PNAS RESEARCH ARTICLE | NEUROSCIENCE

# Orthogonality of sensory and contextual categorical dynamics embedded in a continuum of responses from the second somatosensory cortex

Lucas Bayones[a], Antonio Zainos[a], Manuel Alvarez[a], Ranulfo Romo[b,1], Alessio Franci[c,d,e,1], and Román Rossi-Pool[a,f,1]

Affiliations are included on p. 12.

Contributed by Ranulfo Romo; received September 26, 2023; accepted June 12, 2024; reviewed by Bruno Averbeck and Hugo Merchant

**How does the brain simultaneously process signals that bring complementary information, like raw sensory signals and their transformed counterparts, without any disruptive interference? Contemporary research underscores the brain's adeptness in using decorrelated responses to reduce such interference. Both neurophysiological findings and artificial neural networks support the notion of orthogonal representation for signal differentiation and parallel processing. Yet, where, and how raw sensory signals are transformed into more abstract representations remains unclear. Using a temporal pattern discrimination task in trained monkeys, we revealed that the second somatosensory cortex (S2) efficiently segregates faithful and transformed neural responses into orthogonal subspaces. Importantly, S2 population encoding for transformed signals, but not for faithful ones, disappeared during a nondemanding version of this task, which suggests that signal transformation and their decoding from downstream areas are only active on-demand. A mechanistic computation model points to gain modulation as a possible biological mechanism for the observed context-dependent computation. Furthermore, individual neural activities that underlie the orthogonal population representations exhibited a continuum of responses, with no well-determined clusters. These findings advocate that the brain, while employing a continuum of heterogeneous neural responses, splits population signals into orthogonal subspaces in a context-dependent fashion to enhance robustness, performance, and improve coding efficiency.**

second somatosensory cortex | orthogonal signals | population dynamics | sensory representation | categorical coding

## Significance

An important function of the brain is turning sensation into perception. Yet, how this function is implemented remains unknown. Current research, insights from artificial neural networks, highlights using orthogonal representations as an effective means to transform sensory signals into perceptual signals while separating and simultaneously processing two information streams. Neuronal recordings in S2 while trained monkeys performed the temporal pattern discrimination task (TPDT), revealed that this function is implemented at the population level. While S2 encodes sensory information independently of context, the encoding of categorical information, like task parameters, is only performed when the task demands it. Such distinct and flexible organization, enriched by a spectrum of neural activities, reflects the brain's efficiency, resilience, and purpose for solving cognitive tasks.

The kitchen, a commonplace yet remarkably complex sensorial arena, serves as a vivid illustration of the extraordinary capabilities of the human brain. From the moment one steps into this space with the intent to prepare a meal, a plethora of sensory information floods our neural pathways—the visual recognition of ingredients and utensils, the auditory distinction between background hums and salient sizzles, the tactile feedback of knife against cutting board, the olfactory delight of sautéing garlic, and the gustatory assessment of flavor. Within this sensory tapestry, the brain performs a series of intricate transformations, transducing raw stimuli into meaningful units of information, and translating them into coherent, goal-oriented actions. This process encapsulates the remarkable ability of our neural networks to process diverse sensory information and seamlessly integrate them into a single, coherent experience, an ability that lies at the core of human cognition and behavior. Nevertheless, this raises an intriguing question: How does the brain simultaneously receive raw sensory signals, process them into more abstract transformed ones, and achieve representation and interarea transmission of both kinds of signals in a context-dependent fashion?

Experiments on behaving monkeys have brought forth an attractive hypothesis about the brain's unique encoding mechanisms (1–5). They suggest that the brain utilizes orthogonal encoding subspaces, a kind of cognitive architecture, to reduce interference among signals involved in different processes. It has been seen that this smart arrangement allowed the brain to decouple mnemonic coding with dynamics either related to stimulus representation or comparison (4, 5). It also permitted the network to separate signals associated with motor planning and motor execution (1–3). A recent work in the human motor cortex has also found that motor execution and imagining the movement without executing it evolve with similar dynamics at orthogonal subspaces (6). Orthogonal subspaces were also found in tasks involving sequential memory representation (7) and to differentiate signals that encode either reward or confidence (8). Further research in monkeys has revealed that the neocortical neuronal network can represent multiple related task variables without cross-interference (7). Additionally, studies conducted in the posterior parietal cortex of

Reviewers: B.A., NIH; and H.M., Universidad Nacional Autonoma de México.

The authors declare no competing interest.



[1]To whom correspondence may be addressed. Email: ranulfo.romo@gmail.com, alessio.franci83@gmail.com, or romanr@ifc.unam.mx.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2316765121/-/DCSupplemental.

Published July 11, 2024.

rodents demonstrated that the activity patterns of neuronal groups during evidence accumulation are significantly influenced by past events (9). Here, orthogonal subspace encoding is used by the network to distinguish and reduce interference between learned information and new experiences. Complementarily, recent work in rodents suggested that in order to reduce correlation in the neural representations of related task variables, the population must code the different parameters into decoupled dynamics (10). Importantly, most of these studies have also found that orthogonal signals emerged in a continuum or responses across the circuit (3–8). Therefore, the notion of orthogonal representation in a continuum of responses has garnered support from a variety of studies. However, it remains unclear whether this type of neural scaffold is effective in distinguishing between raw sensory signals, originating from stimulus input, and transformed signals, which represent sensory information that has been processed and abstracted for use by downstream areas in the hierarchical processing pathway.

Research in AI, particularly with artificial neural networks (ANN), has substantially deepened our understanding of orthogonal representations in signal discrimination. Much like their biological counterparts, ANNs have been observed to use orthogonal representations for context-dependent tasks (11). Recent studies in spiking neuronal networks indicate that leveraging latent orthogonal subspaces can boost learning versatility (12). ANNs have also adeptly mimicked outcomes from biological data across several tasks. This demonstrates their capability to distinctly showcase orthogonal differentiation in input, scaling, and gain subspaces (13–15), as well as in category dynamics during delay periods (16). A notable contemporary finding is that both biological and artificial networks encode neural noise and stimulus coding in orthogonal subspaces. This maximizes the decorrelation between population dynamics, thereby enhancing the precision of relevant coding (17). In essence, the employment of artificial networks has harnessed the power of orthogonal representations, revealing promising avenues for enhanced signal processing and task-specific performance.

While the concept of orthogonal encoding and its role in cognitive processes have been extensively studied, little attention has been given to investigating how faithful sensory signals interact with transformed signal representations. Among the variety of brain areas, the secondary somatosensory cortex (S2) constitutes a promising circuit to examine this issue. In our prior research (18), involving macaques trained to undertake a temporal pattern discrimination task (TPDT), we revealed that S2 showcases a broad spectrum of neuronal responses. These ranged from faithful or phase-locked responses to more abstract representations, wherein sensory information from stimuli was integrated and transformed by neurons into direct category encoding. Employing information theory at the single-cell level, we noted that these responses span a continuum: some neurons exhibited a purely sensory response, some others a purely transformed one, but many others exhibited a graded superposition of the two kinds of signals. Building on this single-cell analysis, in this study, we delve into how these signals might interact at the population level. This investigation could yield critical insights into the mechanisms underlying sensory representation and its transformation, aiming to minimize interference among dynamics at the network level. Our exploration of S2's dynamics will not only contribute to a broader knowledge of how the brain converts sensory inputs but also shed light on the context-dependent recruitment of decorrelated population signals during cognitive tasks. This research offers understanding that could potentially overcome the constraints of previous studies that used single units in S2 (18–20). By inspecting the data collected in the TPDT under a population paradigm we found that faithful and categorical dynamics appeared in orthogonal subspaces in the S2 network state space. Hence, decorrelated combinations of neural responses are necessary to decode sensory inputs and categorical and more abstract signals. Population-level analysis, performed through the Uniform Manifold Approximation and Projection (UMAP) method, confirmed the single-cell level observation that neuronal responses exhibit a grade continuum, from purely sensory to purely categorical, rather than discrete clusters of activity types. This finding underscores the seamless integration of neuronal activity during the active task, highlighting the continuous nature of the neural encoding rather than the presence of isolated neuronal groups. The observed continuity disappeared in a less-demanding version of the task (the light control variant), where the categorically distinct population signal faded, creating a discontinuity in the population response patterns that was impossible to visualize with the traditional linear projection methods. This indicates that such responses, and the associated continuum, manifest exclusively under necessary conditions. To provide mechanistic insights, we present a computational model that serves as an example of the possible dynamical mechanisms underlying this phenomenon in a way that is compatible with the orthogonalization of coding subspaces found in the S2 network.

## Results

**Individual Neuronal Responses in S2 during Task Performance.** Two monkeys were trained in the TPDT (21), in which they had to report whether two temporal patterns with a vibrotactile flutter stimuli architecture (P1 and P2), were the same (P1 = P2) or different (P1 ≠ P2) (Fig. 1*A* and *Materials and Methods*). Two possible temporal patterns were presented: extended (E), made of five pulses uniformly spaced in time, and grouped (G), with three pulses grouped at the middle of the stimulus. There were therefore four classes of possible temporal pattern pairs: G-G (c1, red color), G-E (c2, orange), E-G (c3, green), and E-E (c4, blue). Mean stimulus frequency (5 Hz) and duration (1 s) were maintained constant throughout the task. The average performance in the TPDT across S2 recording sessions was 84% (±7%), and this was also consistent across classes (Fig. 1*B*).

