

COGNITIVE NEUROSCIENCE

Multistable bimodal perceptual coding within the ventral premotor cortex

Bernardo Andrade-Ortega¹, Héctor Díaz¹, Lucas Bayones¹, Manuel Alvarez¹, Antonio Zainos¹, Natsuko Rivera-Yoshida^{1†}, Alessio Franci^{2,3*}, Ranulfo Romo^{4*}, Román Rossi-Pool^{1,5*}

How neurons in the ventral premotor cortex (VPC) integrate and process multimodal sensory information remains unclear. We recorded VPC neuronal activity in two monkeys performing a tactile or acoustic bimodal detection task (BDT). Single-cell analyses revealed purely tactile, purely acoustic, bimodal, and decision-mnemonic responses. Population analyses of the VPC network and a recurrent neural network trained on the BDT uncovered notably similar latent dynamics. In both networks, neural trajectories diverged sharply from rest along modality-specific directions during stimulus presentation, while remained at rest during stimulus-absent trials. Subsequently, trajectories rotated toward a mnemonic subspace, where perceptual decisions (acoustic, tactile, and absent) were maintained in three distinctly different stable attractors. Using low-dimensional modeling, we propose a universal mechanism for the transition from sensory processing to memory retention explaining both biological and artificial latent neural dynamics in a simple dynamical motif. These results show that VPC neurons encode bimodal information, can integrate competing sensory inputs, and maintain multimodal perceptual decisions.

INTRODUCTION

In a noisy restaurant, you can still hold a conversation with friends despite the ambient clamor. The brain effectively tunes into one speaker and filters out background noise—partly by integrating auditory signals with visual cues like lip movements (1). This natural interplay between sensory modalities demonstrates how the brain routinely combines information from different senses. While some cortical areas are specialized for multimodal integration (1), others resolve competition among heterogeneous sensory inputs. Understanding how the brain processes and integrates such information remains a vibrant area of research.

Historically, multisensory processing was thought to rely on extensive unimodal processing within separate sensory cortices (2, 3). However, recent studies suggest that even primary sensory areas—once believed to handle only a single type of input—contribute to multisensory integration (4–6). Evidence for this comes from investigations using various measures of neural activity, ranging from low (electroencephalography and functional magnetic resonance imaging) (7–9) to medium [local field potential (LFP)] (5, 10) to high (single-unit recordings) (5, 11, 12) spatial resolution. Although multisensory processing can occur in early cortical regions, the full encoding of perceptual events appears to emerge in higher-order areas within the processing hierarchy (13–16).

In our previous work, we addressed multimodal sensory integration by recording single-unit activity in nonhuman primates performing a bimodal detection task (BDT) to determine whether areas 3b and 1 of the primary somatosensory cortex (S1) exhibit

multisensory processing (17). In this task (Fig. 1A), monkeys were required to report the presence or absence of a sensory stimulus—which could be either tactile or acoustic—and, if detected, to indicate its modality. Tactile and acoustic stimuli of varying amplitudes were randomly interleaved with stimulus-absent trials so that each trial presented either no stimulus or a stimulus (tactile or acoustic) at above-threshold, near-threshold, or subthreshold levels (18). Thus, the BDT compels the brain to weigh tactile and acoustic signals against each other to generate the correct response, requiring the integration of both modalities. Although we identified hierarchical differences between areas 3b and 1, both predominantly exhibited pure tactile responses. This observation raised the question of which cortical area might encode and integrate both tactile and acoustic inputs.

Previous research has identified the VPC (ventral premotor cortex) as a potential intermediary hub between sensory and frontal cortices (2, 19). The VPC receives projections from sensory regions in the parietal and temporal lobes, as well as from association areas in the prefrontal and premotor cortices (20–23). Recent studies have also shown that coordination between the VPC and the medial prefrontal cortex (mPFC) plays a crucial role in monitoring social actions. While mPFC neurons showed stronger modulation by social context, VPC neurons maintain representations of actions even when their social relevance changes, suggesting a more context-independent role for this area (24). In addition, in naïve (untrained) monkeys, some VPC neurons respond to both tactile and acoustic stimuli (25, 26). During tactile tasks, VPC units display sensory, categorical, and persistent decision-related responses (3, 19, 27), and similar response patterns are observed during acoustic tasks (28). These findings suggest that the VPC is well-equipped to process and relay multimodal sensory information. Yet, it remains unclear whether—and how—the VPC contributes to the comparison of competing sensory modalities as required by the BDT. Several key questions remain unanswered: Do VPC neurons encode both sensory modalities? How diverse are their response patterns? And can we distinguish population-level responses to different types of stimuli?

¹Instituto de Fisiología Celular — Neurociencias, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico. ²Montefiore Institute, University of Liège, Av de la Découverte 10, 4000 Liège, Belgique. ³WEL Research Institute, avenue Pasteur, 6, 1300 Wavre, Belgique. ⁴El Colegio Nacional, 06020 Mexico City, Mexico. ⁵Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico.

*Corresponding author. Email: alessio.franci83@gmail.com (A.F.); ranulfo.romo@gmail.com (R.R.); romanr@ifc.unam.mx (R.R.-P.)

†Present address: Laboratorio Nacional de Ciencias de la Sostenibilidad, Instituto de Ecología, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México.

In this study, we recorded neuronal activity from the VPC of two monkeys performing the BDT and analyzed the data at both single-neuron and population levels. We identified unimodal neurons that encoded the presence of either acoustic or tactile stimuli, as well as bimodal neurons that signaled stimulus presence for both modalities. Notably, the responses of bimodal neurons could also be interpreted as amodal, potentially aiding in the detection of stimulus absence. By applying dimensionality reduction techniques,

we uncovered low-dimensional neural dynamics in the VPC that clearly differentiated responses to each sensory modality during stimulation. Specifically, acoustic and tactile stimuli elicited sharply diverging trajectories within a sensory subspace, while in the same subspace, the population dynamics remained stable during stimulus-absent trials. Later, decision-related dynamics emerged along a subspace orthogonal (uncorrelated) to the sensory one, and exhibiting a rotational pattern toward a memory subspace—presumably to reduce interference between sensory and mnemonic processing (29, 30).

Furthermore, a recurrent neural network (RNN) (31, 32) model trained on the BDT exhibited similar low-dimensional neural dynamics. This finding suggests that the observed diverge-then-rotate dynamics may represent a universal computational strategy for sensory discrimination followed by memory retention with minimal sensory-mnemonic interference. Through low-dimensional dynamical system modeling, we show that these dynamics arise from two saddle-point thresholds—each linked to one stimulus modality—and the geometry of their stable manifolds within a simple, multistable feedforward circuit. Together, our results implicate the VPC as a critical hub for integrating sensory inputs across modalities to guide perceptual decisions.

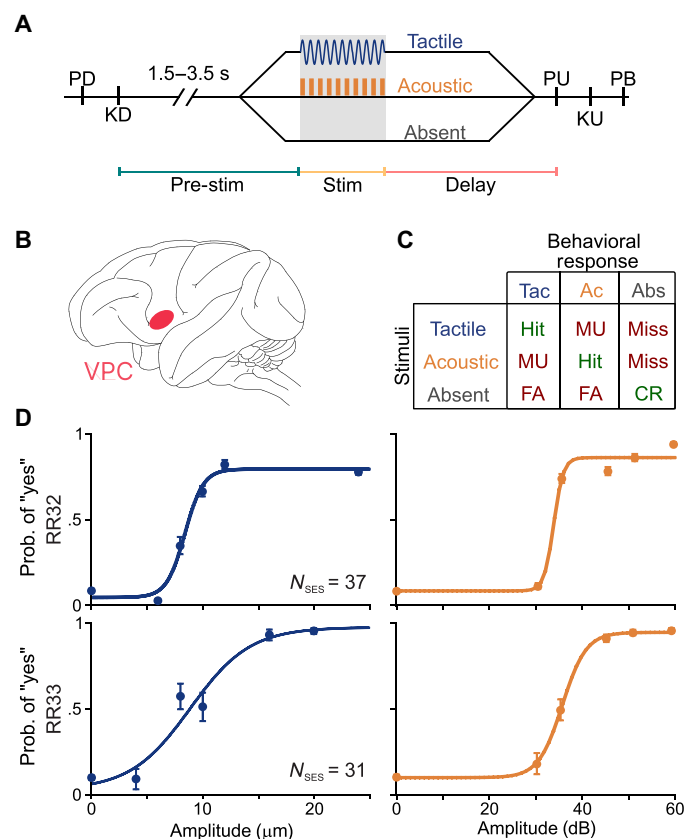


Fig. 1. BDT and performance. (A) Trials began with the lowering of a mechanical probe on one fingertip of the monkey's restrained right hand (Probe Down event, "pd"). The monkeys then placed their left free hand on a nonmoving key (Key Down event, "kd"). After kd, a variable period (1.5 to 3.5 s) ensued to prevent anticipation of the upcoming stimulus. This was followed by a 0.5-s stimulus period, which could be tactile, acoustic, or absent. Poststimulation a fixed 2-s delay occurred, leading to the probe's retraction (Probe Up event, "pu"), signaling the monkey to release the key (Key Up event, "ku") and report its decision by pressing one of three buttons placed at eye level (Push Button event, "pb"), indicating "tactile stimulus present," "acoustic stimulus present," and "stimulus absent." Correct responses were rewarded with a few drops of juice. (B) Depiction of the left cerebral hemisphere with the VPC delineated in red. (C) Matrix categorizing trial outcomes by stimuli and behavioral responses. Correct identifications are denoted in green, while incorrect responses are colored red. In the absence of a stimulus, a stimulus absent response results in a correct rejection (CR), while an incorrect "stimulus present" response indicates a false alarm (FA). With a stimulus present, a "present" response that aligns with the stimulus modality constitutes a hit. A present response that does not match the modality is a misunderstood trial (MU), and an "absent" response is a miss. (D) Psychometric curves for tactile (in blue) and acoustic (in orange) stimuli. Each modality features six classes of stimulus intensity from 0 to 24 μ m (tactile) and 0 to 59 dB (acoustic), over 37 sessions for subject RR32 and 31 sessions for subject RR33.

RESULTS

Bimodal detection task

Two rhesus monkeys (*Macaca mulatta*) were trained to perform BDT in which they determined the presence or absence of sensory stimuli—either vibrotactile (Tac) or acoustic (Ac)—ranging from barely detectable to clearly perceptible. Vibrotactile stimuli were delivered at 20 Hz for 0.5 s (see Fig. 1A), whereas acoustic stimuli were 1-kHz pure tones modulated at 20 Hz, also lasting 0.5 s. In tactile trials, the amplitude of the skin vibrations was precisely controlled; similarly, the intensity of the tones was carefully regulated during acoustic tests (17). During each recording session, animals were presented with multiple sets of 150 trials, where each set contained randomly interleaved trials featuring tactile (Tac, all variants), acoustic (Ac, all variants), and absent (Abs) stimuli, equally distributed across the three conditions (50 trials each). Following a decision-maintenance period, a "probe up" signal (PU) indicated that the animal could raise its hand (KU) and press the decision button (PB). The monkeys signaled their detection (or nondetection) of the stimulus by pressing one of three push buttons labeled "Tac," "Ac," or "Abs," which were conveniently positioned along their left midline. VPC's neuronal activity was recorded throughout the sessions (Fig. 1B). Trials were meticulously classified on the basis of stimulus type and the corresponding behavioral response (Fig. 1C), and psychometric analysis confirmed that the monkeys consistently performed near their psychophysical thresholds (Fig. 1D).

Individual neuron responses

We recorded neural activity from the VPC in animals RR32 ($n = 256$ from the right hemisphere and $n = 310$ from the left hemisphere) and RR33 ($n = 345$ from the left hemisphere), yielding a total of 911 units. In contrast to our previous findings in areas 3b and 1 of the S1—where neurons exhibited highly faithful responses exclusively to tactile stimuli (17)—the VPC population displayed transformed responses to both modalities, indicating a distinct integration of acoustic and tactile inputs (Fig. 2 and fig. S1). Some units showed

