactivation in LIP and FEF surprisingly produced
348 negligible effect (Figure 5D and 5E), although neurons in these two areas also mainly
349 preferred contralateral choice (Figure S2B). Other than PSE, muscimol application
350 did not significantly affected perceptual sensitivity (Figure S6A), or history
351 information usage strategy (Figure S6B and S6C).

352 Because CN is strongly innervated by midbrain dopaminergic inputs,[32] we wonder
353 how selectively modulating dopamine receptors of CN neurons may affect the
354 animals' task performance. Thus, we then applied antagonist of D1 receptor,
355 SCH23390, unilaterally into CN. We found that this manipulation induced
356 significantly more contralateral choice (Figure 5G), opposite to muscimol-mediated
357 bias direction, implying that dopaminergic inputs into CN could modulate evidence
358 accumulation process. Finally, similar to muscimol, blocking D1 receptors neither
359 impacted heading sensitivity (Figure S6D) nor trial-history usage strategy
360 significantly (Figure S6E and S6F).

361 These chemical-inactivation experiments suggest that CN plays a critical role in
362 sensory-motor decision task, yet one possibility is that the observed causal effects
363 may be mainly due to motor-report instead of cognitive perception, particularly when
364 considering that CN is well-known for modulating movement executing.[10,33] We think
365 this is less likely to be the case in our study based on a few observations. First,
366 chemical inactivation mainly affected difficult trials when heading directions were
367 around the reference, which required more cognitive load compared to easy trials with
368 large heading directions (Figure 5B). Second, we further trained the animals to
369 perform a simple color-selection task that required much reduced cognitive load than
370 the fine heading discrimination task (Figure 5H). In this controlled task, neither
371 muscimol nor SCH23390 significantly influenced saccadic behavior any more (Figure
372 5I and 5J), supporting that perturbation of CN activity mainly affected the animals'
373 cognitive process.

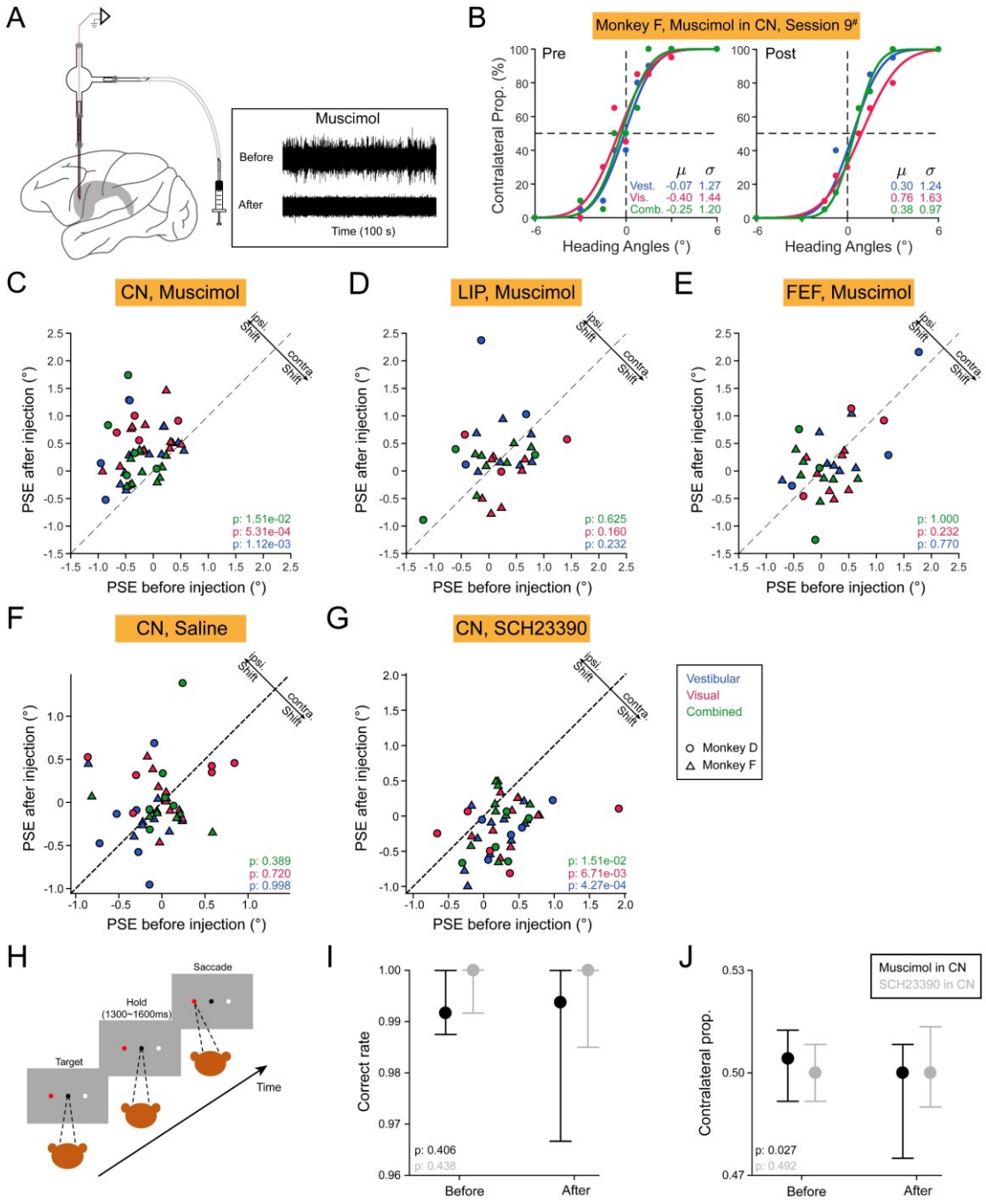

Figure 5. Unilateral inactivation of CN biases animal choice

(A) Chemical injection into targeted areas including CN, LIP and FEF. Inset shows an example case of diminished neural activity after muscimol delivery.

(B) Example of one inactivation experiment, showing performance before (left plot) and after (right plot) muscimol unilateral injection in CN.

(C-G) Comparison of PSE before and after unilateral muscimol injections in CN (C), LIP (D), FEF (E), saline-injection control (F), and SCH23390 injections in CN (G). Shapes represent data from different monkeys. Colors are different modality conditions.

(H-J) Color guided saccade task (H). Comparison of accuracy (I) and PSE (J) before and after muscimol (black symbol) or SCH23390 (gray symbol) injection. Dot and error bar indicate

median and 95% confidence interval. P-values are from Wilcoxon signed-rank test; significance level, 0.05.

Electrical microstimulation in CN indicated sufficient role

The third method for causal-link test was to use electrical microstimulation (50~80 μA, 300 Hz, cathode-leading biphasic, Figure 6A). Microstimulation was typically thought to identify sufficient contributions by artificially activating neural activity,[34,35] although it still remains a debate about its exact impact on evoked neural dynamics.[36,37] Specifically, in each experimental session, microstimulation was applied unilaterally in CN during the stimulus duration period, randomly for half of the total trials. Significance of PSE and threshold change in the psychometric functions were assessed by bootstrap test (Figure 6B and 6C). 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%, Figure 6D). This proportion could be underestimated because of adaption of the currents[38] (Figure S7F). The effect was fairly consistent across sensory modality (Figure S7B). In contrast to PSE, microstimulation only modestly affected psychophysical threshold (Figure S7A), indicating that microstimulation in CN mainly introduced signals instead of noise,[39] supporting that CN sufficiently contributes to the animals' perceptual choice.

We noticed, however, some inconsistence in PSE shift direction between the two animals. Specifically, microstimulation tended to induce more contralateral choice in monkey D (Figure 6D, top row), but more ipsilateral choice in monkey F (Figure 6D, bottom row). This individual difference could not be explained by factors such as anatomical location (Figure S7C and S7D), or coding properties of the stimulated sites (Figure S7E). To test other possibilities, we applied microstimulation during the delay period after stimulus offset and before saccadic choice (Figure 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, Figure 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).