We recorded the activity from 1,646 neurons in S2 (Fig. 1*C* and *Materials and Methods*) during the TPDT (Monkey RR17, n = 1,035; Monkey RR20, n = 611). The responses of 12 and 24 exemplary S2 neurons are displayed in Fig. 1*D* and *SI Appendix*, Fig. S1, respectively. As we previously showed (18), neurons from this area exhibit a broad repertoire of responses with distinguishable neuronal dynamics. Several neurons responded only when sensory stimulation was present and their response faithfully represented the stimulus patterns (some examples in Fig. 1*D* is *a*, and in *SI Appendix*, Fig. S1 are *a* & *b*). These neurons were therefore cataloged as "sensory," "faithful," or "phase-locked." Other recorded neurons, cataloged as "categorical," exhibited a nonsensory kind of response. In those neurons, sensory information was transformed to encode stimuli categories or some other features of the task, rather than the faithful temporal course of stimulus. If we look at these categorical units, they tend to respond more strongly to specific stimulus patterns (G or E) or particular classes (some examples in Fig. 1*D* are *c, e, f,* and *h*, and in *SI Appendix*, Fig. S1 are *I, j, p,* and *q*). For example, unit *b* in Fig. 1*D* responded more strongly when the G pattern was presented during P1, whereas during P2 the same neuron responded selectively to classes c1 & c3. Furthermore, looking at unit *f* in Fig. 1*D*, one can notice that during P2 this neuron responded preferentially to classes c3 & c4 rather than c1 & c2. Moreover, some neurons exhibited a mixture of sensory and categorical responses and were cataloged as "mixed". For instance, neuron *c* in *SI Appendix*, Fig. S1, faithfully represented all classes during

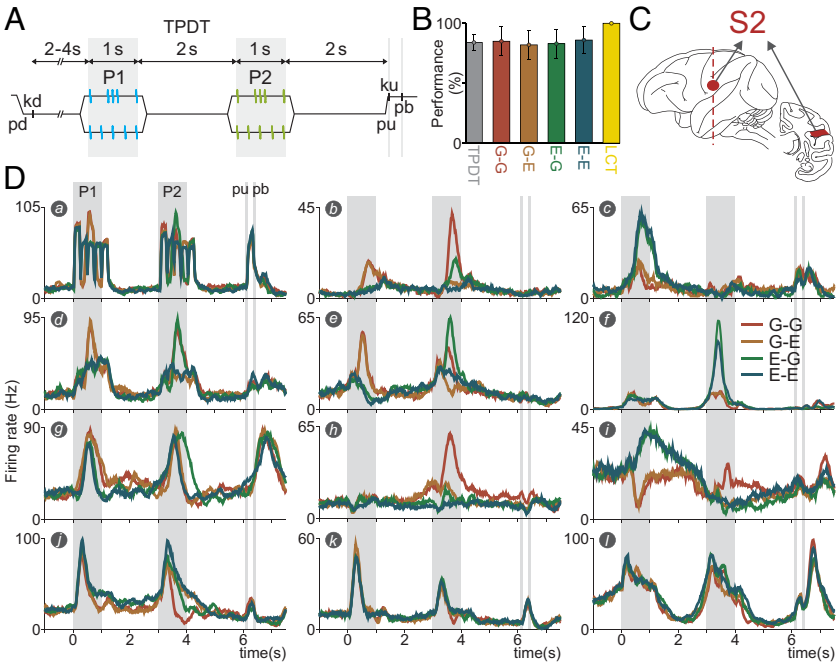


**Fig. 1.** Task, performance, and exemplary responses of S2 neurons. (*A*) Schematic of the TPDT with trial event sequence. A mechanical probe was lowered (pd), indenting the glabrous skin of one fingertip of the monkey's restrained right hand. Consequently, the animal responded by placing its left free hand on a fixed key (kd). After a variable prestimulus period (2 to 4 s), the probe executed the first vibrotactile stimulus (P1, 1 s duration and mean frequency of 5 Hz) with two possible patterns: grouped (G) or extended (E). After a first fixed delay of 2 s (from second 1 to 3), the second stimulus (P2) was presented, with the same configuration of P1 (1 s duration and G or E patterns). Then, a second fixed delay (2 s) was presented (from second 4 to 6), which ended with the probe-up event (pu). After "pu," the monkey released its free left hand from the key (ku) and pressed one of two buttons (pb) to report its decision, lateral to indicate that P1 and P2 were the same (P1 = P2) and medial to indicate that P1 and P2 were different (P1 ≠ P2). (*B*) Performance for the whole TPDT (gray, n = 423 sessions), for each class [G-G (red), G-E (orange), E-G (green), and E-E (blue)] and the whole light control task (LCT) (n = 76 sessions). (*C*) Illustration of the brain's left hemisphere (*Left* figurine) and a coronal brain slice (*Right* figurine) with S2 (recorded area) highlighted in red. Recordings were made contralateral and ipsilateral to the stimulated fingertip. (*D*) Average firing rates of twelve exemplary neurons, separated per class, plotted in the whole TPDT. Color traces indicate the four possible classes: G-G (red), G-E (orange), E-G (green), and E-E (blue). Note that unit *a* displays purely sensory dynamics whereas units *c, h,* and *f* show highly categorical response. Some neurons displaying mixed signals (sensory and categorical) and temporal responses, can also be seen in units *d, e* and *k, l*, respectively.

P1 but selectively diminished its activity in response to class c1 during P2. More neurons exhibiting a mixture of sensory and categorical responses are unit *d*, *g,* and *j* in Fig. 1*D* and units *m, f,* and *s* in *SI Appendix*, Fig. S1. If a sensory or categorical neuron firing rate increased (or decreased) in response to stimulation, it was labeled as positive (or negative, *SI Appendix*, Figs. S1 *k* and *u* and S2*c*). It has been previously reported that the integration of positive and negative S2 neuronal pools improves stimulus information decoding (22, 23). Finally, a fourth set of neurons exhibited strong temporal dynamics (Fig. 1*D*, *k* and *l* and *SI Appendix*, Fig. S1 *v, w*, and *x*). These units behaved like a timekeeper, responding independently of the stimuli. It has been speculated that these temporal responses constitute an essential underlying process for the network to predict the arrival of relevant sensory events (24). In the light of this heterogeneous collection of neuronal activity, one thing becomes clear: S2 exhibits a broad range of neuronal signals, representing sensory signals in disparate ways, from purely sensory to purely categorical and in between. This observation makes S2 a natural candidate for a place where sensory and categorical representations encounter, and perception begins.

All the neurons recorded in the light control task (LCT, n = 313, of which n = 189 from Monkey RR17 and n = 124 from Monkey RR20) were also recorded in the TPDT. This pool of neurons, recorded in both experimental conditions, provides a direct comparison of how cognitive demand (high in TPDT and low in LCT) modulates neuronal activity at the single-cell level (18). In each trial, monkeys received the same stimuli as in the TPDT, but the correct choice was indicated by a visual cue (*Materials and Methods*). The animal's performance was 100% in the control variant (Fig. 1*B*, yellow bar), indicating a strongly reduced cognitive demand. *SI Appendix*, Fig. S2 displays 16 neurons recorded both in TPDT and LCT. It is possible to observe that purely sensory units were not affected by the change of context (*SI Appendix*, Fig. S2 *a* and *c*), in line with what we recently reported (18). On the other hand, the response of purely categorical units in TPDT became completely unresponsive to task parameters in LCT (*SI Appendix*, Fig. S2 *g, k, n*, and *p*). This is also in agreement with our analysis of single-unit activity in the dorsal premotor cortex (DPC), where categorical units silenced significatively their responses during LCT (21). Finally, mixed units only partially altered their coding during LCT. Specifically, they tended to retain the sensory component while losing the categorical one. For example, unit *e* in *SI Appendix*, Fig. S2 responded selectively to the G pattern during P2 in the TPDT but completely lost its categorical representation in LCT, leaving only the sensory trace of the neuron's response. Summarizing, S2 neurons can modify or silence their categorical representation in a context-dependent manner, independently of whether they also present a sensory response. This supports the hypothesis that categorical coding is predominantly activated in contexts with high cognitive requirements (TPDT), highlighting its role in information processing that requires abstract interpretation and decision-making based on perceived stimuli. Conversely, sensory coding, which directly reflects the physical properties of stimuli, is active independently of the cognitive demands required by the context. This suggests that sensory coding operates independently of task complexity, providing a stable foundation for perception regardless of the context.

**Cognitive Demanding Dynamics Represented in S2.** To evaluate response variability associated with task parameter coding we computed the population instantaneous coding variance during TPDT and LCT (*Var Cod*, Fig. 2 *A* and *B*, blue trace). This metric evaluates the variability in neural response rates among different classes and individual neurons at each time point. A high variance in coding (*Var Cod*) indicates that distinct task parameters elicit significantly distinct neuronal activities. We examined the response variance elicited by different task parameters, namely, the first stimulus identity (*Var P1*), the second stimulus identity (*Var P2*), and the decision (*Var Dec*) (Fig. 2 *A* and *B*, light blue, purple, and pink, traces, respectively). *Var P1* therefore measures the response variability emerging during the first stimulus period, *Var P2* the response variability during the second stimulus or comparison period, and *Var Dec* the response variability emerging during the decision-making period. The variances in the peristimulus period reflect the basal fluctuations across trials (~2 sp/s). Parameter variances higher than their respective basal values are associated with relevant coding. To facilitate the comparison across parameters, values associated with basal variance were subtracted in each kind of variance computed. In the TPDT, *Var Cod* and *Var P1* traces showed significant coding for sensory representation during P1. Importantly, the *Var P1* wax and wane during the

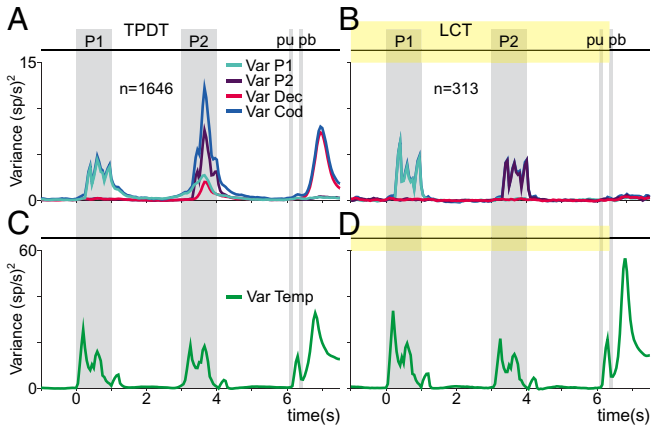


**Fig. 2.** Population variance during active and control tasks. (*A* and *B*) Population variances as a function of time during the TPDT (n = 1,646) and LCT (n = 313). Traces refer to P1 (*Var P1*, light blue trace), P2 (*Var P2*, purple trace), decision (*Var Dec*, pink trace), and coding (*Var Cod*, blue trace) variances. Note that basal variances (associated with residual fluctuations) were subtracted in all variances computed. Akin to frontal lobe areas, categorical and perceptual responses are abolished when LCT is employed, but sensory representations related with stimulation persist. (*C* and *D*) Temporal (or noncoding) variance during TPDT and LCT (*Var Temp*, green trace). This metric captured the sequential events of the task rather than across classes. Note how under LCT the temporal variance increases with respect to TPDT, meaning that noncoding signals constitute an essential part of the network dynamics.

delay period, exhibiting a recall of P1 information during the comparison period. In addition to *Var P1*, *Var P2* and *Var Dec* increase significantly during P2. As opposed to what we observed in DPC (24), where only a small percentage of neurons encoded the second stimulus, S2's neurons showed a marked coding of the identity of P2. The sharp increment in *Var Cod* during P2 is caused by the appearance of specific class and decision coding. Particularly, as we have shown previously, several neurons respond differently for a particular class during P2 (see neurons *n* and *p*, in *SI Appendix*, Fig. S2). Finally, immediately after "pb", significant decision-related variance emerged (*Var Dec*). Notably, all these cognitive-dependent fluctuations vanished in the LCT (Fig. 2*B*). Given that the correct report is indicated by a persistent visual cue, only variances related to precise stimulus identity representation remained in LCT, depicted by increments of *Var P1* & *Var P2* in its corresponding periods. Further, *Var P1* dynamic across the delay and P2 vanished; *Var P2* was reduced noticeably; and *Var Dec* faded away. These results suggest that while phase-locking responses remained during the control condition, categorical and decision coding disappeared (18). This evidence reveals once again the true nature of S2, a colorful palette of sensory and cognitive demanding categorical dynamics, with a striking mixture in between.

