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

Lucas Bayones[1], Antonio Zainos[1], Manuel Alvarez[1], Ranulfo Romo[a,1,2*], Alessio Franci[3,4*], Román Rossi-Pool[1,5*]

[1]Instituto de Fisiología Celular ─ Neurociencias, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico

[2]El Colegio Nacional, 06020 Mexico City, Mexico

[3]Montefiore Institute, University of Liège, Av de la Découverte 10, 4000 Liège, Belgique

[4]WEL Research Institute, avenue Pasteur, 6, 1300 Wavre, Belgique

[5]Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, Mexico City, Mexico.

*Correspondence: ranulfo.romo@gmail.com (R.R.); afranci@uliege.be (A.F.); romanr@ifc.unam.mx (R.R.-P.)

[a]Present address: El Colegio Nacional, 06020 Mexico City, Mexico

Abstract

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 (ANNs) 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 discrimination pattern task (TDPT) 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 non-demanding version of the 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.

SIGNIFICANCE STATEMENT

An important function of the brain is turning sensation into perception. Yet, how this function is implemented remains unknown. Current research, coupled with insights from artificial neural networks, highlights using of orthogonal representations as an effective means to transform sensory signals into perceptual signals while separating and simultaneously processing the two information streams. Neuronal recordings in second somatosensory cortex (S2), obtained from monkeys performing a demanding vibrotactile discrimination task, revealed that this function is implemented at the population level. While S2 consistently encodes sensory inputs, the representation of transformed signals is contingent on task demand. Such distinct and flexible organization, enriched by a broad spectrum of neural activities, reflects the brain's efficiency, resilience, and overall purpose for solving cognitive tasks.

Introduction

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 inter-area transmission of both kinds of signals in a context-dependent fashion?

Research on behaving monkeys has brought forth an attractive hypothesis about the brain's unique encoding mechanisms (1–5). It suggests 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 related to stimulus representation and comparison (4, 5). Furthermore, it also permitted the network to separate signals associated with motor planning and motor execution (1–3). Moreover, orthogonal subspaces were also found in tasks involving sequential memory representation (6) and between signals that encode reward and confidence (7). Further research in monkeys has revealed that the neocortical neuronal network can represent multiple related task variables without cross-interference (6). Additionally, studies conducted in the posterior parietal cortex of rodents, demonstrated that the activity patterns of neuronal groups during the evidence accumulation process are significantly influenced by past events (8). Here, the neuronal network employs a degree of orthogonalization 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 (9). Therefore, the notion of orthogonal representation 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 artificial intelligence, 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-specific tasks (10). Recent studies in spiking neuronal networks indicate that leveraging latent orthogonal subspaces can boost learning versatility (11). 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 (12–14), as well as in category dynamics during delay periods (15). 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 (16). 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 and categorical dynamics. Among the variety of brain areas, the second somatosensory cortex (S2) constitutes a promising circuit to examine this issue. It has been shown that S2 exhibits a wide variety of responses, ranging from faithful/phase lock to transformed/categorical ones, and exploring how these signals might interact during demanding and non-demanding tasks could provide valuable insights about the mechanisms underlying sensory representation and its transformation to reduce interference between dynamics. 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 insights that could potentially overcome the constraints of previous studies that used single units in S2 (17–19). Here, we employed a temporal discrimination pattern task (TDPT), in which monkeys had to compare whether the temporal pattern of two stimuli sequentially presented were equal or different. By using population analysis, we found that faithful and transformed dynamics appeared decoupled in the S2 network. Hence, decorrelated combinations of neural responses are necessary to decode sensory inputs and categorical and more abstract signals. Moreover, we found that these dynamics in the S2 network emerged from a continuum of responses rather than from isolated clusters of neurons. Remarkably, during the control and non-demanding task, the categorical population signal vanished. This means that orthogonal transformed population responses only emerge when it is required. 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 (20), 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. 1A, see Methods). Two possible temporal patterns were presented: extended (E), made of 5 pulses uniformly spaced in time, and grouped (G), with 3 pulses grouped at the middle of the stimulus. There were therefore 4 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. 1B).

We recorded the activity from 1646 neurons in S2 (Fig. 1C, see Methods) during the TPDT (Monkey RR17, n=1035; Monkey RR20, n=611). The responses of 12 and 24 exemplary S2 neurons are displayed in Fig. 1D and Supplementary Fig. 1, respectively. As we previously showed (17), 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. 1D is *a*, and in Fig. S1 are *a* & *b*). These neurons were therefore cataloged as “sensory” or “phase locked”.

Other recorded neurons, cataloged as "categorical", exhibited a transformed non-sensory kind of response coding some abstract features of the task. If we look at these units, they tend to respond more strongly to specific stimulus patterns (G or E) or particular classes (some examples in Fig. 1D are *c, e*, *f* & *h*, and in Fig. S1 are *l, j, p* & *q*). For example, unit *b* in Fig. 1D 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. 1D, 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 Fig. S1, faithfully represented all classes during 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* & *j* in Fig. 1D and units *m, f,* and *s* in Fig. S1. If a sensory or categorical neuron firing rate increased (or decreased) in response to stimulation, it was labeled as positive (or negative, Fig. S1 *k* & *u* and Fig. S2 *c*). It has been previously reported that the integration of positive and negative S2 neuronal pools improve stimulus information decoding (21, 22). Finally, a fourth set of neurons exhibited strong temporal dynamics (Fig. 1D *k* & *l*, and Fig. S1 *v, w* & *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 (23). 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.

Many of the S2 neurons that were recorded during TPDT, were also recorded during the light control task (LCT, n=313, Monkey RR17, n=189; Monkey RR20, n=124), a non-demanding variant of the active task. In each trial, monkeys received the same stimuli as in the TPDT, but the correct choice was indicated by a visual cue (see "Methods"). The animal's performance was 100% in the control variant (Fig. 1B, yellow bar), confirming the lack of any cognitive challenge. Supplementary Figure 2 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 (Fig. S2 *a* & *c*), in line with what we recently reported (17). On the other hand, the response of purely categorical units in TPDT became completely unresponsive to task parameters in LCT (Fig.S2, *g, k, n* & *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 (20). 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 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 provides direct support for the hypothesis that categorical coding emerges only under a cognitive-demandant environment (TPDT), while sensory coding remains unaltered and therefore, independently of the context. In other words, these processes might be related to perception.

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 & B, blue trace).

This metric quantifies how different firing rate activity is among classes and neurons at each time point: a high coding variance means that different task parameters lead to sharply different neuronal activity. We also calculated the response variance associated with each task parameter: P1 (*Var P1*, Fig. 2 A & B, light blue trace), P2 (*Var P2*, Fig. 2 A & B, purple trace) and decision period (*Var Dec*, Fig. 2 A & B, pink trace). 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 to 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 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 (23), 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 Fig. S2). Finally, immediately after "pb", significant decision-related variance emerged (*Var Dec*). Notably, all these cognitive-dependent fluctuations vanished in the LCT (Fig. 2B). 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 (17). 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 Figs. 1 and 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 time-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 means that non-coding (or temporal) signals constitute an essential part of the network dynamics.

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=1646; LCT, n=313). We used the classical meta-population approach, in which the firing rate of separately recorded neurons are combined into a population vector of firing rates (4–7, 24, 25). 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 non-coding 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, 26). 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 seconds of the stimulus periods: P1 and P2, as well as the last 1.5 seconds of the task, which corresponded to the push button period (Fig. 3 A, B & 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 analysis 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 (see Methods). Notice that in the control variant only one PC is significant for both stimulus periods. When we projected the whole TPDT (-1 to 7.5s) on P1-PCs and P2-PCs, we noticed that the first PCs (Fig. 3 A & B, P1-PC1 & P2-PC1, respectively) captured strong sensory (phase-locked) signals. Conversely, both second PCs (Fig. 3 A & B, P1-PC2 & P2-PC2, respectively) captured categorical responses. It is important to stress that, in concordance with Fig. 2, the 2nd P2-PC exhibited a categorical separation among classes (Fig. 3B). Therefore, while 2nd P1-PC is related to P1 and P2 identity during their corresponding stimulus period, the 2nd P2-PC is related to class identity during P2. 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 these two signals as orthogonal dynamics: the faithful sensory input and its transformed–response. Notice how, during the control task the categorical responses disappeared and only the sensory dynamics remained (right panels in Fig. 3 A & B).

To further corroborate this separation between sensory and categorical dynamics, we applied PCA to the covariance matrix quantified from the whole (-1s to 7.5s) TPDT and LCT (Fig. S3 A & 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 a categorical transformation (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, 6). 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 (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.

Focusing on the last period of the task, pb-PCs (Fig.3C, pb-PC1 & pb-PC2) displayed responses akin to PC2 & PC4 in Fig. S3. These components demonstrate a marked decision-related response after the pb event, which is in line with the significant increase in *Var Dec* observed in Fig. 2A. In

the LCT, only the temporal component of the decision remained during the push-button period (Fig. 3C & Fig. S3B), which confirms the lack of any cognitive effort in the LCT.

Overall, the PCs resulting from the period-restricted PCA were similar to those in 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 principal component analysis (dPCA). Unlike traditional PCA, dPCA both maximizes the captured variance and effectively separates the population response variability among the different task parameters (5, 7). Applying this approach, we were able to decompose the population dynamics into new dPCs. We performed two kinds of demixing: P1 stimulus identity and decision. For this, we marginalized neural activity with respect to each one of these parameters and for the two 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 (7)] component was subtracted away from each marginalization. Therefore, the non-coding (or temporal) variance is removed to calculate the demixed axes.

In Figure 4A, we show the population trajectories obtained from the projection over the first two dPCs associated with P1 (Fig. 4A, P1-dPC1 and P1-dPC2). Lines above traces highlight time intervals where P1 can reliably be decoded from single trials (Methods). The population demixed responses merged classes with identical first stimulus, i.e., c1 with c2 (P1 = G) and c3 with c4 (P1 = E). In agreement with PCA analysis, P1-dPC1 and P1-dPC2 captured strong sensory and categorical responses, 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-demixed PCA 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.

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, i.e., c1(G-G) with c4(E-E) and c2 (G-E) with c3 (E-G). Projections along decision-dPC1 strictly captured the match (c1 & c4) and non-match (c2 & c3) outcomes after the pb press. This result agrees with the decision coding dynamics observed in Fig. 2A during pb (pink line). We can hypothesize that this inter-trial (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. 2A 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 inter-trial period. It is important to highlight that to obtain these orthogonal signals it was imperative to apply dPCA to emphasize on the decision dynamics. The applying of PCA to the whole task (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 non-coding or temporal dynamics. These results further suggest that S2 employs orthogonality to split decision signals associated with different cognitive processes.

Analogously to Fig. S3 and in agreement to Fig. 2, when we applied dPCA to the LCT population we only observed one significant P1-dPCA associated to sensory responses (analogous to the shown in the left panel of Fig. 4A). That is, categorical P1 responses disappeared. As decision coding was absent during LCT (pink line in Fig. 2B), none dec-dPCs were significant under this experimental condition. For conciseness, we prefer not to show the only significant P1-dPCA associated with sensory responses.

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 their transformed signals are decoupled and maintained in orthogonal subspaces. Nevertheless, at this point the reader might ask (as we did) if 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?

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 Uniform Manifold Approximation and Projection (UMAP) (27). 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 (28, 29) to uncovering the toroidal geometry underlying grid cell response (30). Moreover, we applied this non-linear approach to the analysis of DPC population dynamics during a categorization task (1) and for establishing hierarchical responses across the somatosensory sub-areas (31).