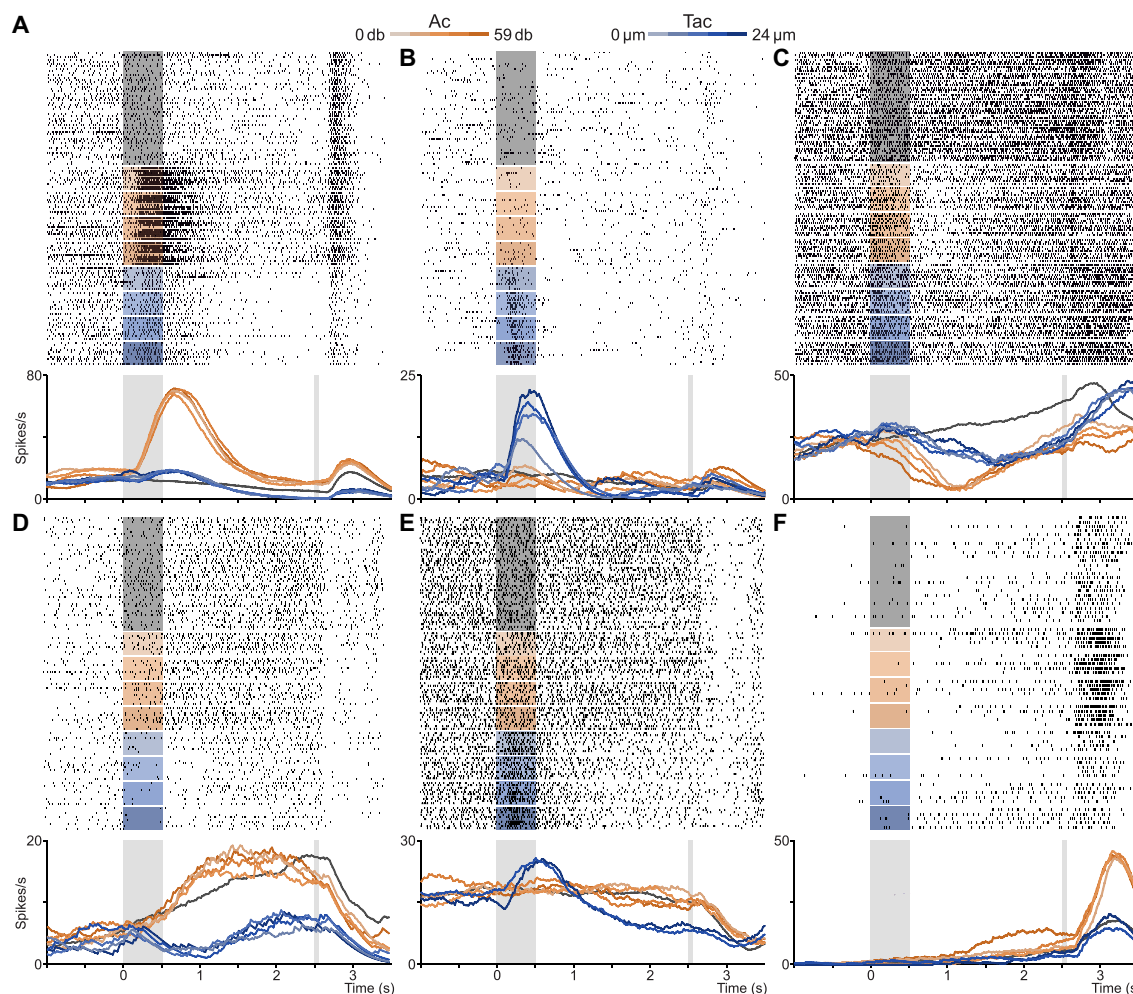


Fig. 2. Individual VPC neuron responses. (A to F) Raster plots and activity profiles of six representative neurons recorded in the VPC. Tactile and acoustic trials are represented by shades of blue and orange, respectively, according to stimulus amplitude, while a deep-gray shade indicates stimulus-absent trials. Black ticks represent neuronal spikes, and light-gray rectangles indicate the stimulation period and PU event. Colored traces depict averaged neuronal activities aligned within trials sharing identical stimulus classes, which enables comparisons across different stimulus intensities and sensory modalities.

selective responses to a single sensory modality, responding exclusively to either tactile or acoustic stimuli. For example, the unit depicted in Fig. 2A demonstrates a strong preference for acoustic stimuli, with its response remaining independent of stimulus amplitude. This characteristic suggests that the neuron encodes the presence of acoustic stimuli rather than variations in intensity or other stimulus attributes. Similarly, the neurons illustrated in Fig. 2 (B and D) and fig. S1B responded selectively to tactile stimuli. Conversely, units such as the one shown in Fig. 2C specifically encoded stimulus absence by exhibiting a reduced firing rate during both tactile and acoustic trials. However, most units displayed a variety of response patterns that varied with stimulus type, as exemplified by the neurons in Fig. 2E and fig. S1G. In addition, some neurons (e.g., those shown in Fig. 2F and fig. S1I) exhibited decision-making-related responses during the movement period following the pu signal. In summary, the VPC exhibits a wide variety of response patterns, with most of the units primarily encoding the presence of either tactile or acoustic stimuli, both modalities simultaneously, or stimulus absence.

Single-unit cognitive dynamics

To obtain a population-level overview of these heterogeneous responses, we first computed the variance across different amplitude levels for tactile (Var_{Tac}) and acoustic (Var_{Ac}) stimuli (Fig. 3A, blue and orange traces, respectively), as well as the variance across the three decision outcomes (Var_{Dec} ; Fig. 3A, purple trace; see Materials and Methods). After subtracting the baseline variance (from $t = -2$ to -1 s), we found that all metrics peaked shortly after stimulus presentation—with a notably larger increase for the tactile condition—and remained stable during the delay until the PU event, after which all metrics rose again, likely reflecting decision execution. The condition-independent variance (Var_{Temp} ; Fig. 3B, green trace) followed a similar temporal profile but was consistently higher than the condition-specific variances, suggesting that most variability captured the sequential task events rather than differences in modality or amplitude conditions. Likewise, the analysis of the stimulus-related information (Fig. 3C) showed that coding began shortly after stimulus presentation, persisted throughout the delay, and increased again after PU, highlighting robust VPC population coding throughout the task.

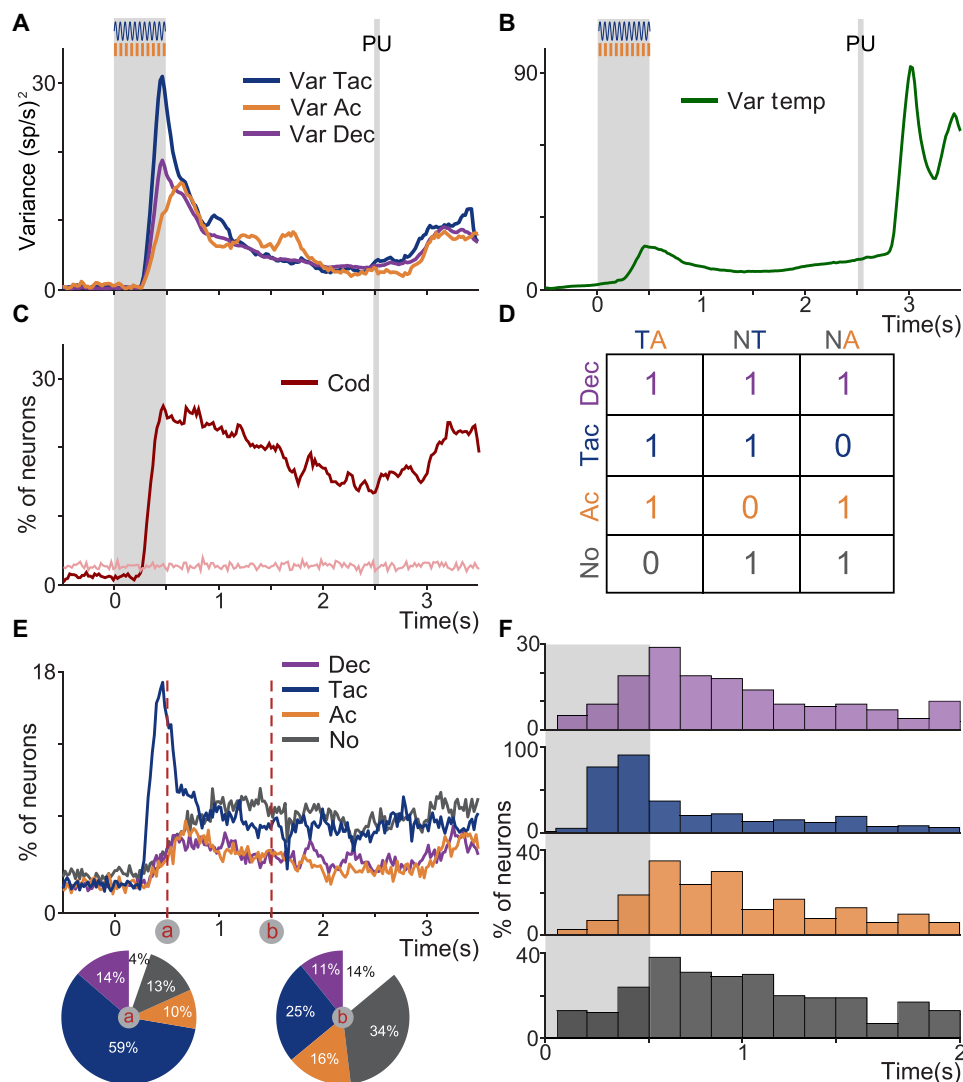


Fig. 3. Population coding dynamics and latencies. (A) Cognitive population variances as a function of time for the average response to each tactile (Var_{Tac} , blue trace) and acoustic (Var_{Ac} , orange trace) stimulus class, as well as for the mean response to each decision type (Var_{Dec} , purple trace). The gray-shaded rectangles denote the periods of stimulation and the PU event. Basal variances (associated with residual fluctuations) were subtracted in all computed variances. (B) Temporal variance (Var_{Temp} , green trace), capturing the sequential events of the task rather than variations across classes. (C) The percentage of neurons conveying significant information about distinct decision outcomes. The comparison evaluated was: tactile (T) versus acoustic (A), versus absent (N). The pink trace shows the chance level ($P = 0.05$, corrected), derived from permutation testing based on baseline activity, indicating the threshold percentage of neurons for statistical significance. (D) Matrix categorizing neurons by their type of encoded information within the indicated comparisons TA, NT, and NA. For each comparison, each neuron received a “1” for significant information encoding or a “0” for nonsignificant. Units were designated as “decisional” (Dec) for scoring 1 in all comparisons, “tactile” (Tac) for 1s in TA and NT, “acoustic” (Ac) for 1s in TA and NA, and “absent” (No) for 1s in NA and NT. (E) Percentage of neurons carrying significant information according to the classification made in (D). Immediately after stimulus onset, most neurons predominantly encode tactile information, as shown in pie chart “a” (final time-bin of the stimulation period). As the delay epoch progresses, encoding becomes more uniform across the four types: Dec, Tac, Ac, and No (“b” pie chart at 1.5 s). The white area indicates neurons with significant coding not classifiable in (D). (F) Coding latency across labeled encoding types. Tactile neurons encoded information faster than others.

To further examine how VPC neurons represented task-related information, we segmented neural responses into 200-ms bins and classified each response into one of four groups—decisional (Dec), tactile (Tac), acoustic (Ac), and stimulus absent (No)—based on significant mutual information comparisons ($P < 0.05$, permutation test). Fig. 3D: TA for tactile-acoustic; NT for absent-tactile; NA for absent-acoustic). For instance, neurons with significant differences in all comparisons were deemed decisional, while those significant only for TA and NT were categorized as tactile, and so on. This method

revealed that tactile encoding dominated and was the first to emerge (Fig. 3E, pie chart “a”), although during the delay period, coding became evenly distributed across categories (Fig. 3E, pie chart “b”). In addition, coding latency analysis (Fig. 3F) showed that tactile neurons responded earlier than neurons encoding other information, underscoring VPC’s integrative capacity. Splitting the data by hemispheres and subject (fig. S2, A to C) confirmed that similar coding signals were present across all groups, even when tactile processing occurred in one hemisphere and motor commands in the other.

We also examined how this neural coding is altered during error trials. Among the six potential errors categories (fig. S3A), only acoustic and tactile omissions occurred frequently enough to allow a robust analysis. For this analysis (fig. S3, B to D), neural coding was assessed starting at 250 ms poststimulus, where the proportion of informative neurons significantly exceeds chance levels. Error trial responses were normalized using z scores across neurons, allowing comparisons across three coding conditions in both correct (hits) and error (misses) trials. In decision-related coding (fig. S3B), correct trials exhibited robust dynamics (aligned with Fig. 3E, purple trace), while coding in error trials was notably reduced. This reduction indicates that neural representations linked explicitly to decision-making are strongly suppressed during errors, irrespective of stimulus modality.

Sensory coding analyses revealed distinct patterns across modalities. Tactile stimulus coding during correct trials (fig. S3C) showed rapid dynamics (like those shown in Fig. 3E, blue trace), largely maintained in error trials, indicating a relative resilience of tactile sensory representations, even during errors. In contrast, auditory stimulus coding during correct trials (fig. S3D) exhibited a delayed onset, consistent with the observations in Fig. 3E (orange trace). During error trials, auditory coding remained stable but notably diminished, highlighting reduced sensitivity to errors compared to tactile coding. This aligns with psychophysical findings, where larger perceptual shifts occurred in auditory tasks for both subjects (Fig. 1D). Last, fig. S3E illustrates the average mutual information of decision-selective neurons per time bin. Consistent with earlier observations, decision coding is effectively absent during error trials as compared to the stable sensory responses, further emphasizing the decoupling between sensory processing and decision-related neural dynamics during incorrect responses.

Overall, these findings indicate that the VPC initially exhibits a strong tactile response and then shifts to a bimodal coding regime during the delay period where neurons encode tactile, acoustic, stimulus absent, and decisional information. This dynamic coding reflects the VPC's flexible role in integrating sensory inputs to guide perceptual decisions.

Population dynamics during BDT

To elucidate the global interactions underlying the diverse VPC responses, we applied dimensionality reduction techniques to reveal the latent dynamics of multisensory integration. We represented the VPC population as a point in an n -dimensional space ($n = 911$ neurons) by combining firing rates from different experiments into a population vector (30, 33–35). Although this space is high-dimensional, the neural activity is assumed to evolve near a low-dimensional neural manifold that provides insight into population-level computation (36–39). Here, we used several methods to identify, visualize, and interpret the structure of the VPC neural manifold during the BDT.

We first used principal components analysis (PCA) to characterize population dynamics over the entire task (Fig. 4A), the sensory (stimulus arrival) phase (Fig. 4B), and the memory (delay) (Fig. 4C). For each phase, covariance matrices were computed (see Materials and Methods), and significant PCs were identified by comparing their variance to that from a noise covariance matrix ($P < 0.01$) (34, 40). Analysis across the entire task revealed that only the first seven PCs were significant (Fig. 4D, top). Projecting activity onto the first six PCs (Fig. 4A) showed that the first PC primarily

captured tactile responses, the second PC mainly reflected acoustic (though partially bimodal) responses, and later PCs were linked to decision-related signals during the delay. Notably, the third PC alone could decode the absence of stimuli, while the fourth distinguished among the three possible decisions.