Single-unit activities reported in Fig. 1 and *SI Appendix*, Fig. S1 revealed that a significant number of neurons exhibited a purely temporal, i.e., independent of task parameters, behavior. To investigate the amount of population variance associated with this kind of dynamics during TPDT and LCT, we calculated the instantaneous variance with respect to the temporal-averaged response (*Var Temp*, Fig. 2 *C* and *D*, green trace). Whenever the population activity deviates from its average value across the task period the temporal variance will increase. Interestingly, during TPDT, the instantaneous temporal population variance was much higher than the cognitive variances at any time bin. This indicates that most of the population variance over the course of a trial captured the variability across sequential events of the task rather than across classes. Additionally, under the LCT condition, the instantaneous temporal variance is even higher in comparison to TPDT. This indicates that noncoding (or temporal) signals are a crucial component of the network's behavior.

To further explore the nature of the coding signals, we categorized the recorded neurons based on their location in either the left (n = 1,162) or right (n = 484) brain hemispheres and then calculated the same population variance measures as Fig. 2 (*SI Appendix*, Fig. S4 *A* and *C*). Population variance measures were similar in the two hemispheres. There were however small differences as well. The left hemisphere (ipsilateral to stimulation) exhibited more pronounced sensory-shaped dynamics compared to the right hemisphere. In contrast, the right hemisphere displayed a notable increase in *Var Dec* relative to the left, which aligns with expectations given its ipsilateral position to the motor execution. The evaluation of condition-averaged dynamics (*Var Temp*) revealed no substantial differences between hemispheres, except for a minor increase in this metric during stimulus periods for the left hemisphere, and a more consistent activity after the pb event in the right hemisphere.

This evidence further underscores the multifaceted essence of S2, portraying it as a vibrant mosaic of sensory inputs and cognitively demanding categorical dynamics. The findings highlight a striking interplay between these elements, accentuated by the nuanced differences in hemisphere involvement, which adds depth to our understanding of S2's role in integrating and processing complex sensory and decision-related information.

**Orthogonal Sensory and Categorical Dynamics in S2.** To further explore the interaction between different types of coding across the heterogeneous responses of the neuronal population of S2, we applied dimensionality reduction methods to unravel the latent dynamics behind these responses. The S2 population response can be represented as a point evolving in time in a *n*-dimensional Euclidean space, where *n* is the number of recorded neurons (TPDT, n = 1,646; LCT, n = 313). We used the classical metapopulation approach, in which the firing rate of separately recorded neurons are combined into a population vector of firing rates (4, 5, 7, 8, 25, 26). As neuronal activity evolves over time, the point moves through the *n*-dimensional space, creating a trajectory that represents the population response. Principal component analysis (PCA) can be used to identify the linear low-dimensional subspace capturing the most salient features (principal components, PCs) of the population dynamic response. However, as we have pointed out, much of this population variance is related to temporal dynamics. Therefore, to study the different types of relevant coding across the population, it might be useful to find a way to diminish the impact of those noncoding signals. For this, one possible solution is to apply PCA to the population responses but limiting it to a narrow period of the task (5, 27). In this case, the temporal variance is much smaller, and the covariance matrix is determined by the coding correlation across neurons throughout this period. Therefore, we applied PCA to the population covariance matrices computed during the final 0.6 s of the stimulus periods: P1 and P2, as well as the last 1.5 s of the task, which corresponded to the push button period (Fig. 3 *A*–*C*, pink marker on top). Again, by focusing on a small period of the task, we can effectively reduce temporal variance and isolate key aspects of the dynamics. Subsequently, we projected neural activity during TPDT and LCT onto these PCs that were ordered according to their explained total variance.

This period-restricted PCA returned three sets of PCs, respectively associated with P1 (P1-PCs), P2 (P2-PCs), and push-button (pb-PCs). Only the significant PCs that explained more variance than a permuted covariance matrix are shown in Fig. 3 (*Materials and Methods*). Notice that in the control variant, only one PC is significant for both stimulus periods. When we projected the whole TPDT (–1 to 7.5 s) on P1-PCs and P2-PCs, we noticed that the

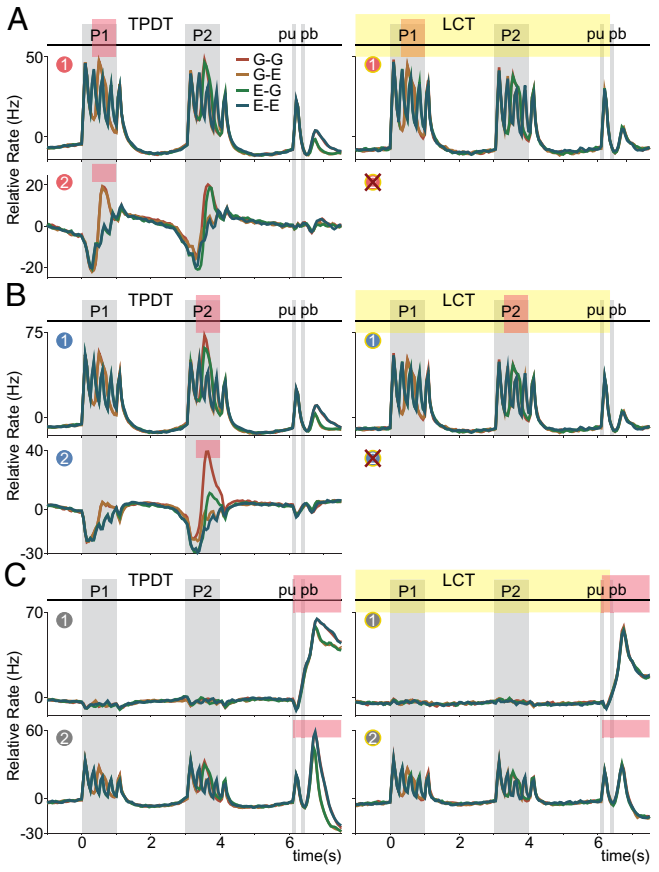


**Fig. 3.** Contextual orthogonality in population dynamics. (*A*–*C*) PCA was applied to the covariance matrices computed on neural activity during the last 0.6 s of P1 and P2, and 1.5 s of the push button period (pink markers on top), to derive associated PCs P1-PCs, P2-PCs, and pb-PCs. The activity from the whole (−1 to 7.5 s) TPDT (n = 1,646) and LCT (n = 313), sorted by class identity, was projected onto the resulting PC axes and ordered by their explained variance (ETV). In TPDT, only the first two PCs for each parameter (1 & 2 encircled numbers) were above the noise level, while in LCT this condition was met only for P1-PC1, P2-PC1, and the first two dec-PCs. (*A* and *B*) TPDT projections with P1-PC1 and P2-PC1 exhibited strong sensory dynamics, which were barely unchanged in the LCT, whereas dynamics obtained with P1-PC2 and P2-PC2 displayed more transformed and categorical-related responses. The latter ones completely disappeared in the LCT. (*C*) TPDT projections onto pb-PC1 and PC2 showed dynamics purely related with decision during the intertrial period. In LCT only the temporal component of this dynamics remained.

first PCs (Fig. 3 *A* and *B*, P1-PC1 & P2-PC1, respectively) captured strong sensory (phase-locked) signals. Conversely, the second principal components in both cases (Fig. 3 *A* and *B*, P1-PC2 & P2-PC2, respectively) encapsulated responses that do not faithfully represent sensory stimuli but rather discern some of their categorical properties. In particular, P1-PC2 discerned the extended or grouped identity of both the first and second stimulus, whereas P2-PC2 mostly discerned the class of the second stimulus by exhibiting a large response only in the G-G class. It is crucial to emphasize the dynamics driven by P2-PC2, which are closely associated with class coding (categorical coding), particularly during the P2 period (Fig. 3*B*). This pattern aligns with the increase in *Var Cod* observed in Fig. 2*A*, suggesting that these dynamics likely emerge for the purposes of comparison. Remarkably, despite the axis having been computed with an unsupervised approach, sensory and categorical signals manifested in different components. This means that the population of S2 neurons intrinsically represent the faithful and categorical responses as orthogonal dynamics. Notice, in particular, that during LCT categorical responses disappeared and only the faithful ones remained (*Right* in Fig. 3 *A* and *B*).

To further corroborate this separation between sensory and categorical dynamics, we applied PCA to the covariance matrix quantified from the whole (−1 s to 7.5 s) TPDT and LCT (*SI Appendix*, Fig. S3 *A* and *B*). At first glance, PCA revealed a complex picture of population dynamics, with both sensory and categorical signals distinctly emerging. Akin to those observed in Fig. 3, the first PC demonstrated a faithful sensory response, while the third PCs exhibited categorical dynamics (*SI Appendix*, Fig. S3*A*). This result supports the hypothesis that sensory and categorical signals are decoupled at the population level. In agreement with previous research, our results suggest that orthogonality in population dynamics might be employed by the S2 network during the TPDT (3, 4, 7). The resulting decoupled dynamics might act as an efficient mechanism to transform sensory inputs into categories while protecting from sensory interference. Further, when PCA was applied to the LCT (*SI Appendix*, Fig. S3*B*), only the sensory component and the dynamic during PB were significant with respect to noise. Analogously to what we observed in single units, PCs in LCT displayed either purely sensory (PC1) or temporal responses (PC2). Thus, the orthogonal categorical signals arise only when the task is cognitively demanding. Finally, when focusing on the last period of the task, pb-PCs (Fig. 3*C*, pb-PC1 & pb-PC2) displayed responses akin to PC2 & PC4 in *SI Appendix*, Fig. S3. These components exhibit a pronounced decision-related response, delineated by categorical (matched/nonmatched classes) distinctions following the pb event, which is in line with the significant increase in *Var Dec* observed in Fig. 2*A*. In the LCT, only the temporal component of the decision remained during the push-button period (Fig. 3*C* and *SI Appendix*, Fig. S3*B*), which confirms the reduced cognitive effort in the LCT.

The organization of sensory, categorical, and decision-related dynamics into orthogonal subspaces, as depicted in Fig. 3 and *SI Appendix*, Fig. S3, is similarly maintained across brain hemispheres at the population level (*SI Appendix*, Fig. S4 *B* and *D*). In both hemispheres, the 1st and 3rd PCs predominantly capture sensory and categorical dynamics, respectively, whereas the second PC reflects a combination of decision-related responses with a minor sensory component. Although no significant regional differences were observed in the types of emergent signals, the left hemisphere showed a modest enhancement in the intensity of both sensory and categorical dynamics. Conversely, the right hemisphere demonstrated more pronounced decision-related dynamics following the pb event. This observation aligns with the variance analysis in *SI Appendix*, Fig. S4 *A* and *C*, revealing a coherent picture across different levels of analysis.