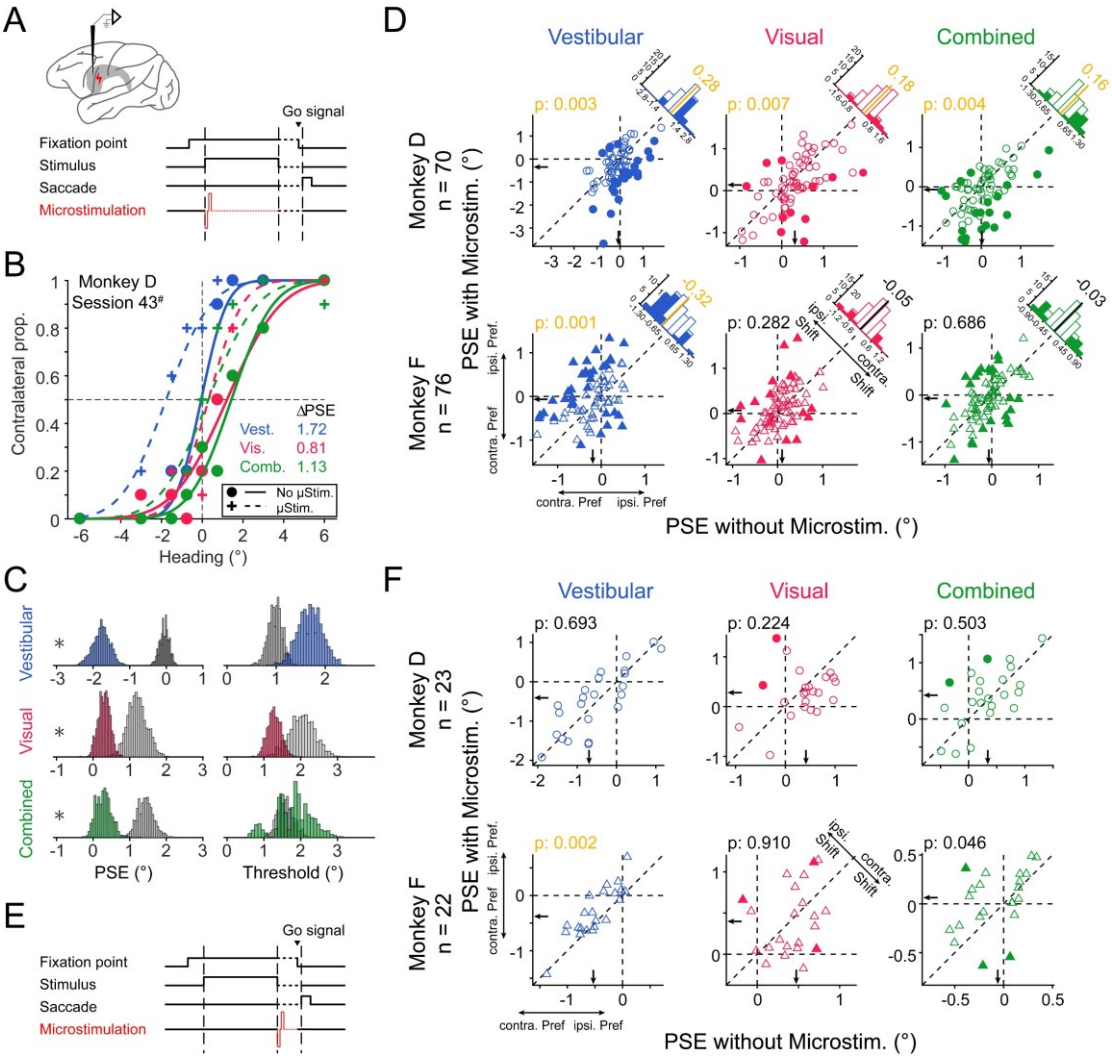

Figure 6. Unilateral microstimulation in CN affects animal choice

(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).

(C) Distributions of PSE (left column) and threshold (right column) constructed from bootstrap for microstimulation (color) and without microstimulation (gray) trials. Asterisk indicates significant change of the PSE or threshold ($p < 0.01$, bootstrap test).

(D) Comparison of PSE with (ordinate) and without (abscissa) microstimulation. Black arrows mark the median PSE. Marginal histogram counts frequency of ΔPSE. Solid bars indicate significant shift and open bars are insignificant change.

(E-F) Microstimulation applied during the delay period when stimulus is absent (E). Comparison of PSE with (ordinate) and without (abscissa) microstimulation (F).

Discussion

It is common to see many brain functions are mediated by a distributed network involving multiple regions that share seemingly similar properties, leading to the

434 "redundant" hypothesis.[2-4] One obvious benefit of this redundancy is that it helps deal
435 with dysfunction when some of the nodes within the network are hampered.
436 Alternatively, the seemingly redundant nodes may actually carry distinctive traits that
437 can be unraveled under more complex contexts. Here we provide clear evidence,
438 showing that for perceptual decision-making, by adding an additional dimension of
439 modality input, the subcortical area CN could become dramatically different from the
440 frontal and parietal association cortices in a low-dimension manifold subspace
441 embedded in higher-dimensional neural state space. In particular, neural trajectory
442 under bimodal condition is largely biased towards vestibular in CN, in contrast to a
443 clear visual-bias in LIP and FEF. Aided by RNN simulations, our results suggest that
444 on top of the role as a relay station for the cortico-striatal circuitry as in conventional
445 view, CN could receive sensory evidence from elsewhere, for example, thalamus,[40] or
446 other sensory cortices,[41-43] leading to subcortical-cortical heterogeneity.

447 **Non-visual dominant CN may reflect animal decision strategy**

448 Our perturbation experiments point out that the primate CN plays critical roles,
449 both essential and sufficient, in the perceptual decision-making task, consistent with
450 results from studies on rodents.[44] Considering that CN shows non-visual (vestibular)
451 dominant neural state during bimodal condition when both single cues overall share
452 similar cue-reliabilities, neither similar to the frontal (FEF) and parietal (LIP) sensory-
453 motor cortex, nor to the polysensory area (MSTd), a question is what does this imply
454 for CN in multisensory perception and decision-making? We think this finding is
455 consistent with a few behaviors observed in previous studies.

456 First, we found CN population dynamics are consistent with prediction using
457 parameters regressed from the generalized drift-diffusion model (GDDM). It's
458 noticeable that the fitted drift rate, reflecting speed of evidence accumulation, is
459 higher for vestibular-only than visual-only condition, indicating an intrinsically higher
460 sensitivity to vestibular stimulus.[31] Second, the vestibular bias may explain
461 "vestibular-overweighting", as reported previously in both human and non-human
462 primate studies by using spatially conflict visuo-vestibular heading stimuli.
463 Specifically, subjects' PSE during bimodal condition with conflict unimodal cues is
464 biased towards vestibular, even when the two unimodal cues carry matching
465 reliability and analogous perceptual sensitivity (psychophysical threshold).[24,45] Future
466 experiments could be conducted by using cue-conflict task and examine CN's role in
467 the animals' behavior. Third, the vestibular bias is consistent with the difference in
468 response time under different modality conditions. Specifically, researchers have
469 shown under a reaction-time version of task, human subjects tend to make choice
470 faster in vestibular-only and bimodal stimulus conditions compared to that in the
471 visual-only condition.[31] This is also exactly what we see in our current study by
472 training one of the animals to perform the reaction-time task.

In addition to the vestibular bias in the neural state space, another feature in CN is that the Euclid distance between different conditions ultimately decreases at the late stage of trials, unlike the sustained plateau in frontal and parietal cortices. This descending Euclid 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[46-48]. 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 reaction-time context, the animal made saccadic choice much earlier (vestibular: ~820 ms, visual: ~1150 ms, bimodal: ~850 ms) than the whole 1.5 second as required in the fixed-duration context. Thus, the dynamic choice signals in CN may be consistent with previous view that compared to cortex, CN is particularly important for decision formation at early stage.[16] In addition to 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 sustain consequently.

Dopaminergic modulation in CN is essential for perceptual decision

It has been recognized that dopaminergic regulation in 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.[49-52] Here, we showed that suppressing CN D1 receptors using antagonist consistently biased animals' choice toward contralateral side, opposite to muscimol inactivation. Out result is thus concordant with the finding that low dose of D1 receptor antagonist typically elevates neuronal activity.[53,54]

It is well known that the basal ganglia circuit contains two pathways of direct and indirect that is associated with D1 and D2 receptor, respectively. The two pathways are likely responsible for complimentary and opposite control of decision behavior.[55] It would be interesting to further test D1 and D2 receptor pathway in future experiments, as has been performed in rodents using state-of-art techniques.[11,55] For example, a recent work developed a recombinase-free retrograde adeno-associated virus (AAV) dependent strategy to precisely modulate CN direct MSNs.[56] It would be promising to identify primate CN circuit-specific contributions to cognitive functions in future studies.

Implications of microstimulation experiment

Electrical microstimulation has been widely used in primates to identify causal roles of neural activity in perceptual choice.[57-59] The technique is quite successful when applied in sensory areas, as evident by induced PSE shift in the direction consistent with the labelled-line of the stimulated neurons (encoded feature).[39] 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).[39] 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.[58] Indeed, in the current study we found 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 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.[58,60] Other factors may also confound microstimulation result. For example, the electrical currents are easily spread, and subsequently activate passing fibers, leading to nonlocal effects.[36,37] Furthermore, electrical currents lack specific targeting, and may impact direct and indirect pathway differently, consequently generating heterogeneous result. Optogenetic or chemogenetic tools, that are typically much more relieved from these issues, may be applied in future experiments to further test sufficient roles of certain CN populations in perceptual decision tasks.

