Please cite this article in press as: Tye et al., Mixed selectivity: Cellular computations for complexity, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.017

Neuron

CellPress
OPEN ACCESS

Perspective

Mixed selectivity: Cellular computations for complexity

Kay M. Tye,[1,2,5,11,*] Earl K. Miller,[3,*] Felix H. Taschbach,[1,4,5,*] Marcus K. Benna,[5,*] Mattia Rigotti,[6,*] and Stefano Fusi[7,8,9,10,*]

[1]Salk Institute for Biological Studies, La Jolla, CA, USA
[2]Howard Hughes Medical Institute, La Jolla, CA
[3]The Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4]Biological Science Graduate Program, University of California, San Diego, La Jolla, CA 92093, USA
[5]Department of Neurobiology, School of Biological Sciences, University of California, San Diego, La Jolla, CA 92093, USA
[6]IBM Research, Zurich, Switzerland
[7]Center for Theoretical Neuroscience, Columbia University, New York, NY, USA
[8]Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA
[9]Department of Neuroscience, Columbia University, New York, NY, USA
[10]Kavli Institute for Brain Science, Columbia University, New York, NY, USA
[11]Kavli Institute for Brain and Mind, San Diego, CA, USA
*Correspondence: tye@salk.edu (K.M.T.), ekmiller@mit.edu (E.K.M.), ftaschbach@salk.edu (F.H.T.), mbenna@ucsd.edu (M.K.B.), mr2226@columbia.edu (M.R.), sf2237@columbia.edu (S.F.)
https://doi.org/10.1016/j.neuron.2024.04.017

SUMMARY

The property of mixed selectivity has been discussed at a computational level and offers a strategy to maximize computational power by adding versatility to the functional role of each neuron. Here, we offer a biologically grounded implementational-level mechanistic explanation for mixed selectivity in neural circuits. We define pure, linear, and nonlinear mixed selectivity and discuss how these response properties can be obtained in simple neural circuits. Neurons that respond to multiple, statistically independent variables display mixed selectivity. If their activity can be expressed as a weighted sum, then they exhibit linear mixed selectivity; otherwise, they exhibit nonlinear mixed selectivity. Neural representations based on diverse nonlinear mixed selectivity are high dimensional; hence, they confer enormous flexibility to a simple downstream readout neural circuit. However, a simple neural circuit cannot possibly encode all possible mixtures of variables simultaneously, as this would require a combinatorially large number of mixed selectivity neurons. Gating mechanisms like oscillations and neuromodulation can solve this problem by dynamically selecting which variables are mixed and transmitted to the readout.

OVERVIEW

Not all brain functions are complex. They do not need to be. Simple functions can be performed by simple architectures or single layers. Seeing an object approaching, tasting a poison, detecting food in your esophagus and swallowing it, and recoiling from something that causes tissue damage are all simple functions that simple circuits and cells can accomplish. The architectures that underpin these functions can have straightforward properties or simple combinations of properties that serve this function, much like a railroad track providing a straightforward, direct route. The lack of flexibility makes these functions quick, efficient, and stereotyped.

By contrast, the neural systems responsible for complex thought and behavior mandate flexibility. Intelligent thought *is* flexible thought. All creatures can react reflexively to the environment. But animals with more complex nervous systems can change how they behave by integrating more parameters into the decision-making process. They tailor ongoing behavior to the current situation and to an ever-shifting set of subgoals and goals. They also take into account an accumulating history of events that bias decision thresholds. This capacity for generalizing context-dependent behavior is crucial for our ability to project our behavior into the future, allowing us to make and execute plans.

The neural substrate for this flexibility can be seen in many places, but it is especially prevalent in cortical areas known to be critical for flexible behavior, such as the prefrontal cortex (PFC). Individually, PFC neurons have adaptive and multivariate response properties,[1,2] referred to as "mixed selectivity."[3–5] They wear many hats, showing different patterns of selectivity in different behavioral contexts. Information is often widely distributed across them. The signal-to-noise ratio of responses for each individual neuron is low, but decodable information for the population is high. PFC neurons have moderately high basal firing rates and high proportions of neurons responding to many stimuli—relying on mixing more diverse inputs that give each individual neuron many jobs encoding many variables. Our view of

Please cite this article in press as: Tye et al., Mixed selectivity: Cellular computations for complexity, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.017

Box 1. The history of the theory of mixed selectivity

Since the time that artificial neural networks were conceived, their typical units have always performed two basic operations: sum multiple signals coming from different neurons and then compute a nonlinear function to this sum. It is then not surprising that the units of these networks exhibit mixed selectivity, as it is almost unavoidable if the input neurons encode different signals. What is less obvious is an understanding of the computational role of mixed selectivity neurons. This has always been one of the important roles of theoretical neuroscience. From the early days of neural network theory, it has been clear that some representations simply do not work and need to be changed. David Marr often speaks about recoding,[8] invoking it when patterns of activities that need to be discriminated are too similar. The first one to realize the importance of mixing is probably Rosenblatt,[9] the father of modern learning neural networks. His perceptron, in the original version (not in one discussed in the book *Perceptrons*[10]), had an intermediate layer of neurons that were randomly connected to the inputs. These neurons had mixed selectivity, and they were necessary to make the representations linearly separable and hence classifiable by a simple linear readout. Other more recent recurrent neural networks were using the same ideas (random projections) to generate mixed selectivity representations with higher dimensionality[11] (echo state machine[12], liquid state machine[13]). The idea behind the SVMs (support vector machines) with nonlinear kernels[14] is again the same: to implicitly transform the representations to make them linearly separable. SVMs with nonlinear kernels are equivalent to a simple neural network with a very large intermediate layer of nonlinear units that mix the inputs nonlinearly, expanding the dimensionality. When we studied attractor neural networks that implement finite-state machines,[3] we realized that we had to use a similar approach and, for the first time, we related dimensionality expansion and nonlinear mixed selectivity in clear terms.

Besides this important but also basic general idea of dimensionality expansion, there were several other theoretical works discussing the importance of mixing in specific problems. For example, in discussions of parallel distributed processing and specifically distributed representations, the point of conjuctive coding is highly relevant to mixed selectivity. Here, Rumelhart et al. point out that the "binding problem" can be solved using neurons capable of local tuning[15]—a capability potentially endowed by differential neuromodulation at each dendritic segment that may be electrically compartmentalized.

Another example: mixed selectivity to the retinal location of a visual stimulus and the position of the eyes can be used to generate a representation of an object's position and then determine the changes in joint coordinates needed to reach the object.[16–19] Neurons with mixed selectivity to the identity of a visual stimulus and its ordinal position in a sequence have been used to model serial working memory.[20] Mixed selectivity to stimulus identity and to a context signal have been used to model visuomotor remapping.[21] More generally, complex nonlinear functions of the sensory inputs, like motor commands, can be expressed as a linear combination of basis functions.[22]

the PFC has evolved: rather than trains of thought on railroad tracks,[6] the PFC more closely resembles cars on the road and highway system. Neurons and their axons provide the highways—the anatomical architecture over which thoughts, feelings, sensations, and motor commands can travel. But the way the roads are used is vastly different based on the immense complexity of the electrical and chemical impulses being trafficked all around the brain. Vehicles of information can take different paths and different destinations. They might share a common path with some vehicles at some times and others at other times. This autonomy provides maximal freedom and dimensionality.

With this comes depth of thought. The multivariate neuronal properties can increase the representational dimensionality of the population, allowing more complex computations.[4,5] Further, the brain tailors the dimensionality to the task at hand. Dimensionality expands and contracts to focus processing along relevant dimensions.[7] This keeps processing on-task and goal directed.

In short, mixed selectivity gives the brain the processing power needed for complexity and flexibility. The cost of this opportunity for flexibility is the inability to take advantage of the regularities of the world to generalize to novel situations. Mixed selectivity allows representation of a large number of different situations in every detail, but sometimes we need to discard or ignore some information to make the right decision.

In this perspective piece, we consider the impact of mixed selectivity on our understanding of neural processing, with the goal of grounding it in a biological implementation. To maximize the transparency of the concept of mixed selectivity, we will first define and describe mixed selectivity and then discuss its importance.

WHAT IS MIXED SELECTIVITY?

There has been some confusion around this term, which has emerged from a number of related concepts (Box 1), but the definition is simple. At its core, mixed selectivity involves a single cell showing consistent activity, which is modulated by multiple statistically independent variables. This single-cell behavior eventually has broad implications for the activity of neuronal ensembles, enabling them to process and integrate a range of independent inputs.

Pure selectivity

In many ways, the notion of pure selectivity can be traced to Hubel and Wiesel's description of the "simple and complex cells" in the primary visual cortex. This pioneering work was a leap forward as it was one of the first demonstrations of the functionality of single neurons. Simple and complex cells spiked to single features. For some time, most theories of cortical function considered pure selectivity.

Please cite this article in press as: Tye et al., Mixed selectivity: Cellular computations for complexity, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.017

CellPress
OPEN ACCESS

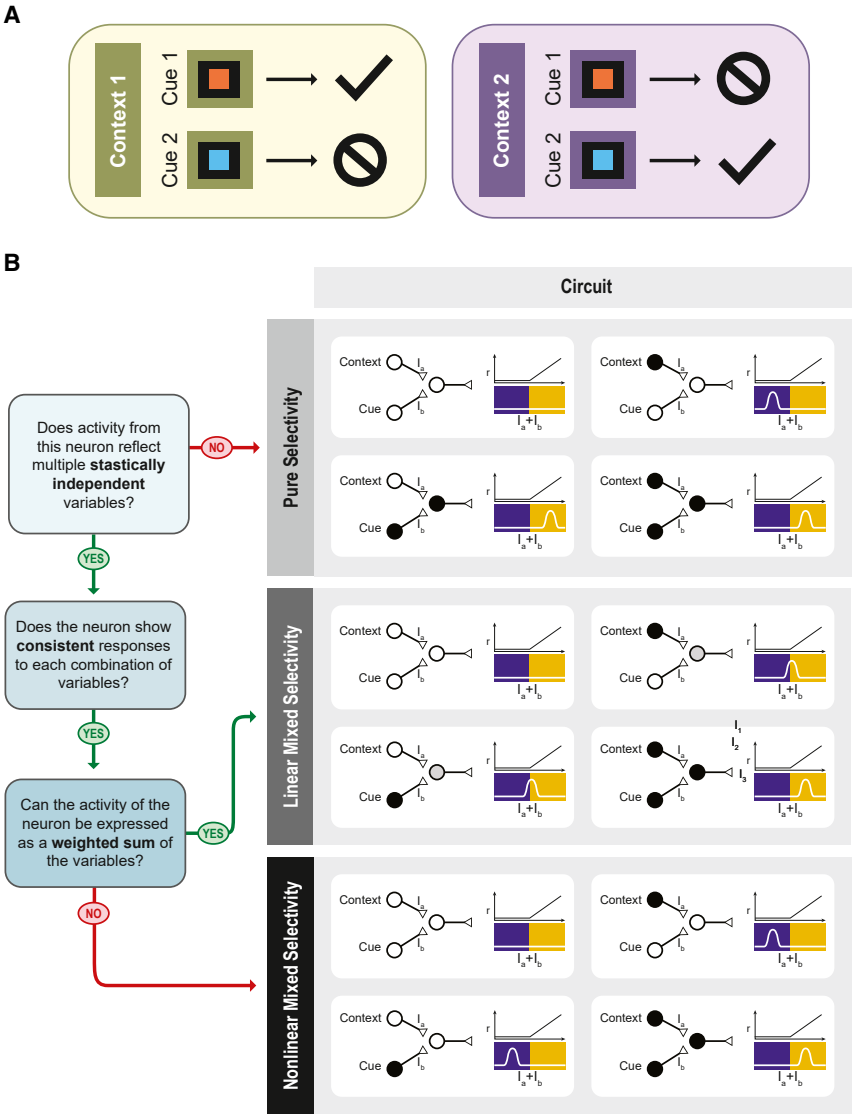

Figure 1. Delineating pure, linear mixed, and nonlinear mixed selectivity using a context-dependent task

(A) In the task, a subject is presented with one of two context signals and subsequently makes a choice from two cues. Rewards are determined based on the initial context.