Overall, the PCs resulting from the period-restricted PCA were similar to those in *SI Appendix*, Fig. S3 but they also amplified and untangled more clearly the salient population dynamic features in each period. This observation confirms that the orthogonal splitting of sensory and categorical responses is a robust and intrinsic feature of S2, and not an artifact of the chosen set of projection axes. These results are remarkable evidence for the organizing role of S2, in which a tangle of neuronal responses is tidied up at the population level to minimize interference and maximize codification.

**Orthogonal Dynamics in S2 Associated with the First Stimulus and Decision-Making.** To further dissect the low-dimensional orthogonal decoupling between different coding dynamics in a way that is more sensitive to task parameters, we performed a demixed PCA (dPCA). Unlike traditional PCA, dPCA both maximizes the captured variance and effectively separates the population response variability among the different task parameters (5, 8). Applying this approach, we were able to decompose the population dynamics into new dPCs. We performed four kinds of demixing: P1 and P2 stimulus identity, decision, and class. For this, we marginalized neural

activity with respect to each one of these parameters, and for the four marginalized data, we computed the covariance matrices throughout the whole TPDT (−1 to 7.5 s). It is important to stress that the condition-independent (see ref. 8) component was subtracted away from each marginalization. Therefore, the noncoding (or temporal) variance is removed to calculate the demixed axes.

In Fig. 4*A*, we show the population trajectories obtained from the projection over the first two dPCs associated with P1 (Fig. 4*A*, P1-dPC1 and P1-dPC2). Lines above traces highlight time intervals where P1 can reliably be decoded from single trials (*Materials and Methods*). The population demixed responses merged classes with identical first stimulus, i.e., c1 with c2 (P1 = G, orange and red traces) and c3 with c4 (P1 = E, blue and green traces). Arbitrarily, we chose c2 (orange) and c4 (blue) to be on top in Fig. 4*A*. In agreement with PCA, P1-dPC1 and P1-dPC2 captured strong sensory and categorical responses (class coding), respectively. Notice that due to the subtraction of the condition-independent signal, the P1-dPC1 sensory response displayed a fast fluctuation around zero. Interestingly, the categorical signal (P1-dPC2) encoding P1 identity, re-emerged during P2, arguably for comparison purposes. Thus, the same axis can be used to decode categorical information during the first stimulus and to get information about P1 during the second stimulus. P1-dPCA demonstrates how S2 can split encoding dynamics, i.e., by keeping the faithful representation of P1 and its abstract representation into orthogonal subspaces. Furthermore, this different method was able to demonstrate that P1 abstract coding during the first stimulus and its recalling during P2 are both decoupled with the pure sensory dynamics. When we analyzed the neural data marginalized with respect to P2 and identified the most significant axes (P2-dPC 1 & 2), we observed a scenario akin to the one described for P1-dPCs. In this scenario, the dynamics merge classes with identical second stimuli, i.e., c1 with c3 (P2 = G, red and green traces) and c2 with c4 (P2 = E, orange and blue traces). The choice of plotting c3 (green) and c4 (blue) on top in *SI Appendix*, Fig. S5*A* was totally arbitrary. P2-dPC1 predominantly captured sensory dynamics, which were especially pronounced during the P2 period. Meanwhile, P2-dPC2 revealed a more transformed response during the same task epoch (*SI Appendix*, Fig. S5*A*). It's worth noting that while these findings echo those seen with P1-dPC representations, they provide a sharper understanding of the task.

Marginalization with respect to decision (Fig. 4*B*), on the other hand, obviously lead to dPCs (dec-dPC1 and dec-dPC2) that merged classes with the same decision outcome (classes matched/nonmatched), i.e., c1(G-G) with c4(E-E) and c2 (G-E) with c3 (E-G). Thus, classes c1 and c4 (red and blue traces) as well as c2 and c3 (orange and green traces), overlapped with blue and green lines being on top without premeditation. Projections along decision-dPC1 strictly captured the match (c1 and c4) and nonmatch (c2 and c3) outcomes after the pb press. This result agrees with the decision coding dynamics observed in Fig. 2*A* during pb (pink line). We can hypothesize that this intertrial (after pb) signal might be relevant to learning processes and history-dependent changes in the population. Interestingly, dec-dPC2 revealed a decision coding that emerged during the comparison period and remained significant during the first period of the second delay. Notice that the decision variance in Fig. 2*A* became significant during an analogous period of the task. Notably, the fact that these dynamics exists in orthogonal subspaces would mean, again, that the S2 network employed decoupled population responses to code the decision during P2 and across the intertrial period. It is important to highlight that to obtain these orthogonal signals it was imperative to apply dPCA to emphasize the decision dynamics. The applying of PCA to the whole task (*SI Appendix*, Fig. S3) or to a narrow period (Fig. 3), would not have allowed us to visualize these decision responses since they are diluted with the noncoding or temporal dynamics. These results further suggest that S2 employs orthogonality to split decision signals associated with different cognitive processes.

To further elaborate on the previous concept and examine the transition of population response from P1 to decision coding, we investigated the dynamic interaction from each class throughout the task (5). By marginalizing the neural data according to classes (G-G, G-E, E-G, E-E), we identified the most significant components (dPCs) related to differences in classes—class-dPC 1 & 2 (Fig. 4*C*). Consistent with P1-dPCA and dec-dPCA, class-dPC1

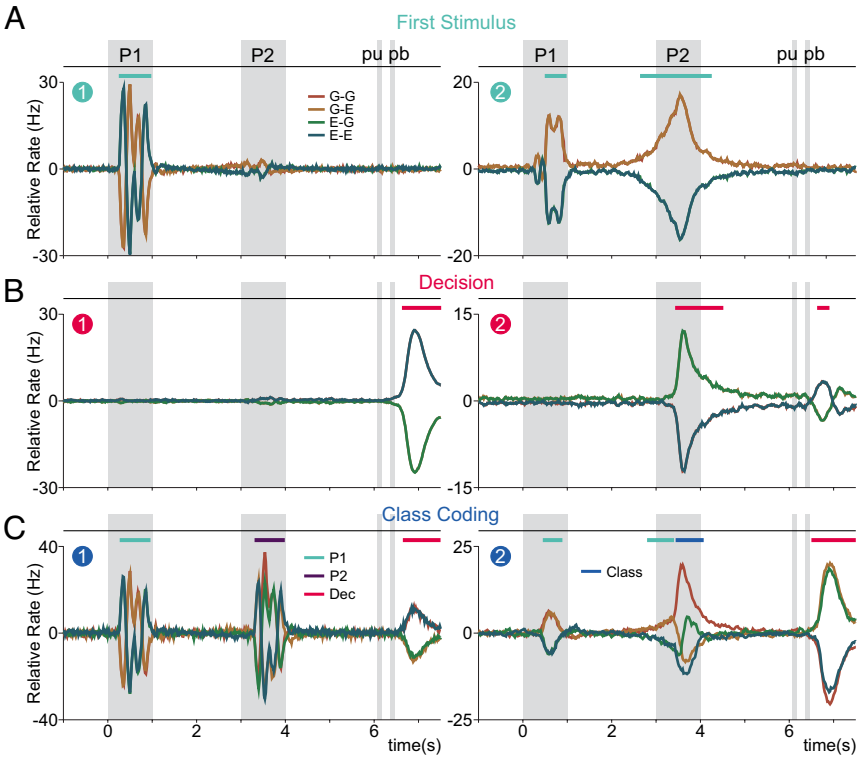


**Fig. 4.** Orthogonal population responses associated with task parameters. dPCA was applied to the marginalized covariance matrices from the whole TPDT (−1 to 7.5 s), with respect to P1 (*A*), decision (*B*), and class (*C*), to derive their associated dPCs. Population activity for TDPT (n = 1,646), sorted by class identity, was projected onto the resulting dPCs axes, and ordered by their EV (P1-dPC1: 6.2%, P1-dPC2: 4.8%; dec-dPC1: 9.1%, dec-dPC2: 4.4%; class-dPC1: 12.2%; class-dPC2: 11.3%). Only the first two dPCs for all marginalized parameters were above noise level. Color lines on top of each projection represent significant coding periods. With respect to P1, the projections exhibited purely sensory dynamics in the P1-dPC1 and categorical ones in the orthogonal component (P1-dPC2). When projecting onto dec-dPCs, signals purely related to decision outcome emerged after pb (dec-dPC1) and during P2 (dec-dPC2). Coherently, projections onto class-dPCs showed sensory-like dynamics in class-dPC1 (during P1 and P2) and a splitting of categorical signals in the orthogonal component (class-dPC2).

predominantly captured sensory dynamics, which were especially pronounced during the P1 and P2 periods. Different classes converge based on the identity of either the first or second stimulus, with traces of varying colors overlapping according to the chosen arbitrary plotting order. On the other hand, class-dPC2 illustrated the divergence of categorical signals, enabling us to observe the interaction of these dynamics within a distinct orthogonal subspace to assist the S2 network in task resolution. Through this analysis, we aim to underline the intrinsic segregation between sensory and categorical signals within the S2, showcasing this dynamic interplay in its clearest form.

In line with Fig. 2 and *SI Appendix*, Fig. S3, the application of dPCA to the LCT population revealed only one significant dPC associated with the stimuli P1 and P2, and class (*SI Appendix*, Fig. S5*B*). The resulting dynamics demonstrated sensory responses akin to those depicted for 1st P1-dPC, 1st class-dPC (Fig. 4 *A* and *C*), and 1st P2-dPC (*SI Appendix*, Fig. S5*A*). All traces matched and colors overlapped as previously described. This indicates the absence of categorical signals during LCT. Given that decision coding was not present during LCT, no decision dPCs were significant in this experimental context (*SI Appendix*, Fig. S5*B*).