Acknowledgements

We thank Wenyao Chen for monkey care and training, and Ying Liu for C++ software programming. We thank Jan Drugowitsch for kindly offering GDDM code. This work was supported by grants from the National Science and Technology Innovation 2030 Major Program (2022ZD0205000) 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.

Declaration of interests

The authors declare no competing interests.

Supplemental information

Document S1. Figure S1-S7

STAR Methods

Animals and surgeries

Three adult monkeys (*Macaca mulatta*), monkey D, F and M, weighting 7~10 kg, were trained in this work. Details about animal surgery and training procedures have been described previously.[21] In brief, monkeys were chronically implanted with a 6-cm lightweight plastic ring, which was served as 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 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 a 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,[61] 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 weak correlation between strength of choice signal and spatial selectivity.[9] 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-second, linear forward self-motion stimulus, defined by either visual, vestibular, or bimodal 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. Correct answer led to reward of a drop of juice. For dead ahead trials (0° heading), 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,[22,62] all stimuli followed a Gaussian velocity with a biphasic acceleration profile to simulate transient self-motion in the environment. Note that to best see a cue combined effect, we have downgraded the visual cue reliability through coherence[63] (10%-30%) to match the vestibular, so that psychophysical threshold under the two single cue conditions would be roughly similar.[21] 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 for more than 20 times, leading to more than 500 trials.

Other than the fixed-duration task (1.5 second), a reaction-time 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 second).

Color selection task. The aim of this task is to test motor execution and bias. The animals were only required to saccade to 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 structure 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 Figure S1G, the recording sites mainly cover head and body parts of CN (AC: -12~8 mm), similar to previous works.[9,64]

617 ***On-line recording and Off-line sorting.*** Once encountering a single neuron in CN,
618 we identified its response field using MGS task. Three epochs were focused: (1)
619 visual response period, 75~400 ms after appearance of the peripheral visual target; (2)
620 memory period, 25~900 ms after the peripheral target disappeared; (3) pre-saccade
621 period, 300~50 ms before saccade initiation. If this neuron showed significant spatial
622 selectivity during any of the three windows, we then used it to guide placing the two
623 choice targets during the following heading discrimination task (IN and OUT). If a
624 neuron did not show significant modulation during the MGS task, we also ran heading
625 discrimination task, albeit with the choice targets placed in the horizontal meridian.
626 All online recorded data were performed off-line sorting, using Spike2 for single-
627 channel data and Kilosort-2.0 for linear-array data.

628 ***Selection of putative medium spiny neuron.*** Three subtypes, including the median
629 spiny neuron (MSN), tonically active neuron (TAN) and fast-spiking neuron (FSN),
630 account for more than 95% of CN neurons.[65] Among them, MSNs are the projection
631 neurons that we aim for, yet they are not easily targeted especially for primate
632 electrophysiology. Thus, we tried to identify MSNs from TANs and FSNs according
633 to a number of indicators.[25] (1) Waveform. Compared to MSNs, FSNs tend to have
634 narrower spike waveforms, measured by the shorter peak-to-trough time. (2)
635 Spontaneous activities. While spontaneous activity of MSNs and TANs are often as
636 low as less than 10 Hz, FSNs usually reach a very high level. We calculated the
637 spontaneous firing rates for each neuron using mean activities of 600~800 ms after
638 reward feedback in the main decision-making task. (3) Firing patterns. While TANs
639 usually fire tonically, MSNs and FSNs often present phasic firing. We measured this
640 through distributions of the inter-spike interval, calculating $ln(ISI_{median}/ISI_{mean})$.

641 The single-channel data were recorded using two systems, AlphaOmega SnR and
642 CED, and outlier detection was conducted separately for each system to avoid system
643 difference (Figure S1A). Note that inclusion of all recorded neurons didn't change our
644 conclusions, because MSNs occupies majority of the CN neurons.

645 **Reversible chemical inactivation**

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

654 **Electrical microstimulation**

655 For a traditional microstimulation experiment particularly applied in sensory
656 cortices, it's necessary to find the neuronal cluster with similar electrophysiological
657 properties near the electrode tip.[57,59,66] However, it has been reported that there is no
658 such cluster in CN.[9] So we did not systematically measure clustering to guide
659 microstimulation, similar to previous work.[58] For the sake of comparison with others'
660 work,[58,60] the current was set to be negative-leading bipolar pulses, 300 Hz, 50~80
661 μA. These parameters would not evoke saccades.[58,67] In a microstimulation block,
662 50% trials were randomly chosen to deliver the electric current, making it unlikely for
663 the animals to quickly adapt strategy.

664 **Data analysis**

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

672 $$\frac{1}{\sigma_{predct.}} = \sqrt{\frac{1}{\sigma_{vest.}^2} + \frac{1}{\sigma_{vis.}^2}}$$

673 where $\sigma_{vest.}$ and $\sigma_{vis.}$ indicate psychometric thresholds measured in vestibular-only
674 and visual-only conditions, respectively.[21]

675 ***Choice and modality preference.*** To construct peri-stimulus time histogram
676 (PSTH), raw spike trains were aligned to events of stimulus onset, saccade onset, and
677 feedback onset. Firing rates were computed in every 10-ms time window and
678 smoothed with a Gaussian kernel (σ = 50 ms). Next, all correct trials were grouped
679 according to animals' choice (IN versus OUT target). When averaging, PSTHs were
680 normalized for each cell across time and modalities by

681 $$r_{norm} = \frac{r - r_{min}}{r_{max} - r_{min}}$$

682 where r_{max} and r_{min} were respectively maximum and minimum firing rates for
683 each neuron. To quantify strength of choice signals, we used receiver operator
684 characteristic (ROC) analysis to define an index of choice divergence (CDiv).[22] CDiv
685 ranged from -1 to 1, with 1 meaning always stronger firing when IN target is chosen,
686 and vice versa for -1. CDiv = 0 indicates no choice preference. Grand CDiv was also
687 computed by using firing rates across the whole stimulus duration to identify choice
688 preference for each neuron in each modality condition.

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

Direction discriminative index (DDI). To estimate the strength of neuronal selectivity to directions in MGS task and passive heading task, DDI[22] was computed as:

$$DDI = \frac{R_{max} - R_{min}}{R_{max} - R_{min} + 2\sqrt{\left(SSE/(N-M)\right)}}$$

where R_{max} and R_{min} respectively indicate mean firing rates on the direction with strongest and weakest activities; SSE is the sum-squared error around mean responses; N is the total number of observations (trials), and M is the number of stimulus directions. Value of DDI ranges from 0 to 1, with the larger value indicating stronger selectivity.

Dimensional reduction of neural activities. To better visualize neural trajectories in low-dimensional subspace and demix task variables, we used three dimensionality reduction methods: principal component analysis (PCA), targeted dimensionality reduction (TDR), and demixed principal component analysis (dPCA).

For PCA, we first constructed a matrix $\boldsymbol{r}$ of size $N_{unit} \times (N_{T*3modalities*2choices})$, whose columns contained firing rates of all single neurons during a specific time bin in each stimulus condition when one of targets was chosen. Then, standard singular value decomposition (SVD) was conducted to get left singular vectors, $\boldsymbol{U}$. The first three columns of $\boldsymbol{U}$ (3 PCs) were chosen to span a denoising subspace, onto which the population activities $\boldsymbol{r}$ were projected. In this three-dimension space, six neural trajectories (3 modalities and 2 choices) were presented as a function of elapsed time.

For TDR, a subspace of $\boldsymbol{D}$ 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:

$$\boldsymbol{r}_{i,t} = \beta_{i,t,1}.\boldsymbol{heading} + \beta_{i,t,2}.\boldsymbol{choice} + \beta_{i,t,3}.\boldsymbol{modality} + \beta_{i,t,4}$$

where $\boldsymbol{r}_{i,t}$ is the z-scored firing rates of neuron i at time t in all trials, $\boldsymbol{heading}$ represented heading directions of all trial, $\boldsymbol{choice}$ included animal decisions (-1: OUT choice; 1: IN choice), and $\boldsymbol{modality}$ contained 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 denoising subspace $\boldsymbol{D}$, 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 de-noised 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.[30]

For dPCA, we utilized toolbox published previously.[27] Briefly, we first built the neural data matrix $\boldsymbol{r}$ of size $N_{unit} \times N_{modality} \times N_{choice} \times T \times N_{rep.}$), where N_{unit}, $N_{modality}$, N_{choice} and $N_{rep.}$ respectively represent number of neurons, modality conditions, choices and repetitions. Then, decomposing the neural activities into five parts: condition-independent, modality-dependent, decision-dependent, modality-decision interaction-dependent, and noise through marginalization. The principal components relative to corresponding parts were found by minimizing the difference between the marginalized data and the reconstructed full data.

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:

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

where $p_{T1}(t)$ represents the probability of choosing T1 target (IN 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) was computed. 80 trials within each group were randomly selected for training, and the rest trial (>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 (IN) choice if the $p_{T1}(t)$ > 0.5 and T2 (OUT) choice if $p_{T1}(t)$ < 0.5 at any moment. This step was repeated 1000 times to assess significance over chance level.