Restricting the analysis to the sensory period (Fig. 4B) yielded three significant PCs (Fig. 4D, middle). Projections onto these PCs revealed clear divergence of neural trajectories according to presented modality (tactile, acoustic, or absent), with tactile trials diverging more rapidly (fig. S6F). Similar dynamics were observed when using full-task PCs for the stimulation period (fig. S4A), consistent with the higher variance noted during sensory processing (Fig. 3A). Furthermore, these results were preserved when PCA was applied to a covariance matrix with equal weighting of absent, tactile, and auditory classes (Materials and Methods), confirming the robustness of the observed dynamics (fig. S5, A to C).

Focusing on the delay period (Fig. 4C), projection onto the resulting first three significant PCs (Fig. 4D, bottom) revealed that trajectories remained well separated throughout the delay period. These findings indicate that while sensory PCs capture dynamics during stimulation, delay-related axes reflect persistent, decision-related signals. Further analysis of mnemonic signals using established working memory methods (34, 41) (fig. S4B) confirmed that these population signals remained stably segregated during the delay.

Orthogonality between sensory and decision mnemonic dynamics

To further examine VPC population dynamics during stimulation and delay, we applied a generalized linear model (GLM) to the acoustic and tactile responses (Fig. 5). As before, tactile and acoustic trajectories diverged during stimulation, aligning with their respective coding axes (Fig. 5A). Similar results were obtained using sensory coefficients from a multilinear regression (MLR) (42–44). Notably, the sensory axes encoding amplitude from each modality evolved in nearly orthogonal directions (fig. S4C), a finding confirmed by a simplified approach (44, 45) based on the difference between suprathreshold and absent activity (fig. S4D). These results indicate that a simple linear decoder can reliably predict the presence of either acoustic or tactile stimulation from VPC neural activity.

During the delay period, neural trajectories remained distinct, although they converged slightly, likely because of decreasing population variance (Fig. 5B). Neural activity during absent trials, in contrast, remained in a stable state throughout both the stimulus and delay periods, consistent with earlier observations (Fig. 4C and fig. S4B).

Next, we analyzed the similarity of population states throughout the task by calculating the cosine similarity (CS) between population activity vectors at different time points for each stimulus condition (41, 46). We averaged these similarities across all conditions (Fig. 5C), as well as separately for tactile (Fig. 5D, top), acoustic (Fig. 5D, middle), and absent trials (Fig. 5D, bottom). This analysis revealed a clear dissociation between sensory processing and the remainder of the task: The VPC population exhibited a distinct state during sensory arrival, which then stabilized during the decision delay. Sorting trials by modality confirmed that tactile responses emerged faster than acoustic responses (in line with Fig. 3F) and

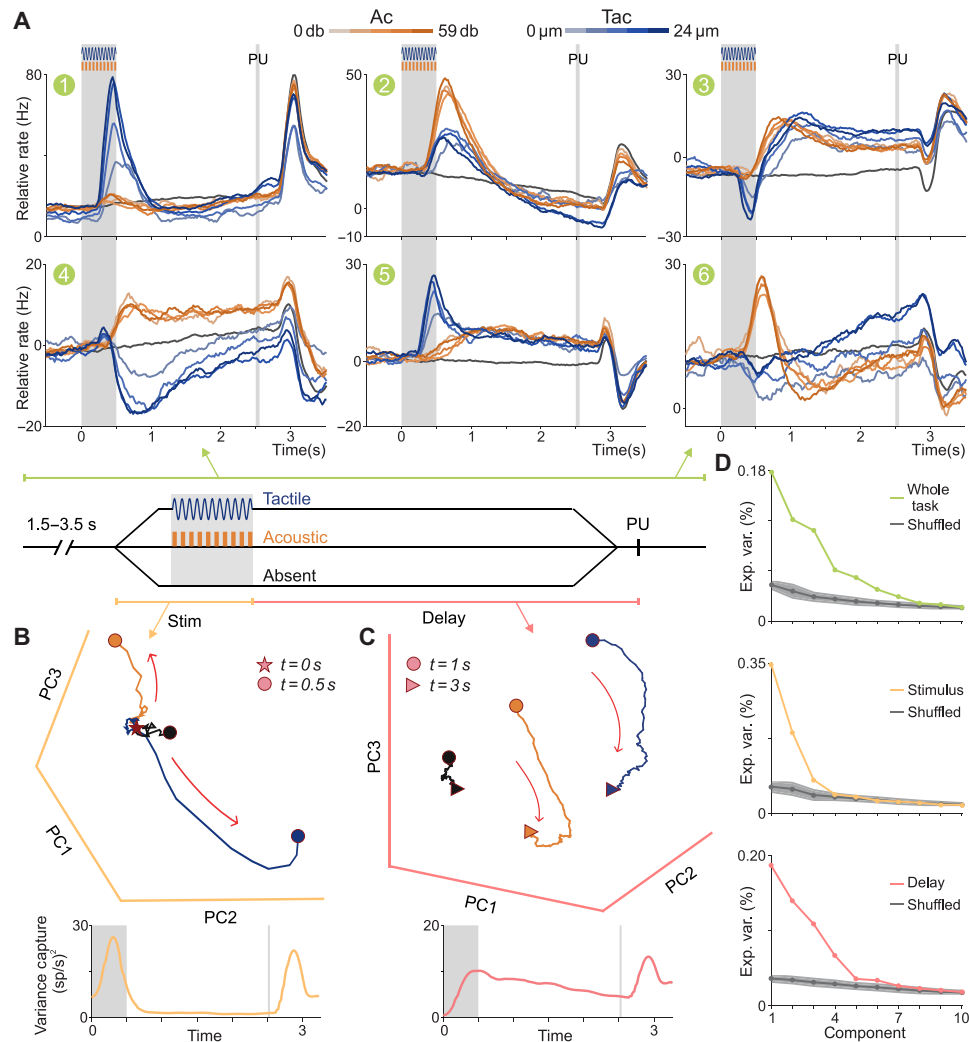


Fig. 4. Population latent dynamics. (A) First six PCs (EV, PC1, 17.8%; PC2, 12.1%; PC3, 10.8%; PC4, 6.1%; PC5, 5.2%; PC6, 3.7%) obtained from PCA of the covariance matrices derived from neural activity throughout the BDT (−1 to 3.5 s, green line on top of BDT schematic). Neural activity sorted by sensory modality and class was projected onto these axes. Activity associated with tactile (acoustic) stimulation is depicted by the traces in different tones of blue (orange), with darker tones indicating stronger stimulation. Activity during stimulus-absent trials is depicted by the dark gray traces. (B) Top, similar to (A) but applying PCA on the covariance matrix during the stimulation period (0 to 0.5 s, yellow trace on BDT scheme). Neural responses associated with the same period and averaged by decision (tactile, acoustic, and stimulus-absent), were projected onto the first three PCs (stars represent $t = 0$ s while filled dots $t = 0.5$ s). Bottom, variance over time captured by the three PCs derived from the stimulation period. (C) Top, similar to (B) but PCs were derived from the covariance matrix during the delay period (0.5 to 1.5 s, pink line in the BDT schematic). Neural activity between 1 and 2.5 s of the BDT, averaged by decision, was then projected onto the first three PCs (filled dots represent $t = 1$ s while filled triangles $t = 2.5$ s). Bottom, variance over time captured by the three PCs derived in the delay period. (D) Explained variance of the first 10 PCs derived in (A) top, (B) middle, and (C) bottom. The dark-red traces, along with the SE in each plot, represent the explained variance of the PCs derived from applying PCA to shuffled covariance matrices, which were used to assess significance.

that the tactile sensory responses closely resemble the network state after the pu signal.

During absent trials, the neural state consistently maintained a high similarity throughout the task despite the presence of some neurons showing ramping activity (Fig. 2, C and D, and fig. S1, A, G, E, and I). This suggests that while ramping modulates firing rate magnitude, it does not alter the overall direction—or “angle”—of the population vector, preserving a stable representation in the absence of sensory input.

To further investigate neural dynamics during absent-stimulus trials, we performed whole-task PCA by equally weighting absent,

tactile, and auditory conditions. This analysis revealed a slow but steady increase in activity during stimulus-free trials (fig. S5B). The trajectory corresponding to the absent-stimulus condition (black trace) departed markedly from stability throughout the task along a fixed direction, consistent with the CS analysis in Fig. 5D (bottom). To isolate this pattern, we computed a covariance matrix using only absent-stimulus trials and extracted the corresponding PCs, identifying a single significant one (fig. S5D). This component exhibited a strong ramping dynamic during the delay period. Notably, when we projected suprathreshold tactile and auditory trials onto this absent-stimulus component, they revealed similar ramping behavior. These

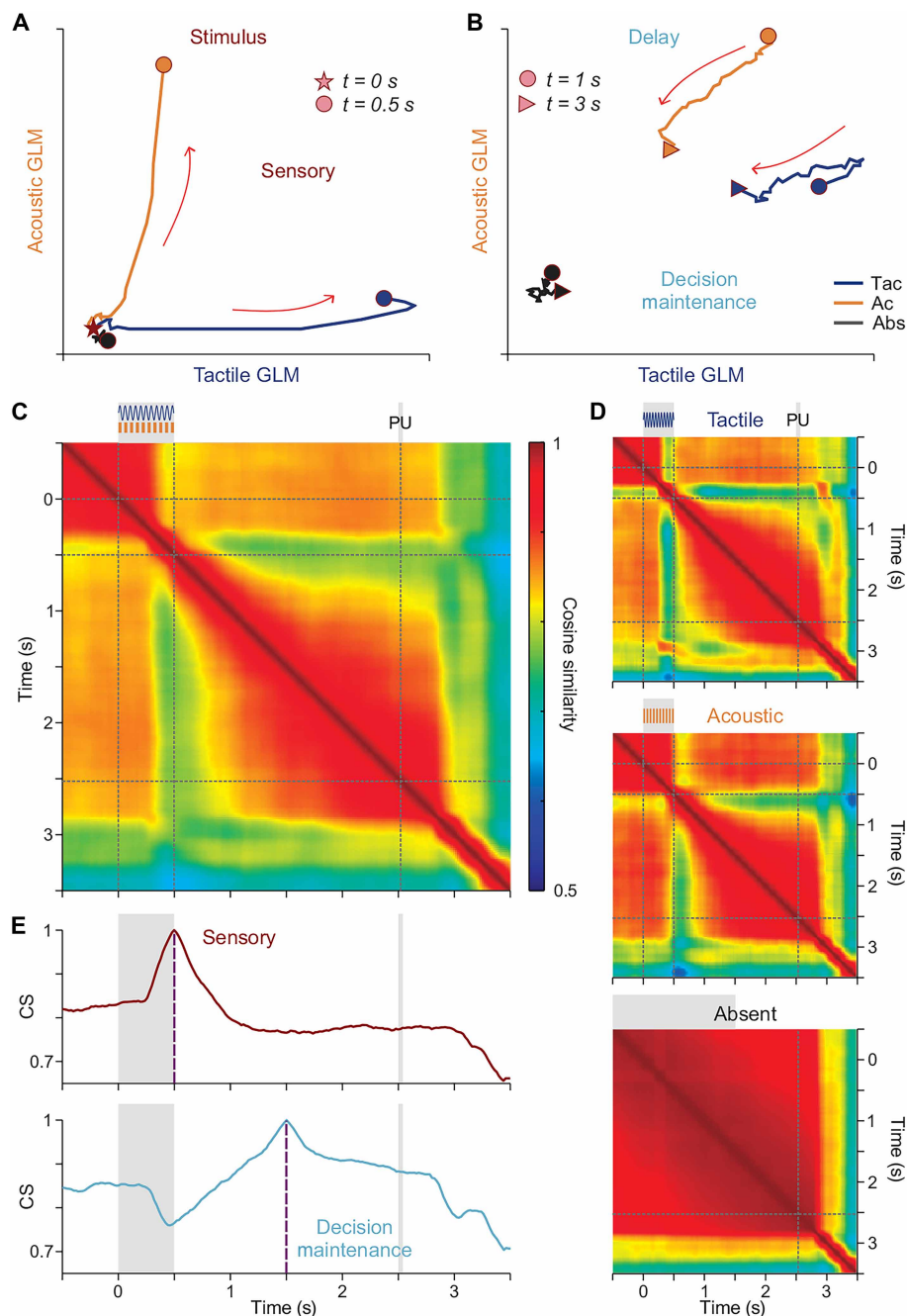


Fig. 5. Bimodal sensory versus mnemonic dynamics. (A and B) Neural activity sorted and averaged by sensory modality was projected over the axes of a GLM computed in the stimulus period (0 to 0.5 s, filled stars and dots mark start and end, respectively) (A), and in the delay period (1 to 2.5 s, filled stars and triangles denote start and end, respectively) (B). Similarly to Fig. 4 (B and C), activity for tactile, acoustic, and stimulus-absent decision are shown as blue, orange, and black traces, respectively. (C and D) CS matrix of the population states obtained by grouping (C) neural data from different modalities or sorted (D) by sensory stimulation (tactile on top, acoustic in the middle) and stimulus-absence (bottom). Dotted lines in each matrix indicate the stimulation period and PU event. Red and blue tones indicate high and low similarity among populational states, respectively. (E) Sensory (top) and decision-maintenance (bottom) CS over time computed using population vectors at $t = 0.5$ s (stimulation epoch) and 1.5 s (delay epoch), respectively.

results point to a shared temporal structure across stimulus modalities, potentially supporting timing and motor-execution (pu) prediction mechanisms.