Until now we have provided enough evidence about how the S2 network intertwines the mix of heterogeneous single responses into a complex and organized space where sensory and categorical signals are decoupled and maintained in orthogonal subspaces. Nevertheless, at this point, the reader might ask (as we did) whether some relation exists between those orthogonal components. In other words, do orthogonal encoding dynamics emerge from a common neuronal substrate or from separate groups of neurons? To initially address this matter, we plotted the neuronal weights from the most significant axes derived by applying PCA (PC1 & 2) or dPCA (P1-dPC1 and class-dPC1 & 2) to the entire TPDT and LCT datasets (*SI Appendix*, Fig. S6 *A* and *B*, respectively). We focused on class-dPCs1 & 2, given their ability to vividly capture the division of sensory and categorical dynamics into distinct orthogonal subspaces. We also selected the P1-dPC1 axis due to its relevance in showcasing sensory responses. In TDPT, the weights for both 1st and 2nd PCA and dPCA (class-dPC1 vs. class-dPC2 or class-dPC1 vs. P1-dPC1) revealed a lack of clusterization in neuronal activity, instead displaying a continuous, roughly uniform distribution. This continuity persisted even when combining axes from different dimensionality reduction methods, such as PC1 vs. class-dPC1, which capture the most pronounced sensory dynamics. Importantly, the differences between sensory class-dPC1 and PC1 are consequences of the marginalization with respect to the condition-independent signals (5, 8, 24), which explain most of the network variance (Fig. 2) and therefore are strongly represented in the PCs but not in the dPCs. Similarly, for the LCT, a continuum of responses was observed when plotting the weights of the most significant axes against one another, reinforcing the absence of any evident clustering in the neuronal responses across both tasks according to linear projection methods.

Up to this point, the results suggest that sensory and categorical dynamics in the S2 network might emerge from a continuous substrate of neuronal activity rather than separate clusters in the TPDT and LCT. Nevertheless, linear methods such PCA and dPCA, may lack the requisite power to discern distinct clusters within the neuronal activity. In light of this, we shifted our analytical focus toward a nonlinear method known for its robust capability in identifying clusters (28).

**A Continuum of Coding Responses in S2 Underlying Orthogonal Coding Dynamics.** To further address this problem, we used a nonlinear dimensionality reduction technique known as UMAP (29). Whereas linear dimensionality reduction methods like PCA and dPCA can only identify the most relevant linear subspaces underlying the organization of a set of *n*-dimension points, UMAP uses differential-geometric ideas to identify linear as well as nonlinear structures, that is topological manifolds, in the projected data. By inspection of the projected data, UMAP can thus be used to study the topology of a set of points and determine the number of connected components, that is, groups or clusters, in which the set is organized. This method was recently used in a variety of contexts, from sorting neuron types (30, 31) to uncovering the toroidal geometry underlying grid cell response (32). Moreover, we applied this nonlinear approach to the analysis of DPC population dynamics during a categorization task (1) and for establishing hierarchical responses across the somatosensory subareas (33). In this context, it is crucial to emphasize that our application of UMAP in this study is specifically aimed at examining the continuity within the neural substrate and identifying potential clusters of neurons responsible for the observed sensory and categorical dynamics. Unlike linear methods such as PCA and dPCA, UMAP is not and cannot be used to define explicit projection axes. Instead, UMAP's utility lies in its provable ability to detect clusters (or lack of), a capability where linear methods necessarily fall short in general (28).

When we applied UMAP to our S2 population, we either focused on the entire TPDT (−1 to 7.5 s, *SI Appendix*, Fig. S7*A*) or on ruling out the movement and intertrial periods (−1 to 6 s, Fig. 5*A*). Given the high variance observed at the end of each trial (Fig. 2), the latter option allowed us to better focus on responses associated with P1, P2, and class coding. To employ UMAP, we mapped each neuron to a point in dimension 4* *nt*, where *nt* is the number of time points, by concatenating its normalized firing rate vectors in the four classes. We then used UMAP to project the resulting set of points in two dimensions (Fig. 5 and *SI Appendix*, Fig. S7). The resulting projections revealed a connected, single-component continuum of responses, with no discrete groups or clusters. To characterize how responses varied across this continuum, we calculated a double-gaussian ($\sigma = 0.4$) weighted average of the activity of all neighboring neurons at certain points (colored X-marks) throughout the 2-D projection (Fig. 5*B* and *SI Appendix*, Fig. S7*B*). Although sensory, categorical, and decision response phenotypes can clearly be detected, they all belong to the same continuous deformation of responses across the projection plane. Notice that in the continuum, categorical signals (averages with purple and yellow crosses in Fig. 5*A*), are surrounded by different kinds of sensory and noncoding responses. Although these signals are averages in specific locations of the UMAP plane, they are strikingly akin to some of the exemplary neurons shown in Fig. 1 and *SI Appendix*, Figs. S1 and S2 (some neurons are *b*, *e*, *f*, and *h* in Fig. 1 and *d*, *f*, *n*, and *p* in *SI Appendix*, Fig. S1 and *e*, *g*, *l*, and *p* in *SI Appendix*, Fig. S2).

Conversely, UMAP projection of neuronal responses in the LCT (Fig. 6 and *SI Appendix*, Fig. S8) revealed a disconnected two-component structure where all that remained were decisional and sensory responses, akin to those seen in representative neurons in *SI Appendix*, Fig. S2 (LCT case; some neurons are *a*, *c*, *k*, and *m*). The disappearance of the categorical response region is seemingly the main cause underlying the disconnection of the neuronal response manifold in LCT. The contribution of neurons with categorical coding in the S2 network during TPDT thus appears to be crucial for the formation and maintenance of the observed continuum of neuronal responses. Here, we wish to emphasize the distinct advantage of UMAP in identifying context-dependent variations in the continuity of the neural substrate—a task at which linear methods do not (and in general, cannot) perform as effectively, as demonstrated in the scatter plots of neuronal weights in *SI Appendix*, Fig. S6*B*.

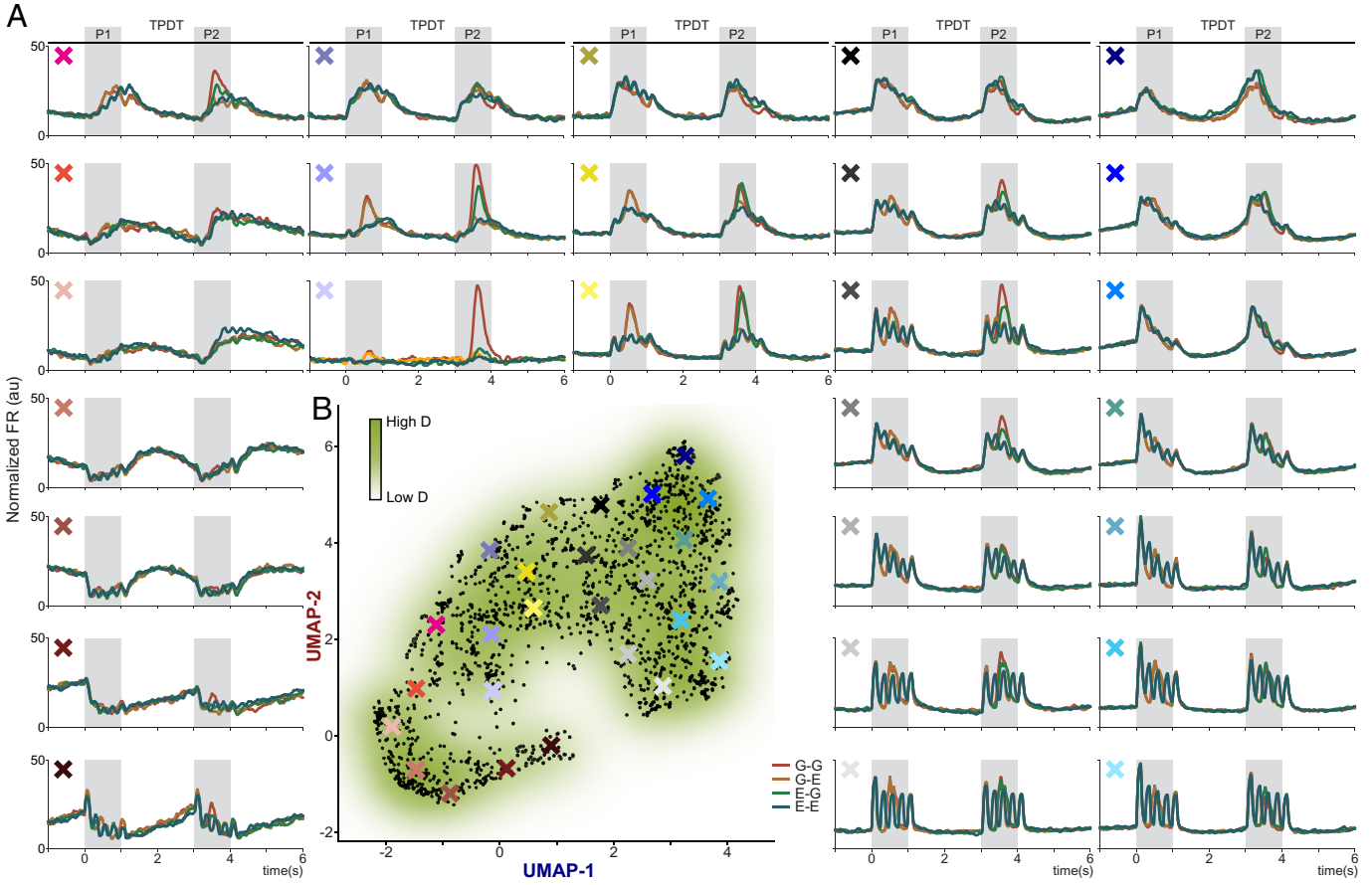


**Fig. 5.** A continuum of responses in S2: from sensory to categorical neurons during the TPDT. To focus on the sensory and categorical responses, we excluded the movement and intertrial periods. (*A*) Each subpanel represents a weighted average of neuronal activity and is presented in backtransform arbitrary units that roughly correspond to firing rate values (Hz). Notice that the exemplary averages range from sensory to more categorical and temporal dynamics (*Bottom*, *Middle*, and *Top* panels, respectively). (*B*) 2-D nonlinear UMAP density projection of the normalized activity of all units (n = 1,646, each black dot) recorded during TPDT (restricted from −1 to 6 s of the task) exhibited a continuum of responses. The colored X marks represent the centers of the double-Gaussian ($\sigma = 0.5$) weighted averages that are displayed in the panels in *A*.

In conclusion, UMAP revealed that the orthogonal representation identified by PCA and dPCA in the S2's population responses emerged from a unique continuous substrate which connects noncoding (temporal), categorical, and sensory single-unit responses. Notably, the disappearance of categorical signals broke up this substrate into distinct components during LCT. Our work underscores the importance of nonlinear analysis methods to understand the neural encoding of physical stimuli. Going deeper, we asked whether there is a network mechanism that could explain the modulation of the neuronal continuum between the TPDT and LCT. We approached this question using qualitative nonlinear dynamical systems modeling of neuronal population activity.