Generalized drift-diffusion model (GDDM). To predict firing dynamics in bimodal condition in reaction-time task, we used GDDM proposed previously.[31] In brief, the model assumes that both sensory and evidence accumulation processes are modulated by cue reliability:

$$x(t) = r(t) k sin(h) + \eta(t)$$

$$X(t) = \int r(t)x(t)dt$$

where $x(t)$ and $X(t)$ are respectively sensory evidence and decision variable at time t; $r(t)$ indicates moment-to-moment reliabilities 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.[31] Cue-reliability during bimodal condition was formulated as:

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

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

$$k_{predct.}^2 = k_{vest.}^2 + k_{vis.}^2$$

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

$$X_{comb.}(t) = \frac{k_{vest.}}{k_{comb.}} X_{vest.}(t) + \frac{k_{vis.}}{k_{comb.}} X_{vis.}(t)$$

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 drift rate; θ indicates decision threshold; b is heading bias; t_{nd} represents non-decision time; p_{lapse} means lapse rate.

Normalized shifts of point of subjective equity. To compare the effects of microstimulation across modal conditions, we normalized the change of PSE (ΔPSE) in each session by dividing the psychometric threshold under the condition without microstimulation. This generated ΔPSE_{norm}, which was unitless.

$$\Delta PSE_{norm} = \frac{\Delta PSE}{\sigma_{no_stim}}$$

Recurrent neural network simulations

We trained randomly initiated recurrent neural networks (RNN), 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:

$$\boldsymbol{x}(t) = \tanh\left(\boldsymbol{W_r}\boldsymbol{x}(t-1) + \boldsymbol{W_i}\boldsymbol{I} + \boldsymbol{B} + \boldsymbol{Z}\right)$$

where matrices $\boldsymbol{W_r}$ and $\boldsymbol{W_i}$ denote recurrent and input weights, respectively; $\boldsymbol{B}$ represents activity offset; $\boldsymbol{Z}$ 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 0.05-s. 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 condition. Indeed, after training, the network produced analogous performance (psychophysical threshold) in visual and vestibular condition.

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

$$x_o(t) = \sum_{i=1}^{256} w_i x_i(t)$$

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

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, RNN-output are initialized through He initialization and updated simultaneously using Adam Optimizer and the back-propagation through time (BPTT). Loss in each time step is calculated using BCELoss combined with a Sigmoid layer (BCEWithLogitsLoss function):

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

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

Supplemental Information

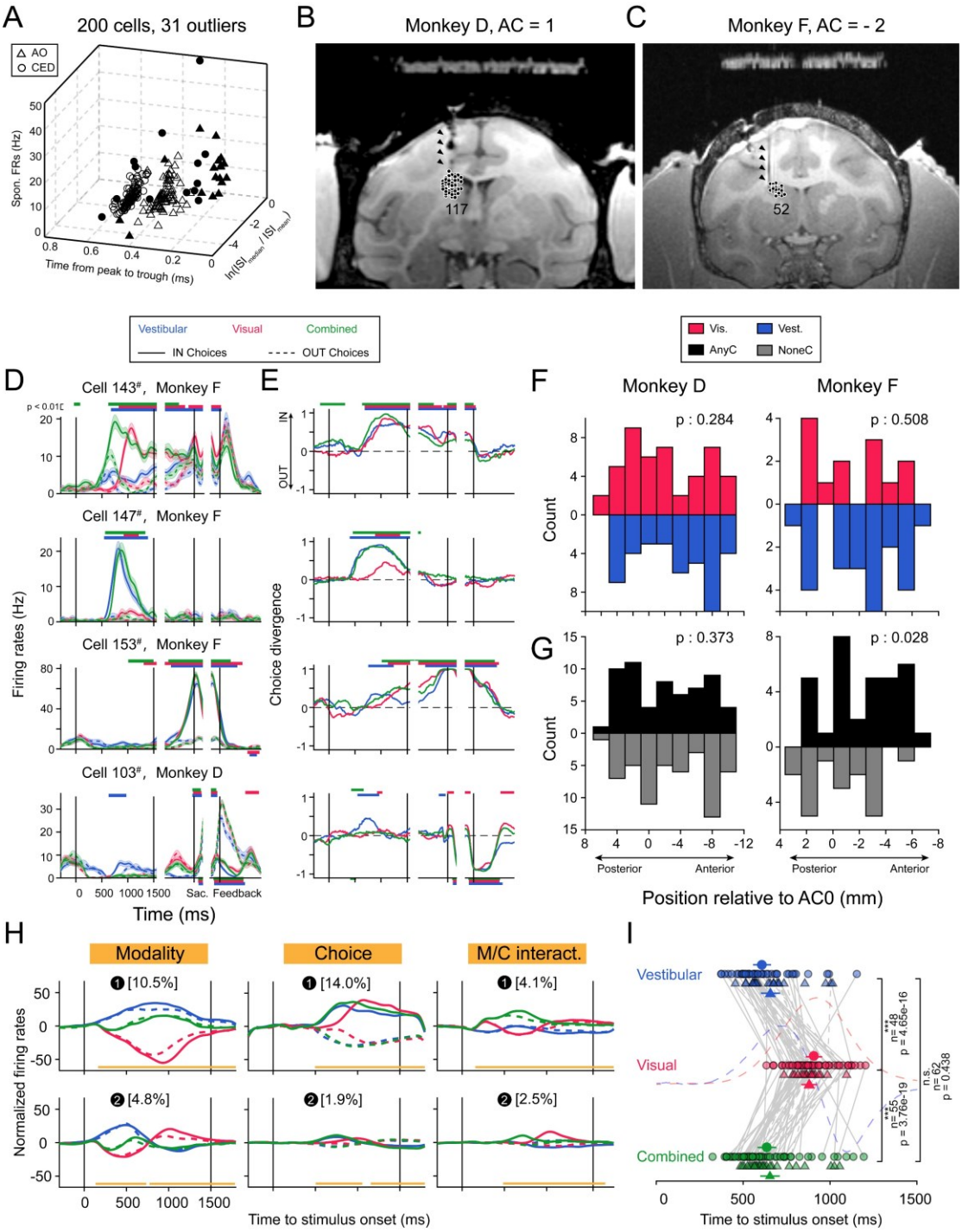

Figure S1. Related to Figure 2. Anatomical distribution of CN MSNs and neuronal properties (A) Selecting MSNs using robust covariance estimate method. The 3-d coordinate indicates that electrophysiological properties are quantified through three metrics: x axis, time from peak to trough of spike waveform; y axis, spontaneous firing rates; z axis, median over mean of the inter-spike interval distribution. Shapes of the data points indicate recording systems (triangle: AlphaOmega SnR; Circle: CED); solid dots represent outliers detected by the algorithm.

(B-C) Representative coronal sections of monkey D (left panel) and monkey F (right panel). All recorded neurons were superimposed onto this section. Triangles mark the artifacts of electrodes from MRI scanning.

(D) PSTHs of four example neurons. For each cell, responses were respectively aligned to the stimulus onset (left subpanels), saccade onset (middle subpanels), and feedback onset (right subpanels). Blue: vestibular-only; Red: visual-only; Green: bimodal condition. Solid and dashed lines respectively represent trials where monkeys choose the IN and OUT targets with respect to the response field of the recorded neuron. Horizontal colored bars mark the time period when IN and OUT trials significantly differ in firing rates ($p < 0.01$, two-tailed t-test). Bar's location indicates the relationship between their responses (IN > OUT, top; IN < OUT, bottom).

(E) CDivs of the same four neurons in (D).

(F) Spatial locations for neurons with different modality preferences in the two animals. Only neurons with significant modality preference were included. Red and blue bars indicate visual and vestibular preferences, respectively. Abscissas represent the location relative to AC0 along the AP axis (positive, posterior to AC0; negative, anterior to AC0), and the ordinates count the frequency. All p values are from Mann-Whitney U-test.

(G) Spatial locations of neurons with and without significant choice preferences. All markers are same as (F), except that colors here denote whether neurons have significant choice preferences (black, significant; gray, not significant).

(H) De-mixing task variables from CN population through dPCA analysis. Columns and rows respectively represent task-related parameters (modality, choice, and modality-choice interaction) and their dPCs (sorted by the proportion of explained total variances, ETV). For each subpanel, the percentage above means proportion of ETV; orange horizontal bars mark the period when corresponding variable can be significantly decoded from that dPC.

(I) Across-modality comparison of divergence time. Data points with error bars denote mean and SEM; gray lines connect data from the same neurons; *** and n.s. respectively represent $p < 1e-03$ and no significance; two-tailed paired t-test; significance level is 0.01.