(B) A flowchart accompanied by circuits elucidates the neuron response categories: pure, linear mixed, or nonlinearly mixed selectivity. Within each circuit, two initiating neurons are distinctly selective for either the context or the cue, directing their specific signals (I_a and I_b) to a third downstream neuron. Alongside each circuit, diagrams exhibit the neuron's firing rate in relation to the cumulative input current. The neuron initiates activity when the combined current (here depicted as a histogram) surpasses a certain threshold, leading to a linear increase in the firing rate. If a neuron's activity is solely representative of a single input, it is termed "purely selective," exemplified in the top circuits where activation is exclusively due to the cue signal. Neurons that represent a linear combination of the independent inputs, consistently responding to their variable combinations, are termed "linear mixed selective," demonstrated by the middle circuits. Here, the neuron has diminished firing rates for individual signals but is most active when both signals combine. Conversely, "nonlinear mixed selective" neurons, represented in the bottom circuits, cannot be described as a linear sum of inputs. Here, the downstream neuron only becomes active when both the cue and context signals are present simultaneously.

The focus then was, understandably, on the first-order questions and parameters we could confidently measure—how neurons represent simple sensory features or produce motor outputs. This, after all, was one of our first forays into how cortical neurons process information. This was studied in experimental paradigms designed to be high fidelity, robust, and intentionally redundant. The purpose was to reduce variability ("noise") so that we could be confident about the processes that sense and move our bodies within our environments.

As we built on this foundational work, our view of cortical neurons has become more dexterous and multifunctional. The more we learn, the more noise is a misnomer. It is rather factors that experimentalists did not or could not measure.[23] Initially, the idea that neurons could do more than one thing was blasphemy (one of us was accused of "turning cortex into a bowl of porridge" at a major meeting). But evidence kept amassing for multifunctionality, including observations such as neurons releasing different neurotransmitters for different functions.[24,25]

It is now evident that pure selectivity, as important as it is, is more of an exception than the rule, especially in the cortex. Its habitats tend to be at the input edges of sensory systems and at the output edges of motor systems. In much of the rest of the cortex, the preponderance of neurons show multifunctional selectivity.[7,26–43]

Linear mixed selectivity

Within mixed selective neurons, there are either nonlinear or linear mixed selectivity neurons. Linear mixed selectivity[4,44–46] neurons are typically the result of a process of abstraction and can facilitate generalization.[34] Linear mixed selectivity neurons show activity that can be expressed as a weighted sum of the responses to each variable (Figure 1B). Consider, for instance, a neuron that responds to both written praise in an email and oral praise.

How does this happen? How do neurons that respond to reading and listening to praise find their way to mutually synapse on the same cell? The likely explanation is that many cells have many inputs, and those associated with environmental information that have an eventual impact on our survival, including acceptance by our social group, will be strengthened, while

Please cite this article in press as: Tye et al., Mixed selectivity: Cellular computations for complexity, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.017

less significant ones will be weakened or pruned. Moreover, the surviving connections are tuned in such a way that the neuron operates in a linear regime.

Nonlinear mixed selectivity

Neurons that respond to multiple, statistically independent variables, where the activity tracking these two variables cannot be expressed as a weighted sum, are considered to be nonlinear mixed selective neurons (Figure 1B). These neurons can be modeled as computing a weighted sum of the activity of their presynaptic neurons and then passing this sum through a nonlinearity. The input neurons must encode different variables, and they can do so in multiple ways: they can be pure selectivity, nonlinear, or linear mixed selectivity neurons. Moreover, any of the numerous nonlinear mechanisms that are involved in the normal functioning of a neuron can contribute to generating nonlinear mixed selectivity: the frequency-input (f-I) curve of a neuron is typically nonlinear,[47] and that would be sufficient. However, there are also nonlinearities in the dendritic integration[48–57] and in synaptic transmission.[58,59]

Rather than abstraction, nonlinear mixed selective neurons offer the capacity for complexity and depth. Because of their versatility, these neurons require a wide array of inputs. To have neurons be readily repurposed in different contexts also suggest a hierarchical organization. Timescales, locations, probabilities, and internal states are all variables that can contribute to context. Context-modulation of cue responding is one of the key applications of this capability. Cues are presented on discrete timescales, while contexts are represented on a longer timescale. Different contexts can influence the value or meaning of cues and, therefore, the same cue should have distinct neural representations under different contexts, as this would be the only way to produce outputs that would yield different behaviors. For example, imagine a mouse encountering an unfamiliar type of berry in its environment. If the mouse is alone and has never experienced this particular berry, it might be hesitant to eat it due to the potential risks of poisoning or illness. However, if the mouse then encounters a conspecific whose breath carries the scent of that same berry, it may interpret this as a sign that the berry is safe to consume. In this context, the smell of the berry on the breath of a fellow mouse serves as a form of social verification, indicating that the food is likely safe and worth the effort to eat. This social cue effectively shifts the mouse's evaluation of the unfamiliar berry from potentially risky to likely safe, all based on the social transmission of food preference[60] (Figure 3A).

WHY MIX NONLINEARLY?

Representations based on nonlinear mixed selectivity have specific computational properties: in particular, they can be high-dimensional and hence usable by a simple linear readout like a downstream neuron. These high-dimensional representations confer enormous flexibility to the neural circuit that reads them out. To understand them, it is instructive to consider what is called representational geometry (Figure 2). It is easier to start from the representations that do not contain mixed selectivity: say, for example, that a visual stimulus can appear in one of two possible contexts. We denote the stimulus by S and the context by C. Both variables have only two values as there are only two stimuli and two contexts. Imagine that all the neurons have pure selectivity to either S or C. For example, we consider three neurons: the firing rate of the first one, r_1, is equal to S (pure selectivity to the stimulus); the second one, r_2, responds only to context C; and the activity of the third one, r_3, depends on S only, but in a more complicated way, i.e., $r_3 = 1-S$. We now consider the activity space in which r_1, r_2, and r_3 are represented along the coordinate axes. In this simple example, we consider only three neurons, but typically this is a high-dimensional space with many axes. Each point in this space represents the population response in one experimental condition (e.g., the first stimulus is presented in the second context). We will have four points in the activity space as we have two contexts and two stimuli.

The different types of selectivity yield different geometries. As the neurons have only pure selectivity, the four points will define a relatively low-dimensional object: the number of dimensions will be equal to the number of task-relevant variables. In our case, the variables are two, and the four points form a rectangle (Figure 2A). The arrangement of these points defines the geometry of the representation. In the case in which we replace the third neuron with a linear mixed selectivity neuron (Figure 2B), the geometry does not change substantially: the flat object that we saw in the case of pure selectivity neurons will rotate, but it will not change its dimensionality. The geometry changes significantly when introducing a nonlinear mixed selectivity neuron (Figure 2C). Now, the four points define a three-dimensional (3D) object called a tetrahedron.

Why is the dimensionality of the representation important? To understand it, we need to take the perspective of a downstream neuron reading out the representation. What this neuron can or cannot do depends on how the points of the representation are arranged. Consider, for example, the situation in which a readout neuron has to be trained to respond to a particular stimulus always in the same way, ignoring the context in which the stimulus is presented. We consider this a simple task. This can be visualized in the activity space by coloring the points according to the desired response of the readout neuron (Figure 3B). For example, we color purple the points corresponding to the conditions where the readout neuron should be active (S = 1) and orange the points for which the readout neuron should be inactive (S = 0).

Is it possible to connect the readout neuron to the pure selectivity neurons that we discussed in Figure 2A in such a way that it can solve this simple task? To answer this question, we need to make further simplifying assumptions. Say that the readout neuron can perform a simple operation: compute the weighted sum of the inputs and compare it with an activation threshold (linear readout). Biological neurons are complex enough to perform this operation. Graphically, we can visualize this operation by drawing a plane (Figure 3B) that separates the inputs in the activity space that activate the readout neuron (the points above the plane) from the inputs that do not reach the activation threshold (the points below the plane). For this geometry, it is possible to draw a plane separating the purple points from the orange points; in other words, a simple linear readout can

Please cite this article in press as: Tye et al., Mixed selectivity: Cellular computations for complexity, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.017

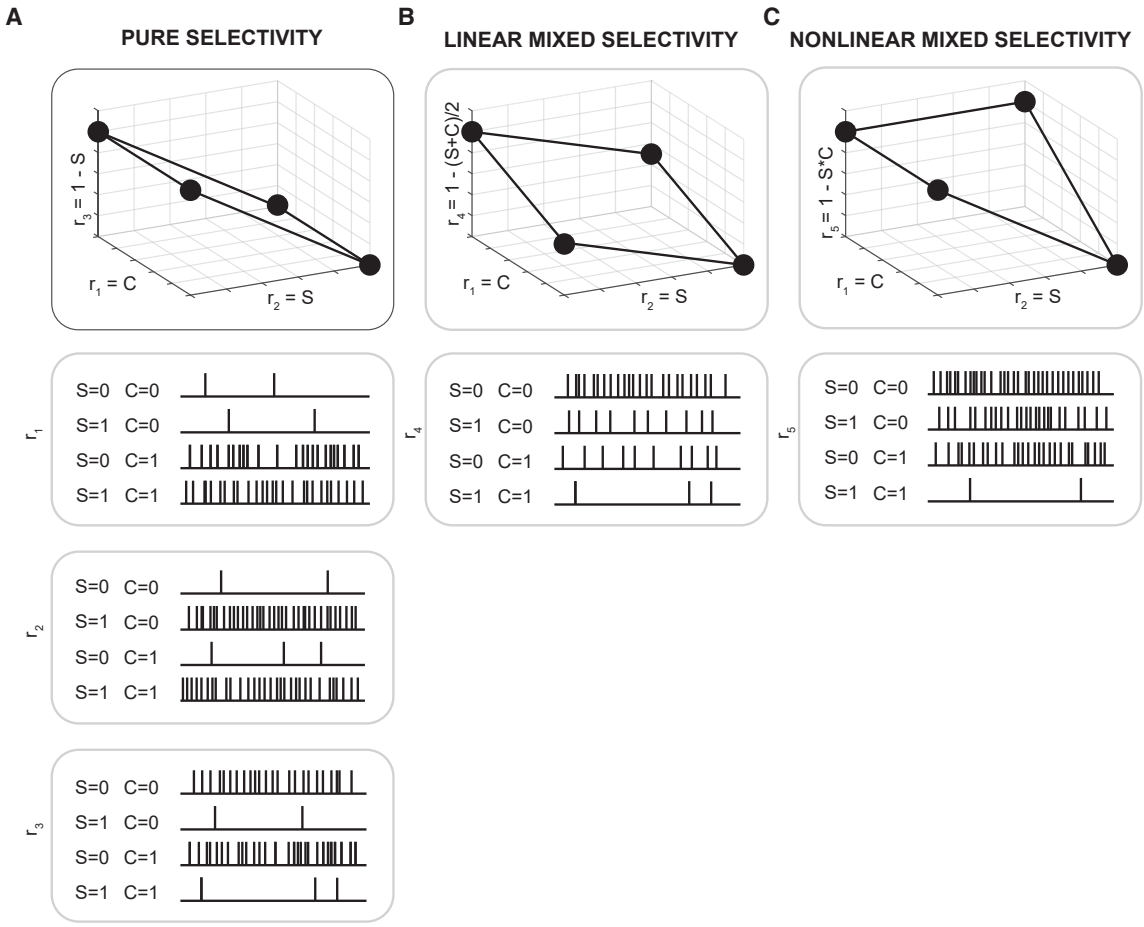

Figure 2. Changes in the representational geometry due to linear and nonlinear mixed selectivity

The plots showcase the firing rates of three neurons in response to different combinations of stimuli (S) and contexts (C), both holding two possible values. For simplicity, two of the firing rates, r_1 and r_2, are depicted as being purely selective to either the context or the stimuli.