**Feedback Gain Modulation in a Network Model of the S2 As a Biologically Plausible Mechanism for the Emergence of Context-Dependent Information Representation.** We used modeling to uncover which mechanisms could explain context-dependent information representation in the S2. We relied on the established fact that the nonlinear recurrent neural network (RNN) response is largely determined and can be modulated by the configuration of its eigenmodes and their dynamical properties (34, 35). Eigenmodes are independent directions in the neural state space of a recurrently interconnected neuronal population that can be used to decompose the network dynamics into a sum of characteristic modal responses. Locally around a dynamical equilibrium of the network, eigenmodes are determined by the eigenvectors of linearized dynamics. Each eigenmode is associated with a time constant $\tau = 1/|\lambda|$ that determines its characteristic timescale of information integration (*SI Appendix*, Fig. S9), where $\lambda$ is the eigenvalue associated with the eigenvector generating the eigenmode. Fast (or faithful) eigenmodes (small $\tau$, large negative $\lambda$) respond rapidly to input and have a short memory, that is, their response converges back to a strongly attractive equilibrium once inputs are removed. Slow (or temporal) eigenmodes (large $\tau$, small negative $\lambda$) integrate inputs slowly and have a long memory, that is, their response largely outlasts input presentation. Finally, metastable/multistable (or categorical) eigenmodes (large $\tau$, small positive $\lambda$) are slow eigenmodes characterized by the existence of a threshold that distinguishes different input categories. The threshold of categorical eigenmodes is determined by the local dynamical instability brought by the small positive mode eigenvalue.

We hand-built (no automatic training used) a nonlinear RNN model with fully tunable eigenmodes (*Materials and Methods*). The model equations are provided in *SI Appendix*, Eq. **S11**. Each node $i$ in the network is associated with a state variable $x_i$ representing the overall activity of a neuron (or a neuronal population). A first network interaction matrix $A$ models how neurotransmitter liberation affects the node dynamics. In particular, $A_{ij}$ models the effect of the liberation of neurotransmitter $j$ on the activity of node $i$. A second network interaction matrix $B$ models how node activation affects neurotransmitter liberation. In particular, $B_{jk}$ models the effect of the activity of node $k$ on the liberation of neurotransmitter $j$. Because of the qualitative nature of our modeling approach, we did not explicitly distinguish between

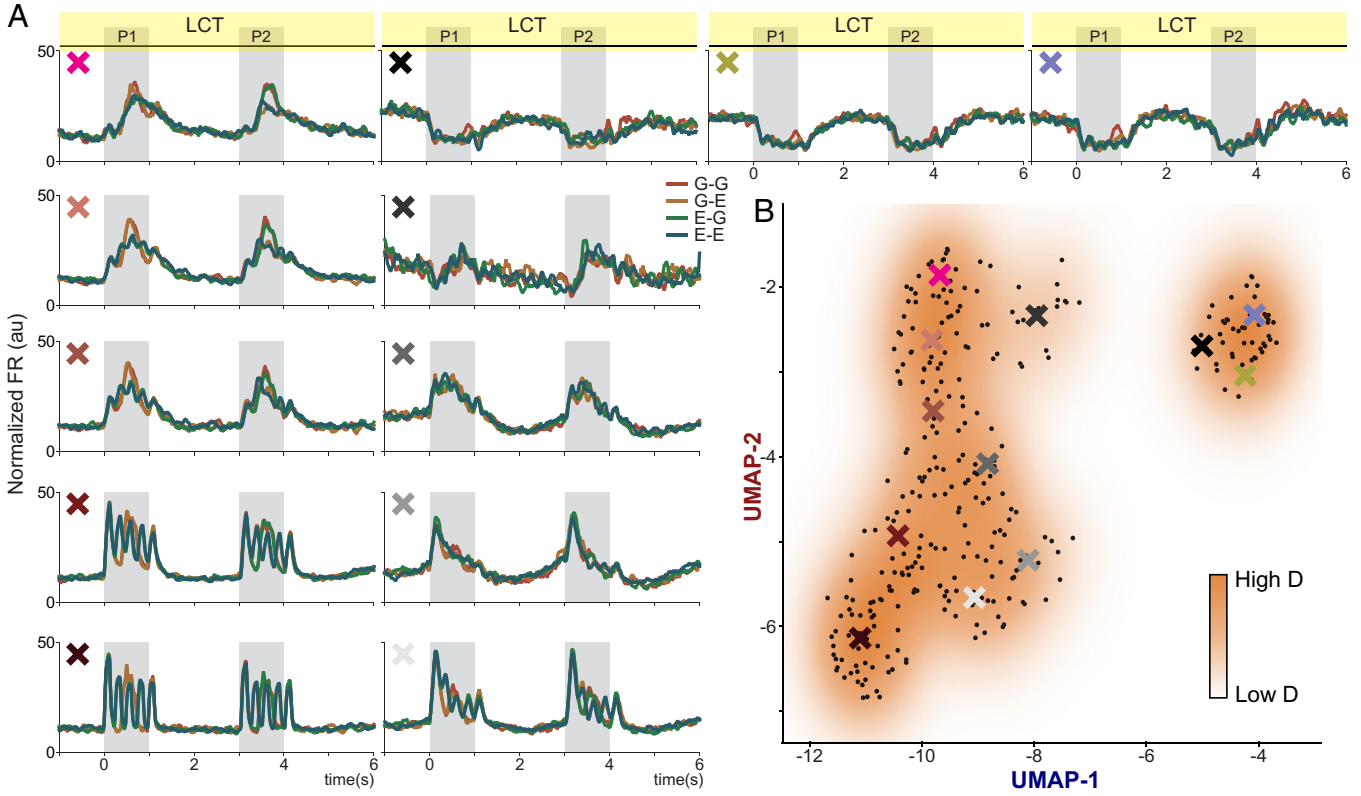


**Fig. 6.** Breaking the continuum in S2 during the nondemanding task. To focus on the sensory and categorical responses, we excluded the movement and intertrial periods. (*A*) Each subpanel represents the weighted average of neuronal activity in arbitrary units that approximately correspond to firing rate values (Hz). Given that LCT is a nondemanding task, the neurons that exhibited categorical dynamics during TPDT disappeared, leaving only sensory and temporal representations. (*B*) UMAP density projection of the normalized activity of all units (n = 313, each black dot) recorded during LCT (restricted from −1 to 6 s of the task) showed a breakup in the continuum that emerged during TPDT, exhibiting only sensory and temporal dynamics. The colored X marks the centers of double-Gaussian (σ = 0.5) weighted averages that are displayed in the panels in *A*.

excitatory and inhibitory neurotransmission, a subtlety that will be explored in future works. Finally, gains $k_j, j = 1, \ldots, N$, model presynaptic modulation: The larger $k_j$, the larger the release of neurotransmitter $j$ for the same presynaptic neural activation. With a minor modification to the model, the gains $k_j, j = 1, \ldots, N$, can equivalently model postsynaptic modulation, i.e., the larger $k_j$, the larger the effect that neurotransmitter release has on the activity of neuron (or neuronal population) $j$.

By tuning the strength of recurrent interconnections, gains $k_j, j = 1, \ldots, N$, are feedback gains that shape the dynamics and thus the information-processing of the network eigenmodes. In particular, by controlling the strength of recurrent amplification, these feedback gains control the local stability of the network eigenmodes, which in turn determines the mode encoding properties: A small feedback gain is associated with highly stable (fast/faithful) eigenmodes; a larger feedback gain is associated with weakly stable (temporal) or metastable/multistable (categorical) eigenmodes.

We hand-built a family of networks with a mixture of faithful, temporal, and categorical eigenmodes and tested their response to extended or grouped stimuli. The obtained responses were then merged into a metapopulation for analysis, as done in experiments. To introduce heterogeneity and variability, we generated faithful, temporal, and categorical eigenmodes by sampling the associated feedback gains according to distributions centered around a different characteristic value $k_0$ for each mode type (Fig. 7*A*). The components of the eigenvector associated with each eigenmode determined the selectivity (pure or mixed) of each modeled neuron. The final network consisted of a mixture of neurons with either pure (one third) or mixed (two third) selectivity (see *Materials and Methods* for details). Of the neurons with pure selectivity, half were purely temporal and half purely categorical. Neurons with mixed selectivity varied between predominantly faithful and predominantly categorical. Mixed selectivity between temporal and faithful and between temporal and sensory responses was not enforced because part of the faithful neurons, those associated with larger feedback gain, and part of the categorical neurons, those associated with smaller feedback gain, already exhibited purely temporal responses.

To model the transition from active (TPDT) to control (LCT) context as observed in experimental data (*SI Appendix*, Fig. S2), we shifted the distribution of feedback gains associated with categorical representations toward smaller values and overlapped it both with sensory and temporal feedback gain distributions (Fig. 7*A*), thus letting categorical representation effectively disappear from the modeled population dynamics. When we ran UMAP over the generated neuronal population in the active and control contexts we obtained the projections in Fig. 7*B*. Similarly to what was obtained on the S2 neuronal population (Figs. 5 and 6 and *SI Appendix*, Figs. S7 and S8), in the active context the simulated neuronal population activity was distributed along a connected manifold forming a continuum of activity types connecting purely faithful, temporal, and categorical responses. Conversely, in control context, the population activity manifold broke into two well-separated clusters due to the disappearance of the categorical response. It is important to stress that the breaking of the response continuum in control was not enforced ad hoc. Indeed, the feedback gain distributions determining the eigenmode family were more contiguous in the modeled control context as compared to active. It is the functional disappearance of the categorical response, bridging purely faithful to purely temporal responses in the active context, that breaks the response continuum in control.

Our model suggests a simple and general biological mechanism to explain the emergence of context-dependent information representation in the S2, namely, modulation of the feedback gains determining the strength of recurrent interconnection-mediated nonlinear amplification through pre- or postsynaptic modulation. More specifically, a neuromodulation-mediated decrease of the feedback gain of categorical eigenmodes is sufficient to remove the local dynamical instability determining their categorization threshold and transform them into either temporal or faithful eigenmodes. Both pre- and postsynaptic modulation are plausible biological candidates to realize such a feedback modulation mechanism. Both in biology and in modeling, the disappearance of categorical eigenmodes is the key factor determining the breaking of the response continuum.