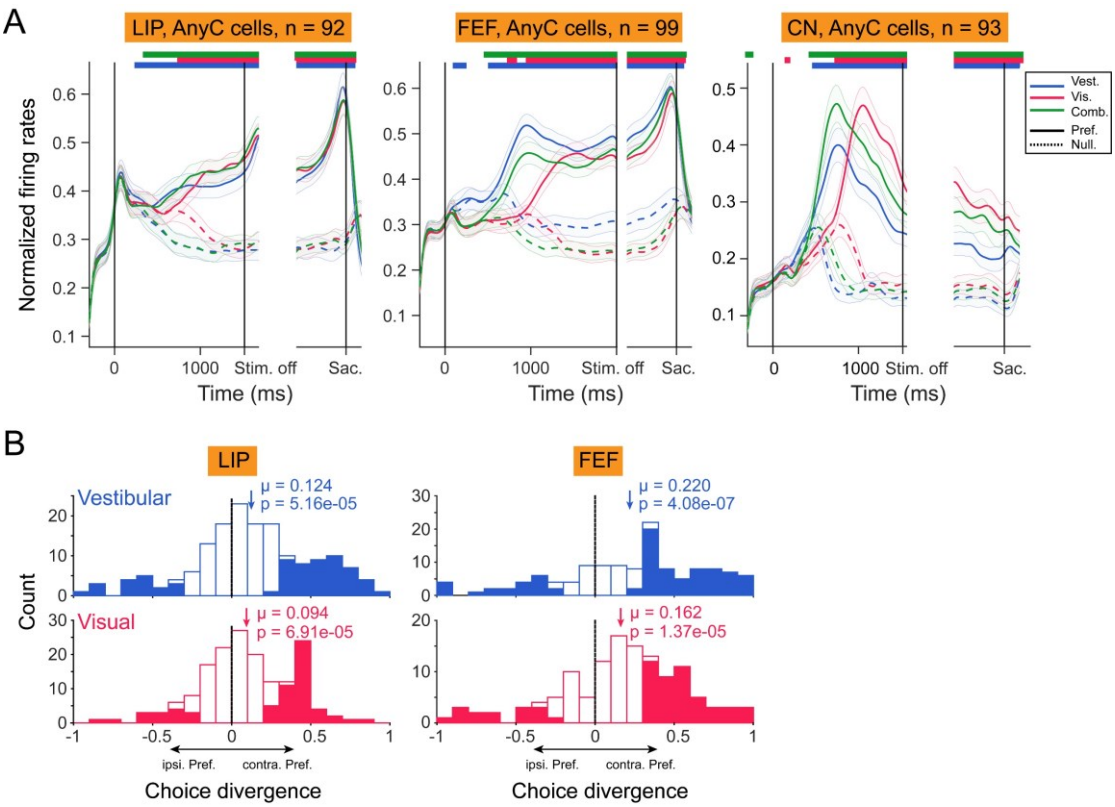

Figure S2. Related to Figure 2. Choice related signals across areas

(A) Population normalized PSTHs of decision-related neurons in LIP (left), FEF (middle), and CN (right). Only cells encoding choice signals in at least one modality conditions are included. Trials are sorted according to modality context and choice. Labels are same as Figure 2B.

(B) Distributions of choice signals under single modality context in LIP (left column) and FEF (right column). Labels are same as Figure 2A.

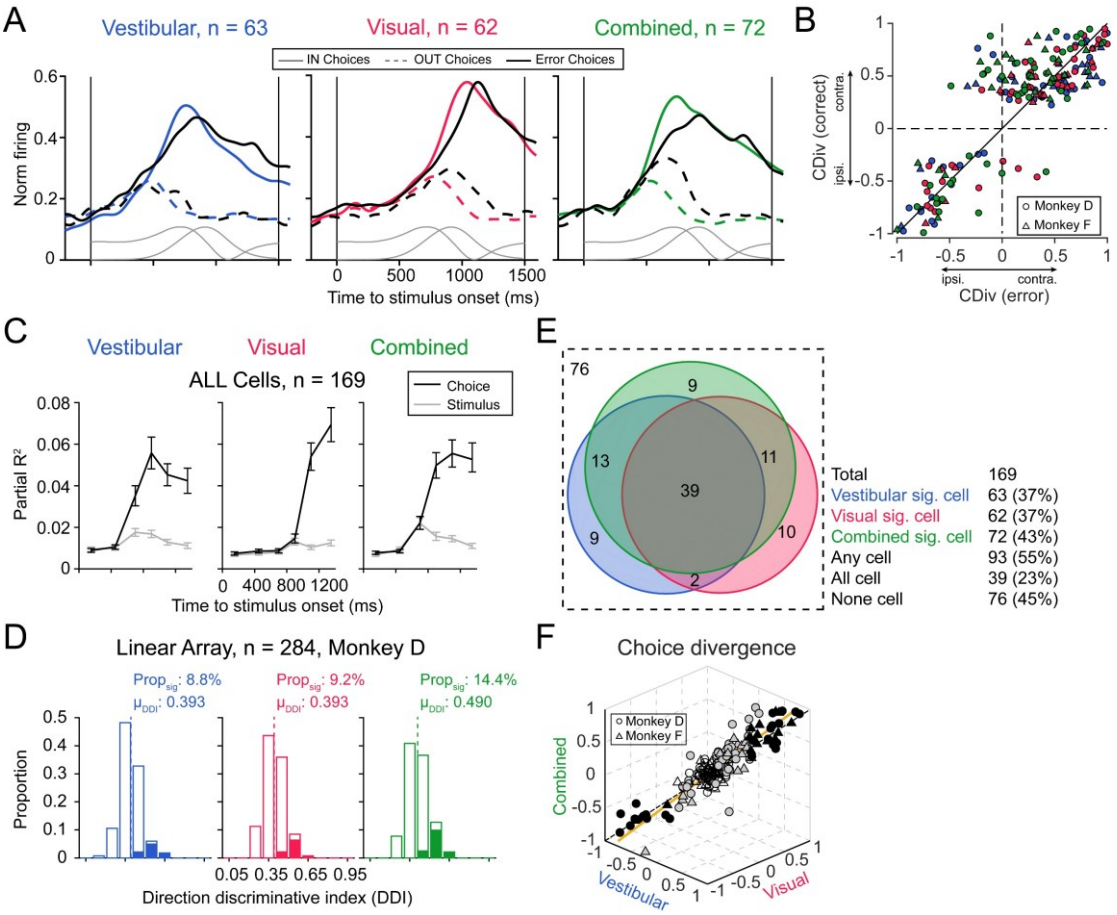

Figure S3. Related to Figure 2. CN neurons are largely modulated by internal decision rather than external physical stimulus

(A) Population average of normalized firing rates in correct (color curves) and error trials (black curves). Solid and dashed lines respectively denote trials where IN and OUT targets are chosen. Gray curves: motion stimulus profile.

(B) Comparing CDivs in correct versus error trials. Sign represents direction (positive: preferring contralateral choice; negative: preferring ipsilateral choice). Circles and triangles are data from monkey D and F, respectively.

(C) Partial correlation of neural responses with choice (black curves) and with heading (gray curves). Error bars represent SEM.

(D) Distributions of neuronal DDI measured during passive, fixation-only heading task. Solid bars indicate significantly tuned neurons (one-way ANOVA, $p < 0.05$).

(E) Classification of CN neurons according to grand CDiv. Number within each colored region marks the count of cells with significant grand CDiv under the corresponding combination of modal conditions.

(F) Correlation of CDivs across modality conditions. Circle and triangle dots are data from monkey D and F, respectively. Colors indicate how many modality conditions in which CDivs are significant (black, three; gray, one or two; empty, none). Orange line is the first principal component of PCA analysis; the first PC captures more than 88% variances; dashed black line is the diagonal.

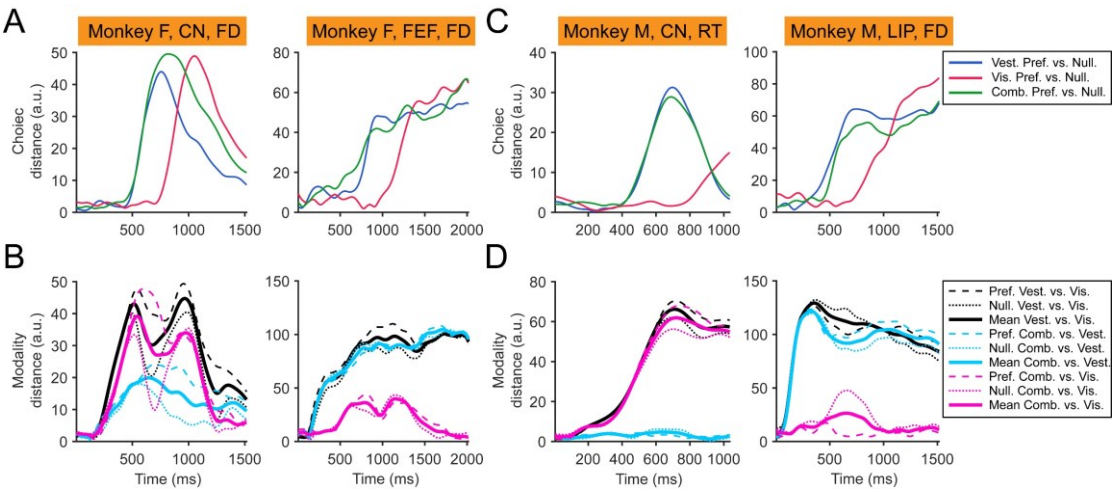

Figure S4. Related to Figure 2. Comparing neural states across regions within the same monkey

(A-B) Temporal dynamics of choice (A) and modality (B) distance in CN (left column) or FEF (right column) of monkey F performing the fixed-duration task. Labels are same as Figure 3D. (C-D) Temporal dynamics of choice (C) and modality (D) distance in (left column) CN of monkey M performing the reaction-time task, or in (right column) LIP of the same monkey but performing the fixed-duration task.