(A) Pure selectivity: here, r_3's activity is inversely related to S ($r_3 = 1-S$). The three-dimensional plot reveals that, in the activity space, the four combinations of S and C outline a rectangle. Corresponding raster plots below display the firing patterns of these neurons for each S and C combination, revealing distinct activation based on their selectivity.

(B) Linear mixed selectivity: with r_4 responding to a linear combination of the context and stimuli the quadrangle outlined by the four combinations is rotated in this activity space.

(C) Nonlinear mixed selectivity: with r_5 demonstrating nonlinear mixed selectivity, the activity space transforms, with the four points now constituting a 3D tetrahedron.

perform this task. The same applies to the high-dimensional representation that involves a nonlinear mixed selectivity neuron. But consider now the same geometries when the task is more complex: the readout has to respond only to stimulus 1 when it appears in context one and to stimulus 0 when it appears in context zero, but not in the other two cases. Now, the coloring of the points is different and, interestingly, it is possible to separate the purple from the orange points in the case of the high-dimensional representation, but not in the case of the low-dimensional one (Figure 3C). This problem is equivalent to the well-known exclusive-OR (XOR) problem, and no plane can separate the points of the low-dimensional representation as required.

In this simple example with only four points, this coloring is the only one that does not have a linear solution. However, as the number of points increases, the colorings that require high dimensionality grow exponentially and become the majority. So, the more complex the task, the more important nonlinear mixed selectivity becomes. Note also that when the number of points increases, the maximal dimensionality increases as well. If one wants to achieve maximal dimensionality, nonlinear mixed selectivity is not sufficient, we also need an additional ingredient: the responses of the neurons have to be *diverse* enough. This is possible only when neurons respond to different combinations of the task-relevant variables. Maximal dimensionality confers a great flexibility to simple linear readouts, as they can separate the points in any arbitrary way or perform many different tasks. However, the maximal dimensionality can only be achieved when neurons mix nonlinearly and have diverse responses.

Diversity does not necessarily mean "completely unstructured." If neuronal responses are characterized by a vector of regression coefficients, and one plots these vectors as points

Please cite this article in press as: Tye et al., Mixed selectivity: Cellular computations for complexity, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.017

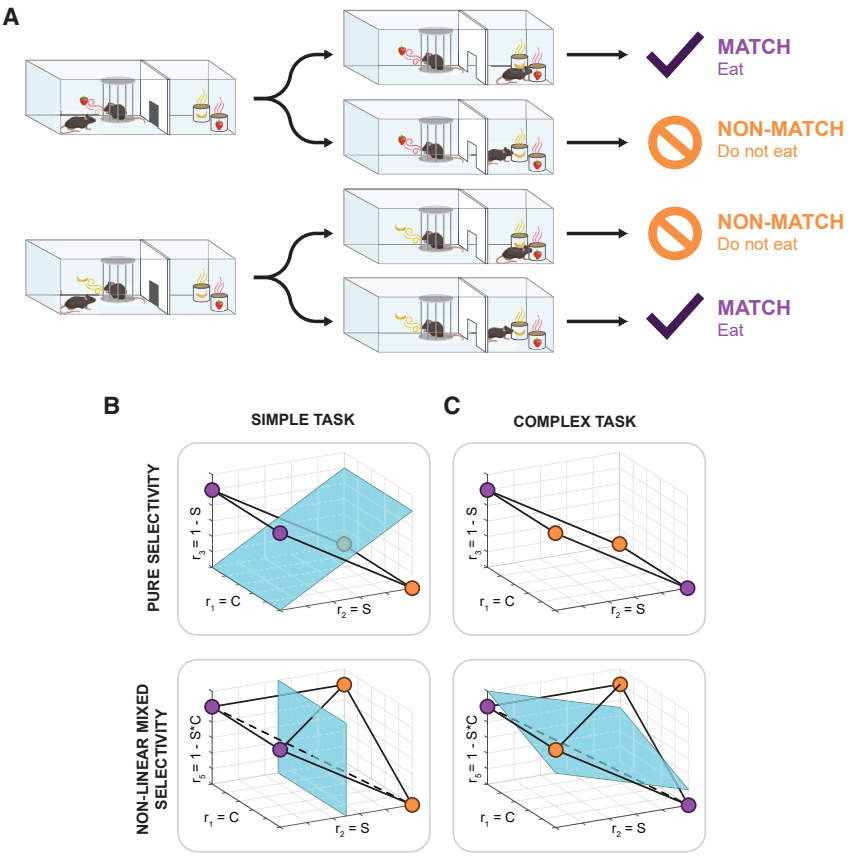

Figure 3. Understanding the importance of dimensionality in neuronal readouts

(A) Social transmission of food preference is a more ethological variant of the delayed match to sample task, showcasing context-dependent decision-making. Initially, a mouse meets a conspecific with a breath scent of an unfamiliar fruit, either a banana (C = 0) or a strawberry (C = 1). Subsequently, it finds two new fruits—a banana (S = 0) and a strawberry (S = 1). Informed by the prior interaction, the mouse's choice is steered by the social transmission of food preference, leaning toward the fruit it detected earlier.

(B and C) The role of dimensionality in solving complex tasks. (B) A downstream neuron receiving activity from r1, r2, and r3 can implement a linear readout to perform simple tasks, such as distinguishing between a banana and a strawberry (S = 0 versus S = 1). (C) For more complex tasks, such as firing when the context smell matches the stimulus (C = 0, S = 0 or C = 1, S = 1), it is not possible for a downstream neuron to implement a linear readout to solve this task. However, a downstream neuron that receives the activity of r1, r2, and r5 as input can easily perform a linear readout due to the higher dimensionality achieved by the nonlinear selectivity of r5. Created with BioRender.com.

in a "selectivity" space, we often do not see any structure and the neurons seem to respond to random combinations of the task variables.[34,37,61–63] However, there are also experiments in which it is possible to observe some structure in the form of clustering: there are groups of neurons that tend to respond in a similar way.[64,65]

Single-neuron versus population-coding properties

The above classification into pure, linear, or nonlinear mixed selectivity categories is done at the level of single neurons. Crucially, this first requires us to decide on a set of relevant external variables (describing, e.g., stimuli or behaviors) that we are interested in—for a different choice of this set of variables the classification of a neuron will differ in general. Let us consider again the earlier example of a neuron that responds to both written and verbal praise. If this neuron's activity corresponds to the total amount of praise received, regardless of its mode of communication, then it could be said to exhibit pure selectivity for praise in general rather than mixed selectivity for written and verbal versions of it, even though the latter description is also valid.

We can establish the selectivity properties of a single neuron by fitting a linear regression model to predict its firing rate from the external variables that we have chosen to investigate. If the fitted coefficient of only one variable is significantly different from zero, we can describe the neuron as having pure selectivity for that variable, whereas multiple non-zero coefficients would correspond to linear mixed selectivity. Furthermore, we can generalize the regression model by including features that are nonlinear combinations (e.g., pairwise products) of the external variables of interest. If this extended model can achieve a significantly higher (cross-validated) accuracy for predicting the firing rate of the neuron by assigning non-zero coefficients to these nonlinear features, the neuron can be said to exhibit nonlinear mixed selectivity. Such a regression analysis is closely related to ANOVA.[4,34,37,61–63]

However, beyond these single-neuron properties, we can also examine the neural representations that the activity patterns of these neurons create at the population level, and many of their properties and computational implications do not depend on the way we choose the external variables of interest. As discussed above, one particularly important geometric property of such a representation is its dimensionality. Pure and linear mixed selectivity lead to low dimensionality, equal to the number of relevant variables. Nonlinear mixed selectivity can increase the dimensionality of the representation beyond this value.

Many computational problems can only be solved by neural networks with hidden units, i.e., they require nonlinear mixing of inputs. A network that has learned to correctly execute such a task will therefore contain at least some neurons that exhibit nonlinear mixed selectivity with respect to the variables encoded in the network inputs. However, if the network was trained for executing a particular task, it may only require very specific nonlinear mixed selectivity neurons that combine task-relevant variables in a certain way that supports the chosen task rather than a variety of different nonlinear mixed selectivity neurons with a diverse set of coding properties. In other words, even if

Please cite this article in press as: Tye et al., Mixed selectivity: Cellular computations for complexity, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.017

the network performs a complex task that requires nonlinear mixing, its representations do not necessarily have to be very high dimensional.

However, representations with (close to) maximal dimensionality, which are achieved only when the responses of different neurons are sufficiently diverse, are an important coding scheme in situations in which a population of neurons does not have sufficient information about which task it is meant to support (perhaps because the neurons are not provided with rich and individually tailored error signals as they would be, e.g., during learning via backpropagation) or if the animal needs to potentially execute a large class of different tasks and flexibly switch between them. In such scenarios, the neural population cannot shape its representations to specifically subserve a particular task, but it can nevertheless try to form a neural representation of the task-relevant variables that is generally useful for executing many possible tasks. One concrete way to achieve this is to create a high-dimensional representation that enables a downstream neuron implementing a linear readout (or linear classifier) to correctly extract many functions of the input (perhaps even all possible dichotomies for a moderate number of binary inputs). In this sense, high dimensionality is associated with the (cognitive) flexibility of the animal—performing new tasks is then simply a matter of finding a different readout, which can be done using a simple perceptron or delta rule, but does not require changing the internal representations, which would necessitate more sophisticated learning algorithms.