Last, Fig. 5E presents the CS across all time points with respect to two key task moments: the end of the sensory period ($t = 0.5$ s) and the middle of the memory maintenance delay ($t = 1.5$ s). Consistent

with the sensory-memory decoupling observed in Fig. 4B (bottom), the population state during sensory processing is markedly different from that during the delay, as evidenced by the green (low CS) band (between $t = 0.3$ and $t = 0.7$ s) in the TxT matrix (Fig. 5C) and the separate, persistent red (high CS) rhomboid-shaped region observed during the delay.

Rotational dynamics from stimulation to decision maintenance

To disentangle condition-averaged temporal dynamics from modality-specific responses, we segregated neural activity into modality-dependent and modality-independent components. Prior studies have shown that condition-independent signals account for much of the network's response (33, 40, 47). To achieve this, we marginalized the covariance matrices into modal and amodal parts, and then we applied demixed PCA (dPCA) (33) to extract the latent responses corresponding to each. In Fig. 6A, the first six dPCA components are shown, ordered by explained variance (Fig. 6B). The first three components—capturing most of the variance—were amodal, exhibiting rich dynamics during both stimulation and response periods regardless of the stimulus modality and also displayed ramping during absent trials, consistent with earlier observations.

In contrast, the next two components were modality-specific, capturing the bulk of the cognitive dynamics during stimulus arrival and decision maintenance. These modal components were active during stimulation and the decision delay but not after the PU period. When projecting the population activity onto these two components (Fig. 6C), we observed a clear separation of tactile and acoustic trajectories during stimulation, followed by a rapid rotation at the beginning of the delay and a slow convergence toward a subspace orthogonal to the initial sensory response. These findings suggest that the VPC population exhibits rotational dynamics that may help reduce interference between sensory inputs and decision representations—a mechanism previously observed in rodents during learning to minimize conflicts between sensory and mnemonic signals (29).

Common multistability mechanism across artificial and biological networks

We investigated whether the dynamics observed in the VPC could be compared to those in an RNN trained to solve the BDT (Fig. 7A).

This network architecture has been successfully used to reproduce population dynamics during other cognitive tasks (31, 32, 42, 48, 49) and can switch among many previously learned cognitive tasks (49). Notably, these models exhibited comparable dynamics to biological networks (42, 44, 48, 50, 51). We recreated the BDT in an RNN with three input nodes, each representing a sensory channel of a distinct modality, and one additional node for the go cue, as well as three output nodes corresponding to the response buttons used in the macaques' task. Following standard methodology (31), the network was trained until it reached good performance (Materials and Methods and Fig. 7, B and C).

In concordance with VPC population responses (Fig. 4B), the RNN showed a divergence of population trajectories during stimulation (Fig. 7D, left). Sensory representation was followed by rotational tactile and acoustic responses that remained segregated during the delay period (Fig. 7, D, right, and E), similarly to what was observed in the VPC (Fig. 6C). Again, in line with the biological network response, in absent trials, the RNN population response did not change its angle until the response onset (Fig. 7F). The similar dynamics found in the RNN model and the VPC networks suggest that the mechanisms studied in the VPC allow for general bimodal information integration and bimodal perceptual decision maintenance.

A general dynamical mechanism for the transition from sensory to mnemonic representations

To propose a general mechanism underlying the results obtained with biological and artificial networks, we derived a low-dimensional dynamical system model that captures the separation between sensory and maintenance axes (Fig. 8A). The proposed dynamical system (Eqs. 21 and 22 and Fig. 8B) includes two saddle points whose stable manifolds act as sensory thresholds for detecting each sensory modality. When the stimulus intensity exceeds a certain threshold, the acoustic (orange) and tactile (blue) responses cross

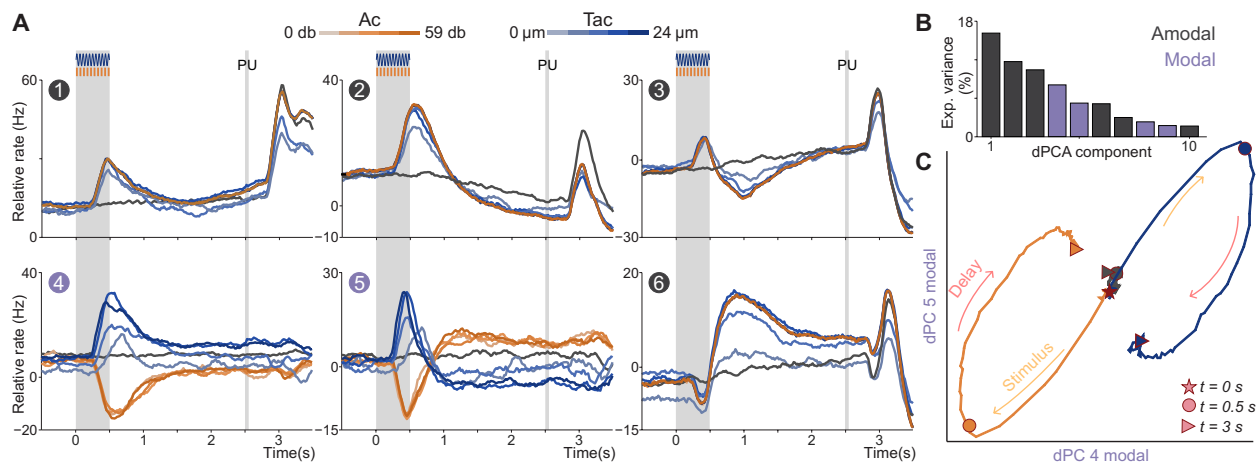


Fig. 6. Neural states and bimodal rotational dynamics. (A) dPCA applied to marginalized covariance matrices from the whole BDT (0 to 3.5 s) with respect to class identity (tactile and acoustic) to derive the most prominent dPCs. Population activity for the entire BDT (−1 to 3.5 s, $n = 911$), sorted by class averages within each sensory modality (tactile and acoustic, tones of blue and orange traces, respectively) or absence of stimulation (absent, dark-gray trace) projected onto the first six dPCs ordered by their EV (dPC1, 16.1%; dPC2, 11.6%; dPC3, 10.3%; dPC4, 8.05%; dPC5, 5.2%; dPC6, 5.1%). Projection onto most dPCs showed amodal coding behavior, while projections onto dPC4 and dPC5 displayed modal dynamics with clear tactile, acoustic, and stimulus-absent encoding and maintenance. Gray-shaded rectangles represent the stimulation period and PU. (B) EV of the first 10 dPCs with modal and amodal dPCs represented by purple and gray rectangles, respectively. (C) Neural activity across the whole task (0 to 3.5 s) sorted by modality [tactile (blue), acoustic (orange), and stimulus-absent (black)] projected over the first two modal dPCs (dPC4 and dPC5) and plotted in a 2D phase diagram. Filled stars, dots, and triangles mark moments $t = 0$, 0.5 , and 2.5 s, respectively.

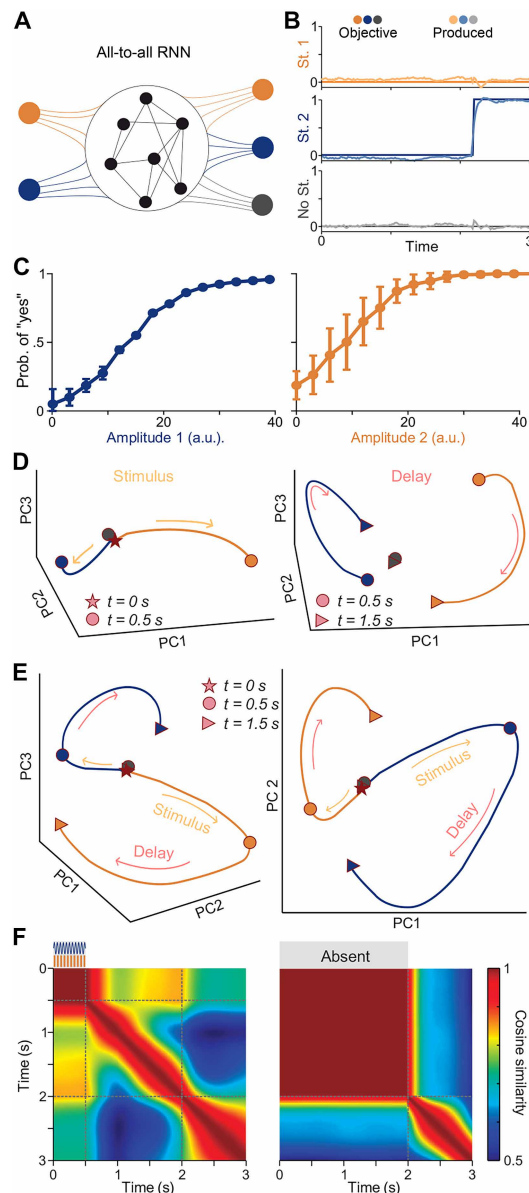


Fig. 7. Bimodal detection on an RNN. (A) Schematic of an all-to-all RNN trained for BDT. The network receives input stimuli—St. 1 (blue) and St. 2 (orange), shown as the left nodes—and generates corresponding outputs: tactile, acoustic, or absent (blue, orange, and gray right-side nodes). (B) Example supraliminal trial responses after training for St. 1 (tactile, blue), St. 2 (acoustic, orange), and No St. (absent, gray). Saturated and unsaturated traces correspond to target and produced outputs, respectively. After stimulus presentation, the RNN maintains quiescence until the PU (go-cue), when the correct output channel is expected to activate (shown: St. 2 trial). (C) Posttraining “psychometric” curves for each stimulus type, trained over a limited amplitude range. Values are displayed in arbitrary units (a.u.). (D) PCA applied to covariance matrices from stimulus (0 to 0.5 s) and delay (0.5 to 2 s) epochs to extract prominent PCs. Simulated neural activity from these epochs, sorted by modality [tactile (blue), acoustic (orange), and absent (black)], is projected onto the first three PCs, revealing strong similarity to biological data. (E) PCA of the full BDT simulation (0 to 2 s, excluding response period). Dynamics diverge by sensory input (3D plot, left), resembling modal axes dPC4 and dPC5 (2D plot, right). (F) CS matrices from all trial types (left) or only stimulus-absent trials (right). The patterns resemble those from biological neural activity. Dotted lines indicate stimulus period and PU event.

the respective sensory threshold (green line) and subsequently relax toward stable equilibrium points associated with memory retention for the two sensory modalities. However, if the stimulus is too weak, the response fails to reach the threshold and returns to the stable equilibrium associated with the absence of sensory inputs (Fig. 8C). See also fig. S7 for the model’s response to increasingly stronger tactile and acoustic stimuli.

The reduced model in Eqs. 21 and 22 provides a system-level description of the dynamical interaction between sensory and mnemonic representations. According to this model, sensory representation dynamics (Eq. 21) act as a simple linear low-pass filter on sensory inputs. The output of these sensory dynamics feeds into the mnemonic representation dynamics (Eq. 22), which are nonlinear and exhibit multistability among attractors corresponding to distinct perceptual decisions. The rotational geometry observed during the transition from sensory to mnemonic representation (Fig. 6C and Fig. 7E) is explained by the phase-plane geometry of the sensory-mnemonic dynamical system described by Eqs. 21 and 22: a cascade of a vanishing-memory low-pass filter for transient sensory input representation, followed by a nonlinear multistable system for persistent memory retention of relevant sensory input features.

Because of the universality of low-pass filtering and multistable dynamics, the proposed model offers a general dynamical mechanism likely shared across sensory modalities and animal model tasks. This mechanism elucidates the transition from sensory to mnemonic representations within neural populations, contributing to a deeper understanding of how sensory information is retained and transformed into memory across different contexts.

DISCUSSION

Our investigation of the BDT revealed that the VPC is a natural candidate for the bimodal integration of tactile and acoustic sensory inputs. Neurons in this region exhibited diverse response patterns, encoding stimulus presence, absence, and modality. Tactile responses emerged first, but the VPC consistently represented and maintained information of both sensory modalities, underscoring its role in integrating multimodal inputs for decision-making. These findings, replicated across subjects and hemispheres, highlight the VPC’s contribution to sensory differentiation and cognitive processes.

Population-level analyses distinguished sensory responses from decision maintenance, revealing low-dimensional neural trajectories that initially diverged in response to sensory stimuli and then rotated into a subspace orthogonal to the sensory one. This same behavior was replicated in an RNN trained on the BDT, supporting the hypothesis of a shared computational mechanism. Using a simplified two-dimensional (2D) model, we demonstrated that when sensory input was sufficiently strong, the system crossed a separatrix and then stabilized in an attractor state that maintained the decision. This suggests that the VPC is not only involved in sensory encoding but also in sustaining decisions through distinct neural mechanisms. Further evidence for this dissociation comes from analyses comparing hit versus error trials. These analyses consistently showed that the neural signals associated with hits and errors differentiated more strongly during the decision retention period instead of during the initial sensory processing. This observation suggests the existence and importance of a marked temporal separation between distinct cognitive processes linked to correct performance.