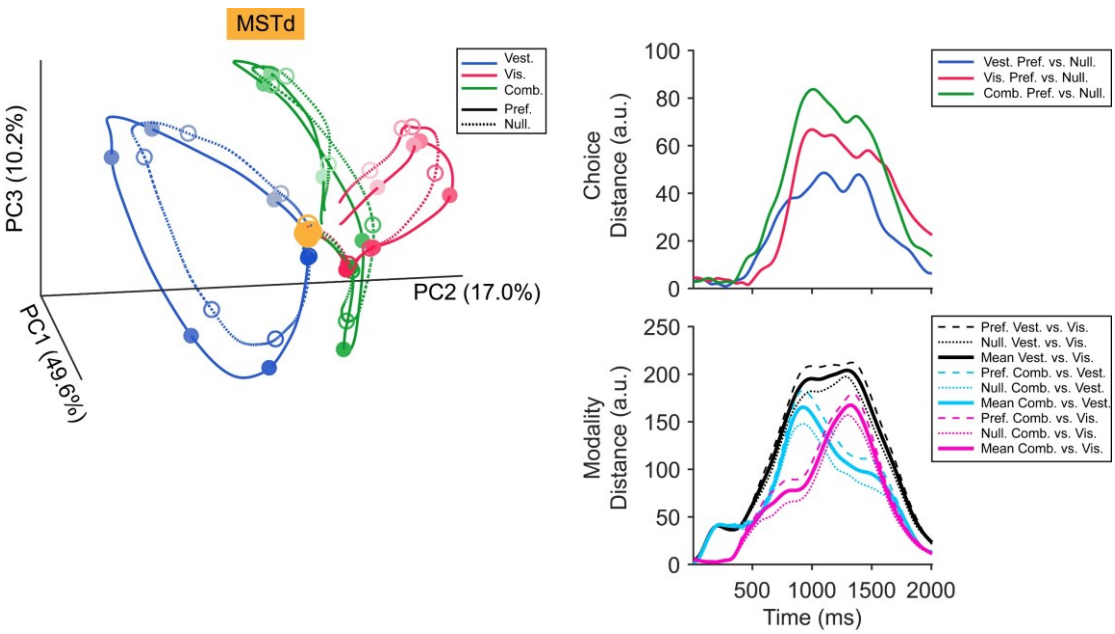

Figure S5 Related to Figure 2. Neural state trajectories of area MSTd

Labels are same as Figure 2.

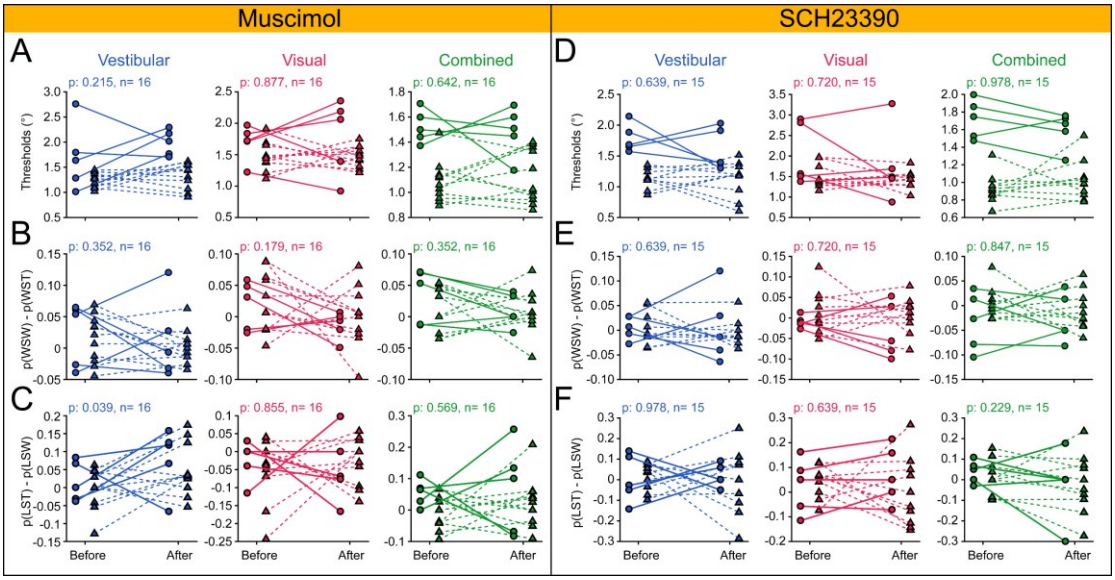

Figure S6. Related to Figure 5. Effects of unilateral chemical inactivation in CN on perceptual sensitivity and history strategy

(A) Psychometric thresholds before and after injecting muscimol. Different symbols represent data from the two animals. P-value is from Wilcoxon signed-rank test; significance level is 0.05.

(B) Win-switch versus win-stay proportion before and after injecting muscimol. Labels are same as (A).

(C) Lose-stay versus lose-switch proportion before and after injecting muscimol. Labels are same as (A).

(D-F) Same format as in (A-C), but for D1-receptor antagonist SCH23390 injection in CN.

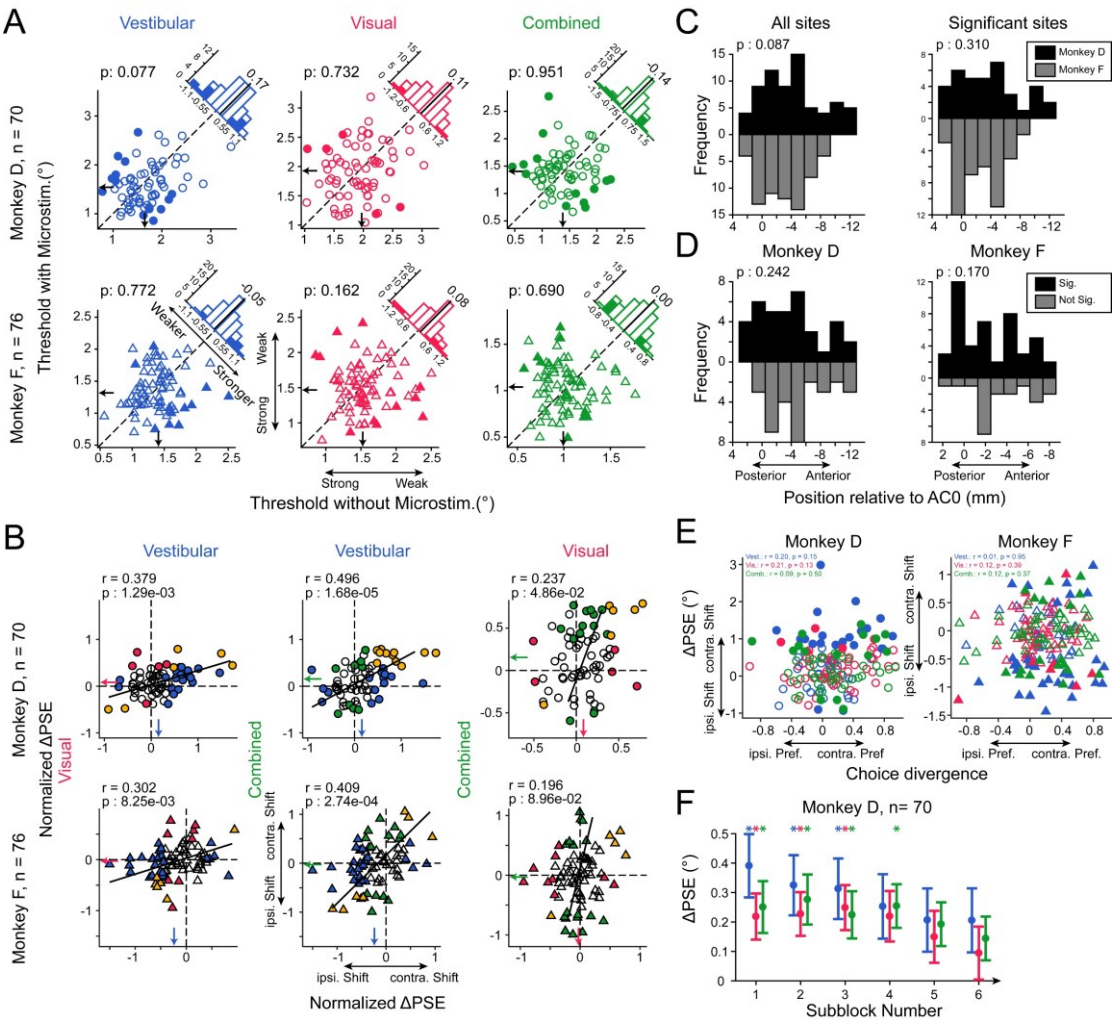

Figure S7. Related to Figure 6. Effects of unilateral microstimulation in CN on perceptual sensitivity and other aspects

(A) Effects the microstimulation during the stimulus period induces on monkeys' psychometric thresholds. Labels are same as Figure 6D.