Expression of mixed selectivity across circuits and structures

To appreciate what our brains might look like without mixed selectivity, we can examine which functions and structures at the micro- and macrocircuit level rely on mixed selectivity. Few functions are completely devoid of circuit flexibility, and although the labeled lines circuit motif is the oldest in neuroscience, it is not the most prevalent.[66] Conversely, mixed selectivity was discovered in the context of higher cognitive functions, allowing the brain to learn new rules and switch between different rules in different contexts.[1,4,5,67] These computations for complexity can be seen in cells and circuits down to the olfactory glomeruli[68] and in the auditory[69] and somatosensory cortex.[70] The fact that we can learn to like poisons (like caffeine and alcohol) and learn to hate calorie-rich foods (pistachio ice cream after food poisoning) suggests that even our sensory systems have flexibility.[70,71] There are many circuit motifs that can give rise to flexibility, circuits that diverge depending on gating, that converge to be integrated, or that compete to orchestrate competing mechanisms,[66] and the principles of mixed selectivity are foundational to them all.

Labeled lines circuits have their perks. Signal-to-noise ratio is maximized, as crosstalk is minimized. The lack of crosstalk guarantees high fidelity. Moreover, pure selectivity neurons can have other advantages in terms of energy consumption and number of needed connections. Although the representations based on pure and linear mixed selectivity are completely equivalent from the point of view of a linear readout (they have the same geometry), when one imposes that the neural activity can only be positive, the pure selectivity neurons consume less energy.[72]

Let us begin by considering sensory association systems wherein the goal is to produce appropriate reliable responses to stimuli. Of all types of information, getting reliable sensory information about our dynamic environment as we navigate it is paramount. Bipolar cells relaying information from photoreceptors to retinal ganglion cells do not require complexity or depth as much as fidelity—they know what information that photoreceptor is providing, which makes decoding trivial.

Initially, our brains filter incoming sensory information, forwarding for full processing only what is important to send along. Neurons within the thalamic nuclei handle this initial filtration, sending massive spikes that increase ~500%[73] from their basal activity for the sharpest, most fleeting of signals to be unambiguously broadcast to multiple distributed systems. In parallel, other sensory processing systems through the auditory cortex, for example, will undergo plasticity based on experience, behavioral state, hormonal signals, etc.[71,74–76]

The thalamus sends information to many places throughout the corticolimbic system[77]—even the basal ganglia,[78] including the basolateral amygdala (BLA) and the PFC.

The BLA is a hub for integrating sensory information and rapidly converting that into a behavioral response.[79] The amygdala is a relatively primitive structure with the capacity for plasticity to form new associations and lies in the middle of the spectrum. The amygdala has been demonstrated to have a substantial amount of hard-wiring[79–82] in terms of certain neurons driving specific functions irrespective of state or context, although there is still some evidence for mixed selectivity for some variables (see, e.g., O'Neill et al.[83]). Further support for the notion of fixed functions existing in the BLA include work demonstrating the existence of a long-lasting engram or memory trace that is stored long term in the same cells in which it was formed.[84–86]

Both the BLA and PFC receive sensory information from many inputs, including the thalamus, but process it differently.[87] "Structured" as a divergence point for positive and negative valence representation, the amygdala has some neurons selectively responding to positive valence and others to negative,[81,82,88–91] though it mixes some other variables.[83] The striatum then collates these different signals, weighs them, and decides on a single motor plan from several rehearsed or innate motor sequences.[92–94]

Conversely, the PFC approaches the received information with precisely the opposite strategy as it has diametrically distinct functional goals from the sensorimotor chain. Its computational aim is to weigh choices and delve deeper rather than produce fixed motor responses. Thus, a different set of tools are needed. This requires a different cellular and synaptic architecture that deviates from hardwired functions, opting for maximal flexibility. This flexibility is achieved by sending diffuse signals everywhere, each dimension represented by a combination of decentralized signals, allowing any number of readily decipherable messages to be selected for readout at any given time.

Unlike the strong signals sent out by the amygdala, the PFC communicates with a hum of whispers. Classic plots of

Please cite this article in press as: Tye et al., Mixed selectivity: Cellular computations for complexity, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.017

peri-stimulus time rasters or histograms show high baseline noise with modest signals —but an extraordinarily high number of variables can be decoded from PFC activity. Importantly, high dimensionality is not random in the sense that these neurons are always processing different signals; they are reliable and perform consistently within the same context, showing different responses in varying contexts. Information is not represented by individual cells but is more prevalent in the population.[67] It is routed in different ''subspaces'' through population dynamics rather than individual neurons.[34,38,95–98]

The hippocampus and dentate gyrus utilize a sparse coding strategy, with most neurons silent at any given moment[99,100] and this coding scheme densely packed into layer CA1 still involves neurons with mixed selectivity to position, head direction, and speed of an animal that freely explores an environment.[37] Remapping can be interpreted as another expression of this coding scheme: hippocampal neurons encode sequences, the delivery of reward, and the encounter with other animals.[63,101] All these variables affect the neural representations and they can easily result in some form of remapping when mixed together with position. This is not surprising if one assumes that the hippocampus is fundamentally a memory system that compresses the complex sensory representations by taking advantage of their regularities (see, e.g., Gluck and Myers[102] and Benna and Fusi[103]). The resulting representations are sparser, more decorrelated, higher dimensional, and involve mixed selectivity neurons that encode all the different aspects of sensory experiences that are memorized.

Unlike sparse coding systems like the hippocampus, the PFC employs a more proliferative coding strategy. For instance, in the amygdala, approximately 10%–15% of neurons will respond to a salient stimulus, like a footshock-predicting stimulus, at low basal firing rates (~1 Hz).[84,85,90,104,105] However, in the PFC, the same predictive cue will be encoded by around 30%–40% of neurons, albeit with a lower signal-to-noise ratio (basal firing rates ~10 Hz).[106–108]

These differences in coding strategies in different micro- and macrocircuits across the brain endow the brain with its versatility and may help illustrate the functional utility of mixed selectivity when it is expressed to greater or lesser degrees within the mammalian brain.

NOT EVERYTHING CAN BE MIXED

We often observe that neural representations supporting a specific task exhibit the maximal dimensionality enabled by the task[4] and, hence, that the task-relevant variables are mixed in all possible ways. However, the experiments performed in a laboratory in a highly controlled environment typically involve a relatively small number of variables, sometimes only one or two. What happens when the subjects perform a real-world task? In a simple task, like the one considered in Figure 1 with only two binary variables, context, and stimulus, the total number of possible conditions is 4. In general, however, as a function of the number of binary variables V, the number of task conditions grows exponentially as 2^V. For a complex, real-work task expressing a sizable number of task-relevant variables V, this number can therefore be huge, meaning that reaching maximal dimensionality would require a correspondingly large number of mixed selectivity neurons. Because then the maximal dimensionality is bounded by the number of neurons and the number of conditions,[5,109] we might reach the upper limit imposed by the size of the neural population before all variable combinations are possibly encoded in mixed selectivity neurons.

Fortunately, the world is highly structured and solving a realistic task typically departs from the worst-case scenario that would require mixing all variables. For example, behavioral contexts are often compartmentalized such that we might have to mix a subset of variables in one task and distinct sets of variables in different tasks but never have to worry about mixing variables that are relevant in different tasks. If we can selectively gate the variables we are interested in, then the number of required neurons would be significantly smaller. In a simple example, where 2 tasks need to be implemented and each only involves a distinct subset of V/2 variables, we would need $2^{V/2} + 2^{V/2} = 2^{V/2+1}$ neurons as opposed to 2^V. The plot in Figure 4A shows how many neurons we need when we have to mix only a fraction, F, of the variables. As the world becomes more structured (i.e., 1/F increases), the number of needed neurons decreases very rapidly (note that the scale of the y axis is logarithmic). How can we implement such a gating mechanism that would allow our population to exploit this structure? One simple answer is to choose properly the connections and the response properties of the input neurons. Indeed, properly connected nonlinear neurons can readily implement the most general form of gating and solve arbitrarily complex problems. However, it is sometimes difficult to learn these connectivity schemes, and there are other forms of gating that can complement those based on the careful choice of the circuit architecture. For example, there are at least two other mechanisms that implement some form of dynamic gating (see Figure 4B), selecting which neurons to listen to depending on the task or the context. There are two implementational motifs for gating responses or signaling to a given neuron that it should participate in one ensemble or another: electrical and chemical. Oscillatory dynamics naturally organize neural activity into functional patterns (Figure 4C). Neuromodulatory signals rely on the diverse distribution of receptor expression profiles to signal the appropriate ensemble to amplify depending on cues, contexts, or internal states (Figure 4D). These two mechanisms likely work together to dynamically form functional networks.

Oscillatory dynamics flexibly organize mixed selectivity neurons

Like a stadium crowd doing ''the wave,'' mixed selectivity neurons can be organized on-the-fly, shifting their participation in different ensembles and networks for different functions. This organization has to occur in real-time at a scale large enough to produce function. One can see such organization in the oscillations of local field potentials (LFPs).[110,111] They reflect coordinated changes in neuronal excitability at the mesoscale, involving millions of neurons.[112] It is at this scale that the brain focuses attention, makes decisions, executes actions, and retrieves memories—processes that necessitate the coordination of numerous neurons.[6,113]

There is ample evidence for a role for neural oscillations in the dynamic organization of functional networks. Top-down

Please cite this article in press as: Tye et al., Mixed selectivity: Cellular computations for complexity, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.017