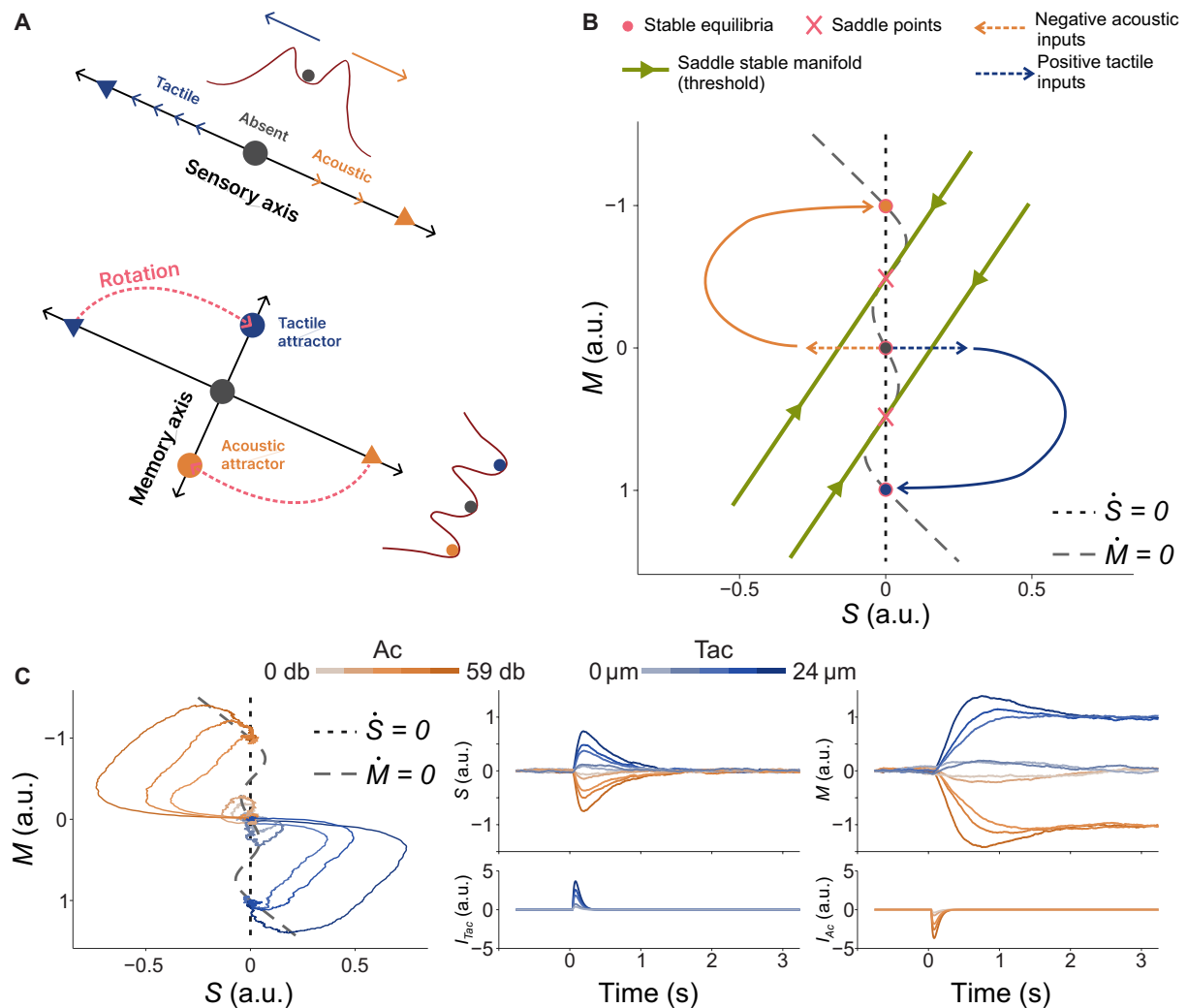


Fig. 8. Attractor-based model of bimodal detection. (A) Sketch of a low-dimensional dynamical model performing bimodal detection. When receiving a sufficiently strong stimulus from one of the two modalities, the dynamics integrate it along the sensory axis. After the stimulation ends, the state converges with a rotational motion toward a memory subspace that preserves the memory of the stimulus's presence and modality in distinct attractors. (B) Sketch of the phase plane of the suggested bimodal detection dynamics. The fixed points of the system are represented as filled pink circles for stable equilibria and crosses for saddle points. The S and M nullclines are depicted with dotted and dashed gray lines, respectively. Saddle stable manifolds are sketched in green. (C) Example runs for three suprathreshold stimuli and two subthreshold stimuli of distinct amplitudes. Left, system's phase plane. Middle, sensory integration variable over time. Right, memory retention variable over time. Bottom, sensory input used for tactile and acoustic runs. The proposed 2D model qualitatively reproduces and explains the dynamics shown in Figs. 6C and 7E.

The separation between sensory and maintenance dynamics is crucial because it allows the brain to retain information using a different coding scheme than the one used during sensory reception. This strategy enables continuous processing and attention to incoming stimuli without disrupting the maintenance of previously encoded information (29) or impairing communication with other brain areas (52). Although the separation of sensory and decision responses is a primary organizational principle, projecting the population activity onto axes optimized to uncover temporal dynamics revealed shared ramping behaviors across all conditions, including absent trials. Our work thus reveals a bimodal perceptual mechanism in the VPC that provides unique insight into how two competing sensory modalities can be integrated.

Furthermore, it is important to consider how sensory integration in the VPC might change under constructive interference rather

than competition. Constructive interference could enhance the fidelity and robustness of encoded information, potentially leading to more accurate and faster decision-making processes (12, 53–55). Such interference might allow the VPC to combine inputs from different sensory modalities in a manner that strengthens the overall signal, facilitating clearer perceptual decisions. Understanding whether and how VPC leverages constructive interference will be essential for clarifying its role in multimodal sensory processing.

Several questions remain. For example, why does tactile encoding emerge faster than acoustic encoding? One possibility relates to the limitations of our experimental design, which may inherently favor faster neural responses to tactile stimuli. Specifically, the acoustic stimuli used here might not fully exploit the capacity of VPC neurons to encode complex auditory features such as spatial location (25). Thus, our task parameters—for instance, the acoustic

stimulus type or presentation method—could inadvertently lead to slower or less robust acoustic responses compared to tactile ones. Unlike primary sensory areas [S1 (17) and A1 (13)] the VPC exhibits bimodal encoding, raising the possibility that it might be the first cortical area to feature neurons with distinct modality-specific responses (2, 19). While not conclusive, the VPC's rapid and robust responses make it a strong candidate for this role.

Recent research has also implicated the VPC in abstract language processing. Traditionally, Broca's area has been considered the hub for language, but emerging evidence suggests that the VPC contributes notably to language-related representations (56). Language is inherently multimodal—integrating acoustic, visual, and tactile information (57)—and our findings support the notion that the VPC plays a vital role for encoding and maintaining such multimodal sensory information, a capacity that could be critical for higher-order cognitive functions like language.

The role of oscillatory activity in bimodal integration also remains an open question. It is possible that the bimodal coding we observed is reflected in LFP frequency bands, with the beta band being a strong candidate (58, 59). Although the specific role of beta oscillations in the VPC is not fully understood, their established role in other brain regions suggests that they could facilitate communication between different sensory areas and the VPC through spikes synchronization (60). Such synchronization may regulate sensory attention and enhance the integration of task-relevant inputs.

Attention likely regulates neural activity within the VPC. One plausible hypothesis is that VPC responses are attentional—that is, when a sensory modality is behaviorally irrelevant, the corresponding neural activity will reduce or cease (2, 14). Future experiments involving manipulations of attentional demands can clarify how VPC neurons dynamically modify their responses according to task context (42, 61) and advance our understanding of how multisensory integration and attention interact. A recent experiment in our laboratory demonstrated that VPC networks have a remarkable capacity to adjust to temporal demands, showing context-dependent specialization of interval processing (62). Considering this, the tactile-auditory asymmetries observed in our experiment could be specific properties of the stimuli used and not a bias of VPC processing per se. In line with findings in humans, it is plausible that VPC selectively represents sounds with social, motoric, or ethologically relevant characteristics—characteristics not explicitly addressed by our task design (56). Examination of how VPC representations shift across more naturalistic or socially meaningful auditory contexts will be a major focus of future research.

In summary, when compared with previous studies (2, 19, 27, 28, 62), our findings demonstrate that the VPC plays a key role in segregating sensory inputs and sustaining multimodal perceptual alternatives. The pronounced segregation of sensory and maintenance responses within a low-dimensional neural space—in both biological and artificial systems—illustrates the versatility of VPC dynamics. This research not only shows that a mechanism previously thought to be restricted to unimodal contexts is effective in bimodal conditions but also paves the way for further investigation into how the brain processes and retains multisensory information.

MATERIALS AND METHODS

Bimodal detection task

Two monkeys (*M. mulatta*) were trained to detect the presence or absence of either a vibrotactile stimulus or acoustic stimulus

(Fig. 1A) (17). Vibrotactile stimuli ranged from 0 to 24 μm , while acoustic stimuli ranged from 0 to 59 dB, covering subthreshold, threshold, and suprathreshold detection levels for both modalities. The vibrotactile stimuli were delivered via a computer-controlled stimulator with a 2-mm round tip (BME Systems) applied to the glabrous skin of the distal segment of one fingertip on the monkey's restrained right hand, while the acoustic stimuli were presented through a speaker. Each task's trial began when the stimulator probe (PD) lowered down and indent the monkey's skin surface on one finger from its right restrained hand. This event signaled the start of each single trial. The monkey immediately responded by grasping an immovable key with its left free hand (KD). After a variable prestimulus delay (indicated in Fig. 1A), the monkey was presented with one single stimulus randomly selected from these three conditions: tactile (Tac) stimuli, acoustic stimuli (Ac), or absent (Abs) stimuli. After a 2-s delay, the stimulator was lifted off from the fingertip (PU), signaling the monkey to release the lever (KU) and press a button (PB) to indicate the perceived stimulus modality or its absence. Both types of stimuli consisted of sinusoidal signals at a frequency of 20 Hz. For the acoustic modality, each signal was a pure tone at 1 kHz, while for the vibrotactile modality, the stimuli were trains of mechanical-sinusoidal pulses. The monkeys indicated their detection by pressing one of three buttons: "vibrotactile stimulus present," "acoustic stimulus present," or "stimulus absent." Correct responses were rewarded with a few drops of fruit juice. Performance was quantified using psychometric techniques (Fig. 1E). Both animals were handled in accordance with the institutional standards of the US National Institutes of Health and the Society for Neuroscience. All protocols were approved by the Institutional Animal Care and Use Committee of the Instituto de Fisiología Celular at the National Autonomous University of Mexico.

Recording sites and datasets

Neuronal recordings were obtained with an array of seven independent, movable microelectrodes (2 to 3 megohm) inserted into VPC (Fig. 1B) (20–22, 27). In monkey RR32, recordings were made ipsilaterally (right hemisphere) and contralaterally (left hemisphere) to the stimulated hand, while in monkey RR33, recordings were made only ipsilaterally to the stimulated hand (right hemisphere). The electrode locations were confirmed through standard histological techniques. In each recording session (1 session = 1 day of recordings), multiple sets of trials were presented to the animals. Each set consisted of 150 trials randomly interleaved across tactile (Tac), acoustic (Ac), and absent (Abs) stimulus conditions. The sets were balanced, presenting exactly 50 trials for each of these three conditions. Recording sites (location of the electrodes in the same brain region) varied across sessions, completing approximately 30 sessions (30 days) per brain area. In the VPC, a total of 42 recording sessions were conducted for monkey RR32, and 27 for RR33. The electrode penetration locations were mapped onto the cortical surface by marking the edges of the small chamber (7 mm in diameter) placed over each cortical area. Across the sessions, we recorded 566 neurons in monkey RR32 (256 from the right hemisphere and 310 from the left hemisphere) and 345 neurons in RR33. Neurons displaying a preference for encoding one of the modalities (tactile or acoustic) were distributed throughout the VPC, making the identification of subgroup responses highly unlikely (see Fig. 3E and fig. S2, A to C, top and middle).

Data analysis

Except for the Mistake Analysis section, all analyses and results presented in this study were conducted exclusively on hit trials, and subliminal classes were excluded when plotting the class averages, unless specified. This approach ensures that the data reflect accurate sensory detection and response, focusing solely on trials where the monkeys successfully identified the presence of stimuli.

Firing rate

For each neuron, we calculated a time-dependent firing rate for each trial using overlapping rectangular causal windows with a length of 200 ms and steps of 20 ms. For visualization purposes only, as shown in Fig. 2 and fig. S1, we applied a causal Gaussian convolution with a σ of 200 ms. This smoothing technique allows for a clearer representation of firing rate dynamics over time, without distorting the temporal relationships in the neuronal activity.

Instantaneous coding variance across the population

To evaluate the influence of each type of sensory stimulus coding (tactile or acoustic) and decision outcome, we calculated the instantaneous variance in a time-dependent manner using the computed firing rate of each neuron. At each time bin, the population's instantaneous tactile variance (Var_{Tac} ; Fig. 3A and fig. S2, A to C, middle, blue trace in all plots) was calculated as the quadratic sum of the firing rate fluctuations across different tactile stimulus identities (classes) and neurons

$$Var_{Tac}(t) = \frac{1}{N} \frac{1}{6} \sum_{i=1}^N \sum_{Tac=1}^{c=6} [r^i(t, Tac) - r^i(t)]^2 \quad (1)$$

This metric captures the variability in neural responses specifically linked to the tactile stimuli at any given moment in time, providing insight into how the population dynamically encodes this sensory information.

Analogously, the population's instantaneous acoustic variance (Var_{Ac} ; Fig. 3A and fig. S2, A to C, middle, orange trace in all plots) was calculated to measure the fluctuations in firing rate across acoustic stimulus classes and neurons

$$Var_{Ac}(t) = \frac{1}{N} \frac{1}{6} \sum_{i=1}^N \sum_{Acu=1}^6 [r^i(t, Ac) - r^i(t)]^2 \quad (2)$$

Next, the population's instantaneous decision variance (Var_{Dec} ; Fig. 3A and fig. S2, A to C, middle, purple trace in all plots) was calculated to measure the fluctuations in firing rate across different decision identities and neurons

$$Var_{Dec}(t) = \frac{1}{N} \frac{1}{3} \sum_{i=1}^N \sum_{Dec=1}^3 [r^i(t, Dec) - r^i(t)]^2 \quad (3)$$

This metric quantifies the variability in neural responses associated with the different decision outcomes over time.