(B) Cross-modality correlation of microstimulation effects on PSE shift. PSE shift was normalized with respect to perceptual sensitivity (threshold). Colors of dots mean in which cue condition microstimulation significantly biases animal choices (blue, just in vestibular-only condition; red, just in visual-only condition; green, just in combined condition; orange, in both conditions; empty, in neither condition). In each plot, black line was fitted using least sum of squared perpendicular distances; p values and correlation coefficient rare from Spearman correlation analysis.

(C-D) Spatial locations of microstimulation sites across monkeys (C) and within individual animal (D). In each panel, ordinate and abscissa respectively indicate number of cases and the location relative to AC0 on the AP axis (positive, posterior to AC0; negative, anterior to AC0); color means different monkeys (black, monkey D; gray, monkey F); the p value is from Mann-Whitney U-test.

(E) Correlation between microstimulation effects on PSE (ordinate) and choice preferences (abscissa) of the local multi-units. Filled dots indicate cases with significant microstimulation induced PSE shift; p value and correlation coefficient rare from Spearman correlation.

(F) Adaptation of microstimulation effects. The ΔPSE was calculated in a sliding 5-repetition windows (step is 1 repetition) for all cases from monkey D. Dots with error bars represent mean with SEM. Asterisks mark the subblocks where microstimulation significantly biases PSE (two-tailed t-test; significance level is 0.01).

REFERENCES

1. Mizusaki, B.E.P., and O'Donnell, C. (2021). Neural circuit function redundancy in brain disorders. Curr Opin Neurobiol *70*, 74-80. 10.1016/j.conb.2021.07.008.
2. Lee, D., Seo, H., and Jung, M.W. (2012). Neural basis of reinforcement learning and decision making. Annu Rev Neurosci *35*, 287-308. 10.1146/annurev-neuro-062111-150512.
3. Christophel, T.B., Klink, P.C., Spitzer, B., Roelfsema, P.R., and Haynes, J.D. (2017). The Distributed Nature of Working Memory. Trends Cogn Sci *21*, 111-124. 10.1016/j.tics.2016.12.007.
4. Steinmetz, N.A., Zatka-Haas, P., Carandini, M., and Harris, K.D. (2019). Distributed coding of choice, action and engagement across the mouse brain. Nature *576*, 266-273. 10.1038/s41586-019-1787-x.
5. Siegel, M., Buschman, T.J., and Miller, E.K. (2015). Cortical information flow during flexible sensorimotor decisions. Science *348*, 1352-1355. 10.1126/science.aab0551.
6. Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J Neurophysiol *86*, 1916-1936. 10.1152/jn.2001.86.4.1916.
7. Kim, J.N., and Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. Nat Neurosci *2*, 176-185. 10.1038/5739.
8. Horwitz, G.D., Batista, A.P., and Newsome, W.T. (2004). Representation of an abstract perceptual decision in macaque superior colliculus. J Neurophysiol *91*, 2281-2296. 10.1152/jn.00872.2003.
9. Ding, L., and Gold, J.I. (2010). Caudate encodes multiple computations for perceptual decisions. J Neurosci *30*, 15747-15759. 10.1523/JNEUROSCI.2894-10.2010.
10. Hikosaka, O., Sakamoto, M., and Usui, S. (1989). Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. J Neurophysiol *61*, 780-798. 10.1152/jn.1989.61.4.780.
11. Kravitz, A.V., Freeze, B.S., Parker, P.R., Kay, K., Thwin, M.T., Deisseroth, K., and Kreitzer, A.C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature *466*, 622-626. 10.1038/nature09159.
12. Cai, X., Kim, S., and Lee, D. (2011). Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. Neuron *69*, 170-182. 10.1016/j.neuron.2010.11.041.
13. Kim, Hyoung F., and Hikosaka, O. (2013). Distinct Basal Ganglia Circuits Controlling Behaviors Guided by Flexible and Stable Values. Neuron *79*, 1001-1010. 10.1016/j.neuron.2013.06.044.
14. Mello, G.B., Soares, S., and Paton, J.J. (2015). A scalable population code for time in the striatum. Curr Biol *25*, 1113-1122. 10.1016/j.cub.2015.02.036.

15. Monteiro, T., Rodrigues, F.S., Pexirra, M., Cruz, B.F., Goncalves, A.I., Rueda-Orozco, P.E., and Paton, J.J. (2023). Using temperature to analyze the neural basis of a time-based decision. Nat Neurosci *26*, 1407-1416. 10.1038/s41593-023-01378-5.
16. Ding, L., and Gold, J.I. (2013). The basal ganglia's contributions to perceptual decision making. Neuron *79*, 640-649. 10.1016/j.neuron.2013.07.042.
17. Lo, C.C., and Wang, X.J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nat Neurosci *9*, 956-963. 10.1038/nn1722.
18. Peters, A.J., Fabre, J.M.J., Steinmetz, N.A., Harris, K.D., and Carandini, M. (2021). Striatal activity topographically reflects cortical activity. Nature *591*, 420-425. 10.1038/s41586-020-03166-8.
19. Chen, S., Liu, Y., Wang, Z.A., Colonell, J., Liu, L.D., Hou, H., Tien, N.W., Wang, T., Harris, T., Druckmann, S., et al. (2024). Brain-wide neural activity underlying memory-guided movement. Cell *187*, 676-691 e616. 10.1016/j.cell.2023.12.035.
20. Van Hoesen, G.W., Yeterian, E.H., and Lavizzo-Mourey, R. (1981). Widespread corticostriate projections from temporal cortex of the rhesus monkey. J Comp Neurol *199*, 205-219. 10.1002/cne.901990205.
21. Gu, Y., Angelaki, D.E., and Deangelis, G.C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. Nat Neurosci *11*, 1201-1210. 10.1038/nn.2191.
22. Hou, H., Zheng, Q., Zhao, Y., Pouget, A., and Gu, Y. (2019). Neural Correlates of Optimal Multisensory Decision Making under Time-Varying Reliabilities with an Invariant Linear Probabilistic Population Code. Neuron *104*, 1010-1021 e1010. 10.1016/j.neuron.2019.08.038.
23. Zheng, Q., Zhou, L., and Gu, Y. (2021). Temporal synchrony effects of optic flow and vestibular inputs on multisensory heading perception. Cell Rep *37*, 109999. 10.1016/j.celrep.2021.109999.
24. Butler, J.S., Smith, S.T., Campos, J.L., and Bulthoff, H.H. (2010). Bayesian integration of visual and vestibular signals for heading. J Vis *10*, 23. 10.1167/10.11.23.
25. Friedman, A., Homma, D., Gibb, L.G., Amemori, K., Rubin, S.J., Hood, A.S., Riad, M.H., and Graybiel, A.M. (2015). A Corticostriatal Path Targeting Striosomes Controls Decision-Making under Conflict. Cell *161*, 1320-1333. 10.1016/j.cell.2015.04.049.
26. Raposo, D., Kaufman, M.T., and Churchland, A.K. (2014). A category-free neural population supports evolving demands during decision-making. Nat Neurosci *17*, 1784-1792. 10.1038/nn.3865.
27. Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C.E., Kepecs, A., Mainen, Z.F., Qi, X.L., Romo, R., Uchida, N., and Machens, C.K. (2016). Demixed principal component analysis of neural population data. Elife *5*. 10.7554/eLife.10989.
28. Rigotti, M., Barak, O., Warden, M.R., Wang, X.J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature *497*, 585-590. 10.1038/nature12160.
29. Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Foster, J.D., Nuyujukian, P., Ryu, S.I., and Shenoy, K.V. (2012). Neural population dynamics during reaching. Nature *487*, 51-56. 10.1038/nature11129.
30. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78-84. 10.1038/nature12742.