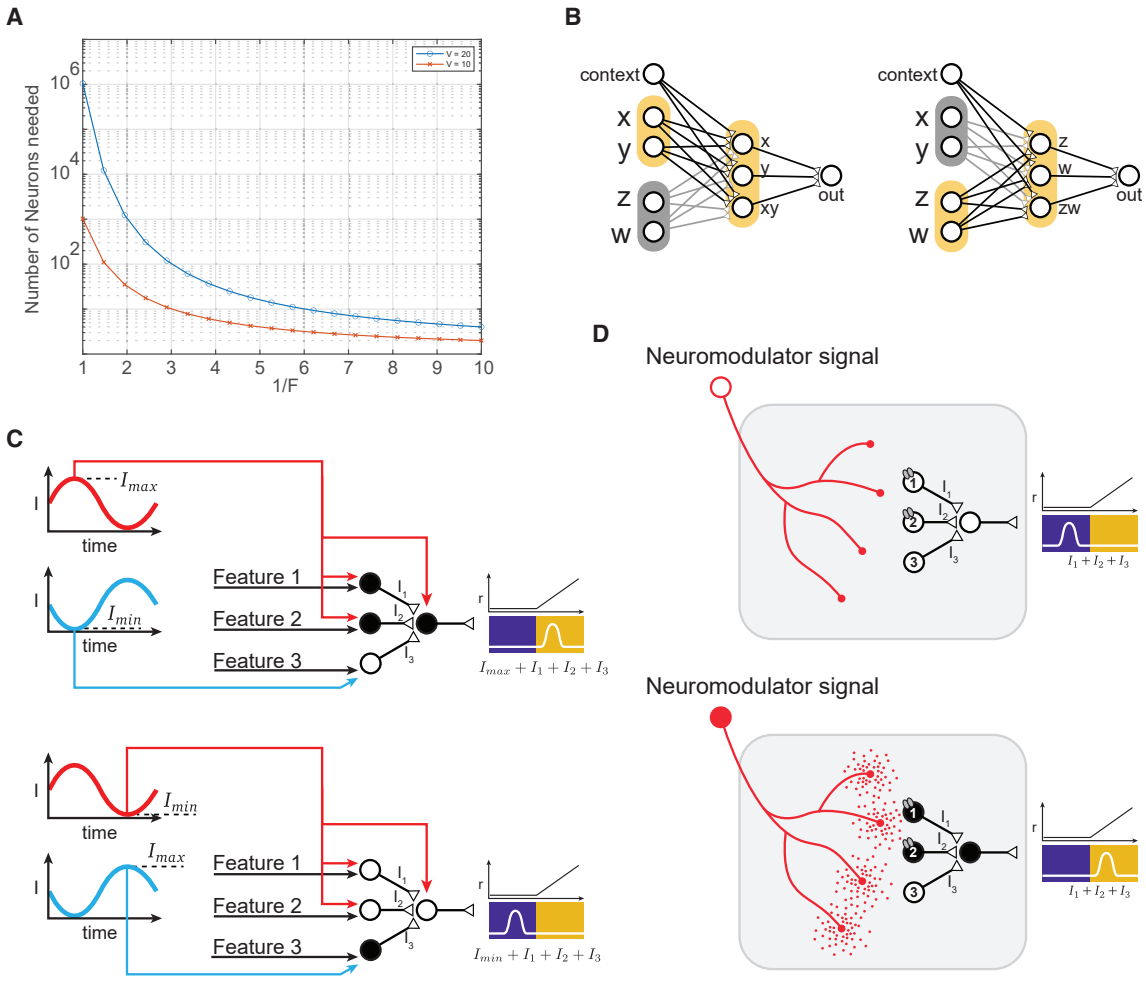

Figure 4. Strategies to orchestrate gating

(A) The number of neurons that are needed as a function of the number of variables that should be mixed for each task, when the number of tasks 1/F increases (where F is the fraction of all the variables that have to be mixed per task). The red curve is for a relatively simple task, with 10 binary variables, and the blue curve is for a more complex task, with 20 variables.

(B) The basic principle of dynamic gating. Here, we have a simple three-layer feedforward neural network that can compute the exclusive-OR function of the input. With a three-unit hidden layer, the network achieves maximum dimensionality with two inputs and can easily compute the exclusive-OR. With four input neurons, as we have in this example, we need to either increase the number of hidden units or gate half of the input using a context signal such as is shown here.

(C) Dynamic mixing using oscillations. One way in which this gating can be implemented biologically is through some of the input and downstream neurons receiving the same oscillatory input. In context 1, when the oscillatory input to the neurons coding for feature 1 and 2 is high, it will be easier for the neurons to spike, thus making it easier for the downstream neuron to integrate their signals and fire. Similarly, if that oscillatory input is low, it will be harder for these neurons to spike.

(D) Dynamic gating using neuromodulation. Here, the context signal takes the form of a neuromodulatory signal that makes it easier for neurons with the correct receptors to become excited, thus facilitating the integration of the firing of neurons 1 and 2 by the downstream neuron in the bottom picture.

information is reflected in patterns of LFP coherence.[114–117] Changes in oscillatory dynamics track changes in attentional focus and state.[118–125] Oscillations help route information and form network assemblies in the hippocampus and cortex.[126–129] Further, LFPs serve as reliable sources of information, unaffected by neuronal representation drift.[130] Further, oscillations form traveling waves[131–133] that can have precise influences on networks and impact function.[134–138] The theory of spatial computing ties this together to explain how the brain applies rhythms to physical patches of the cortex to selectively control just the right neurons at the right times to do the right things.[139]

Importantly, the fluctuating electric fields not only reflect organization but can also create organization by having a causal influence. There are numerous instances of ephaptic coupling (i.e., the causal influence of electric fields) in the brain.[140–148] When cortical neurons are not spiking (which is much of the time) their membrane potentials are oscillating below the spiking threshold. This reflects and contributes to the surrounding electric fields. Thus, cortical neurons spend much of their time "teetering" on the edge of spiking in a sea of fluctuating electric fields. Even small changes in the fields can "push" them one way or another. Electric fields are an ideal "orchestra conductor" for coordinating neural activity. They spread influences at the speed of light. It would

be remarkable if evolution did not take advantage of this phenomenon to use it for organizing neural activity.

Oscillations likely interact with neuromodulation. Neurons corelease fast neurotransmitters as well as neuromodulators, which are packaged into either vesicles or dense core vesicles.[149,150] Neuromodulation affects excitability. Plus, certain oscillatory frequencies can cause preferential release of dense core vesicles containing neuromodulatory signals. This offers another layer of control.

Neuromodulatory signals tune ensemble volume, orchestrating mixed selectivity

Another organizational push may come from neuromodulation, which can have effects in hundreds of milliseconds and last for hours.[151] Neuromodulation can trigger the transition between behavioral states,[152] often by activating a neuromodulatory or neuropeptidergic nucleus that can modulate the release of a neuromodulator. These neuromodulators have broad but unique innervation patterns across the brain. Their signals will be read out through axonal innervation patterns and downstream receptor expression profiles. The receptor expression profiles are predominantly predetermined but subject to experience-dependent plasticity. They can act as a weighted filter, allowing different neuromodulators to generate varied brain states. From ghrelin inducing hunger to oxytocin-stimulating prosocial behavior, neuromodulatory systems offer a wide spectrum of programs.[153,154]

There are different uses for neuromodulation and they can have specific effects. In the BLA, at baseline conditions, there is a bias toward prioritizing negative valence—which is adaptive given that predation is a more immediate threat to survival than not obtaining food or water. The relative bias can be altered or even flipped by changes in internal or external conditions. Food restriction can shift the balance between positive and negative valence processing circuits in the BLA,[155] which may facilitate the prioritization of food-seeking via changes in hormonal, peptidergic, or modulatory signals. Additionally, increasing the concentration of neurotensin in the BLA gates reward learning, effectively shifting the bias toward reward learning.[91]

Multivariate and specific effects of neuromodulation are evident in the PFC, where mixed selectivity is expressed by most neurons. Dopamine can serve as the "switch operator" for directing information flow.[6,66,156] In the mPFC, increasing dopaminergic tone amplifies the signal-to-noise ratio for information about aversive, but not appetitive, stimuli in a specific projection to a brainstem region, the periaqueductal gray (PAG).[108] Importantly, dopamine concentration does not act uniformly on different PFC neurons; mPFC neurons projecting to the nucleus accumbens (Nac) showed a suppression of activity, while mPFC-PAG neurons showed a selective amplification of information about punishments, such as air puff or foot shock, but not rewards such as sucrose.[108]

Importantly, dendritic nonlinearities are a key subcellular component that provide a possible mechanistic explanation for how mixed selectivity can be implemented in the context of neuromodulation. Various neuromodulatory receptors may be expressed and multiplexed on a single cell, allowing a given cell to be recruited (or suppressed and effectively excluded) from a given ensemble. Neuronal dendrites offer a high level of electrical compartmentalization onto which different functional classes of receptors (ranging from excitatory or inhibitory to different neuromodulatory receptors) may be segregated into different compartments,[50,54] allowing the cell to receive and transmit information differently in the presence of different concentrations of neuromodulators—such as dopamine, serotonin, and norepinephrine—as demonstrated in acetylcholine.[55] For example, a single cell may have a proximal dendritic compartment receiving bottom-up sensory input and a distal dendritic compartment receiving top-down, predictive coding inputs, and allow a cell to integrate both of these signals.

Finally, the chemical signal of neuromodulation binds to GPCRs, initiating second-messenger cascades that can contribute to shifting cells into different membrane states, thus altering the electrical signals. By biasing resting membrane potentials, and using different receptors to induce different changes to the cell, neuromodulators can shift the synchrony and the oscillatory patterns of the brain.

CONCLUDING REMARKS

When the first description of mixed selectivity was articulated over a decade ago, the initial focus was on the population dynamics of the PFC and a high-level cognitive function, working memory. Now, we know that mixed selectivity is not a rare feature of neurons of certain brain areas, organisms, and functions. Rather, mixed selectivity is ubiquitous.[4,157,158] It is present across species and across functions from high-level cognition to "automatic" sensorimotor processes such as object recognition[159] and even to homeostatic processes.[160] The widespread presence of mixed selectivity underscores its fundamental role in providing the brain with the scalable processing power needed for complex thought and action.[4]

ACKNOWLEDGMENTS

Special thanks to Amy Cao for her invaluable contributions to the design and artistic rendering of the figures. K.M.T. is an HHMI Investigator, member of the Kavli Institute for Brain and Mind, and the Wylie Vale chair at the Salk Institute for Biological Studies, and this work was supported by funding from Salk, HHMI, Kavli Foundation, Dolby Family Fund, R01-MH115920 (NIMH), R37-MH102441 (NIMH), and Pioneer Award DP1-AT009925 (NCCIH). E.K.M. was supported by the Office of Naval Research N00014-22-1-2453, NIMH 1R01MH131715-01, The JPB Foundation, and The Picower Institute for Learning and Memory. S.F. was supported by the Simons Foundation, the Gatsby Charitable Foundation, and the Swartz Foundation. M.K.B was supported by R01NS125298 (NINDS) and the Kavli Institute for Brain and Mind. M.R. was supported by the European Union's Horizon Europe research and innovation programme under grant agreement no. 101070408 (SustainML).

DECLARATION OF INTERESTS

K.M.T. and E.K.M. are members of the *Neuron* advisory board.

REFERENCES

1. Duncan, J., and Miller, E.K. (2002). Cognitive focus through adaptive neural coding in the primate prefrontal cortex. In Principles of frontal lobe function (Oxford University Press), pp. 278–291. https://doi.org/10.1093/acprof:oso/9780195134971.003.0018.

2. Duncan, J., and Miller, E.K. (2013). Adaptive Neural Coding in Frontal and Parietal Cortex. In Principles of Frontal Lobe Function, J. Duncan, E. Koechlin, D.T. Stuss, and R.T. Knight, eds. (Oxford University Press), pp. 292–301. https://doi.org/10.1093/med/9780199837755.003.0023.

3. Rigotti, M., Ben Dayan Rubin, D., Wang, X.-J., and Fusi, S. (2010). Internal Representation of Task Rules by Recurrent Dynamics: The Importance of the Diversity of Neural Responses. Front. Comput. Neurosci. *4*, 24. https://doi.org/10.3389/fncom.2010.00024.

4. Rigotti, M., Barak, O., Warden, M.R., Wang, X.-J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature *497*, 585–590. https://doi.org/10.1038/nature12160.

5. Fusi, S., Miller, E.K., and Rigotti, M. (2016). Why neurons mix: high dimensionality for higher cognition. Curr. Opin. Neurobiol. *37*, 66–74. https://doi.org/10.1016/j.conb.2016.01.010.

6. Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167.

7. Brincat, S.L., Siegel, M., von Nicolai, C., and Miller, E.K. (2018). Gradual progression from sensory to task-related processing in cerebral cortex. Proc. Natl. Acad. Sci. USA *115*, E7202–E7211. https://doi.org/10.1073/pnas.1717075115.