The tactile and acoustic variances were obtained by using only the means of the stimulus classes within their respective modalities. The decision variance, on the other hand, was calculated by averaging across all classes for each decision type (tactile, acoustic, and no stimulus). The overall class variance, which reflects variability across the entire range of stimuli, was computed using all classes, including all amplitudes for acoustic, tactile, and no stimulus trials.

The values of Var_{Tac} , Var_{Ac} , and Var_{Dec} , computed from the activity within the time window $t = -1$ to 0 s [before the arrival (or not) of stimulation], represent the inherent residual noise in the firing rate estimates (~ 2 spikes/s). To interpret these values as a measure of population coding, Var_{Tac} , Var_{Ac} , and Var_{Dec} must exceed this

resting state variance (basal variance). In our analysis, to ensure meaningful comparisons across parameters, the basal variance was subtracted from each type of variance computed. This adjustment allows us to more accurately assess the degree of coding attributable to sensory stimuli and decision processes, beyond the inherent fluctuations.

Instantaneous temporal variance across the population

The population instantaneous temporal variance (Var_{Temp} ; Fig. 3B and fig. S2, A to C, bottom, green trace in all plots) was computed at each time point as compared to the mean firing rate. This variance is a reflection of the time variation with respect to the baseline period. In particular, it is defined as the quadratic sum of differences between the average firing rate at each time point [$r^i(t)$] and the average baseline firing rate r_{bas}^i (from -1 to 0 s) across neurons

$$Var_{Temp}(t) = \frac{1}{N} \sum_{i=1}^N [r^i(t) - r_{bas}^i]^2 \quad (4)$$

This metric gives an account of the dynamics in the population firing rate over time relative to its baseline activity, emphasizing the temporal dynamics in the neural population throughout the task.

Mutual information

For each neuron, we calculated four types of mutual information for each time window using Shannon's information theory (63, 64): information modality coding (I_{COD} , depicted as Cod in Fig. 3C) and information derived from the combinations of possible decisions: tactile-acoustic (TA), tactile-no stimulus (NT), and acoustic-no stimulus (NA). To compute I_{COD} , we considered the firing rate probability distributions associated with the different stimulus classes within each modality [$P(r|s)$] as well as the overall firing rate probability distribution [$P(r)$], which combines all stimulus classes and modalities. For each time bin, we generated three conditional distributions $P(r|s)$: one accounted for each one of the tactile classes [$P(r|Tac)$], one for each of the acoustic classes [$P(r|Ac)$], and one for stimulus absent class [$P(r|Abs)$]. In addition, we computed the unconditioned probability distribution [$P(r)$], which reflects the general likelihood of any neural response across all classes and modalities. The coding information was then quantified using mutual information

$$I_{COD}(t) = \sum_{r,s} P(r|s) P(s) \log_2 \left(\frac{P(r|s)}{P(r)} \right) \quad (5)$$

This measures the amount of information the neural response r provides about the stimulus s .

To compute the information metric for the combinations I_{TA} , I_{NT} , and I_{NA} , we followed the same procedure as for I_{COD} (Eq. 5) but used the corresponding firing rate probability distributions $P(r|s)$ based on the trials of stimuli classes and modalities involved in each pair of combinations. For each time bin, we generated three distributions (same as I_{COD}): tactile and acoustic distributions for I_{TA} , tactile and absent distributions for I_{NT} , and acoustic and absent distributions for I_{NA} (17). In all cases, the unconditioned probability distribution $P(r)$ was computed (in each time) by combining the relevant trials from the stimuli classes and modalities. Neurons were then categorized (as shown in the chart in Fig. 3D) based on whether they showed significant information for at least two of the three possible information metrics (I_{TA} , I_{NT} , and I_{NA}). They were classified as decisional (Dec) if they scored significantly in all combinations (I_{TA} , I_{NT} , and I_{NA}), tactile (Tac) if they scored in I_{TA} and I_{NT} , acoustic

(Ac) if they scored in I_{TA} and I_{NA} , and, finally, absent (No) if they carried relevant information for the I_{NT} and I_{NA} pair. The sorted neuron subgroups were plotted as the percentage of neurons over time (represented by traces and pie charts in Fig. 3E). The pie chart considered the total number of significant neurons. Neurons that showed significant information for only one of the possible combinations were not classified into a specific group but were included in the total neuron count. This subgroup is shown as the white portions of the pie charts (Fig. 3E).

For all information metrics computed, significance was calculated per neuron in each time bin. A permuted value of mutual information was obtained by performing 1000 permutations of the trials. A mutual information value was considered significant if the probability of a permutation yielding the same or a greater value was below 0.05 ($P < 0.05$). The percentage of neurons with significant mutual information was then corrected using a multiple comparison test to account for the number of neurons and time bins analyzed. In addition, to address the issue of finite sampling, we applied the correction proposed in (54). For the significance test, we also applied a correction for multiple comparisons using a clustering method as described in (65), retaining only the sets of significance-connected time bins whose size exceeded a defined threshold.

Coding latency

This metric denotes the time at which the encoded signals first become significant (2). We calculated coding latencies for the four subgroups of categorized neurons (Dec, Tac, Ac, and No). For each subgroup, latency was defined as the first-time window (after stimulus presentation) that was followed by three consecutive time windows with a significant P value. To determine the number of consecutive bins (three in this case) required to establish significance, we applied a significance-by-size approach, ensuring that the cluster size was unlikely to occur by chance ($P < 0.05$).

Mistake analyses

In fig. S3A, trial outcomes (hits, misses, false alarms, misunderstood, and correct rejections) are represented by their frequency across conditions, providing a probability landscape for each outcome. Threshold and suprathreshold trials were analyzed to assess the mutual information about decision outcome (fig. S3B), tactile (fig. S3C), and acoustic (fig. S3D) modalities, for hit and miss trials corresponding to neurons with relevant coding in each time bin.

To quantify neural responses during these trials, we calculated a z score for each neuron (n) at each time bin (t), relative to its basal activity (bas)

$$z_n(t) = \frac{r_n(t) - \hat{r}_{n,bas}}{\sigma_{n,b}} \quad (6)$$

We defined the sign of coding ($\eta_{n,tac}$, $\eta_{n,acu}$) at each time bin as the difference between the mean neuronal response in tactile or acoustic trials and the mean response in no-stimulus (absent) trials. Using this sign, we computed the signed z score

$$\bar{z}_{n,tac}(t) = \eta_{n,tac}(t) z_n(t) \quad (7)$$

From this calculation, we derived the z -score distributions across all coding neurons, averaged across trials, separately for correct (hits) and incorrect (pooling misses, FA and MU) responses (fig. S3E).

Population states similarity

To analyze the similarity of population states across different times of the task (t_1, t_2), within each modality (tactile or acoustic; Fig. 5D, top and middle) or for the absence of stimulation (stimulus absent; Fig. 5D, bottom), we computed a population vector for each time window of the BDT. The vector takes the form

$$\mathbf{r}_c(t) = [\mathbf{r}_c^1(t), \mathbf{r}_c^2(t), \dots, \mathbf{r}_c^N(t)] \quad (8)$$

where each element of this vector corresponds to the mean firing rate of all trials for each neuron i under condition c , at the specific time t [$\mathbf{r}_c^i(t)$]. This vector, $\mathbf{r}_c(t)$, captures the population activity patterns across neurons for a given condition at different time points. In addition, to compare the population states across conditions (modalities and amplitudes) (Fig. 5C), a total population vector was calculated by concatenating the vectors for each condition c at each time window t

$$\mathbf{r}(t) = [\mathbf{r}^1(t), \mathbf{r}^2(t), \dots, \mathbf{r}^4(t)] \quad (9)$$

This represents the combined neural activity from all conditions and neurons, allowing for direct comparison of population states across different stimulus classes and modalities at various time points. The comparison between population states was assessed by computing the CS metric (34, 41, 46). This measure calculates the cosine of the angle between two vectors, without considering their magnitudes. The CS is 1 for two parallel vectors and 0 for orthogonal vectors (90° difference). CS can be mathematically represented as

$$CS(t_i, t_j) = \frac{\mathbf{r}(t_i) \cdot \mathbf{r}(t_j)}{|\mathbf{r}(t_i)| |\mathbf{r}(t_j)|} \quad (10)$$

where $\mathbf{r}(t_i) \cdot \mathbf{r}(t_j)$ is the dot product of two vectors at times t_i and t_j , respectively, and $|\mathbf{r}(t_i)|$ and $|\mathbf{r}(t_j)|$ are the magnitudes of the vectors at those same time points. Therefore, $CS(t_i, t_j)$ defines a matrix of dimension $T \times T$, which quantifies the similarity of population states at two different times. In this work, we quantified these matrices by considering all conditions and modalities (Fig. 5C) or by separating the trials according to their modality (tactile, acoustic, or absent; Fig. 5D). Analogous analyses were applied to the artificial networks in Fig. 7.

Principal components analysis

The main aim of PCA is to find an alternative coordinate system in which the data can be represented more succinctly and compactly. In other words, the idea is to define a low-dimensional subspace that captures most of the variance of the high-dimensional neural state space (30, 40). Typically, the number of significant dimensions ($P < 0.05$) is reduced substantially, going from as many possible dimensions as there are neurons (N) to just a few axes that concentrate most of the variance (38). To characterize how the population activity covaries across classes over time, we performed PCA (Fig. 4 and figs. S4, A and B, and S5) using liminal (2) and supraliminal (2) classes from the two stimulus modalities (tactile and acoustic), along with the stimulus-absent condition (1 class). We compute covariance across classes and time: from -1 to 3.5 s for the entire BDT in Fig. 4A and figs. S4A and S5 (A and B, green line on task schematic); for the stimulation period (0 to 0.5 s, yellow line) in Fig. 4B; and for the delay period (0.5 to 2.5 s, pink line) in Fig. 4C and fig. S4B.

PCA yields an alternative coordinate system for the N -dimensional data, where the first coordinate accounts for the most variance in

the neural population. The second coordinate accounts for as much of the remaining variance as possible, and so on, with each subsequent axis constrained to be orthogonal to the previous ones. PCA was computed from the firing rate covariance matrix, averaging over time bins t and classes c

$$C_{ij} = \frac{1}{T} \frac{1}{C} \sum_{t=0}^T \sum_{c=1}^C [r_c^i(t) - \langle r^i \rangle] [r_c^j(t) - \langle r^j \rangle] \quad (11)$$

where C_{ij} represents the covariance between neuron i and neuron j , T is the number of time windows in the period being considered, C is the number of conditions (nine in our task), $r_c^i(t)$ is the firing rate of neuron i at time t for class c , and $\langle r^i \rangle$ is the mean firing rate of neuron i across all time bins and conditions.

For analysis shown in fig. S5, PCA was applied to covariance matrices calculated by weighing each stimulus class according to its probability of appearance (Eq. 12)

$$C_{ij} = \frac{1}{T} \sum_{t=0}^T \sum_{c=1}^C \frac{1}{P(C)} [r_c^i(t) - \langle r^i \rangle] [r_c^j(t) - \langle r^j \rangle] \quad (12)$$

The significant PCs were determined by comparison against noise-corrected covariance matrices derived from shuffled data (Fig. 4D and fig. S5C, see Noise and Dimensionality in Materials and Methods). In addition, for isolating absent-stimulus dynamics (fig. S5D), covariance matrices were calculated exclusively from absent-stimulus trials (Eq. 13)

$$C_{ij} = \frac{1}{T} \sum_{t=0}^T [r_{Abs}^i(t) - \langle r^i \rangle_{Abs}] [r_{Abs}^j(t) - \langle r^j \rangle_{Abs}] \quad (13)$$

The diagonalization of the covariance matrix, $C = UDU^T$, produces an alternative coordinate system given by the columns of matrix U , which we call the derived axes or PCs. On the other hand, D is a diagonal matrix of positive values. The diagonal elements of D represent the amount of variance in the population activity captured by the corresponding PCs. We then order the PCs according to the amount of variance captured. The projection of the N -dimensional data onto the k th PC is given by

$$PC_k(t, c) = \sum_{i=1}^N U_k^i r_c^i(t) \quad (14)$$

where U_k^i is the i element of the k PC (U_k). Therefore, the PCs are linear readouts of the population activity; in other words, they are orthogonal linear combinations of the firing rates of the individual neurons, with no correlation between components (fig. S6, A to D). Thus, the contribution of each neuron to a given PC_k is given by the i th element of U_k . These PCs can be thought of as a low-dimensional description of the population activity in this coding subspace.

Mnemonic PCs shown in fig. S4 were computed by applying PCA to covariance matrices obtained from time-averaged neural population vectors during the delay period (0.5 to 2.5 s of BDT). For each experimental condition c , the elements for each neuron were averaged over the different time bins i , effectively collapsing the temporal dimension (34, 41).

Speed of population dynamics

To calculate the speed of the population dynamics (fig. S6F), we used the embedding space of the first significant PCs. After projecting the trajectories of each class to this space, we measured their evolution at each time step with the Euclidean distance, obtaining their speed as a function of time.