When we applied UMAP to our S2 population, we either focused on the entire TPDT (-1 to 7.5 seconds, Fig. S4A) or on ruling out the movement and inter-trial periods (-1 to 6 s, Fig. 5A). 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 firing rate vectors in the four classes. We then used UMAP to project the resulting set of points in two dimensions (Fig. 5 & S4). 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 (σ=0.4) weighted-average of the activity of all neighboring neurons at certain points (colored X-marks) throughout the 2D projection (Fig. 5B & S4B). 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. 5A), are surrounded by different kinds of sensory and non-coding responses. Despite these signals are averages in specific locations of the UMAP plane, they are strikingly akin to some of the exemplary neurons shown in Figs. 1, S1 and S2 (some neurons are *b*, *e*, *f*, *h* in Fig. 1, *d*, *f*, *n*, *p* in Fig. S1 and *e*, *g*, *l*, *p* in Fig. S2).

Conversely, UMAP projection of neuronal responses in the LCT (Fig. 6 & S5) revealed a disconnected two-component structure where all that remained were decisional and sensory responses, akin to those seen in representative neurons in Fig. S2 (LCT case; some neurons are *a*, *c*, *k*, *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.

In conclusion, UMAP revealed that the orthogonal representation identified by PCA and dPCA analysis from in the S2 population responses emerged from a unique continuous substrate which connects non-coding (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 if 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.

A network model reproduces the emergence in S2 of the continuous neuronal manifold during TPDT and its breakup during LCT through feedback gain modulation

The model we used to describe S2 neuronal population activity is a standard recurrent neural network (RNN) with sigmoidal activation function (see Methods). Close to equilibrium, the network response to stimuli is largely determined by the network eigenmodes (32) and by the qualitative properties of the reduced dynamics inside each eigenmode, like its characteristic timescale (33). To model S2 neural population dynamics, which does not show any prominent oscillatory activity, and to keep the presentation simple, we solely considered three kinds of one-dimensional eigenmodes (Fig. S6). Eigenmodes with monostable dynamics in which the only equilibrium state is strongly attractive are associated with fast and convergent (faithful) responses. Eigenmodes with monostable dynamics in which the only equilibrium state is weakly attractive are associated with slow and smoothed (temporal) responses. Eigenmodes with bistable or metastable dynamics are associated with threshold-like (categorical) responses. In our model, the qualitative dynamics of each eigenmode was determined by the strength k of the positive feedback arising from recurrent interconnections inside that mode (see Methods). When positive feedback was absent or small, the dynamics were monostable and strongly attractive (faithful). As positive feedback was increased, its destabilizing effect made the dynamics decreasingly attractive and slower (temporal), and eventually bistable or metastable (categorical). We introduced heterogeneity and variability in our model by generating a family of faithful (sensory coding), temporal, and categorical eigenmodes by sampling k according to distributions centered around a different characteristic value k0 for each mode type (Fig. 7 A). In the active (TPDT) context, k0=0.975 for faithful, k0=1.0 for temporal, and k0=1.05 for categorical. To model the transition to the control context in a way that was compatible with experimental data (Fig. S2), we shifted the distribution of positive feedback gains associated to categorical representations toward smaller values and overlapped it both with sensory and temporal distributions (Fig. 7A), thus letting categorical representation disappear from the modeled

population dynamics. The final network consisted of a mixture of neurons with either pure (one third) or mixed (two third) selectivity (see Methods for details). Of the neurons with pure selectivity, half were purely temporal and half purely categorical. Neurons with mixed selectivity formed a continuum between faithful and categorical. Mixed selectivity between temporal and faithful and between temporal and sensory was not enforced because part of the faithful neurons, i.e., those with large enough k, already exhibited slow temporal dynamics and, similarly, part of the categorical neurons, i.e., those with small enough k, also exhibited purely temporal responses. We stress that with our model we solely aimed at reproducing and explaining qualitatively the experimental data during the P1 period. Extending the model to capture the full task is non-trivial, given the needed complex interactions between working memory, sensation, and decision-making. We leave this objective for future work.

When we ran UMAP over the generated neuronal population in the active and control contexts we obtained the projections in Figure 7B. As similarly obtained on the S2 neuronal population (Figures 5,6 and Suppl. Figures 4,5), 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. Thus, our model was able to reproduce the key qualitative experimental observation that the neuronal manifold breaks apart during control trials. It did so through the simple modulation of the positive feedback gain associated to the recurrent interconnections inside categorical modes. The importance of recurrent interconnections for the emergence of slow categorical dynamics was discussed previously (see review (34)). Our model points to the strength of these recurrent interconnections as a suitable target for attention mechanisms aiming at tuning cognitive effort as a function of context.

Discussion

Our research provides five critical insights into the functions and capabilities of the neuronal population S2. First, in line with our previous studies, neurons in S2 displayed a multifaceted array 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 suggests S2's potential for highly efficient computations (35–37). 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 (20), 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 as a whole possesses an innate capability for managing and segregating diverse signals but during a non-demanding 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. While similar mechanisms have been identified in other neuronal populations (2, 4, 6), we believe our findings present a network uniquely designed to differentiate between sensory and transformed dynamics. Fourth, the use of a nonlinear dimensionality reduction methods 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 (non-coding), 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 (38, 39) 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 (40) has been a topic of rigorous discussion (3, 41). A recent study has shown that the orthogonal dynamics associated with sensory and mnemonic coding appeared during implicit task learning (4). Their results in (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 inter-area neural communication (42). 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 (inter-trial 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 post-movement activity might pertain to evaluating the reward associated with the decision (43). 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 neuronal encoding substrate, we employed UMAP. In the 2D 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 (6). 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 (12, 44–47). 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's been 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 discovery of an intrinsic 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 novel methodologies like RNNs, we are not only charting new territories in our understanding of neural dynamics but also shaping the trajectory of future neuroscience research. The intricacies of S2, as showcased, hold profound implications, offering fresh perspectives on the brain's computational prowess and adaptability.

ACKNOWLEDGMENTS

We thank to 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; IBRO Early Career Award 2022 (to R.R.-P.) from International Brain Research Association. L.B. is a postdoctoral student (Postdoctoral fellowship CONACYT-838783).

Methods

Temporal pattern discrimination task (TPDT)

The TPDT used here has been previously described (17, 20). Briefly, two monkeys (*Macaca mulatta*) were trained to report whether the temporal structure of two vibrotactile stimuli patterns (P1 and P2) of equal mean frequency (5 Hz, 5 pulses) were the same (P2 = P1) or different (P2 ≠ P1; Fig. 1A). The temporal structure of each pattern was either grouped (G) or extended (E) with a fixed stimulation period of 1 s. The five pulses were delivered periodically during the extended pattern (E), and three grouped centered pulses with a smaller distance between them as compared to the first and final pulses, during the grouped pattern (G). Monkeys performed the task in blocks of trials in which the two stimulus patterns had a fixed mean frequency. The right arm, hand and fingers were held comfortably but firmly throughout the experiments. The left hand operated an immovable key (elbow at ~90°) and two push buttons in front of the animal, 25 cm away from the shoulder, at eye level. Stimuli were delivered to the skin of one digit from the distal segment of the right restrained hand, via a computer-controlled stimulator (2 mm round tip, BME Systems, Baltimore, MD). The descending of the probe until achieves a skin indentation of 500 µm, marked the beginning of the trial (probe down event, “pd” in Fig. 1A). After pd the animals gently grabbed with their free left hand an immovable key (key down event, “kd” in Fig. 1A). Vibrotactile stimuli consisted of trains of short mechanical pulses; each pulse consisted of a single-cycle sinusoid lasting 20 ms. Time is always referenced to first stimulus onset (0 s corresponds to the start of P1). In a trial, P1 and P2 were delivered consecutively to the glabrous skin of one fingertip, separated by a fixed inter-stimulus delay period of 2 s (1 to 3 s). Each stimulus could be one of the two possible patterns: grouped (G, upper trace of Fig. 1A) or extended (E, lower trace of Fig. 1A) pulses. Therefore, in total there were four possible P1-P2 combinations, denominated as classes: G-G (class 1, c1), G-E (class 2, c2), E-G (class 3, c3) and E-E (class 4, c4). These were presented in pseudo-random order to the monkeys across trials. The monkeys were asked to report whether P2 = P1 (match: combinations E-E and G-G) or P2 ≠ P1 (non-match: combinations E-G and G-E) after a fixed delay period of 2 s (4 to 6 s) between the end of P2 and the mechanical probe rising from the skin (probe up event, “pu” in Fig. 1A). The “pu” functioned as a go signal that triggered the animal’s release of the key (key up event, “ku” in Fig. 1a). The monkey indicated their decision by pressing one of two push

buttons with the left hand (push button event, “pb” in Fig. 1A, lateral push button for P2 = P1, medial push button for P2 ≠ P1). As the two stimulus patterns had equal mean frequency over their full duration (1 s), the decision had to be based on comparison of their temporal structure. The animal was rewarded for correct decisions with a drop of liquid. Animals were handled in accordance with standards of the National Institutes of Health (NIH) and Society for Neuroscience (SfN). All protocols were approved by the Institutional Animal Care and Use Committee of the Instituto de Fisiología Celular (IFC), Universidad Nacional Autónoma de México (UNAM).

Light control task (LCT)

In this task variant, events proceeded exactly as described above and in Fig. 1A, except that when the probe touched the skin (“pd”), one of the two push buttons was illuminated, indicating the correct choice throughout the trial until pb. Identical stimuli were used. The monkey grasped the key until the probe was lifted, but in this case the light was turned off when the probe lifted from the skin. The monkey was rewarded for pressing the illuminated button. Maintaining stimuli and arm movements identical to the TPDT, the decision must be based on the visual stimuli instead.

Task design and performance

The TPDT is not a simple variation of the vibrotactile frequency discrimination task (VFDT) (51). Some cognitive demands and the basic structure of the tasks are similar: both require attention to two separate vibrotactile stimuli (TPDT: P1, P2; VFDT: f1, f2), working memory and a comparison to reach the decision report. Nevertheless, the TPDT requires a very different evaluation of the stimuli; as they only differ by their temporal structure, any computation must be restricted to the internal structure to identify, categorize, and distinguish between them (20). Further, the comparison process is significantly different between the two tasks. Expanding on the necessitated computation, the VFDT can be solved by computing a difference between the parametric representation of the stimulus frequencies to indicate whether f1 > f2 or f1 < f2, whereas the TPDT offers no comparable method of solution (in any trial P1 and P2 always have the same mean frequency). The TPDT demands a match (P2 = P1) vs. non- match decision (P2 ≠ P1). Hence, the comparison employs categorical representations (instead of parametric) of the stimulus patterns. We computed the average performance across S2 recording sessions (p = 84.0%; Monkey RR17, p = 84.5% and Monkey RR20, p = 83.1%). Average performance for TPDT, LCT and each class, is shown in Fig. 1B. Although each animal received around two years of training, this task was difficult enough to impede 100% performance; this reflects the very high-cognitive demands of the TPDT. To provide some context, the average training period to achieve similar performance levels for the VFDT was about six to eight months; for the vibrotactile detection task, the average time was two months (51). After training in the TPDT, the monkeys saturated their average performance around 84% (Fig. 1B, n_{SES} = 423 recording sessions; Monkey RR17, n_{SES} = 281; Monkey RR20, n_{SES} = 142). In addition, the performance was statistically identical for each class (20). Notably, task repetition across recording sessions did not improve performance. However, the performance for the LCT was consistently 100% (Fig. 1B, n_{SES} = 76 recording sessions; Monkey RR17, n_{SES} = 49; Monkey RR20, n_{SES} = 27); this reflects the lack of cognitive demand required for the guided task, as intended by design. As a final observation, the animals were first trained in the LCT, and then gradually introduced to the TPDT. During the recording sessions in S2 (Fig. 1C), animals switched between performing the TPDT and the LCT.