8. Marr, D. (1971). Simple memory: a theory for archicortex. Philos. Trans. R. Soc. Lond. B *262*, 23–81. https://doi.org/10.1098/rstb.1971.0078.

9. Rosenblatt, F. (1962). Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms (Spartan Books).

10. Minsky, M., and Papert, S. (1969). Perceptrons; an Introduction to Computational Geometry (MIT Press). https://doi.org/10.7551/mitpress/11301.001.0001.

11. Buonomano, D.V., and Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. Nat. Rev. Neurosci. *10*, 113–125.

12. Jaeger, H. (2002). Adaptive Nonlinear System Identification with Echo State Networks. In Advances in Neural Information Processing Systems (MIT Press).

13. Maass, W., Natschläger, T., and Markram, H. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. Neural Comput. *14*, 2531–2560. https://doi.org/10.1162/089976602760407955.

14. Cortes, C., and Vapnik, V. (1995). Support-vector networks. Mach. Learn. *20*, 273–297. https://doi.org/10.1007/BF00994018.

15. Rumelhart, D.E., Hinton, G.E., and McClelland, J.L. (1986). A General Framework for Parallel Distributed Processing. In Parallel Distributed Processing: Explorations in the Microstructure of Cognition, *1* (MIT Press), p. 26.

16. Zipser, D., and Andersen, R.A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. Nature *331*, 679–684. https://doi.org/10.1038/331679a0.

17. Pouget, A., and Sejnowski, T.J. (1997). Spatial transformations in the parietal cortex using basis functions. J. Cogn. Neurosci. *9*, 222–237. https://doi.org/10.1162/jocn.1997.9.2.222.

18. Pouget, A., and Snyder, L.H. (2000). Computational approaches to sensorimotor transformations. Nat. Neurosci. *3* (*Suppl*), 1192–1198. https://doi.org/10.1038/81469.

19. Salinas, E., and Abbott, L.F. (2001). Coordinate transformations in the visual system: how to generate gain fields and what to compute with them. In Progress in Brain Research Advances in Neural Population Coding, *130* (Elsevier), pp. 175–190. https://doi.org/10.1016/S0079-6123(01)30012-2.

20. Botvinick, M., and Watanabe, T. (2007). From Numerosity to Ordinal Rank: A Gain-Field Model of Serial Order Representation in Cortical Working Memory. J. Neurosci. *27*, 8636–8642. https://doi.org/10.1523/JNEUROSCI.2110-07.2007.

21. Salinas, E. (2004). Context-dependent selection of visuomotor maps. BMC Neurosci. *5*, 47. https://doi.org/10.1186/1471-2202-5-47.

22. Poggio, T. (1990). A Theory of How the Brain Might Work. Cold Spring Harb. Symp. Quant. Biol. *55*, 899–910. https://doi.org/10.1101/SQB.1990.055.01.084.

23. Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural dynamics are dominated by richly varied movements. Nat. Neurosci. *22*, 1677–1686. https://doi.org/10.1038/s41593-019-0502-4.

24. Vaaga, C.E., Borisovska, M., and Westbrook, G.L. (2014). Dual-transmitter neurons: Functional implications of co-release and co-transmission. Curr. Opin. Neurobiol. *29*, 25–32. https://doi.org/10.1016/j.conb.2014.04.010.

25. Aldrich, S.B. (2019). The Use of Multiple Neurotransmitters at Synapses. In Synaptic Transmission, S.D. Meriney and E.E. Fanselow, eds. (Academic Press), pp. 449–480. https://doi.org/10.1016/B978-0-12-815320-8.00021-1.

26. Buschman, T.J., Siegel, M., Roy, J.E., and Miller, E.K. (2011). Neural substrates of cognitive capacity limitations. Proc. Natl. Acad. Sci. USA *108*, 11252–11255. https://doi.org/10.1073/pnas.1104666108.

27. McKenzie, S., Frank, A.J., Kinsky, N.R., Porter, B., Rivière, P.D., and Eichenbaum, H. (2014). Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. Neuron *83*, 202–215. https://doi.org/10.1016/j.neuron.2014.05.019.

28. Bianco, I.H., and Engert, F. (2015). Visuomotor Transformations Underlying Hunting Behavior in Zebrafish. Curr. Biol. *25*, 831–846. https://doi.org/10.1016/j.cub.2015.01.042.

29. Parthasarathy, A., Herikstad, R., Bong, J.H., Medina, F.S., Libedinsky, C., and Yen, S.-C. (2017). Mixed selectivity morphs population codes in prefrontal cortex. Nat. Neurosci. *20*, 1770–1779. https://doi.org/10.1038/s41593-017-0003-2.

30. Zhang, C.Y., Aflalo, T., Revechkis, B., Rosario, E.R., Ouellette, D., Pouratian, N., and Andersen, R.A. (2017). Partially mixed selectivity in human posterior parietal association cortex. Neuron *95*, 697–708.e4. https://doi.org/10.1016/j.neuron.2017.06.040.

31. Grunfeld, I.S., and Likhtik, E. (2018). Mixed selectivity encoding and action selection in the prefrontal cortex during threat assessment. Curr. Opin. Neurobiol. *49*, 108–115. https://doi.org/10.1016/j.conb.2018.01.008.

32. Parthasarathy, A., Tang, C., Herikstad, R., Cheong, L.F., Yen, S.-C., and Libedinsky, C. (2019). Time-invariant working memory representations in the presence of code-morphing in the lateral prefrontal cortex. Nat. Commun. *10*, 4995. https://doi.org/10.1038/s41467-019-12841-y.

33. Aoi, M.C., Mante, V., and Pillow, J.W. (2020). Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. Nat. Neurosci. *23*, 1410–1420. https://doi.org/10.1038/s41593-020-0696-5.

34. Bernardi, S., Benna, M.K., Rigotti, M., Munuera, J., Fusi, S., and Salzman, C.D. (2020). The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. Cell *183*, 954–967.e21. https://doi.org/10.1016/j.cell.2020.09.031.

35. Cruzado, N.A., Tiganj, Z., Brincat, S.L., Miller, E.K., and Howard, M.W. (2020). Conjunctive representation of what and when in monkey hippocampus and lateral prefrontal cortex during an associative memory task. Hippocampus *30*, 1332–1346. https://doi.org/10.1002/hipo.23282.

36. Enel, P., Wallis, J.D., and Rich, E.L. (2020). Stable and dynamic representations of value in the prefrontal cortex. eLife *9*, e54313. https://doi.org/10.7554/eLife.54313.

37. Stefanini, F., Kushnir, L., Jimenez, J.C., Jennings, J.H., Woods, N.I., Stuber, G.D., Kheirbek, M.A., Hen, R., and Fusi, S. (2020). A Distributed Neural Code in the Dentate Gyrus and in CA1. Neuron *107*, 703–716.e4. https://doi.org/10.1016/j.neuron.2020.05.022.

38. Panichello, M.F., and Buschman, T.J. (2021). Shared mechanisms underlie the control of working memory and attention. Nature *592*, 601–605. https://doi.org/10.1038/s41586-021-03390-w.

39. Ledergerber, D., Battistin, C., Blackstad, J.S., Gardner, R.J., Witter, M.P., Moser, M.-B., Roudi, Y., and Moser, E.I. (2021). Task-dependent mixed selectivity in the subiculum. Cell Rep. *35*, 109175. https://doi.org/10.1016/j.celrep.2021.109175.

40. Zhou, Y., Rosen, M.C., Swaminathan, S.K., Masse, N.Y., Zhu, O., and Freedman, D.J. (2021). Distributed functions of prefrontal and parietal cortices during sequential categorical decisions. eLife *10*, e58782. https://doi.org/10.7554/eLife.58782.

41. Sendhilnathan, N., Goldberg, M.E., and Ipata, A.E. (2022). Mixed Selectivity in the Cerebellar Purkinje-Cell Response during Visuomotor Association Learning. J. Neurosci. *42*, 3847–3855. https://doi.org/10.1523/JNEUROSCI.1771-21.2022.

42. Fu, Z., Beam, D., Chung, J.M., Reed, C.M., Mamelak, A.N., Adolphs, R., and Rutishauser, U. (2022). The geometry of domain-general performance monitoring in the human medial frontal cortex. Science *376*, eabm9922. https://doi.org/10.1126/science.abm9922.

43. Wallach, A., Melanson, A., Longtin, A., and Maler, L. (2022). Mixed selectivity coding of sensory and motor social signals in the thalamus of a weakly electric fish. Curr. Biol. *32*, 51–63.e3. https://doi.org/10.1016/j.cub.2021.10.034.

44. Kaufman, M.T., Churchland, M.M., Ryu, S.I., and Shenoy, K.V. (2014). Cortical activity in the null space: permitting preparation without movement. Nat. Neurosci. *17*, 440–448. https://doi.org/10.1038/nn.3643.

45. Chang, L., and Tsao, D.Y. (2017). The Code for Facial Identity in the Primate Brain. Cell *169*, 1013–1028.e14. https://doi.org/10.1016/j.cell.2017.05.011.

46. Higgins, I., Chang, L., Langston, V., Hassabis, D., Summerfield, C., Tsao, D., and Botvinick, M. (2021). Unsupervised deep learning identifies semantic disentanglement in single inferotemporal face patch neurons. Nat. Commun. *12*, 6456. https://doi.org/10.1038/s41467-021-26751-5.

47. Rauch, S.L., Shin, L.M., and Wright, C.I. (2003). Neuroimaging studies of amygdala function in anxiety disorders. Ann. N. Y. Acad. Sci. *985*, 389–410. https://doi.org/10.1111/j.1749-6632.2003.tb07096.x.

48. Larkum, M.E., Zhu, J.J., and Sakmann, B. (1999). A new cellular mechanism for coupling inputs arriving at different cortical layers. Nature *398*, 338–341. https://doi.org/10.1038/18686.

49. Murayama, M., Pérez-Garci, E., Nevian, T., Bock, T., Senn, W., and Larkum, M.E. (2009). Dendritic encoding of sensory stimuli controlled by deep cortical interneurons. Nature *457*, 1137–1141. https://doi.org/10.1038/nature07663.

50. Xu, N.L., Harnett, M.T., Williams, S.R., Huber, D., O'Connor, D.H., Svoboda, K., and Magee, J.C. (2012). Nonlinear dendritic integration of sensory and motor input during an active sensing task. Nature *492*, 247–251. https://doi.org/10.1038/nature11601.

51. Royer, S., Zemelman, B.V., Losonczy, A., Kim, J., Chance, F., Magee, J.C., and Buzsáki, G. (2012). Control of timing, rate and bursts of hippocampal place cells by dendritic and somatic inhibition. Nat. Neurosci. *15*, 769–775. https://doi.org/10.1038/nn.3077.

52. Bittner, K.C., Grienberger, C., Vaidya, S.P., Milstein, A.D., Macklin, J.J., Suh, J., Tonegawa, S., and Magee, J.C. (2015). Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons. Nat. Neurosci. *18*, 1133–1142. https://doi.org/10.1038/nn.4062.

53. Manita, S., Suzuki, T., Homma, C., Matsumoto, T., Odagawa, M., Yamada, K., Ota, K., Matsubara, C., Inutsuka, A., Sato, M., et al. (2015). A Top-Down Cortical Circuit for Accurate Sensory Perception. Neuron *86*, 1304–1316. https://doi.org/10.1016/j.neuron.2015.05.006.