Noise and dimensionality

To isolate the PCs that capture significant fluctuations in neural activity, we first determined a “noisy” firing rate for each neuron (34, 38, 40). This was achieved by calculating the firing rate of each neuron at time t during a trial within the condition c and subtracting the average firing rate of all trials at the same time point t for that condition c . This process removed the signal component relevant to the neuron’s encoding function across conditions c and time, leaving only the residual fluctuations. Then, for each neuron, we randomly selected 10 trials to compute the covariance with every other neuron. This random selection was repeated in each iteration to ensure a robust estimation of each neuron’s contribution to the overall noise covariance matrix, quantifying the covariance across all possible pairs of neurons. In mathematical notation, this follows

$$C_{ij}^{Noise} = \frac{1}{T} \frac{1}{K} \sum_{t=0}^T \sum_{k=1}^K \eta_i^k(t) \eta_j^k(t) \quad (15)$$

where η_i^k is the noisy firing rate value in time t of trial k that belongs to neuron i . PCA was then performed on the resulting covariance matrix to extract the eigenvalues, which indicate the variance captured by the PCs. We calculated the proportion of explained variance for each PC derived. To ensure statistical robustness, this matrix construction and eigenvalue extraction process was repeated 100 times with varying random seeds, allowing us to determine the mean and SD of the noise-associated explained variance for the different PCs. This analysis enabled the identification of the dimensionality of the system through the identification of the meaningful PCs that had eigenvalues above the noise level, which were statistically significant (Fig. 4D and fig. S5C). The significant PCs were then used to compute the corresponding projections.

Demixed PCA

Theoretical motivation and technicalities of this method are clearly explained in (30, 33). dPCA is divided into two steps: supervised and unsupervised. The supervised component of dPCA involves untangling the neural activity from predefined task variables to build marginalized covariance matrices, similar to variable selection in linear regression. Afterward, the method proceeds with an unsupervised PCA on these matrices, similar to traditional PCA. In our task, we marginalized the population activity ($\mathbf{X}$) across all liminal and supraliminal conditions ($\mathbf{X}_{t,c}$) for the tactile (4) and acoustic (4) modalities, excluding the absent stimulus condition. In this part of the study, we chose not to include absent stimuli when calculating the axes, allowing us to better understand the population dynamics between the two modalities, tactile and acoustic. However, although the axes are computed without these stimuli, it remains possible to project the activity of the absent stimuli onto these axes (Fig. 6).

To compute the demixed axes, the condition-independent component was subtracted during the marginalization process. To calculate the marginalization averages, we used the N -dimensional population activity

$$\bar{\mathbf{X}}(t, c) = \bar{\mathbf{r}}(t, c) - \bar{\mathbf{r}}(t) \quad (16)$$

Then, $\bar{\mathbf{X}}_\phi$ denotes the marginalized matrices with $\phi \in \{t, c\}$ (the marginalization variable used in this study). Further, $\bar{\mathbf{r}}(t)$ represents the firing rate average over all hit trials at each time bin (combining the 10 conditions, 5 tactile and 5 acoustic), while $\bar{\mathbf{r}}(t, c)$ is the firing rate average per condition (separating trials according to the 5 different conditions within each sensory modality: tactile or

acoustic). Once the marginalization is performed, dPCA finds separate decoder and encoder matrices for each $\bar{X}_\phi$ by minimizing the reduced-rank regression term

$$L_\phi = \|\bar{X}_\phi - F_\phi D_\phi \bar{X}\|^2 \quad (17)$$

where $\bar{X}$ is the centered whole population data matrix (i.e., the average activity of each neuron is 0). The solution to this problem can be obtained analytically through singular value decomposition. Each component D_ϕ can be ordered by the amount of explained variance, with the most prominent decoding axis referred to as the 1st dPC (dPC1) of the variable ϕ . To avoid overfitting in dPCA, we introduced a regularization term and performed cross-validation to select the optimal regularization parameter. For Fig. 6A, we projected the N -dimensional data for a given condition $[\bar{r}(t, c)]$ —in this case, the firing rate average per condition, computed by grouping trials according to tactile amplitude, acoustic amplitude, and stimulus-absent modalities—onto the most prominent decoding axis (k) or the variable ϕ .

Here, we emphasize that dPCA finds the decoder and encoder of each marginalized variable, ϕ , minimizing each L_ϕ separately. This allows for the independent calculation of the demixed decoder for the parameter condition. In Fig. 6A, we show the projections of neural activity sorted by condition onto the most prominent condition-dPCs. Most trajectories displayed an amodal encoding of tactile and acoustic stimuli, except along the condition-dPC4 and 5 axes, where the resultant representations exhibited a modality-specific (modal) coding of conditions across tactile and acoustic stimuli. Given that these modal axes show a clear separation between tactile and acoustic stimuli, we refer to them in Fig. 6 as modal-dPCs. Conversely, the components that are independent of the modality are referred to as amodal-dPCs. In Fig. 6C, neural activity was averaged by stimulus modality (tactile or acoustic) or absence of stimulus (stimulus absent) and then projected onto the modal-dPC4 and 5. Unlike PCA, modal dPCA weights had a positive correlation (fig. S6E). The projections were plotted against each other to create a 2D diagram, visualizing the separation of neural trajectories across modalities.

Regression analysis with MLR and GLM

We used MLR (42, 44) and GLM to fit the mean firing rates of neuronal responses from different classes within the same sensory modality (tactile or acoustic) across various time windows. This analysis allowed us to assess the dynamic responses of neurons during specific periods of interest within the BDT.

The MLR model assumes that the mean firing rates are normally distributed. In our model, the regressors included tactile amplitude and acoustic amplitude, and the model was formulated as follows

$$Y_i(t) = \beta_0^i + \beta_1^i(t) \text{Amp}_{Tac} + \beta_2^i(t) \text{Amp}_{Ac} + \epsilon_i(t) \quad (18)$$

where $Y_i(t)$ represents the mean firing rate of neuron i at time t ; Amp_{Tac} (tactile amplitude) and Amp_{Ac} (acoustic amplitude) are the predictors; β_0^i , β_1^i , and β_2^i are the regression coefficients; $\epsilon_i(t)$ is the error term, assumed to be normally distributed. For the GLM, we assumed a Poisson distribution for the firing rates, with the identity link function. The GLM included additional regressors to account for the presence of tactile and acoustic stimuli, which were specified as

$$\begin{aligned} \lambda_i(t) = & \beta_0^i + \beta_1^i(t) \text{Amp}_{Tac} + \beta_2^i(t) \text{Amp}_{Ac} \\ & + \beta_3^i(t) \text{Pres}_{Tac} + \beta_4^i(t) \text{Pres}_{Ac} \end{aligned} \quad (19)$$

where $\lambda_i(t)$ is the expected firing rate, Pres_{Tac} and Pres_{Ac} are binary indicators (0 or 1), representing whether any tactile or acoustic stimulus was present, respectively. Projection axes were constructed using the regression coefficients obtained from both MLR and GLM models. For each regression variable, a vector $\mathbf{r}_m$ was calculated and normalized such that $\|\mathbf{r}_m\| = 1$. The mean firing rate for each condition, modality, and time bin was then projected onto these vectors using the dot product. In Fig. 5 (A and B), we used a GLM, while in fig. S4C, we used an MLR.

Figure S6 illustrates correlations among neural contributions to various PCA and GLM-derived axes, as well as the speed of population dynamics. For each plot (fig. S6, A to E), neural contributions were assessed by computing the correlation coefficients (ρ) between neurons' weights on different axes derived from PCA (fig. S6, A to C), GLM (fig. S6D), and dPCA (fig. S6E).

Recurrent neural network

We used an RNN (31, 32, 42, 44, 48, 49, 51) model to analyze neural response dynamics during BDT. The network architecture was inspired by previous studies using an RNN described by the following dynamic equation

$$\mathbf{r}(t) = \text{Sfp} \left[W_{rec} \mathbf{r}_{t-1} + W_{in} \mathbf{u}_t + \mathbf{b} + \sqrt{2} \sigma N(0, 1) \right] \quad (20)$$

where $\text{Sfp}(x) = \log(1 + e^x)$ is the softplus activation function, ensuring the nonnegativity of the firing rates, $\mathbf{r}(t)$.

Network configuration

The network consisted of 256 units, each integrating signals through two primary pathways: recurrent weights (W_{rec}), a matrix handling the recurrent connections, initialized with ones along the diagonal and other elements drawn from a normal distribution $N\left(0, \frac{0.3}{\sqrt{N}}\right)$, promoting stable dynamics (32); input weights (W_{in}), a matrix for input connections that channels external signals into the network; and bias and noise, each unit incorporated a trainable bias $\mathbf{b}$ and Gaussian noise $N(0, 1)$ at each time step to model variability and stochasticity in neuronal firing.

Input and stimulation

The input $\mathbf{u}_t$ to the RNN was structured to simulate sensory stimuli along with a go cue. Specifically, we defined three input units (as stimulus channels), one for each type of stimulus and one for the go cue. From 1 to 3.2 s post-trial start, one stimulus channel is switched on to one of the four amplitudes (0.01, 0.05, 0.1, and 0.5) and is maintained for 0.5 s. The go cue value remained at 0 until 1 s poststimulus, whereupon it elevated to 1 until the trial's end, half a second later.

Output layer and learning

The network activity was processed by a fully connected output layer $\mathbf{z}_t$, formulated as $\mathbf{z}_t = W_{out} \mathbf{r}_t + \mathbf{b}_{out}$, with each unit in this layer corresponding to a potential response (stimulus 1 present, stimulus 2 present, and stimulus absent). The network was trained to minimize the squared error cost function $C = \sum (\mathbf{z}_t - \mathbf{y}_t)^2$, where $\mathbf{z}_t$ and $\mathbf{y}_t$ represent the activities of output units and target vectors, respectively. Parameters were updated via stochastic gradient descent, with a batch size of 64 and a learning rate of 0.0005. Training of the network proceeded until performance accuracy, which we defined as correct answers produced (31, 32), reached 80%.

Rotating memories

We propose modeling the sensory and memory retention dynamics observed in the VPC as a 2D system of differential equations

with state variables S and M . These variables are intended to capture the dynamics along the two most representative modal dPCs—namely, dPC4 and dPC5 (Fig. 6). The first state variable, S , models the stimulus-driven, early transient sensory response, while the second state variable, M , represents the persistent mnemonic response, which retains the perceptual decision throughout the delay period.

To derive suitable dynamics for S and M , we rely on the idea that sensory dynamics are predominantly faithful and linear, while mnemonic dynamics are categorical and nonlinear. Specifically, sensory responses exhibit exponentially vanishing memory, while mnemonic responses retain memory through multiple stable equilibria. Therefore, we propose that sensory dynamics can be modeled as a first-order low-pass filter driven by sensory inputs. In contrast, mnemonic dynamics are modeled as a first-order nonlinear system driven by the output of the sensory dynamics, featuring multiple stable equilibria: one corresponding to the VPC resting state (stimulus-absent response) and two others associated with the presence of acoustic and tactile stimuli.

The proposed model can be expressed as

$$\dot{S} = -S + I \quad (21)$$

$$\dot{M} = -M + \tanh[k_1(1 + k_2 M^2)M] + k_3 S \quad (22)$$

where I represents the sensory inputs ($I < 0$ for acoustic stimuli and $I > 0$ for tactile stimuli), and k_1, k_2, k_3 are positive gain parameters. In line with our modeling hypothesis, the system described by these equations consists of two key components: a first-order filter with state variable S and a nonlinear multistable system with state variable M . The low-pass filter in equation (Eq. 21) captures the transient nature of sensory responses, while the multistable nonlinear dynamics in equation (Eq. 22) model the categorical and persistent nature of mnemonic retention. The multistable structure of the mnemonic dynamics allows the system to maintain one of three stable states—stimulus-absent, tactile stimulus-present, or acoustic stimulus-present—thereby capturing the decision process observed in the VPC during the BDT (Fig. 8 and fig. S7).

Supplementary Materials

This PDF file includes:

Figs. S1 to S7

REFERENCES AND NOTES

- B. E. Stein, T. R. Stanford, Multisensory integration: Current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* **9**, 255–266 (2008).
- R. Rossi-Pool, J. Vergara, R. Romo, “Constructing perceptual decision-making across cortex” in *The Cognitive Neurosciences*, 6/E, D. Poeppel, G. Mangun, M. Gazzaniga, Eds. (The MIT Press, 2020), pp. 413–427.
- A. Hernández, V. Nácher, R. Luna, A. Zainos, L. Lemus, M. Alvarez, Y. Vázquez, L. Camarillo, R. Romo, Decoding a perceptual decision process across cortex. *Neuron* **66**, 300–314 (2010).
- M. M. Murray, S. Molholm, C. M. Michel, D. J. Heslenfeld, W. Ritter, D. C. Javitt, C. E. Schroeder, J. J. Foxe, Grabbing your ear: Rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb. Cortex* **15**, 963–974 (2005).
- C. Kayser, The multisensory nature of unisensory cortices: A puzzle continued. *Neuron* **67**, 178–180 (2010).
- A. A. Ghazanfar, J. X. Maier, K. L. Hoffman, N. K. Logothetis, Multisensory integration of dynamic faces and voices in Rhesus monkey auditory cortex. *J. Neurosci.* **25**, 5004–5012 (2005).
- J. R. Mahoney, P. C. C. Li, M. Oh-Park, J. Verghese, R. Holtzer, Multisensory integration across the senses in young and old adults. *Brain Res.* **1426**, 43–53 (2011).
- C. Perrodin, C. Kayser, N. K. Logothetis, C. I. Petkov, Natural asynchronies in audiovisual communication signals regulate neuronal multisensory interactions in voice-sensitive cortex. *PNAS* **112**, 273–278 (2015).
- J. S. Butler, J. J. Foxe, I. C. Fiebelkorn, M. R. Mercier, S. Molholm, Multisensory representation of frequency across audition and touch: High density electrical mapping reveals early sensory-perceptual coupling. *J. Neurosci.* **32**, 15338–15344 (2012).
- M. Gori, G. Bertonati, C. Campus, M. B. Amadeo, Multisensory representations of space and time in sensory cortices. *Hum. Brain Mapp.* **44**, 656–667 (2023).
- C. Kayser, C. I. Petkov, M. Augath, N. K. Logothetis, Integration of touch and sound in auditory cortex. *Neuron* **48**, 373–384 (2005).
- H. Atilgan, S. M. Town, K. C. Wood, G. P. Jones, R. K. Maddox, A. K. C. Lee, J. K. Bizley, Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. *Neuron* **97**, 640–655.e4 (2018).
- L. Lemus, A. Hernández, R. Luna, A. Zainos, R. Romo, Do sensory cortices process more than one sensory modality during perceptual judgments? *Neuron* **67**, 335–348 (2010).
- R. Romo, R. Rossi-Pool, Turning touch into perception. *Neuron* **105**, 16–33 (2020).
- A. Betancourt, O. Pérez, J. Gámez, G. Mendoza, H. Merchant, Amodal population clock in the primate medial premotor system for rhythmic tapping. *Cell Rep.* **42**, 113234 (2023).
- J. Vergara, N. Rivera, R. Rossi-Pool, R. Romo, A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron* **89**, 54–62 (2016).
- S. Parra, H. Diaz, A. Zainos, M. Alvarez, J. Zizumbo, N. Rivera-Yoshida, S. Pujalte, L. Bayones, R. Romo, R. Rossi-Pool, Hierarchical unimodal processing within the primary somatosensory cortex during a bimodal detection task. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2302617120 (2022).
- V. de Lafuente, R. Romo, Neuronal correlates of subjective sensory experience. *Nat. Neurosci.* **8**, 1698–1703 (2005).
- V. de Lafuente, R. Romo, Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14266–14271 (2006).
- G. Rizzolatti, G. Luppino, The cortical motor system. *Neuron* **31**, 889–901 (2001).
- M. Gerbella, A. Belmalih, E. Borra, S. Rozzi, G. Luppino, Cortical connections of the anterior (F5a) subdivision of the macaque ventral premotor area F5. *Brain Struct. Funct.* **216**, 43–65 (2011).
- G. Luppino, A. Murata, P. Govoni, M. Matelli, Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Exp. Brain Res.* **128**, 181–187 (1999).
- M.-T. Lu, J. B. Preston, P. L. Strick, Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *J. Comp. Neurol.* **341**, 375–392 (1994).
- T. Ninomiya, A. Noritake, K. Kobayashi, M. Isoda, A causal role for frontal cortico-cortical coordination in social action monitoring. *Nat. Commun.* **11**, 5233 (2020).
- M. S. Graziano, L. A. Reiss, C. G. Gross, A neuronal representation of the location of nearby sounds. *Nature* **397**, 428–430 (1999).
- M. S. Graziano, G. S. Yap, C. G. Gross, Coding of visual space by premotor neurons. *Science* **266**, 1054–1057 (1994).
- R. Romo, A. Hernández, A. Zainos, Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron* **41**, 165–173 (2004).
- L. Lemus, A. Hernandez, R. Romo, Neural encoding of auditory discrimination in ventral premotor cortex. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 14640–14645 (2009).
- A. Libby, T. J. Buschman, Rotational dynamics reduce interference between sensory and memory representations. *Nat. Neurosci.* **24**, 715–726 (2021).
- L. Bayones, A. Zainos, J. Manuel Alvarez, R. Romo, A. Franci, R. Rossi-Pool, Orthogonality of sensory and contextual categorical dynamics embedded in a continuum of responses from the second somatosensory cortex. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2316765121 (2024).
- G. R. Yang, X.-J. Wang, Artificial neural networks for neuroscientists: A primer. *Neuron* **107**, 1048–1070 (2020).
- Z. Bi, C. Zhou, Understanding the computation of time using neural network models. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1053–10540 (2020).
- D. Kobak, W. Brendel, C. Constantinidis, C. E. Feierstein, A. Kepecs, Z. F. Mainen, X.-L. Qi, R. Romo, N. Uchida, C. K. Machens, Demixed principal component analysis of neural population data. *eLife* **5**, e10989 (2016).
- R. Rossi-Pool, A. Zainos, M. Alvarez, J. Zizumbo, J. Vergara, R. Romo, Decoding a decision process in the neuronal population of dorsal premotor cortex. *Neuron* **96**, 1432–1446.e7 (2017).
- B. M. Dekleva, R. H. Chowdhury, A. P. Batista, S. M. Chase, B. M. Yu, M. L. Boninger, J. L. Collinger, Motor cortex retains and reorients neural dynamics during motor imagery. *Nat. Hum. Behav.* **8**, 729–742 (2024).
- E. M. Trautmann, S. D. Stavisky, S. Lahiri, K. C. Ames, M. T. Kaufman, D. J. O’Shea, S. Vyas, X. Sun, S. I. Ryu, S. Ganguli, K. V. Shenoy, Accurate estimation of neural population dynamics without spike sorting. *Neuron* **103**, 292–308 (2019).

37. C. Langdon, M. Genkin, T. A. Engel, A unifying perspective on neural manifolds and circuits for cognition. *Nat. Rev. Neurosci.* **24**, 363–377 (2023).
38. R. Rossi-Pool, R. Romo, Low dimensionality, high robustness in neural population dynamics. *Neuron* **103**, 177–179 (2019).
39. D. L. Barack, J. W. Krakauer, Two views on the cognitive brain. *Nat. Rev. Neurosci.* **22**, 359–371 (2021).
40. C. K. Machens, R. Romo, C. D. Brody, Functional, but not anatomical, separation of “what” and “when” in prefrontal cortex. *J. Neurosci.* **30**, 350–360 (2010).
41. J. D. Murray, A. Bernacchia, N. A. Roy, C. Constantinidis, R. Romo, X.-J. Wang, Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 394–399 (2017).
42. V. Mante, D. Sussillo, K. V. Shenoy, W. T. Newsome, Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
43. M. C. Aoi, V. Mante, J. W. Pillow, Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nat. Neurosci.* **23**, 1410–1420 (2020).
44. F. Carnevale, V. de Lafuente, R. Romo, O. Barak, N. Parga, Dynamic control of response criterion in premotor cortex during perceptual detection under temporal uncertainty. *Neuron* **86**, 1067–1077 (2015).
45. H. K. Inagaki, L. Fontolan, S. Romani, K. Svoboda, Discrete attractor dynamics underlies persistent activity in the frontal cortex. *Nature* **566**, 212–217 (2019).
46. O. Barak, M. Tsodyks, R. Romo, Neuronal population coding of parametric working memory. *J. Neurosci.* **30**, 9424–9430 (2010).
47. R. Rossi-Pool, J. Zizumbo, M. Alvarez, J. Vergara, A. Zainos, R. Romo, Temporal signals underlying a cognitive process in the dorsal premotor cortex. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7523–7532 (2019).
48. W. Chaisangmongkon, S. K. Swaminathan, D. J. Freedman, X.-J. Wang, Computing by robust transience: How the fronto-parietal network performs sequential, category-based decisions. *Neuron* **93**, 1504–1517.e4 (2017).
49. G. R. Yang, M. R. Joglekar, H. F. Song, W. T. Newsome, X.-J. Wang, Task representations in neural networks trained to perform many cognitive tasks. *Nat. Neurosci.* **22**, 297–306 (2019).
50. J. Wang, D. Narain, E. A. Hosseini, M. Jazayeri, Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018).
51. E. D. Remington, D. Narain, E. A. Hosseini, M. Jazayeri, Flexible sensorimotor computations through rapid reconfiguration of cortical dynamics. *Neuron* **98**, 1005–1019.e5 (2018).
52. J. D. Smedo, A. Zandvakili, C. K. Machens, B. M. Yu, A. Kohn, Cortical areas interact through a communication subspace. *Neuron* **102**, 249–259.e4 (2019).
53. G. T. Meijer, P. Marchesi, J. F. Mejias, J. S. Montijn, C. S. Lansink, C. M. A. Pennartz, Neural correlates of multisensory detection behavior: Comparison of primary and higher-order visual cortex. *Cell Rep.* **31**, 107636 (2020).
54. M. T. Lippert, T. Steudel, F. Ohl, N. K. Logothetis, C. Kayser, Coupling of neural activity and fMRI-BOLD in the motion area MT. *Magn. Reson. Imaging* **28**, 1087–1094 (2010).
55. A. P. Khandhadia, A. P. Murphy, L. M. Romanski, J. K. Bizley, D. A. Leopold, Audiovisual integration in macaque face patch neurons. *Curr. Biol.* **31**, 1826–1835.e3 (2021).
56. F. R. Willett, E. Kunz, C. Fan, D. Avansino, G. Wilson, E. Y. Choi, F. Kamdar, L. R. H. Hochberg, S. Druckmann, K. Shenoy, J. Henderson, A high-performance speech neuroprosthesis. *Nature* **620**, 1031–1036 (2023).
57. F. Pulvermüller, How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cogn. Sci.* **17**, 458–470 (2013).
58. E. Rassi, Y. Zhang, G. Mendoza, J. C. Méndez, H. Merchant, S. Haegens, Distinct beta frequencies reflect categorical decisions. *Nat. Commun.* **14**, 2923 (2023).
59. S. Haegens, J. Vergara, R. Rossi-Pool, L. Lemus, R. Romo, Beta oscillations reflect supramodal information during perceptual judgment. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 13810–13815 (2017).
60. J. Aru, J. Aru, V. Priesemann, M. Wibral, L. Lana, G. Pipa, W. Singer, R. Vicente, Untangling cross-frequency coupling in neuroscience. *Curr. Opin. Neurobiol.* **31**, 51–61 (2015).
61. R. Rossi-Pool, A. Zainos, M. Alvarez, G. Díaz-deLeon, R. Romo, A continuum of invariant sensory and behavioral-context perceptual coding in secondary somatosensory cortex. *Nat. Commun.* **12**, 2000 (2021).
62. H. Díaz, L. Bayones, M. Alvarez, B. Andrade-Ortega, S. Valero, A. Zainos, R. Romo, R. Rossi-Pool, Contextual neural dynamics during time perception in the primate ventral premotor cortex. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2420356122 (2025).
63. S. Panzeri, R. Senatore, M. A. Montemurro, R. S. Petersen, Correcting for the sampling bias problem in spike train information measures. *J. Neurophysiol.* **98**, 1064–1072 (2007).
64. V. Koren, G. Bondanelli, S. Panzeri, Computational methods to study information processing in neural circuits. *Comput. Struct. Biotechnol. J.* **21**, 910–922 (2023).
65. E. Maris, R. Oostenveld, Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).

Acknowledgments: We would like to thank N. E. Rodríguez for insightful comments and for proofreading the manuscript. **Funding:** This work was supported by grants PAPIIT-IN210819 (to R.R. and R.R.-P.), PAPIIT-IN205022 (to R.R.-P.), and IN203825 (to R.R.-P.) from the Dirección de Asuntos del Personal Académico de la Universidad Nacional Autónoma de México; CONAHCYT-240892 (to R.R.) and CONAHCYT-319347 (to R.R.-P.) from Consejo Nacional de Ciencia y Tecnología; SECITHI-CBF-2025-I-2054 (to R.R.-P.) from Secretaría de Ciencia, Humanidades, Tecnología e Innovación; and IBRO Early Career Award 2022 (to R.R.-P.) from International Brain Research Association. H.D. is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, UNAM. L.B. is a postdoctoral researcher (Postdoctoral fellowship CONACYT-838783). **Author contributions:** Investigation: H.D., A.Z., L.B., M.A., A.F., R.R., and R.R.-P. Conceptualization: B.A.-O., H.D., A.Z., M.A., and R.R.-P. Methodology: B.A.-O., H.D., A.Z., M.A., N.R.-Y., A.F., R.R., and R.R.-P. Writing—original draft: B.A.-O., H.D., L.B., M.A., A.F., and R.R.-P. Writing—review and editing: B.A.-O., L.B., A.Z., M.A., A.F., R.R., and R.R.-P. Resources: A.Z. and M.A. Data curation: B.A.-O., M.A., and N.R.-Y. Validation: B.A.-O., L.B., A.Z., M.A., and N.R.-Y. Formal analysis: B.A.-O., H.D., L.B., M.A., A.F., and R.R.-P. Software: B.A.-O., H.D., L.B., M.A., A.F., and R.R.-P. Visualization: B.A.-O., L.B., A.Z., M.A., and R.R.-P. Supervision: L.B., R.R., and R.R.-P. Funding acquisition: R.R. and R.R.-P. Project administration: R.R. and R.R.-P. Scientific content of the work: B.A.-O., H.D., L.B., A.Z., M.A., N.R.-Y., A.F., R.R., and R.R.-P. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The data from the Bimodal Detection Task can be accessed at the following link: <https://zenodo.org/records/14361832>. This includes the trial psychophysics, recorded neuronal activity, data reading instructions, and scripts to classify neurons, compute firing rates, analyze coding, perform population dynamics with PCA and dPCA, and generate both recurrent neural networks (RNNs) and a simplified model. In addition, the scripts can also be accessed at: <https://github.com/BernardoAO/BDT-VPC> and <https://github.com/Rossi-Pool-Lab/Bimodal-Detection-analysis>

Submitted 13 February 2025

Accepted 19 August 2025

Published 19 September 2025

10.1126/sciadv.adw5500

Multistable bimodal perceptual coding within the ventral premotor cortex

Bernardo Andrade-Ortega, Héctor Díaz, Lucas Bayones, Manuel Alvarez, Antonio Zainos, Natsuko Rivera-Yoshida, Alessio Franci, Ranulfo Romo, and Román Rossi-Pool

Sci. Adv. **11** (38), eadw5500. DOI: 10.1126/sciadv.adw5500

View the article online

<https://www.science.org/doi/10.1126/sciadv.adw5500>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

Copyright © 2025 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).