Recordings

Neuronal recordings were obtained with an array of seven independent, movable microelectrodes (2–3 MΩ) inserted into S2 (Fig.1 C), either contralateral (left hemisphere) or ipsilateral (right hemisphere) to the stimulated hand. We were careful to record just above the primary auditory cortex (A1), and we tested this by using auditory stimulation to ensure that neurons were only responding to vibrotactile stimuli (17–19). The receptive fields of the neurons recorded were all very large and some were bimanual, and since the monkey´s hand was carefully fixed in the same manner during each recording session, we do not believe it is possible for these neurons to be responding to motor data in a categorical manner, as would be seen in the parietal ventral area (PV). Coherently, categorical decision responses during P2 or after pb, disappeared during the LCT (see Fig. S2). We recorded 20 trials per stimulus pair (c1; c2; c3; c4). Recording sites changed within terms of recording sessions where one term lasted approximately 30 sessions (1 day = 1 session); the locations of the penetrations were used to construct surface maps in S2 by marking the edges of the small chamber (7 mm in diameter). It is important to emphasize that neurons which displayed sensory or transformed responses were recorded across the entire S2 region, and the possibility of identifying subgroups responses is extremely low (see Fig. 5 for continuum of responses). The neuronal recording protocol was identical for both the TPDT and LCT.

Datasets

We recorded 1646 S2 neurons using the TPDT, with the stimulus set with 5 Hz mean frequency (Monkey RR17, n=1035; Monkey RR20, n=611). In addition, 313 neurons (Monkey RR17, n=189; Monkey RR20, n=124) were recorded sequentially with TPDT and LCT, with also the 5 Hz mean frequency set. For each neuron of the datasets (n=1646 and n=313), we calculated a time-dependent firing rate per trial using a 100 ms deterministic square kernel with 50 ms steps, beginning 1 s before stimulus pattern P1 and continuing until the end of the trial (1.5 s after the push button press). In previous work we have seen that this window-width is optimal for decoding pattern information in this same area (Fig. S10 in (17)). Importantly, each dataset is defined by four dimensions: N, number of neurons; C, stimulus conditions (classes, always 4); T, time (-1 to 7.5 s, always 170 bins); K, number of hit trials (for each class). A remarkable feature of this task design is the low number of stimulus conditions (four classes), which were equally demanding for the subject. This design allowed us to have, on average, 15.3 hit trials (and 2.9 error trials) per stimulus class for each studied neuron.

Population Analysis

Single S2 neurons displayed a large repertoire of neuronal responses associated with one or several components of the TPDT (see Fig. 1, S1 & S2). Here, we focus our analysis on the temporal neuronal population signals. For each neuron, we averaged per class the time-dependent firing rate of hit trials (c1, c2, c3 or c4). Using the peri-stimulus time histogram (PSTH) of each neuron, we constructed pseudo-simultaneous population responses by combining neural data mostly recorded separately. For each time and class, the population response is defined by an N-dimensional vector in which each component represents the firing rate from a different neuron. This means that including all the recorded neurons (n=1646), we obtained a 1646-dimensional firing rate vector that depended on the time and class ($\bar{r}(t,c)$). The population firing rates averaged over all hit trials ($\bar{r}(t)$) was an N- dimensional vector that measures the mean response for each neuron ($r^i(t)$) as a function of time. For the LCT, the control condition, the population response was a 313-dimensional firing rate vector.

In the next equations, we employed the following notation: $\bar{r}(t)$ is the firing rate average over all hit trials at each time bin (joining the four classes, c1-c4), $\bar{r}(t,c)$ is the firing rate average per class (grouping trials according to each of the four classes), $\bar{r}(t,P1)$ is the firing rate average per P1 stimuli (splitting trials according to P1-identity), $\bar{r}(t,P2)$ is the firing rate average per P2 stimuli (separating trials according to P2-identity) and $\bar{r}(t,Dec)$ is the firing rate average per decision (dividing trials according to decision). To further explain this measure, the i component, $r^i(t)$, represents the firing rate average of neuron i at time t, across all trials. Similarly, $r^i(t,Dec)$ represent the firing rate average of neuron i at time t across trials with the same decision outcome (P2=P1 or P2≠P1).

Instantaneous Coding Variances across the population

At each time point, the population instantaneous coding variance (Var_{COD}, Figs. 2A and 2B, blue trace) was computed as the quadratic square sum of the firing rate fluctuations among classes and neurons:

$$Var_{COD}(t) = \frac{1}{N}\frac{1}{4}\sum_{i=1}^{N}\sum_{c=1}^{4}(r^i(t,c) - r^i(t))^2 \qquad \text{(Eq. S1)}$$

This metric, normalized per neuron, measures the population's variation of firing rate between classes at each time bin. With this definition, variation will be due to any class- related change in the population activity and due to any residual noise.

To evaluate the influence of each kind of coding, we extended this metric to calculate the instantaneous variance associated with each specific task parameter. At each time bin, the population instantaneous P1 variance (Var_{P1}, Figs. 2A and 2B, light blue trace) was computed as the quadratic square sum of the firing rate fluctuations among P1 identity and neurons:

$$Var_{P1}(t) = \frac{1}{N}\frac{1}{2}\sum_{i=1}^{N}\sum_{P1=1}^{2}(r^i(t,P1) - r^i(t))^2 \qquad \text{(Eq. S2)}$$

Analogously, the population instantaneous P2 variance (Var_{P2}, Figs. 2A and 2B, purple trace) measures the firing rate fluctuations among P2 identity and neurons:

$$Var_{P2}(t) = \frac{1}{N}\frac{1}{2}\sum_{i=1}^{N}\sum_{P2=1}^{2}(r^i(t,P2) - r^i(t))^2 \qquad \text{(Eq. S3)}$$

The population instantaneous decision variance (Var_{Dec}, Figs. 2A and 2B, pink trace) measures the firing rate fluctuations among decision identity and neurons:

$$Var_{Dec}(t) = \frac{1}{N}\frac{1}{2}\sum_{i=1}^{N}\sum_{DEC=1}^{2}(r^i(t,Dec) - r^i(t))^2 \qquad \text{(Eq. S4)}$$

The value of Var_{COD} during the period immediately before P1 onset represented the inherent residual noise in the firing rate estimates (~2 sp/s); to be interpreted as a degree of population coding, Var_{COD} should be higher than this resting state variance (basal variance). The same reasoning applies to the other specific variances. To facilitate the comparison across parameters, values associated to basal variance were subtracted in each kind of variance computed.

Instantaneous Temporal Variances across the Population

At each time point, the population instantaneous temporal variance (Var_{Temp}, Fig. 2C and Fig. 2D) with respect to the mean firing rate, was computed as the quadratic square sum between

the mean firing rate for each time bin ($r^i(t)$) and the mean firing rate across the whole trial r^i (from -1s to 7.5s) among neurons:

$$Var_{Temp}(t) = \frac{1}{N}\sum_{i=1}^{N}(r^i(t) - r^i)^2 \qquad \text{(Eq. S5)}$$

Principal Component Analysis (PCA)

The main aim of PCA is to find a new coordinate system in which the data can be represented in a more succinct and compact manner. In other words, the idea is to define a low-dimensional subspace that captures most of the variance of the high-dimensional neural space. To characterize how the population activity covaries across classes as a function of time, we performed PCA (Figs. 3 & S3) over classes (c1, c2, c3 or c4) where we combined variance over classes and time (from -1s to 7.5s in Fig. S3 and smaller periods in Fig. 3). PCA yields a new coordinate system for the N-dimensional data, in which the first coordinate accounts for as much of the variance of the neural population. The second coordinate accounts for as much of the remaining variance, and so on; however, each subsequent axis is restricted to be orthogonal to all previous axes. PCA was computed from the firing rate covariance matrix, that averages over time bins, *t*, and classes, *c*:

$$C_{ij} = \frac{1}{TC-1}\sum_{t0}^{tF}(r^i(t,c) - r^i(t))(r^j(t,c) - r^j(t)) \qquad \text{(Eq. S6)}$$

where *T* denotes the number of time bins between *tF* and *t0*, and *C* is the number of classes (4 in our task), $r^i(t,c)$ denotes the trial-averaged firing rate of the neuron i, under class c, at time t, and $r^i(t)$ is the firing rate average of neuron i across classes at time t. In Fig. 3 and S3, PCs were calculated for TPDT (n=1646) and LCT (n=313). The diagonalization of the covariance matrix, $C = UDU^T$, yields a new coordinate system given by the columns of the matrix U. We refer to the columns of U as the principal components (PCs). On the other hand, D is a diagonal matrix of positive values. The diagonal elements of D give the amounts of population activity variance captured by the corresponding PCs. We then ordered the PCs depending on this amount of variance captured. The projection of the N- dimensional data onto the kth PC is given by:

$$PC_k(t,c) = \sum_{i=1}^{N} U_k^i(r^i(t,c) - r^i(t)) \qquad \text{(Eq. S7)}$$

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 linear combinations of the firing rates of the individual neurons. Thus, the contribution of each neuron to a given PC_k is given by the ith element of U_k. These PCs can be thought of as a low-dimensional description of the population activity in this coding subspace.

Demixed Principal Component Analysis (dPCA)

The algorithmic details and mathematical justification for this method were outlined in (7). The method has a supervised and unsupervised part. In brief, dPCA decomposes the neural activity by different chosen task variables to compute marginalized covariance matrices (this is the supervised part, similar to choosing the variables to fit a linear model). Afterward, it carries out a principal component-like analysis over those matrices (this is completely unsupervised). In our task, we marginalized the population activity (X) with respect to: P1 identity ($\overline{X_{t,P1}}$), and decision

outcome ($\overline{X_{t,Dec}}$). To calculate the marginalization averages, we use the N-dimensional population activity:

$$\overline{X_{t,P1}}(t,c) = \bar{r}(t,P1) - \bar{r}(t);$$
$$\overline{X_{t,Dec}}(t,c) = \bar{r}(t,Dec) - \bar{r}(t) \qquad \text{(Eq. S8)}$$

Then, $\overline{X_\phi}$ denotes the marginalized matrices with $\phi \in \{\{t,P1\},\{t,Dec\}\}$ (all the marginalization variables used in this study). Further, $\bar{r}(t)$ is the firing rate average over all hit trials at each time bin (joining the four classes, c1-c4), $\bar{r}(t,P1)$ is the firing rate average per P1 stimuli (splitting trials according to P1 identity, G or E) and $\bar{r}(t,Dec)$ is the firing rate average per decision (dividing trials according to decision: P2 = P1 or P2 ≠ P1). All $\overline{X_\phi}$ have the same sizes (7). Once marginalization is performed, dPCA finds separate decoder and encoder matrices for each $\overline{X_\phi}$, by minimizing with reduced-ranked regression the term:

$$L_\phi = \| \overline{X_\phi} - F_\phi D_\phi \bar{X} \|^2 \qquad \text{(Eq. S9)}$$

where $\bar{X}$ is the centered whole population data matrix (i.e., the average activity of each neuron is 0). The solution of this problem can be obtained analytically in terms of singular value decompositions. Each component D_ϕ can be ordered by the amount of explained variance. The most prominent decoding axis is called the 1st demixed principal component (1st dPC) of variable ϕ. To avoid dPCA over fitting, we introduced a regularization term and performed cross-validation to choose the regularization parameter. To obtain Fig. 4A & B, we projected the N-dimensional data for a given class ($\bar{r}(t,c)$, this is the firing rate average per class computed by grouping trials according to each of the four classes) onto the most prominent decoding axis (k) of a given variable ϕ. These projections were given by:

$$dPC_k^\phi(t,c) = \sum_{i=1}^{N} D_{k,\phi}^i (r^i(t,c) - r^i(t)) \qquad \text{(Eq. S10)}$$

where $dPC_{k,\phi}^i$ is the i component of the k most relevant axis of a demixed variable ϕ. Here we emphasize that dPCA finds the decoder and encoder of each marginalized variable, ϕ, minimizing each L_ϕ separately. This means that we can calculate individually the demixed decoders for each parameter: P1-identity and decision outcome. Figs. 4A & B shows the most prominent demixed components for P1 and decision, respectively.