54. Ranganathan, G.N., Apostolides, P.F., Harnett, M.T., Xu, N.-L., Druckmann, S., and Magee, J.C. (2018). Active dendritic integration and mixed neocortical network representations during an adaptive sensing behavior. Nat. Neurosci. *21*, 1583–1590. https://doi.org/10.1038/s41593-018-0254-6.

55. Williams, S.R., and Fletcher, L.N. (2019). A Dendritic Substrate for the Cholinergic Control of Neocortical Output Neurons. Neuron *101*, 486–499.e4. https://doi.org/10.1016/j.neuron.2018.11.035.

56. Takahashi, N., Ebner, C., Sigl-Glöckner, J., Moberg, S., Nierwetberg, S., and Larkum, M.E. (2020). Active dendritic currents gate descending cortical outputs in perception. Nat. Neurosci. *23*, 1277–1285. https://doi.org/10.1038/s41593-020-0677-8.

57. Aru, J., Suzuki, M., and Larkum, M.E. (2020). Cellular Mechanisms of Conscious Processing. Trends Cogn. Sci. *24*, 814–825. https://doi.org/10.1016/j.tics.2020.07.006.

58. Tsodyks, M.V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc. Natl. Acad. Sci. USA *94*, 719–723. https://doi.org/10.1073/pnas.94.2.719.

59. Brunel, N., and Wang, X.-J. (2001). Effects of Neuromodulation in a Cortical Network Model of Object Working Memory Dominated by Recurrent Inhibition. J. Comput. Neurosci. *11*, 63–85. https://doi.org/10.1023/A:1011204814320.

60. Wrenn, C.C. (2004). Social Transmission of Food Preference in Mice. Curr. Protoc. Neurosci. *28*, 8–15. https://doi.org/10.1002/0471142301.ns0805gs28.

61. Raposo, D., Kaufman, M.T., and Churchland, A.K. (2014). A category-free neural population supports evolving demands during decision-making. Nat. Neurosci. *17*, 1784–1792. https://doi.org/10.1038/nn.3865.

62. Kaufman, M.T., Benna, M.K., Rigotti, M., Stefanini, F., Fusi, S., and Churchland, A.K. (2022). The implications of categorical and category-free mixed selectivity on representational geometries. Curr. Opin. Neurobiol. *77*, 102644. https://doi.org/10.1016/j.conb.2022.102644.

63. Boyle, L.M., Posani, L., Irfan, S., Siegelbaum, S.A., and Fusi, S. (2024). Tuned geometries of hippocampal representations meet the demands of social memory. Neuron. *112*, 1358–1371.

64. Hirokawa, J., Vaughan, A., Masset, P., Ott, T., and Kepecs, A. (2019). Frontal cortex neuron types categorically encode single decision variables. Nature *576*, 446–451. https://doi.org/10.1038/s41586-019-1816-9.

65. Hocker, D.L., Brody, C.D., Savin, C., and Constantinople, C.M. (2021). Subpopulations of neurons in lOFC encode previous and current rewards at time of choice. eLife *10*, e70129. https://doi.org/10.7554/eLife.70129.

66. Tye, K.M. (2018). Neural Circuit Motifs in Valence Processing. Neuron *100*, 436–452. https://doi.org/10.1016/j.neuron.2018.10.001.

67. Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78–84. https://doi.org/10.1038/nature12742.

68. Caron, S.J.C., Ruta, V., Abbott, L.F., and Axel, R. (2013). Random convergence of olfactory inputs in the Drosophila mushroom body. Nature *497*, 113–117. https://doi.org/10.1038/nature12063.

69. Downer, J.D., Verhein, J.R., Rapone, B.C., O'Connor, K.N., and Sutter, M.L. (2021). An Emergent Population Code in Primary Auditory Cortex Supports Selective Attention to Spectral and Temporal Sound Features. J. Neurosci. *41*, 7561–7577. https://doi.org/10.1523/JNEUROSCI.0693-20.2021.

70. Nogueira, R., Rodgers, C.C., Bruno, R.M., and Fusi, S. (2023). The geometry of cortical representations of touch in rodents. Nat. Neurosci. *26*, 239–250. https://doi.org/10.1038/s41593-022-01237-9.

71. Marlin, B.J., Mitre, M., D'amour, J.A., Chao, M.V., and Froemke, R.C. (2015). Oxytocin enables maternal behaviour by balancing cortical inhibition. Nature *520*, 499–504. https://doi.org/10.1038/nature14402.

72. Whittington, J.C.R., Dorrell, W., Ganguli, S., and Behrens, T. (2022). Disentanglement with Biological Constraints: A Theory of Functional Cell Types. In Proceedings of The Eleventh International Conference on Learning Representations.

73. Leppla, C.A., Keyes, L.R., Glober, G., Matthews, G.A., Batra, K., Jay, M., Feng, Y., Chen, H.S., Mills, F., Delahanty, J., et al. (2023). Thalamus sends information about arousal but not valence to the amygdala. Psychopharmacol. (Berl.) *240*, 477–499. https://doi.org/10.1007/s00213-022-06284-5.

74. Schafe, G.E., and LeDoux, J.E. (2000). Memory Consolidation of Auditory Pavlovian Fear Conditioning Requires Protein Synthesis and Protein Kinase A in the Amygdala. J. Neurosci. *20*, RC96. https://doi.org/10.1523/JNEUROSCI.20-18-j0003.2000.

75. Tootoonian, S., Coen, P., Kawai, R., and Murthy, M. (2012). Neural Representations of Courtship Song in the Drosophila Brain. J. Neurosci. *32*, 787–798. https://doi.org/10.1523/JNEUROSCI.5104-11.2012.

76. Hindmarsh Sten, T., Li, R., Otopalik, A., and Ruta, V. (2021). Sexual arousal gates visual processing during Drosophila courtship. Nature *595*, 549–553. https://doi.org/10.1038/s41586-021-03714-w.

77. O'Muircheartaigh, J., Keller, S.S., Barker, G.J., and Richardson, M.P. (2015). White Matter Connectivity of the Thalamus Delineates the Functional Architecture of Competing Thalamocortical Systems. Cereb. Cortex *25*, 4477–4489. https://doi.org/10.1093/cercor/bhv063.

78. Barsy, B., Kocsis, K., Magyar, A., Babiczky, Á., Szabó, M., Veres, J.M., Hillier, D., Ulbert, I., Yizhar, O., and Mátyás, F. (2020). Associative and plastic thalamic signaling to the lateral amygdala controls fear behavior. Nat. Neurosci. *23*, 625–637. https://doi.org/10.1038/s41593-020-0620-z.

79. Janak, P.H., and Tye, K.M. (2015). From circuits to behaviour in the amygdala. Nature *517*, 284–292. https://doi.org/10.1038/nature14188.

80. Tye, K.M., Prakash, R., Kim, S.-Y., Fenno, L.E., Grosenick, L., Zarabi, H., Thompson, K.R., Gradinaru, V., Ramakrishnan, C., and Deisseroth, K. (2011). Amygdala circuitry mediating reversible and bidirectional control of anxiety. Nature *471*, 358–362. https://doi.org/10.1038/nature09820.

81. Namburi, P., Beyeler, A., Yorozu, S., Calhoon, G.G., Halbert, S.A., Wichmann, R., Holden, S.S., Mertens, K.L., Anahtar, M., Felix-Ortiz, A.C., et al. (2015). A circuit mechanism for differentiating positive and negative associations. Nature *520*, 675–678. https://doi.org/10.1038/nature14366.

82. Namburi, P., Al-Hasani, R., Calhoon, G.G., Bruchas, M.R., and Tye, K.M. (2016). Architectural Representation of Valence in the Limbic System. Neuropsychopharmacology *41*, 1697–1715. https://doi.org/10.1038/npp.2015.358.

83. O'Neill, P.-K., Posani, L., Meszaros, J., Warren, P., Schoonover, C.E., Fink, A.J.P., Fusi, S., and Salzman, C.D. (2023). The representational geometry of emotional states in basolateral amygdala. Preprint at bioRxiv. https://doi.org/10.1101/2023.09.23.558668.

84. Reijmers, L.G., Perkins, B.L., Matsuo, N., and Mayford, M. (2007). Localization of a Stable Neural Correlate of Associative Memory. Science *317*, 1230–1233. https://doi.org/10.1126/science.1143839.

85. Han, J.-H., Kushner, S.A., Yiu, A.P., Hsiang, H.-L.L., Buch, T., Waisman, A., Bontempi, B., Neve, R.L., Frankland, P.W., and Josselyn, S.A. (2009). Selective Erasure of a Fear Memory. Science *323*, 1492–1496. https://doi.org/10.1126/science.1164139.

86. Ramirez, S., Liu, X., MacDonald, C.J., Moffa, A., Zhou, J., Redondo, R.L., and Tonegawa, S. (2015). Activating positive memory engrams suppresses depression-like behaviour. Nature *522*, 335–339. https://doi.org/10.1038/nature14514.

87. Salzman, C.D., and Fusi, S. (2010). Emotion, Cognition, and Mental State Representation in Amygdala and Prefrontal Cortex. Annu. Rev. Neurosci. *33*, 173–202. https://doi.org/10.1146/annurev.neuro.051508.135256.

88. Paton, J.J., Belova, M.A., Morrison, S.E., and Salzman, C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature *439*, 865–870. https://doi.org/10.1038/nature04490.

89. Shabel, S.J., and Janak, P.H. (2009). Substantial similarity in amygdala neuronal activity during conditioned appetitive and aversive emotional arousal. Proc. Natl. Acad. Sci. USA *106*, 15031–15036. https://doi.org/10.1073/pnas.0905580106.

90. Beyeler, A., Namburi, P., Glober, G.F., Simonnet, C., Calhoon, G.G., Conyers, G.F., Luck, R., Wildes, C.P., and Tye, K.M. (2016). Divergent Routing of Positive and Negative Information from the Amygdala during Memory Retrieval. Neuron *90*, 348–361. https://doi.org/10.1016/j.neuron.2016.03.004.

91. Li, H., Namburi, P., Olson, J.M., Borio, M., Lemieux, M.E., Beyeler, A., Calhoon, G.G., Hitora-Imamura, N., Coley, A.A., Libster, A., et al. (2022). Neurotensin orchestrates valence assignment in the amygdala. Nature *608*, 586–592. https://doi.org/10.1038/s41586-022-04964-y.

92. Bariselli, S., Fobbs, W.C., Creed, M.C., and Kravitz, A.V. (2019). A competitive model for striatal action selection. Brain Res. *1713*, 70–79. https://doi.org/10.1016/j.brainres.2018.10.009.

93. Kimchi, E.Y., and Laubach, M. (2009). Dynamic Encoding of Action Selection by the Medial Striatum. J. Neurosci. *29*, 3148–3159. https://doi.org/10.1523/JNEUROSCI.5206-08.2009.

94. Surmeier, D.J., Plotkin, J., and Shen, W. (2009). Dopamine and synaptic plasticity in dorsal striatal circuits controlling action selection. Curr. Opin. Neurobiol. *19*, 621–628. https://doi.org/10.1016/j.conb.2009.10.003.