31. Drugowitsch, J., DeAngelis, G.C., Klier, E.M., Angelaki, D.E., and Pouget, A. (2014). Optimal multisensory decision-making in a reaction-time task. Elife *3*. 10.7554/eLife.03005.
32. Bjorklund, A., and Dunnett, S.B. (2007). Dopamine neuron systems in the brain: an update. Trends Neurosci *30*, 194-202. 10.1016/j.tins.2007.03.006.
33. Kitama, T., Ohno, T., Tanaka, M., Tsubokawa, H., and Yoshida, K. (1991). Stimulation of the caudate nucleus induces contraversive saccadic eye movements as well as head turning in the cat. Neurosci Res *12*, 287-292. 10.1016/0168-0102(91)90118-i.
34. Nichols, M.J., and Newsome, W.T. (2002). Middle temporal visual area microstimulation influences veridical judgments of motion direction. J Neurosci *22*, 9530-9540. 10.1523/JNEUROSCI.22-21-09530.2002.
35. Yu, X., Hou, H., Spillmann, L., and Gu, Y. (2018). Causal Evidence of Motion Signals in Macaque Middle Temporal Area Weighted-Pooled for Global Heading Perception. Cereb Cortex *28*, 612-624. 10.1093/cercor/bhw402.
36. Histed, M.H., Bonin, V., and Reid, R.C. (2009). Direct activation of sparse, distributed populations of cortical neurons by electrical microstimulation. Neuron *63*, 508-522. 10.1016/j.neuron.2009.07.016.
37. Tolias, A.S., Sultan, F., Augath, M., Oeltermann, A., Tehovnik, E.J., Schiller, P.H., and Logothetis, N.K. (2005). Mapping cortical activity elicited with electrical microstimulation using FMRI in the macaque. Neuron *48*, 901-911. 10.1016/j.neuron.2005.11.034.
38. Jeurissen, D., Shushruth, S., El-Shamayleh, Y., Horwitz, G.D., and Shadlen, M.N. (2022). Deficits in decision-making induced by parietal cortex inactivation are compensated at two timescales. Neuron *110*, 1924-1931 e1925. 10.1016/j.neuron.2022.03.022.
39. Gu, Y., Deangelis, G.C., and Angelaki, D.E. (2012). Causal links between dorsal medial superior temporal area neurons and multisensory heading perception. J Neurosci *32*, 2299-2313. 10.1523/JNEUROSCI.5154-11.2012.
40. Potegal, M., Copack, P., de Jong, J., Krauthamer, G., and Gilman, S. (1971). Vestibular input to the caudate nucleus. Exp Neurol *32*, 448-465.
41. Maunsell, J.H.R., and Vanessen, D.C. (1983). 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. https://doi.org/10.1523/JNEUROSCI.03-12-02563.1983.
42. Reig, R., and Silberberg, G. (2014). Multisensory integration in the mouse striatum. Neuron *83*, 1200-1212. 10.1016/j.neuron.2014.07.033.
43. Borra, E., Biancheri, D., Rizzo, M., Leonardi, F., and Luppino, G. (2022). Crossed Corticostriatal Projections in the Macaque Brain. J Neurosci *42*, 7060-7076. 10.1523/JNEUROSCI.0071-22.2022.
44. Yartsev, M.M., Hanks, T.D., Yoon, A.M., and Brody, C.D. (2018). Causal contribution and dynamical encoding in the striatum during evidence accumulation. Elife *7*. 10.7554/eLife.34929.
45. Fetsch, C.R., Turner, A.H., DeAngelis, G.C., and Angelaki, D.E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. J Neurosci *29*, 15601-15612. 10.1523/JNEUROSCI.2574-09.2009.
46. Nienborg, H., and Cumming, B.G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature *459*, 89-92. 10.1038/nature07821.

1045 47. Yates, J.L., Park, I.M., Katz, L.N., Pillow, J.W., and Huk, A.C. (2017). Functional dissection
1046 of signal and noise in MT and LIP during decision-making. Nat Neurosci *20*, 1285-1292.
1047 10.1038/nn.4611.
1048 48. Okazawa, G., Sha, L., Purcell, B.A., and Kiani, R. (2018). Psychophysical reverse correlation
1049 reflects both sensory and decision-making processes. Nat Commun *9*, 3479. 10.1038/s41467-
1050 018-05797-y.
1051 49. Noudoost, B., and Moore, T. (2011). Control of visual cortical signals by prefrontal dopamine.
1052 Nature *474*, 372-375. 10.1038/nature09995.
1053 50. Ott, T., Jacob, S.N., and Nieder, A. (2014). Dopamine receptors differentially enhance rule
1054 coding in primate prefrontal cortex neurons. Neuron *84*, 1317-1328.
1055 10.1016/j.neuron.2014.11.012.
1056 51. Puig, M.V., and Miller, E.K. (2012). The role of prefrontal dopamine D1 receptors in the
1057 neural mechanisms of associative learning. Neuron *74*, 874-886.
1058 10.1016/j.neuron.2012.04.018.
1059 52. Williams, Z.M., and Eskandar, E.N. (2006). Selective enhancement of associative learning by
1060 microstimulation of the anterior caudate. Nat Neurosci *9*, 562-568. 10.1038/nn1662.
1061 53. Nicola, S.M., Surmeier, J., and Malenka, R.C. (2000). Dopaminergic modulation of neuronal
1062 excitability in the striatum and nucleus accumbens. Annu Rev Neurosci *23*, 185-215.
1063 10.1146/annurev.neuro.23.1.185.
1064 54. Vijayraghavan, S., Wang, M., Birnbaum, S.G., Williams, G.V., and Arnsten, A.F. (2007).
1065 Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory.
1066 Nat Neurosci *10*, 376-384. 10.1038/nn1846.
1067 55. Bolkan, S.S., Stone, I.R., Pinto, L., Ashwood, Z.C., Iravedra Garcia, J.M., Herman, A.L.,
1068 Singh, P., Bandi, A., Cox, J., Zimmerman, C.A., et al. (2022). Opponent control of behavior
1069 by dorsomedial striatal pathways depends on task demands and internal state. Nat Neurosci
1070 *25*, 345-357. 10.1038/s41593-022-01021-9.
1071 56. Chen, Y., Hong, Z., Wang, J., Liu, K., Liu, J., Lin, J., Feng, S., Zhang, T., Shan, L., Liu, T., et
1072 al. (2023). Circuit-specific gene therapy reverses core symptoms in a primate Parkinson's
1073 disease model. Cell *186*, 5394-5410 e5318. 10.1016/j.cell.2023.10.004.
1074 57. Hanks, T.D., Ditterich, J., and Shadlen, M.N. (2006). Microstimulation of macaque area LIP
1075 affects decision-making in a motion discrimination task. Nat Neurosci *9*, 682-689.
1076 10.1038/nn1683.
1077 58. Ding, L., and Gold, J.I. (2012). Separate, causal roles of the caudate in saccadic choice and
1078 execution in a perceptual decision task. Neuron *75*, 865-874. 10.1016/j.neuron.2012.07.021.
1079 59. Salzman, C.D., Murasugi, C.M., Britten, K.H., and Newsome, W.T. (1992). Microstimulation
1080 in visual area MT: effects on direction discrimination performance. J Neurosci *12*, 2331-2355.
1081 60. Doi, T., Fan, Y., Gold, J.I., and Ding, L. (2020). The caudate nucleus contributes causally to
1082 decisions that balance reward and uncertain visual information. Elife *9*. 10.7554/eLife.56694.
1083 61. Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area
1084 during a combined visual discrimination reaction time task. J Neurosci *22*, 9475-9489.
1085 62. Gu, Y., Watkins, P.V., Angelaki, D.E., and DeAngelis, G.C. (2006). Visual and nonvisual
1086 contributions to three-dimensional heading selectivity in the medial superior temporal area. J
1087 Neurosci *26*, 73-85. 10.1523/JNEUROSCI.2356-05.2006.

63. Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci *12*, 4745-4765. 10.1523/JNEUROSCI.12-12-04745.1992.
64. Hikosaka, O., Sakamoto, M., and Usui, S. (1989). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. J Neurophysiol *61*, 814-832. 10.1152/jn.1989.61.4.814.
65. Lanciego, J.L., Luquin, N., and Obeso, J.A. (2012). Functional neuroanatomy of the basal ganglia. Cold Spring Harb Perspect Med *2*, a009621. 10.1101/cshperspect.a009621.
66. Li, W., Lu, J., Zhu, Z., and Gu, Y. (2022). Causal contribution of optic flow signal in Macaque extrastriate visual cortex for roll perception. Nat Commun *13*, 5479. 10.1038/s41467-022-33245-5.
67. Watanabe, M., and Munoz, D.P. (2011). Saccade reaction times are influenced by caudate microstimulation following and prior to visual stimulus appearance. J Cogn Neurosci *23*, 1794-1807. 10.1162/jocn.2010.21554.