Uniform Manifold Approximation and Projection (UMAP)

This algorithm is a non-linear dimensionality reduction technique, based on: (1) approximating the manifold on which we assume the data was sampled from and (2) projecting this manifold into a low dimensional space while maintaining its local and global topological properties (27). For this, we concatenated all the $\bar{r}(t,c)$ vectors horizontally for each neuron, according to the TPDT or LCT set. The firing rate was calculated using a window of 100 ms and steps of 50 ms, yielding a total of 170 time-bins per class when considering the whole TPDT or LCT (-1 to 7.5 s, Figs. S4 and S5) or 140 time-bins per class when considering a restricted TPDT or LCT (-1 to 6 s, without decision, Figs. 5 and 6), and 680 or 560 dimensions once all vectors were concatenated, when considering the whole or restricted task variants, respectively. As a result, the size of the matrix input into UMAP was 1646x680 or 1646x560 for the entire and restricted TPDT, respectively, and 313x680 or 313x560 for the entire and restricted LCT, respectively, where the 1646 (TPDT) or 313 (LCT) neurons were taken as observations, and the 680 or 560 time-point dimensions were reduced to two representative dimensions. We then project the inferred

manifold into the new 2-dimensional space to determine whether a non-linear dimensionality reduction technique could isolate (or not) neural responses with the characteristic population dynamics identified in Figs. 3, S3 and 4. In Figs. 5B, 6B, S4A and S5B, we included a density contour plot calculated with a Gaussian kernel (σ = 0.5). For Figs. 5A, 6A, S4B and S5A we calculated a double-Gaussian (σ = 0.5) weighted-average centered on each of the colored X marks to visualize how the dynamics vary across the UMAP plane: neurons were weighted based on the Gaussian Probability Density Functions and then divided by the total sum of the weights.

Network model

S2 neural population dynamics was modeled as nonlinear recurrent neural network (RNN)

$$\dot{x}_i = -x_i + \sum_{j=1}^{N} A_{ij} S\left(k_j \sum_{k=1}^{N} B_{jk} x_k + I_j(t)\right), \qquad i = 1, \dots, N. \qquad \text{(Eq. S11)}$$

where $x_i(t)$ describes the average activity (firing rate) of a neuron i in a temporal window centered at time t. Depending on the context, $x_i(t)$ can also be interpreted as the average activity of a sub-population i of neurons. In vector form, model (1) becomes:

$$\tau \dot{\boldsymbol{x}}(t) = -\boldsymbol{x}(t) + A\boldsymbol{S}\left(\boldsymbol{k} \odot B\boldsymbol{x}(t) + C\boldsymbol{I}(t)\right) \qquad \text{(Eq. S12)}$$

where $\boldsymbol{x} = (x_1, \dots, x_N)$, $\boldsymbol{I} = (I_1, \dots, I_N)$, $\boldsymbol{k} = (k_1, \dots, k_N)$, $\boldsymbol{S}(\boldsymbol{x}) = (\tanh(x_1), \dots, \tanh(x_n))$, and $\odot$ denotes the Hadamard product. The sigmoidal function $S(\cdot) = \tanh(\cdot)$ models neurotransmitter release and other kind of synaptic and intrinsic neural nonlinearities. Matrices A and B encodes the recurrent interconnection topology. B_{jk} is the effect that the activity of a neuron (or neural sub-population) k has on the liberation of neurotransmitter j. A_{ij} is the effect that the liberation of neurotransmitter j has on the activity of neuron (or neural sub-population) i. The gains $k_j \geq 0$ control the overall strength of neurotransmitter liberation, thus controlling the linearized recurrent interconnection strength $\rho_{ik} = \sum_{j=1}^{N} A_{ij}\, k_j\, B_{jk} \sum_{l=1}^{N} A_{kl}\, k_l B_{li}$ between neurons (or neural sub-populations) i and k. The eigenvector of the effective recurrent interconnection matrix $A(\boldsymbol{k} \odot B)$ are the network eigenmodes determining the qualitative network response stimuli.

Our network design procedure involves the following steps. 1. Fix the gains k_j, which fixes the eigenmode types, as detailed in the main text. 2. Fix a family of (independent) eigenvectors $\boldsymbol{v}_j$, which fix the eigenmode directions in the space state. In particular, the i-th component of eigenvector $\boldsymbol{v}_j$ is the contribution of neuron (or neural sub-population) i to eigenmode j. 3. Define the change of variable matrix $U = [\boldsymbol{v}_1 \dots \boldsymbol{v}_N]$. 4. Let $A = U$ and $B = U^{-1}$ in (2), which enforces that the network model (2) has eigenmode $\boldsymbol{v}_j$ with eigenvalues k_j.

Based on the step above, we generated a meta-population of neurons by running $n_{rep} = 50$ instances of an $N = 6$ neuron/neural sub-population network with the following parameters. Let $\mathcal{U}(a, b)$ denote the uniform distribution on the interval (a, b) and $\mathcal{N}(\mu, \sigma)$ the Gaussian distribution with mean μ and variance σ. Active trial: $k_1, k_3, k_4 \in \mathcal{U}(1.025, 1.05)$, $k_2, k_5 \in \mathcal{U}(0.95, 1.0)$, $k_6 \in \mathcal{N}(1.0, 0.0025)$. Hence, 6 is predominantly temporal. The eigenmode directions were generated according to the same distributions in both active and control trials as follows. Let $\alpha_1, \alpha_2 \in \mathcal{U}(0,1)$ and $\beta_1 = 1 - \alpha_1$, $\beta_2 = 1 - \alpha_2$. Then $\boldsymbol{v}_1 = (\alpha_1, \beta_1, 0, 0, 0, 0)$, $\boldsymbol{v}_2 = (\beta_1, \alpha_1, 0, 0, 0, 0)$, $\boldsymbol{v}_3 = (0, 0, 1, 0, 0, 0)$, $\boldsymbol{v}_4 = (\alpha_2, \beta_2, 0, 0, 0, 0)$, $\boldsymbol{v}_5 = (\beta_2, \alpha_2, 0, 0, 0, 0)$,

$\boldsymbol{v}_6 = (0, 0, 0, 0, 0, 1)$, in such a way that there are 4 mixed eigenmodes $(1, 2, 4, 5)$ and two pure eigenmodes $(3, 6)$. Observe that the generated eigendirections are independent with probability one. In active trials, categorical eigenmodes were also receiving a constant inhibition with intensity $I_{inh} = -0.0075$ that favored meta-stability instead of bistability of the mode. Eigenmode $1, 3, 4$ were receiving the stimuli with intensity $I_{stim} = 0.0666$, eigenmodes $2, 5$ with intensity $I_{stim} = 0.05$, and eigenmode 6 with $I_{stim} = 0.01$. Julia code used to generate the numerical results, is available upon request.

Data availability

Data files are publicly available at Zenodo (DOI: 10.5281/zenodo.4421855); see (52). Source data are provided with this paper.

Code availability

The custom python and C scripts employed in the analysis of this data, as well as the experimental protocols, are available from the corresponding authors on reasonable request.

References

1. G. Diaz-deLeon, *et al.*, An abstract categorical decision code in dorsal premotor cortex. *PNAS* **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. Y. Xie, *et al.*, Geometry of sequence working memory in macaque prefrontal cortex. *Science* **375**, 632–639 (2022).

7. D. Kobak, *et al.*, Demixed principal component analysis of neural population data. *Elife* **5**, e10989 (2016).

8. A. S. Morcos, C. D. Harvey, History-dependent variability in population dynamics during evidence accumulation in cortex. *Nat. Neurosci.* **19**, 1672–1681 (2016).

9. 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).

10. 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).

11. 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).

12. J. Wang, D. Narain, E. A. Hosseini, M. Jazayeri, Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018).

13. 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).

14. 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).

15. 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).

16. O. I. Rumyantsev, *et al.*, Fundamental bounds on the fidelity of sensory cortical coding. *Nature* **580**, 100–105 (2020).

17. 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).

18. 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).

19. 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).

20. R. Rossi-Pool, *et al.*, Emergence of an abstract categorical code enabling the discrimination of temporally structured tactile stimuli. *PNAS* **113**, E7966–E7975 (2016).

21. 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).

22. R. Romo, A. Hernández, A. Zainos, E. Salinas, Correlated Neuronal Discharges that Increase Coding Efficiency during Perceptual Discrimination. *Neuron* **38**, 649–657 (2003).

23. R. Rossi-Pool, *et al.*, Temporal signals underlying a cognitive process in the dorsal premotor cortex. *PNAS* **116**, 7523–7532 (2019).

24. G. F. Elsayed, J. P. Cunningham, Structure in neural population recordings: an expected byproduct of simpler phenomena? *Nat. Neurosci.* **20**, 1310–1318 (2017).

25. M. C. Aoi, V. Mante, J. W. Pillow, Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nat. Neurosci.* **23**, 1410–1420 (2020).

26. J. D. Murray, *et al.*, Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *PNAS* **114**, 394–399 (2017).

27. L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3**, 861 (2018).

28. E. K. Lee, *et al.*, Non-linear dimensionality reduction on extracellular waveforms reveals cell type diversity in premotor cortex. *Elife* **10**, e67490 (2021).

29. E. Becht, *et al.*, Dimensionality reduction for visualizing single-cell data using UMAP.

Nat. Biotechnol. **37**, 38–44 (2019).

30. R. J. Gardner, *et al.*, Toroidal topology of population activity in grid cells. *Nature* **602**, 123–128 (2022).

31. S. Parra, *et al.*, Hierarchical unimodal processing within the primary somatosensory cortex during a bimodal detection task. *PNAS*, 10.1073/pnas.2213847119 (2022).

32. G. R. Yang, X.-J. Wang, Artificial Neural Networks for Neuroscientists: A Primer. *Neuron* **107**, 1048–1070 (2020).

33. A. Chadwick, *et al.*, Learning shapes cortical dynamics to enhance integration of relevant sensory input. *Neuron* **111**, 106-120.e10 (2023).

34. X.-J. Wang, Macroscopic gradients of synaptic excitation and inhibition in the neocortex. *Nat. Rev. Neurosci.* **21**, 169–178 (2020).

35. 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).

36. S. Fusi, E. K. Miller, M. Rigotti, Why neurons mix: high dimensionality for higher cognition. *Curr. Opin. Neurobiol.* **37**, 66–74 (2016).

37. M. Rigotti, *et al.*, The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013).

38. E. Salinas, P. Thier, Gain Modulation: A Major Computational Principle of the Central Nervous System. *Neuron* **27**, 15–21 (2000).