## Discussion

Our research provides five critical insights into the functions and capabilities of the S2 neuronal population. First, in line with our previous studies, neurons in S2 displayed a multifaceted diversity of responses, characterized by a palette of faithful (sensory), categorical, and temporal signals. This diversity of neural substrates positions S2 as a key driver in sensory abstraction, possibly serving as a primary source of inputs for downstream hierarchical processes. Such observable heterogeneity and mixed selectivity also suggest S2's potential for highly efficient computations (36–38). Second, using two different contexts distinguished by their cognitive demand revealed that S2 context sensitivity is different as compared to other areas. Unlike in area 3b of S1, where only faithful and context-invariant responses are recorded (21), or in DPC, where the context-dependent coding primarily aims at comparing stimuli and forming decision-related responses, S2 displayed neurons with combined context-dependent and context-independent coding. While this coding versatility may partially stem from the array of responses mentioned earlier, it also hints at underlying complexities in S2 that go beyond single-unit response cataloging. When the task requires, S2 possesses the capability for managing and segregating diverse signals, but during a nondemanding task, this processing is turned off. Third, our population analysis not only corroborated this hypothesis but also unveiled a mechanism within the S2 network through which sensory and transformed dynamics are separated into orthogonal subspaces. Interestingly, this arrangement of signal splitting was observed in both brain hemispheres, with minor yet not significant differences in their lateralization relative to the stimulation. While similar mechanisms have been identified in other neuronal populations (2, 4, 7), we believe our findings present a network uniquely designed to differentiate between sensory and categorical dynamics. Fourth, the use of a nonlinear dimensionality reduction method tailored to identifying clusters (or lack of) in high-dimensional data structures, revealed that these dynamics arise from a context-dependent continuum of neural activity, rather than from isolated neuronal clusters. Thus, the identified orthogonal representation directions emerged from a single continuous substrate, where sensory, temporal (noncoding), and categorical responses are smoothly interconnected. Fifth, a network model suggested that the context-dependent differences observed in S2 population dynamics can be explained through a simple feedback gain-modulation mechanism. This observation provides well-known attentional mechanisms in the brain (39, 40) with a new mode of action, namely, switching on and off specific coding dynamics and shaping the neural manifold where neuronal population coding happens as a function of context.

The question of whether the brain actually employs orthogonality for dissociating signals at the population level or whether the evidence for orthogonal representations is due to the used dimensionality reduction techniques (41) has been a topic of rigorous discussion (3, 42). A recent study has shown that the orthogonal dynamics associated with sensory and mnemonic coding appeared during implicit task learning (4). Their results in ref. 4 suggest that orthogonal population dynamics are not present in naive animals, but they emerge through familiarization with the task. They also show that this mechanism might have its roots in individual neurons, specifically in those that either maintain ("stable neurons") or reverse ("switching neurons") their selectivity over time. The authors of that study found that the interplay between these distinct neuronal dynamics enables the rotation of population representations. This allows a unified network of neurons to effectively encode both sensory data and short-term memories, thereby optimizing computational efficiency. In S2, the decoupling of sensory and categorical dynamics into well-defined subspaces might have its origin in neurons which purely encode the different signals but also in those that exhibit mix dynamics which probably could mingle their level of coding between sensory and categorical. Thus, the S2 network would be able to dynamically dissociate what has already been transformed from the raw sensory input, reducing interference, and maximizing flexible computations. To support this finding, we applied period-restricted and whole task approaches in addition to parameter marginalized dimensionality reduction analysis. We found similar results with all these different approaches, suggesting that the orthogonal separation of the dynamics is an intrinsic feature of the S2 network, rather than an artifact arising from the selected projection axes.

Importantly, the emergence of the coding orthogonal subspaces may be related to the underlying network communication structure. Prior research proposed that the development of orthogonal, decoupled dynamics may act as "communication subspaces", serving as an effective framework for enhancing interarea neural communication (43). In their study, the authors showed that activity fluctuations in the secondary visual area (V2) are associated with a specific subset of activity patterns in the primary visual area (V1). Our work builds upon the strength of this study too, given that the uncoupled subspaces that emerged from our analysis could play a role to enhance communication with other brain areas. For instance, one might guess that sensory signals are involved in communication with areas of S1, categorical signals with frontal and premotor cortices, and push-button ones with the execution of decisions by motor cortices. Further research employing simultaneous recording among several areas is required to clarify the role of these orthogonal dynamics for communicating signals across the brain in the task employed here.

An intriguing observation that emerged from our population analysis is that orthogonal coding might also be used by the S2 network to differentiate decision signals. While there was significant decision-related coding observed both before and after push button movement (intertrial period), our dPCA results reveal that these two decision signals are also orthogonally arranged. This suggests that they may be involved in distinct functions. For instance, coding during the second stimulus may be related to the comparison, whereas postmovement activity might pertain to evaluating the reward associated with the decision (44). This means that S2 responses could be influenced by decision history, hinting at a more nuanced function for this area than previously suggested. Such a mechanism might form part of a feedback loop that continually refines the network's performance based on recent experiences. This process may also be facilitated by the unique continuum substrate of responses upon which the network is based. This opens an interesting hypothesis to be verified in future experiments studying the changes in S2 dynamics during learning.

Our last hypothesis opens, in turn, an intriguing question: What are the potential benefits of a network built upon a continuous substrate of responses? Might such a feature promote the development of orthogonal subspaces? To investigate the continuity of the

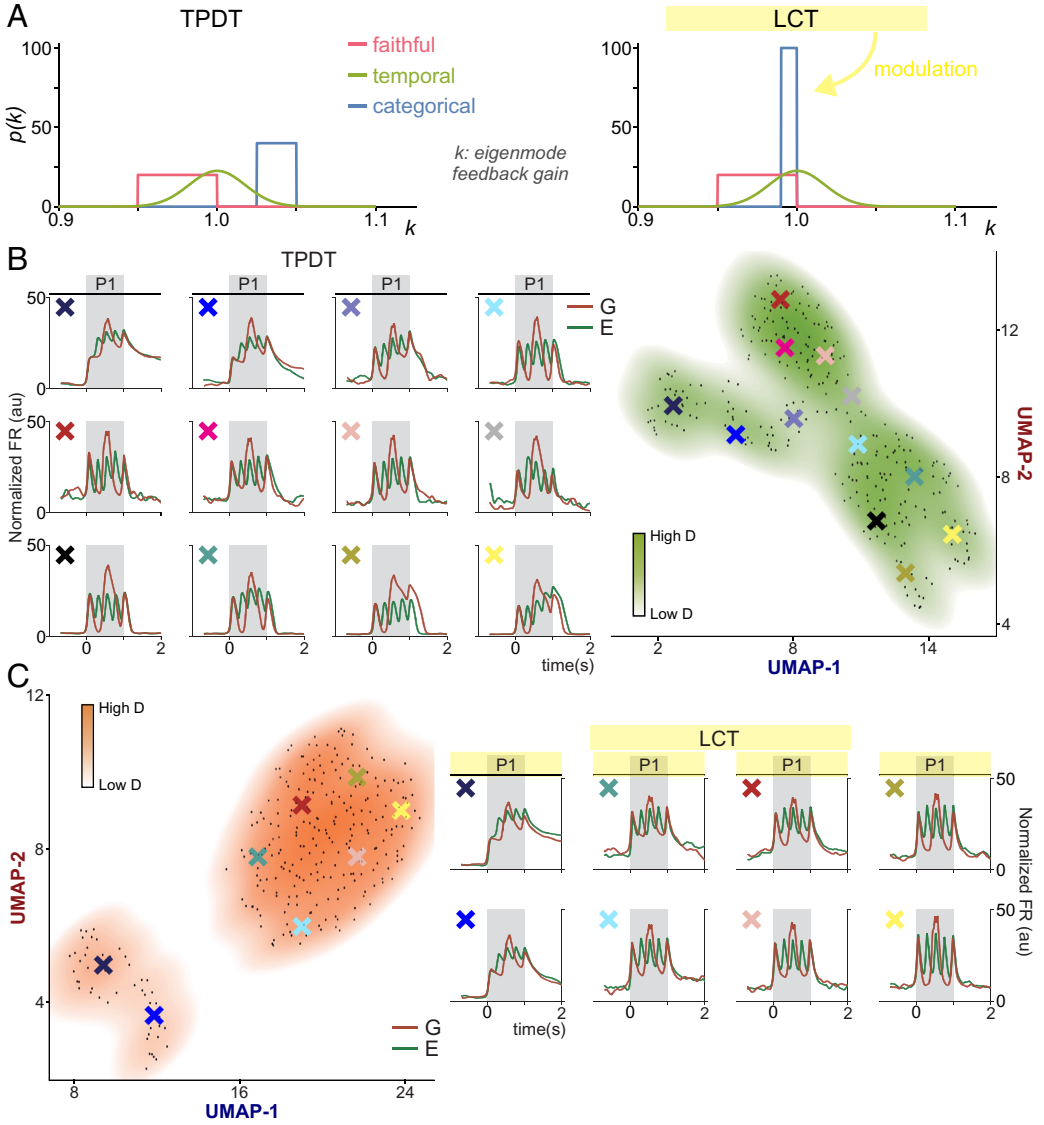


**Fig. 7.** A network model of the S2 neuronal population dynamics. The network reproduces the transition from a continuum of responses in TDPT to clustered responses in LCT. (*A*) Faithful, temporal, and categorical eigenmodes in our network were defined by different characteristic network feedback gains (see main text and *Materials and Methods*). For each eigenmode type, the defining feedback gain was sampled from the distribution sketched in the panel to create a family of heterogeneous eigenmodes over which the network structure was built resulting in a mixture of pure and mixed neurons. The transition from active to control was modeled by modulating the feedback gain distribution of the categorical eigenmodes as drawn. (*B*) Similar to Fig. 5 but for the network model. In active condition, the modeled S2 neural manifold exhibited a continuum of responses connecting faithful, temporal, and categorical modes. (*C*) Similar to Fig. 6 but for the network model. In LCT condition, categorical responses disappeared from the modeled S2 neural manifold, which led to its break-up into two sharply separated clusters of purely temporal and faithful responses.

neuronal encoding substrate, we employed UMAP. In the 2-D visualization, distinct patterns in averaged neural activity were evident (Fig. 5), with sensory and temporal signals flanking the categorical ones. Yet, density-based clustering on these UMAP projections did not identify any clear subpopulations. The property of the S2 network of having ordered but continuously connected patches of neuronal activity types is in line with recent findings where orthogonal encoding strategies were also uncovered. Recent research, where monkeys were trained to memorize sequential spatial locations of three visual stimuli, discovered that the frontal network encoded each stimulus position into a distinct orthogonal subspace (7). The authors analyzed the anatomical and functional organization of individual neurons and found, similar to us, a family of activity patches intertwined within a unified substrate. They posit that such spatial arrangement might offer advantages in positive feedback learning, coordinate transformations, and even in integrating multiple sensory modalities. Even if our recordings preclude spatial analysis of S2's neurons, we hypothesize that the lack of functional clusters may also be related to a structural continuity. We propose this line of study for future exploration.

The use of RNNs has emerged as a groundbreaking tool in cognitive neuroscience, offering deep insights into the mechanisms hidden within neural dynamics (13, 35, 45–47). Like many studies referenced here, our aim was to design an artificial model that mirrors the dynamics triggered by the initial stimulus in our task. By adjusting the strength of positive feedback gains, we successfully shaped the coding dynamics based on context. However, as the reader may also ask: How might this mechanism operate within a biological substrate as the S2 network? One hypothesis is drawn from local field potentials (LFPs). In a visual task, it was observed that distinct propagation directions of alpha and gamma oscillations provide an efficient feedback and feedforward control mechanism that regulates information transmission (48). In that research, the authors discovered that stimulation of the deeper layers prompted the spread of alpha oscillations to V1, dampening activity in this region. This gain modulation associated with LFP bands has been also suggested by subsequent studies (49). Parallel research in a somatosensory frequency discrimination task has also explored the behavior of alpha oscillations across cortical areas (50). The study found that lower alpha power correlated with increased firing rates in several areas, including S2. Particularly, S2 exhibited an increase in alpha power during the beginning of the working memory period which corresponded with a decrease in activity and coding in this network. This means that the modulation in alpha power might serve as a feedback mechanism, facilitated by higher-order areas, optimizing the processing of relevant sensory data. While in this study we did not center on examining oscillatory activity in S2, the findings presented thus far open promising avenues for future research within this task.