39. M. Carandini, D. J. Heeger, Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* **13**, 51–62 (2012).

40. S. Ganguli, H. Sompolinsky, Compressed Sensing, Sparsity, and Dimensionality in Neuronal Information Processing and Data Analysis. *Annu. Rev. Neurosci.* **35**, 485–508 (2012).

41. D. Badre, A. Bhandari, H. Keglovits, A. Kikumoto, The dimensionality of neural representations for control. *Curr. Opin. Behav. Sci.* **38**, 20–28 (2021).

42. 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).

43. 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).

44. V. Mante, D. Sussillo, K. V Shenoy, W. T. Newsome, Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).

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. *PNAS*, 201921609 (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. *PNAS* **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. *PNAS* **108**, 19377–19382 (2011).

51. R. Romo, R. Rossi-Pool, Turning Touch into Perception. *Neuron* **105**, 16–33 (2020).

52. 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 (February 10, 2021).

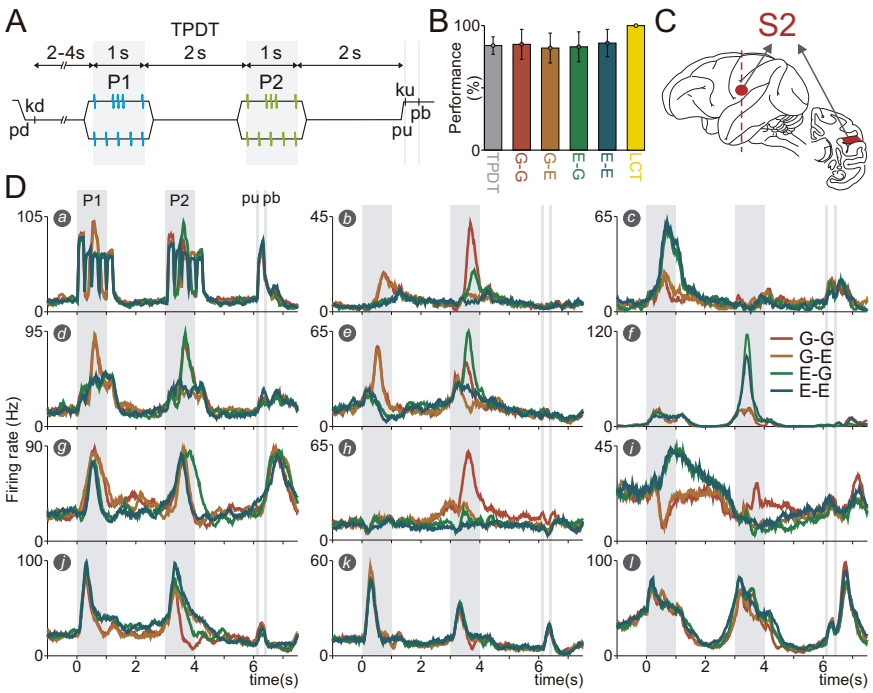


Figure 1. Task, performance and exemplary responses of S2 neurons. (A) Schematic of the temporal pattern discrimination task (TPDT) with trial event sequence. A mechanical probe was lowered (pd), indenting the glabrous skin of one fingertip of the monkey´s restrained right hand. Consequently, the animal 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 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 & 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

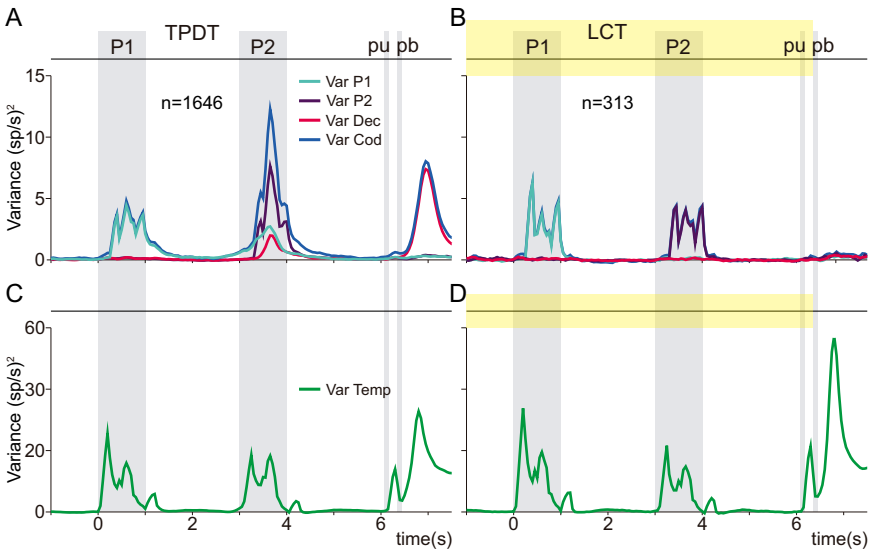


Figure 2. Population variance during active and control tasks. (A, B) Population variances as a function of time during the TPDT (n=1646) and LCT (n=313). Traces refer to P1 (*Var P1*, light blue trace), P2 (*Var P2*, purple trace), decision (*Var Dec*, pink trace) and coding (*Var Cod*, blue trace) 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, D) Temporal (or non-coding) 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 non-coding signals constitute an essential part of the network dynamics.

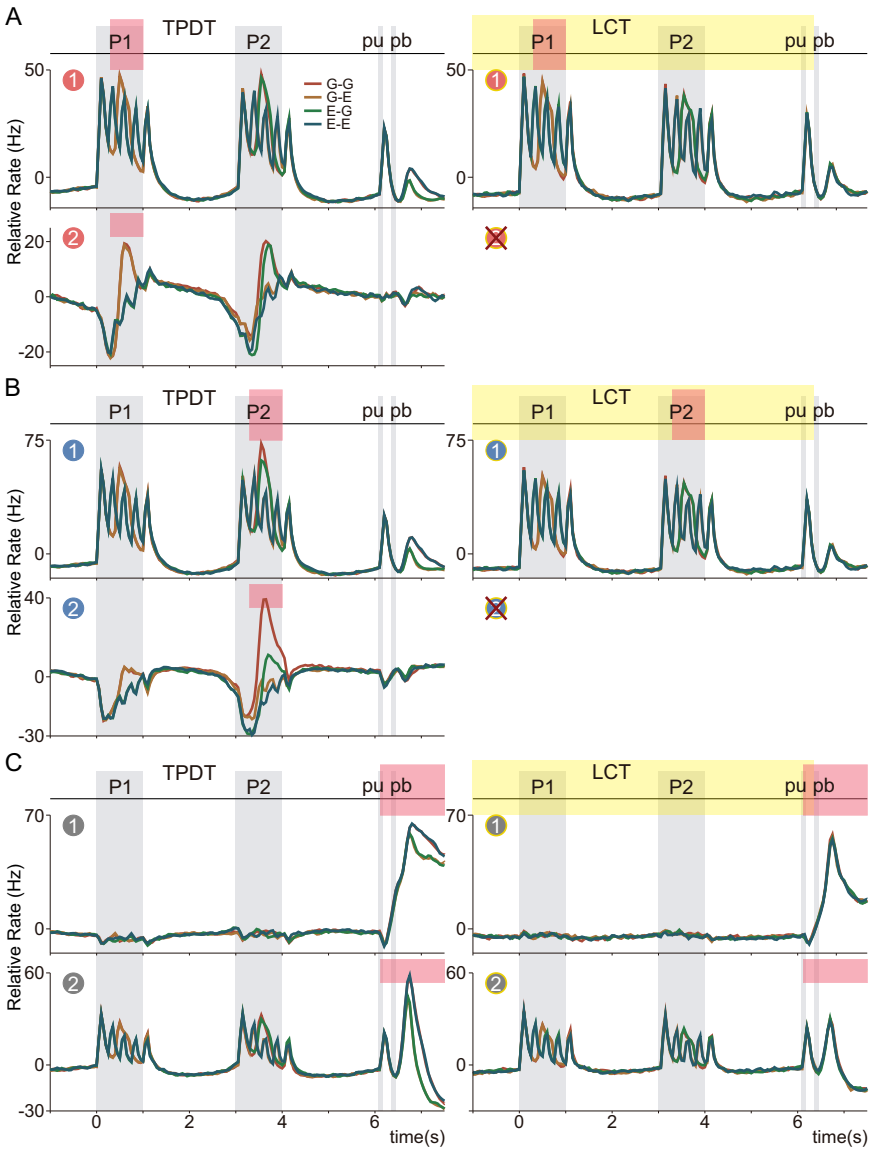


Figure 3. Contextual orthogonality in population dynamics. (A, B and C) Principal component analysis (PCA) was applied to the covariance matrices computed on neural activity during the last 0.6 s of P1 and P2, and 1.5 s of the push button period (pink markers on top), to compute associated principal components (PCs) P1-PCs, P2-PCs and pb-PCs. The activity from the whole (-1 to 7.5 s) TPDT (n=1646) 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, 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 inter-trial period. In LCT only the temporal component of this dynamics remained.

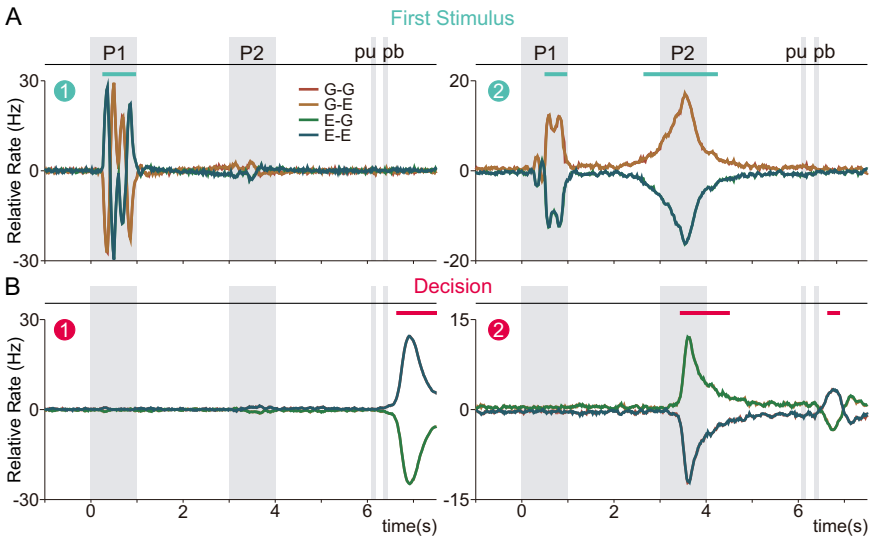


Figure 4. Orthogonal population responses associated with task parameters. Demixed principal component analysis (dPCA) was applied to the marginalized covariance matrices from the whole TPDT (-1 to 7.5 s), with respect to P1 (A) and decision (B), to derive associated demixed principal components (dPCs). Population activity for TDPT (n=1646), 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%). Only the first two dPCs for both parameters were above the 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).

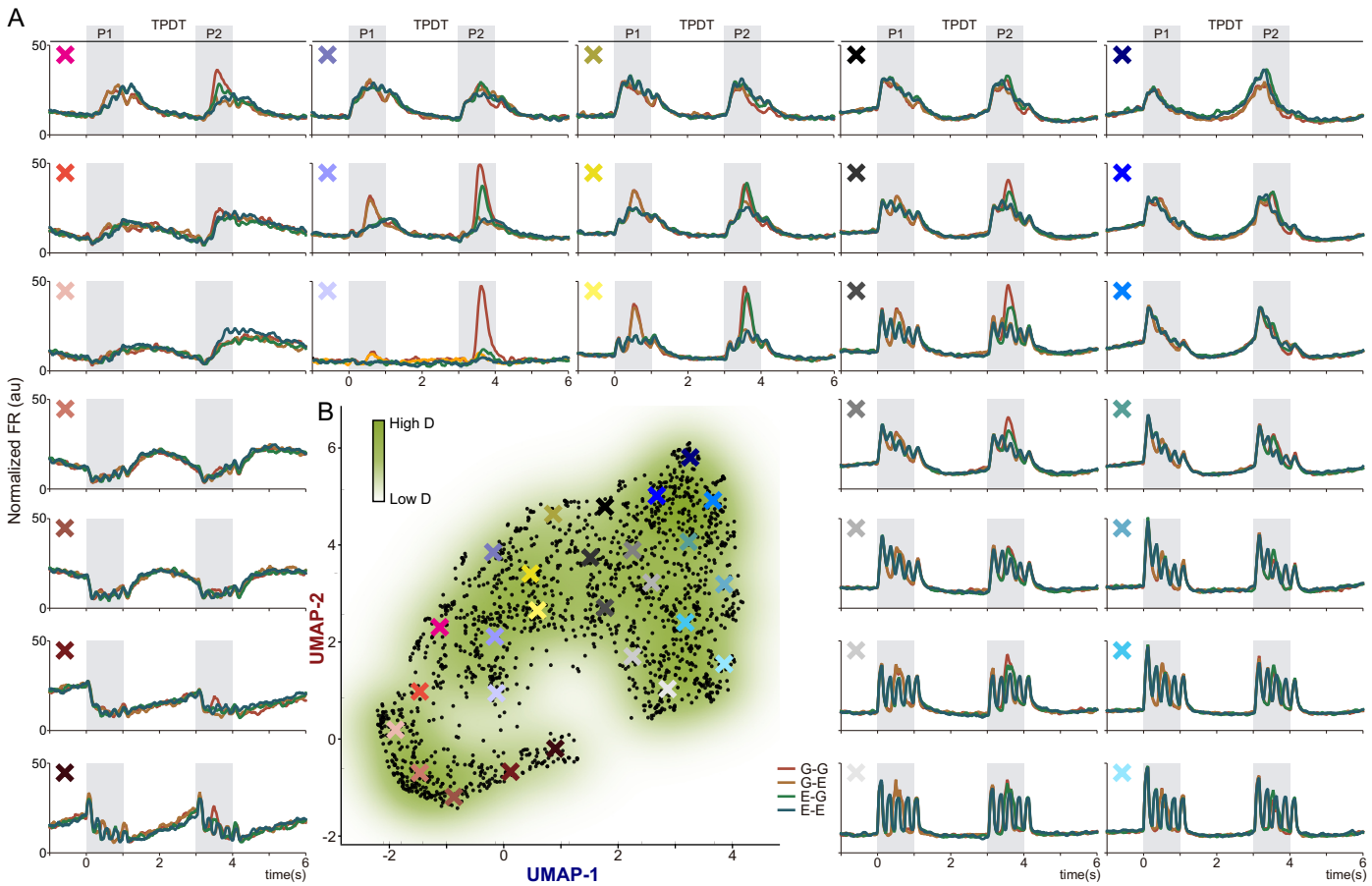


Figure 5. A continuum of responses in S2: from sensory to categorical neurons during the TPDT. To focus on the sensory and categorical responses, we excluded the movement and inter-trial periods. (A) Each subpanel represents a weighted average of neuronal activity and is presented in back-transform arbitrary 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=1646, each black dot) recorded during TPDT (restricted from -1 to 6 s of the task) exhibited a continuum of responses. The color X marks represent the centers of the double-Gaussian (σ = 0.5) weighted averages that are displayed in the panels in A.

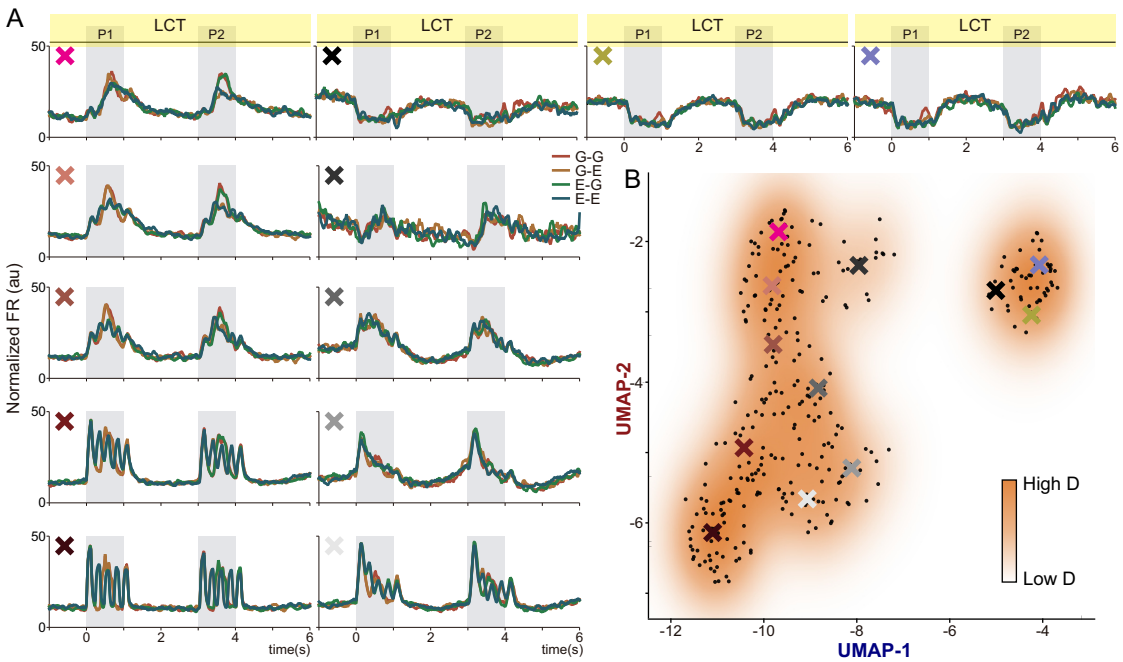


Figure 6. Breaking the continuum in S2 during the non-demanding task. To focus on the sensory and categorical responses, we excluded the movement and inter-trial periods. (A) Each subpanel represents the weighted average of neuronal activity in arbitrary units that approximately correspond to firing rate values (Hz). Given that LCT is a non-demanding 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 color X marks represents the centers of double-Gaussian ($\sigma = 0.5$) weighted averages that are displayed in the panels in A.