In sum, our exploration into the intricacies of S2 has unveiled a dynamic landscape of neural responses, adaptive behaviors associated with the S2 population responses, and shows an innovative mechanism underpinning sensory abstraction and processing. The identification of a inherent coding orthogonalization strategy, the intricate continuum of neural activity, and the nuanced interplay of decision-related signals all attest to the rich complexity of S2. As we build upon prior findings and integrate advanced methodologies like RNNs, we are not only exploring uncharted territories in our understanding of neural dynamics but also shaping the trajectory of future neuroscience research. The complexities of S2, as showcased, hold profound implications, offering fresh perspectives on the brain's computational prowess and adaptability.

## Materials and Methods

Two monkeys were trained to report whether the temporal structure of two vibrotactile stimuli of equal frequency was the same or different (Fig. 1*A* and *SI Appendix*). Neuronal recordings were obtained in the S2 while the monkeys performed TPDT. Animals were handled in accordance with standards of the NIH and Society for Neuroscience. All protocols were approved by the Institutional Animal Care and Use Committee of the Instituto de Fisiología Celular, Universidad Nacional Autónoma de México.

**Data, Materials, and Software Availability.** The custom python and C scripts employed in the analysis of these data, as well as the experimental protocols, are available from the corresponding authors on reasonable request. Data files are publicly available at Zenodo (DOI: 10.5281/zenodo.4421855); see ref. 51. Source data are provided with this paper.

**ACKNOWLEDGMENTS.** We thank Hector Diaz for his technical assistance. This work was supported by grants PAPIIT-IN205022 from the Dirección de Asuntos del Personal Académico de la Universidad Nacional Autónoma de México (to R.R.-P.) and CONAHCYT-319347 (to R.R.-P.) from Consejo Nacional de Ciencia y Tecnología; International Brain Research Organization (IBRO) Early Career Award 2022 (to R.R.-P.) from International Brain Research Association. L.B. is a postdoctoral student (Postdoctoral fellowship CONACYT-838783).

Author affiliations: [a]Instituto de Fisiología Celular, Departamento de Neurociencia Cognitiva, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; [b]El Colegio Nacional, Mexico City 06020, Mexico; [c]Departmento de Matemática, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; [d]Montefiore Institute, University of Liège, Liège 4000, Belgique; [e]Wallon ExceLlence (WEL) Research Institute, Wavre 1300, Belgique; and [f]Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico

Author contributions: R.R., A.F., and R.R.-P. designed research; A.Z., M.A., and R.R. performed research; L.B., A.F., and R.R.-P. analyzed data; and L.B., R.R., A.F., and R.R.-P. wrote the paper.

1. G. Diaz-deLeon *et al.*, An abstract categorical decision code in dorsal premotor cortex. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2214562119 (2022).
2. M. T. Kaufman, M. M. Churchland, S. I. Ryu, K. V. Shenoy, Cortical activity in the null space: Permitting preparation without movement. *Nat. Neurosci.* **17**, 440–448 (2014).
3. G. F. Elsayed, A. H. Lara, M. T. Kaufman, M. M. Churchland, J. P. Cunningham, Reorganization between preparatory and movement population responses in motor cortex. *Nat. Commun.* **7**, 13239 (2016).
4. A. Libby, T. J. Buschman, Rotational dynamics reduce interference between sensory and memory representations. *Nat. Neurosci.* **24**, 715–726 (2021).
5. R. Rossi-Pool *et al.*, Decoding a decision process in the neuronal population of dorsal premotor cortex. *Neuron* **96**, 1432–1446.e7 (2017).
6. B. M. Dekleva *et al.*, Motor cortex retains and reorients neural dynamics during motor imagery. *Nat. Hum. Behav.* **8**, 729–742 (2024).
7. Y. Xie *et al.*, Geometry of sequence working memory in macaque prefrontal cortex. *Science* **375**, 632–639 (2022).
8. D. Kobak *et al.*, Demixed principal component analysis of neural population data. *Elife* **5**, e10989 (2016).
9. A. S. Morcos, C. D. Harvey, History-dependent variability in population dynamics during evidence accumulation in cortex. *Nat. Neurosci.* **19**, 1672–1681 (2016).
10. S. A. Koay, A. S. Charles, S. Y. Thiberge, C. D. Brody, D. W. Tank, Sequential and efficient neural-population coding of complex task information. *Neuron* **110**, 328–349.e11 (2022).
11. T. Flesch, K. Juechems, T. Dumbalska, A. Saxe, C. Summerfield, Orthogonal representations for robust context-dependent task performance in brains and neural networks. *Neuron* **110**, 1258–1270.e11 (2022).
12. B. DePasquale, D. Sussillo, L. F. Abbott, M. M. Churchland, The centrality of population-level factors to network computation is demonstrated by a versatile approach for training spiking networks. *Neuron* **111**, 631–649.e10 (2023).
13. J. Wang, D. Narain, E. A. Hosseini, M. Jazayeri, Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018).
14. 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).
15. M. Beiran, N. Meirhaeghe, H. Sohn, M. Jazayeri, S. Ostojic, Parametric control of flexible timing through low-dimensional neural manifolds. *Neuron* **111**, 739–753.e8 (2023).
16. 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).
17. O. I. Rumyantsev *et al.*, Fundamental bounds on the fidelity of sensory cortical coding. *Nature* **580**, 100–105 (2020).
18. R. Rossi-Pool, A. Zainos, M. Alvarez, G. Diaz-deLeon, R. Romo, A continuum of invariant sensory and behavioral-context perceptual coding in secondary somatosensory cortex. *Nat. Commun.* **12**, 2000 (2021).
19. E. Salinas, A. Hernández, A. Zainos, R. Romo, Periodicity and firing rate as candidate neural codes for the frequency of vibrotactile stimuli. *J. Neurosci.* **20**, 5503–5515 (2000).
20. R. Romo, A. Hernández, A. Zainos, L. Lemus, C. D. Brody, Neuronal correlates of decision-making in secondary somatosensory cortex. *Nat. Neurosci.* **5**, 1217–1225 (2002).
21. R. Rossi-Pool *et al.*, Emergence of an abstract categorical code enabling the discrimination of temporally structured tactile stimuli. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7966–E7975 (2016).
22. F. Carnevale, V. de Lafuente, R. Romo, N. Parga, An optimal decision population code that accounts for correlated variability unambiguously predicts a subject's choice. *Neuron* **80**, 1532–1543 (2013).
23. R. Romo, A. Hernández, A. Zainos, E. Salinas, Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron* **38**, 649–657 (2003).
24. R. Rossi-Pool *et al.*, Temporal signals underlying a cognitive process in the dorsal premotor cortex. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7523–7532 (2019).
25. G. F. Elsayed, J. P. Cunningham, Structure in neural population recordings: An expected byproduct of simpler phenomena? *Nat. Neurosci.* **20**, 1310–1318 (2017).
26. M. C. Aoi, V. Mante, J. W. Pillow, Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nat. Neurosci.* **23**, 1410–1420 (2020).
27. J. D. Murray *et al.*, 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).
28. Y. Yang *et al.*, Dimensionality reduction by UMAP reinforces sample heterogeneity analysis in bulk transcriptomic data. *Cell Rep.* **36**, 109442 (2021).
29. L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
30. E. K. Lee *et al.*, Non-linear dimensionality reduction on extracellular waveforms reveals cell type diversity in premotor cortex. *Elife* **10**, e67490 (2021).
31. E. Becht *et al.*, Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2019).
32. R. J. Gardner *et al.*, Toroidal topology of population activity in grid cells. *Nature* **602**, 123–128 (2022).
33. S. Parra *et al.*, Hierarchical unimodal processing within the primary somatosensory cortex during a bimodal detection task. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2213847119 (2022), 10.1073/pnas.2213847119.
34. A. Chadwick *et al.*, Learning shapes cortical dynamics to enhance integration of relevant sensory input. *Neuron* **111**, 106–120.e10 (2023).
35. V. Mante, D. Sussillo, K. V. Shenoy, W. T. Newsome, Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
36. S. Kira, H. Safaai, A. S. Morcos, S. Panzeri, C. D. Harvey, A distributed and efficient population code of mixed selectivity neurons for flexible navigation decisions. *Nat. Commun.* **14**, 2121 (2023).
37. S. Fusi, E. K. Miller, M. Rigotti, Why neurons mix: High dimensionality for higher cognition. *Curr. Opin. Neurobiol.* **37**, 66–74 (2016).
38. M. Rigotti *et al.*, The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013).
39. E. Salinas, P. Thier, Gain modulation: A major computational principle of the central nervous system. *Neuron* **27**, 15–21 (2000).
40. M. Carandini, D. J. Heeger, Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* **13**, 51–62 (2012).
41. S. Ganguli, H. Sompolinsky, Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis. *Annu. Rev. Neurosci.* **35**, 485–508 (2012).
42. D. Badre, A. Bhandari, H. Keglovits, A. Kikumoto, The dimensionality of neural representations for control. *Curr. Opin. Behav. Sci.* **38**, 20–28 (2021).
43. J. D. Semedo, A. Zandvakili, C. K. Machens, B. M. Yu, A. Kohn, Cortical areas interact through a communication subspace. *Neuron* **102**, 249–259.e4 (2019).
44. A. Akrami, C. D. Kopec, M. E. Diamond, C. D. Brody, Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature* **554**, 368–372 (2018).
45. 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).
46. Z. Bi, C. Zhou, Understanding the computation of time using neural network models. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 10530–10540 (2020).
47. A. Barri, M. T. Wiechert, M. Jazayeri, D. A. DiGregorio, Synaptic basis of a sub-second representation of time in a neural circuit model. *Nat. Commun.* **13**, 7902 (2022).
48. T. van Kerkoerle *et al.*, Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 14332–14341 (2014).
49. J. Ni *et al.*, Gamma-rhythmic gain modulation. *Neuron* **92**, 240–251 (2016).
50. S. Haegens, V. Nácher, R. Luna, R. Romo, O. Jensen, α-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 19377–19382 (2011).
51. A. Zainos *et al.*, Single neuron activity of S1, S2 and DPC of Macaca Mulatta subjects performing the temporal pattern discrimination task. Zenodo (2021). https://doi.org/10.5281/ZENODO.4421855. Accessed 10 February 2021.