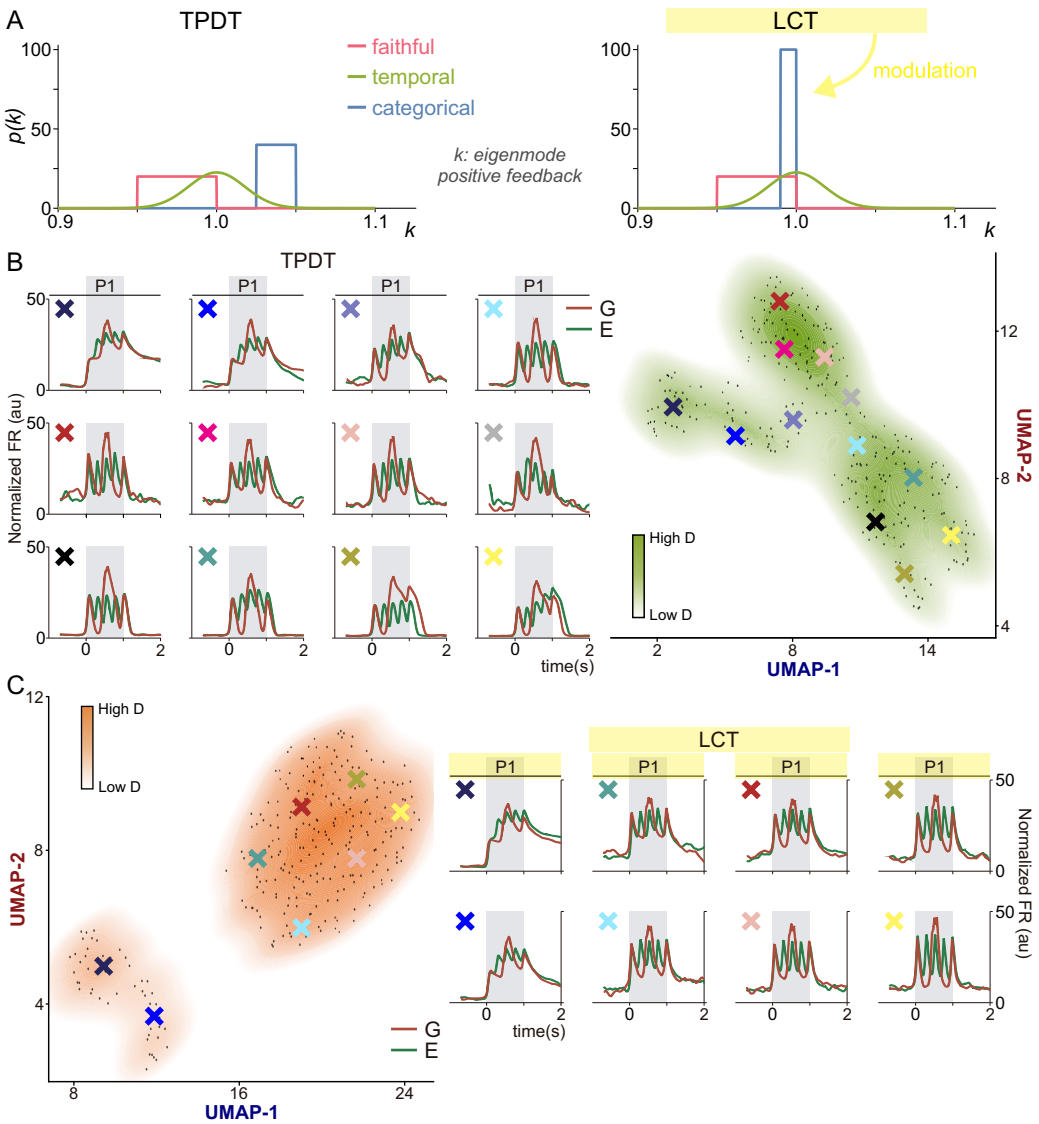
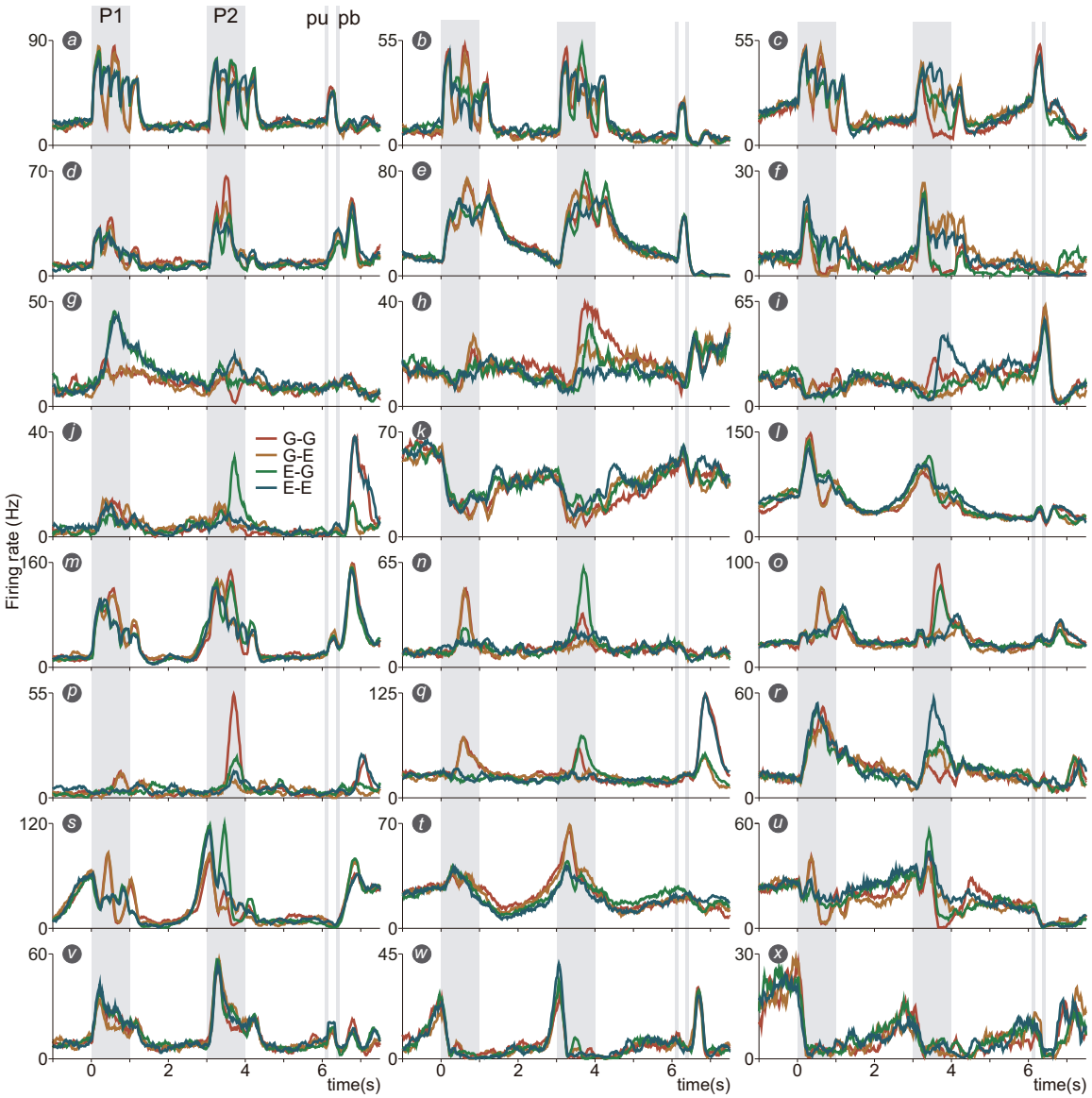
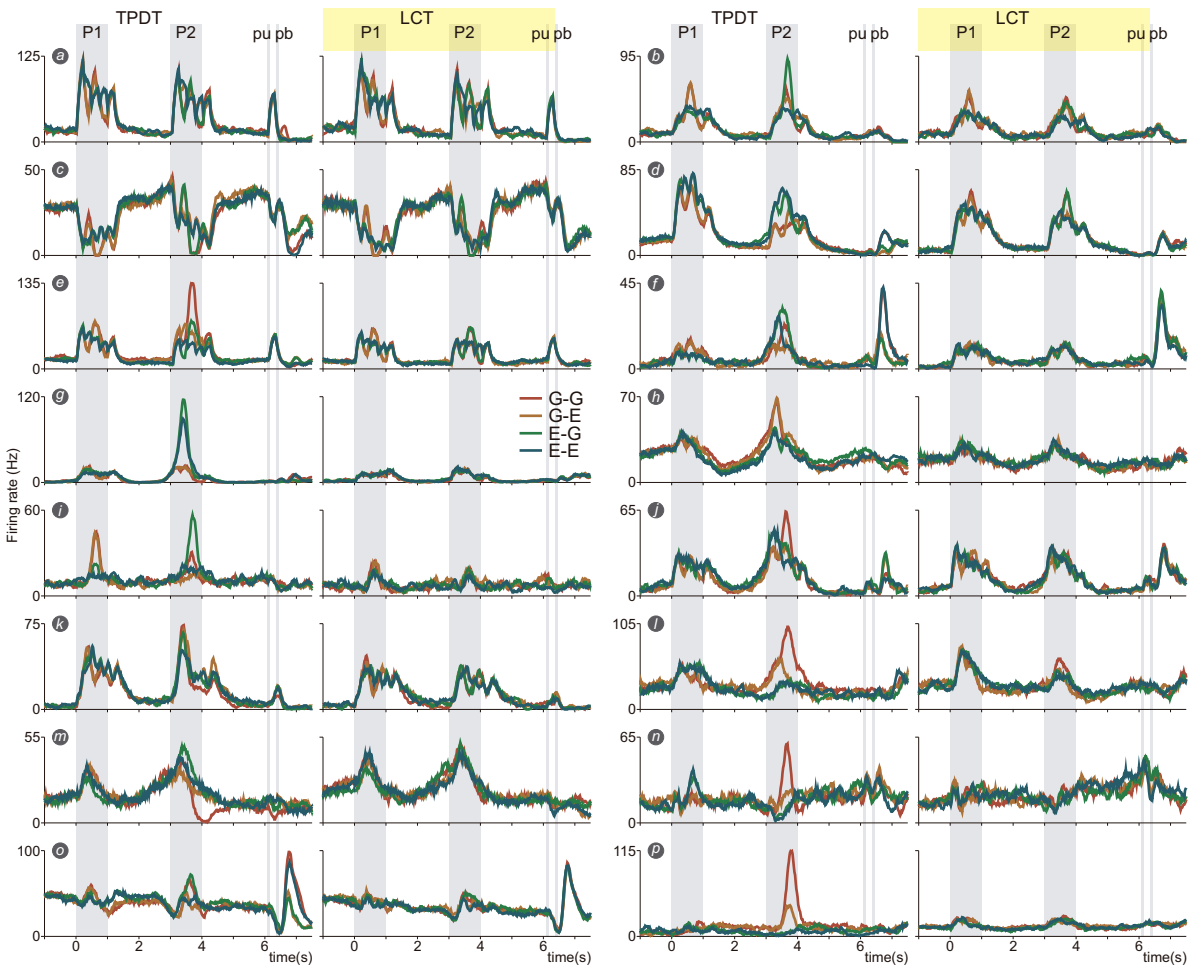


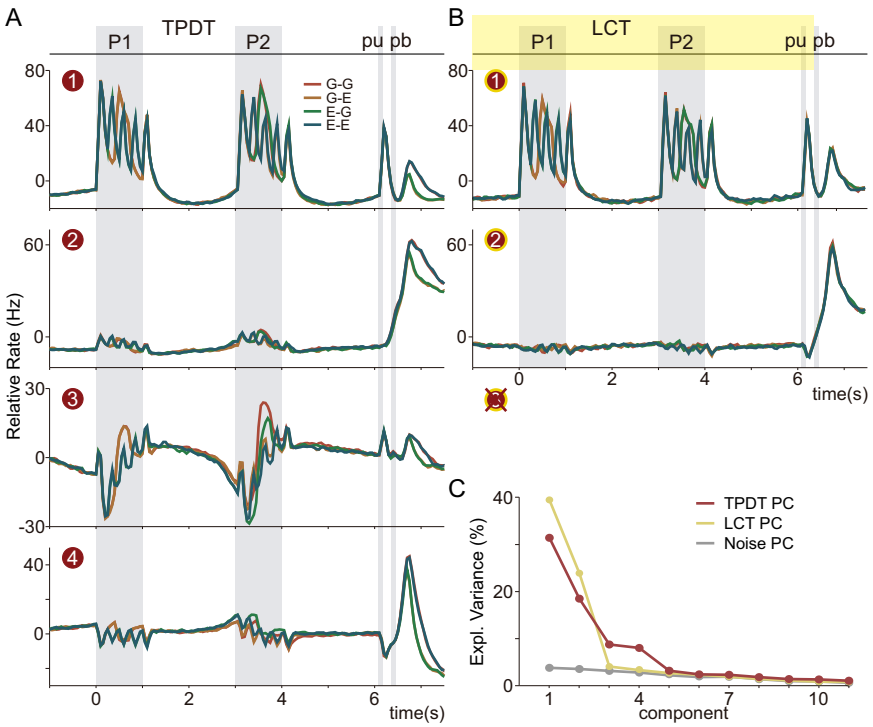
Figure 7. A simplified model of S2 neuronal population dynamics. The network reproduces the transition from a continuum of responses in TDPT to clustered responses in LCT. (A) Faithful, temporal, and categorical eigenmodes in our network were defined by different characteristic positive feedback gain (see main text and Methods). For each eigenmode type, the defining positive 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 positive feedback gain distribution of the categorical eigenmodes as drawn. (B) Similar to Figure 5 but for the computational mode. In active condition, the modeled S2 neural manifold exhibited a continuum of responses connecting faithful, temporal, and categorical modes. (C) Similar to Figure 6 but for the computational 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.



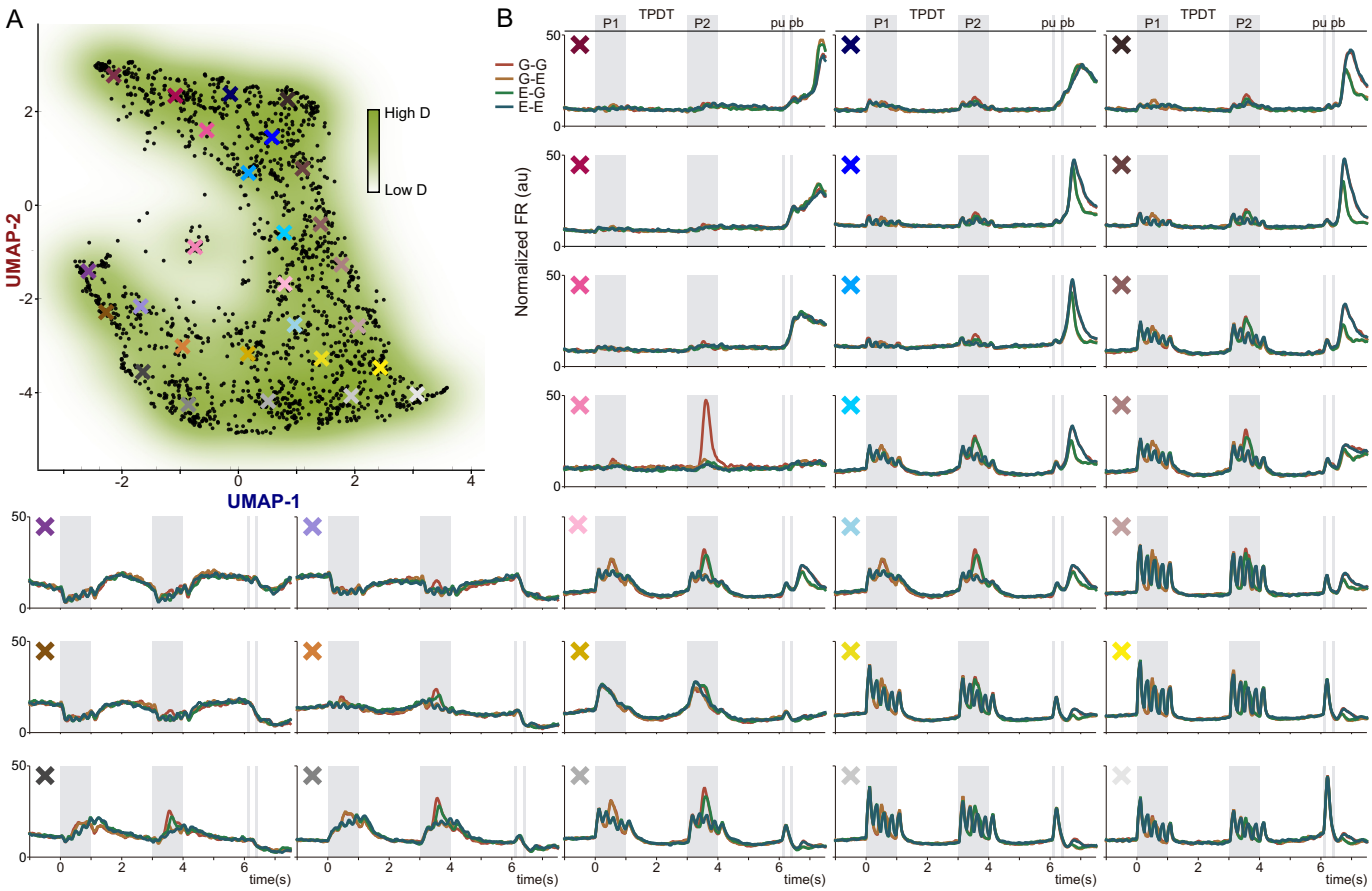
Supplementary Figure 1. Activity of single neurons in S2. A mix of representative neurons (24 units in total, from *a* to *x*), wherein sensory, categorical, and temporal responses, are shown during TPDT. Traces represent firing rate averages per-class for each neuron. Each color refers to one of the four possible pattern pairs (class) of G (grouped) and E (extended): c1 (G-G, red), c2 (G-E, orange), c3 (E-G, green) and c4 (E-E, blue).



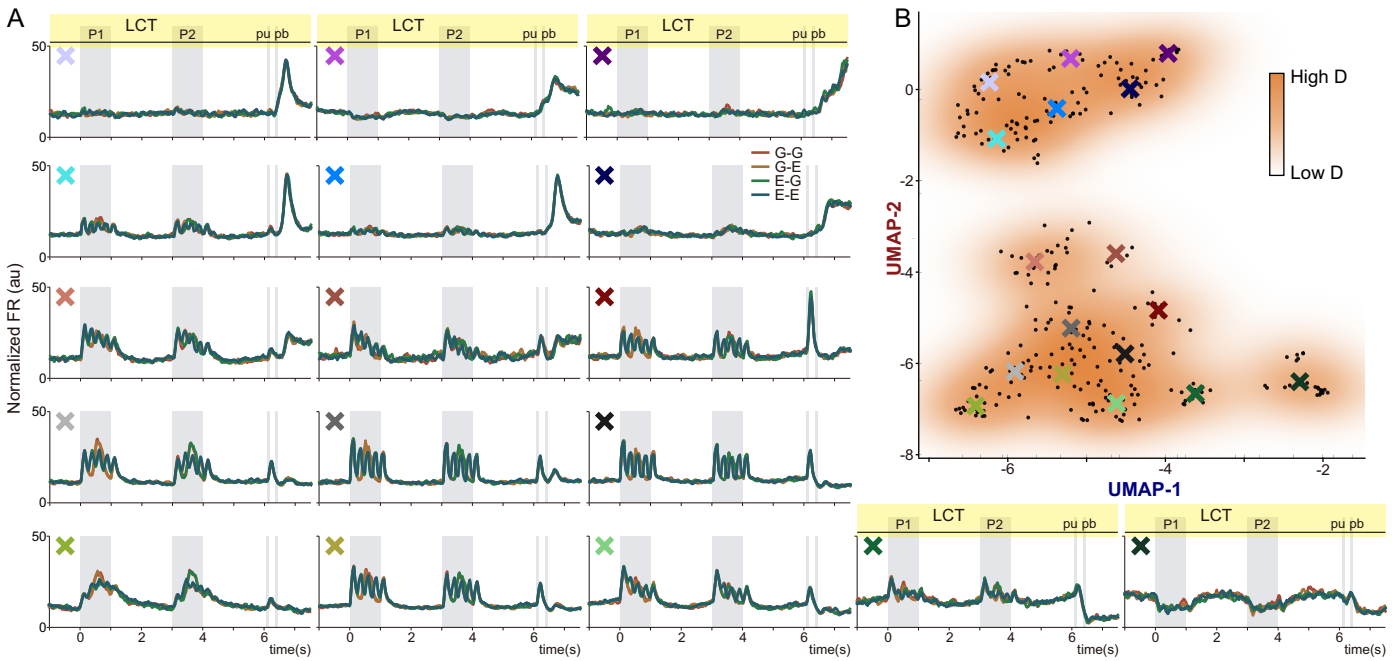
Supplementary Figure 2. Single S2 neurons recorded during the TPDT and LCT. Exemplary neurons (16 units, from *a* to *p*, shown in TPDT and LCT) exhibited sensory, categorical, and temporal responses. Traces represent firing rate averages per-class for each neuron in both tasks. Each color refers to one of the four possible pattern pairs (class) of G (grouped) and E (extended): c1 (G-G, red), c2 (G-E, orange), c3 (E-G, green) and c4 (E-E, blue). Notice how neurons, as units *g and p*, turned off their categorical coding during the control task, leaving only temporal representations.



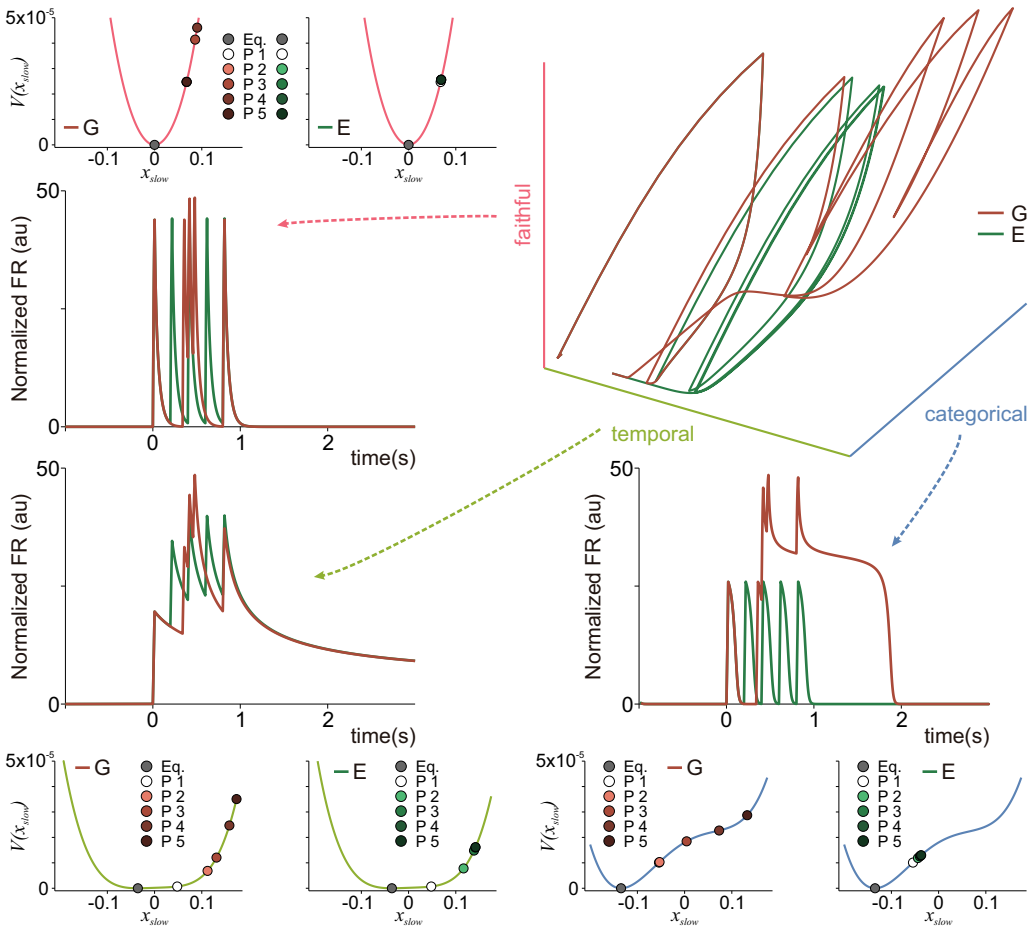
Supplementary Figure 3. Population dynamics in S2 during the whole TPDT and LCT. (A, B) Principal component analysis (PCA) was applied to the covariance matrices computed with the neural activity from the entire active and control task (-1 to 7.5 s), to derive their principal components (PCs). Then the activity from the whole TPDT (n=1646) and LCT (n=313) sorted by class identity were projected onto those PCs and ordered by their explained variance (PC1: 31.44%; PC2: 18.5%; PC3: 8.75%; PC4: 8.02%), which are also resumed in panel C. In TPDT, sensory (PC1), transformed/categorical (PC3) and inter-trial decision (PC3 & PC4) dynamics plainly emerged. On the other hand, in LCT, only sensory (PC1) and temporal inter-trial (PC2) dynamics, remained.



Supplementary Figure 4. Orthogonal dynamics emerged in S2 from continuum responses during the TPDT. Contrary to Figure 5, here, we employed the neural activity from the whole TPDT. (A) 2-D nonlinear UMAP density projection of the normalized activity of all units (n=1646, each black dot) recorded during the whole TPDT (-1 to 7.5 s), exhibited a continuum of responses. The color X marks represents the centers of double-Gaussian (σ = 0.5) weighted averages displayed in the panels in B. (B) Each subpanel represents the weighted average of neuronal activity and is presented in back-transform arbitrary units that roughly correspond to the firing rate values (Hz). Notice that categorical average responses are surrounded by sensory and temporal average signals, as they also are anchored in the middle of the 2D projection. This kind of response might be useful for the S2 network to connect and form the continuum.



Supplementary Figure 5. Continuum of responses in S2 breaks under the control task. Contrary to Figure 6, here, we employed the neural activity from the whole LCT. (A) Each subpanel represents the weighted average of neuronal activity in arbitrary units that approximately correspond to firing rate values (Hz). As LCT is a task variant that does not demand cognitive effort, the representative averages displayed only sensory and temporal responses. (B) UMAP density projection of the normalized activity of all units (n=313, each black dot) recorded during the whole LCT (-1 to 7.5 s), showed a breakup in the continuum that emerged during TPDT, exhibiting discrete subclusters of sensory (bottom) and temporal (top) dynamics. The color X marks represents the centers of double-Gaussian ($\sigma = 0.5$) weighted averages that are displayed in the panels in A.



Supplementary Figure 6. Eigenmodes define network dynamics in S2. Encoding along different network eigenmodes is determined by the qualitative properties of the reduced dynamics inside each eigenmode. The three different eigenmode types of the simulated S2 network: faithful (pink), temporal (orange), categorical (dark blue). For each mode, we show the associated one-dimensional response and sketch the 1D potential function that determines the mode qualitative dynamical response. Faithful modes are characterized by a small positive feedback gain and therefore have a sharply concave potential function which makes the response fast, strongly attractive, and faithfully tracking the presented stimulus. Temporal modes are characterized by an intermediate strength of positive feedback and therefore have much a flatter potential function, which slows down the response and allows for temporal integration of the presented stimuli. Categorical modes are characterized by a large positive feedback gain and therefore have a non-convex potential function, with either a double well (bistable case) or not (metastable case), whose inflection point determines the appearance of a threshold for categorization.