95. Yoo, S.B.M., and Hayden, B.Y. (2020). The Transition from Evaluation to Selection Involves Neural Subspace Reorganization in Core Reward Regions. Neuron *105*, 712–724.e4. https://doi.org/10.1016/j.neuron.2019.11.013.

96. Badre, D., Bhandari, A., Keglovits, H., and Kikumoto, A. (2021). The dimensionality of neural representations for control. Curr. Opin. Behav. Sci. *38*, 20–28. https://doi.org/10.1016/j.cobeha.2020.07.002.

97. Sapountzis, P., Paneri, S., Papadopoulos, S., and Gregoriou, G.G. (2022). Dynamic and stable population coding of attentional instructions coexist in the prefrontal cortex. Proc. Natl. Acad. Sci. USA *119*, e2202564119. https://doi.org/10.1073/pnas.2202564119.

98. Rust, N.C., and Cohen, M.R. (2022). Priority coding in the visual system. Nat. Rev. Neurosci. *23*, 376–388. https://doi.org/10.1038/s41583-022-00582-9.

99. GoodSmith, D., Chen, X., Wang, C., Kim, S.H., Song, H., Burgalossi, A., Christian, K.M., and Knierim, J.J. (2017). Spatial Representations of Granule Cells and Mossy Cells of the Dentate Gyrus. Neuron *93*, 677–690.e5. https://doi.org/10.1016/j.neuron.2016.12.026.

100. Senzai, Y., and Buzsáki, G. (2017). Physiological Properties and Behavioral Correlates of Hippocampal Granule Cells and Mossy Cells. Neuron *93*, 691–704.e5. https://doi.org/10.1016/j.neuron.2016.12.011.

101. Aronov, D., Nevers, R., and Tank, D.W. (2017). Mapping of a non-spatial dimension by the hippocampal-entorhinal circuit. Nature *543*, 719–722. https://doi.org/10.1038/nature21692.

102. Gluck, M.A., and Myers, C.E. (1993). Hippocampal mediation of stimulus representation: a computational theory. Hippocampus *3*, 491–516. https://doi.org/10.1002/hipo.450030410.

103. Benna, M.K., and Fusi, S. (2021). Place cells may simply be memory cells: Memory compression leads to spatial tuning and history dependence. Proc. Natl. Acad. Sci. USA *118*, e2018422118. https://doi.org/10.1073/pnas.2018422118.

104. LeDoux, J.E., Cicchetti, P., Xagoraris, A., and Romanski, L.M. (1990). The lateral amygdaloid nucleus: sensory interface of the amygdala in fear conditioning. J. Neurosci. *10*, 1062–1069. https://doi.org/10.1523/JNEUROSCI.10-04-01062.1990.

105. Quirk, G.J., Repa, C., and LeDoux, J.E. (1995). Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. Neuron *15*, 1029–1039. https://doi.org/10.1016/0896-6273(95)90092-6.

106. Burgos-Robles, A., Vidal-Gonzalez, I., and Quirk, G.J. (2009). Sustained Conditioned Responses in Prelimbic Prefrontal Neurons Are Correlated with Fear Expression and Extinction Failure. J. Neurosci. *29*, 8474–8482. https://doi.org/10.1523/JNEUROSCI.0378-09.2009.

107. Burgos-Robles, A., Kimchi, E.Y., Izadmehr, E.M., Porzenheim, M.J., Ramos-Guasp, W.A., Nieh, E.H., Felix-Ortiz, A.C., Namburi, P., Leppla, C.A., Presbrey, K.N., et al. (2017). Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. Nat. Neurosci. *20*, 824–835. https://doi.org/10.1038/nn.4553.

108. Vander Weele, C.M., Siciliano, C.A., Matthews, G.A., Namburi, P., Izadmehr, E.M., Espinel, I.C., Nieh, E.H., Schut, E.H.S., Padilla-Coreano, N., Burgos-Robles, A., et al. (2018). Dopamine enhances signal-to-noise

ratio in cortical-brainstem encoding of aversive stimuli. Nature *563*, 397–401. https://doi.org/10.1038/s41586-018-0682-1.

109. Gao, P., and Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. Curr. Opin. Neurobiol. *32*, 148–155. https://doi.org/10.1016/j.conb.2015.04.003.

110. Singer, W. (2018). Neuronal oscillations: unavoidable and useful? Eur. J. Neurosci. *48*, 2389–2398. https://doi.org/10.1111/ejn.13796.

111. Buzsáki, G., and Vöröslakos, M. (2023). Brain rhythms have come of age. Neuron *111*, 922–926. https://doi.org/10.1016/j.neuron.2023.03.018.

112. Buzsaki, G. (2021). The Brain from Inside Out (Oxford University Press).

113. Gazzaley, A., and Nobre, A.C. (2012). Top-down modulation: bridging selective attention and working memory. Trends Cogn. Sci. *16*, 129–135. https://doi.org/10.1016/j.tics.2011.11.014.

114. Buschman, T.J., Denovellis, E.L., Diogo, C., Bullock, D., and Miller, E.K. (2012). Synchronous Oscillatory Neural Ensembles for Rules in the Prefrontal Cortex. Neuron *76*, 838–846. https://doi.org/10.1016/j.neuron.2012.09.029.

115. Antzoulatos, E.G., and Miller, E.K. (2014). Increases in functional connectivity between prefrontal cortex and striatum during category learning. Neuron *83*, 216–225. https://doi.org/10.1016/j.neuron.2014.05.005.

116. Pinotsis, D.A., Geerts, J.P., Pinto, L., FitzGerald, T.H.B., Litvak, V., Auksztulewicz, R., and Friston, K.J. (2017). Linking canonical microcircuits and neuronal activity: Dynamic causal modelling of laminar recordings. NeuroImage *146*, 355–366. https://doi.org/10.1016/j.neuroimage.2016.11.041.

117. Bastos, A.M., Lundqvist, M., Waite, A.S., Kopell, N., and Miller, E.K. (2020). Layer and rhythm specificity for predictive routing. Proc. Natl. Acad. Sci. USA *117*, 31459–31469. https://doi.org/10.1073/pnas.2014868117.

118. Eeckman, F.H., and Freeman, W.J. (1990). Correlations between unit firing and EEG in the rat olfactory system. Brain Res. *528*, 238–244. https://doi.org/10.1016/0006-8993(90)91663-2.

119. Fries, P., Reynolds, J.H., Rorie, A.E., and Desimone, R. (2001). Modulation of Oscillatory Neuronal Synchronization by Selective Visual Attention. Science *291*, 1560–1563. https://doi.org/10.1126/science.1055465.

120. Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. Science *320*, 110–113. https://doi.org/10.1126/science.1154735.

121. Kay, L.M., Beshel, J., Brea, J., Martin, C., Rojas-Líbano, D., and Kopell, N. (2009). Olfactory oscillations: the what, how and what for. Trends Neurosci. *32*, 207–214. https://doi.org/10.1016/j.tins.2008.11.008.

122. Buschman, T.J., and Miller, E.K. (2009). Serial, Covert Shifts of Attention during Visual Search Are Reflected by the Frontal Eye Fields and Correlated with Population Oscillations. Neuron *63*, 386–396. https://doi.org/10.1016/j.neuron.2009.06.020.

123. Kay, L.M. (2015). Olfactory system oscillations across phyla. Curr. Opin. Neurobiol. *31*, 141–147. https://doi.org/10.1016/j.conb.2014.10.004.

124. Frederick, D.E., Brown, A., Brim, E., Mehta, N., Vujovic, M., and Kay, L.M. (2016). Gamma and Beta Oscillations Define a Sequence of Neurocognitive Modes Present in Odor Processing. J. Neurosci. *36*, 7750–7767. https://doi.org/10.1523/JNEUROSCI.0569-16.2016.

125. Fiebelkorn, I.C., and Kastner, S. (2019). A Rhythmic Theory of Attention. Trends Cogn. Sci. *23*, 87–101. https://doi.org/10.1016/j.tics.2018.11.009.

126. Rubino, D., Robbins, K.A., and Hatsopoulos, N.G. (2006). Propagating waves mediate information transfer in the motor cortex. Nat. Neurosci. *9*, 1549–1557. https://doi.org/10.1038/nn1802.

127. Davis, Z.W., Dotson, N.M., Franken, T.P., Muller, L., and Reynolds, J.H. (2023). Spike-phase coupling patterns reveal laminar identity in primate cortex. eLife *12*, e84512. https://doi.org/10.7554/eLife.84512.

128. Jutras, M.J., Fries, P., and Buffalo, E.A. (2013). Oscillatory activity in the monkey hippocampus during visual exploration and memory formation. Proc. Natl. Acad. Sci. USA *110*, 13144–13149. https://doi.org/10.1073/pnas.1302351110.

129. Nardin, M., Kaefer, K., Stella, F., and Csicsvari, J. (2023). Theta oscillations as a substrate for medial prefrontal-hippocampal assembly interactions. Cell Rep. *42*, 113015. https://doi.org/10.1016/j.celrep.2023.113015.

130. Pinotsis, D.A., and Miller, E.K. (2022). Beyond dimension reduction: Stable electric fields emerge from and allow representational drift. NeuroImage *253*, 119058. https://doi.org/10.1016/j.neuroimage.2022.119058.

131. Muller, L., Reynaud, A., Chavane, F., and Destexhe, A. (2014). The stimulus-evoked population response in visual cortex of awake monkey is a propagating wave. Nat. Commun. *5*, 3675. https://doi.org/10.1038/ncomms4675.

132. Lubenov, E.V., and Siapas, A.G. (2009). Hippocampal theta oscillations are travelling waves. Nature *459*, 534–539. https://doi.org/10.1038/nature08010.

133. Takahashi, K., Saleh, M., Penn, R.D., and Hatsopoulos, N.G. (2011). Propagating Waves in Human Motor Cortex. Front. Hum. Neurosci. *5*, 40. https://doi.org/10.3389/fnhum.2011.00040.

134. Davis, Z.W., Muller, L., Martinez-Trujillo, J., Sejnowski, T., and Reynolds, J.H. (2020). Spontaneous travelling cortical waves gate perception in behaving primates. Nature *587*, 432–436. https://doi.org/10.1038/s41586-020-2802-y.

135. Alamia, A., and VanRullen, R. (2019). Alpha oscillations and traveling waves: Signatures of predictive coding? PLoS Biol. *17*, e3000487. https://doi.org/10.1371/journal.pbio.3000487.

136. Muller, L., Chavane, F., Reynolds, J., and Sejnowski, T.J. (2018). Cortical travelling waves: mechanisms and computational principles. Nat. Rev. Neurosci. *19*, 255–268. https://doi.org/10.1038/nrn.2018.20.

137. Takahashi, K., Kim, S., Coleman, T.P., Brown, K.A., Suminski, A.J., Best, M.D., and Hatsopoulos, N.G. (2015). Large-scale spatiotemporal spike patterning consistent with wave propagation in motor cortex. Nat. Commun. *6*, 7169. https://doi.org/10.1038/ncomms8169.

138. Xu, Y., Long, X., Feng, J., and Gong, P. (2023). Interacting spiral wave patterns underlie complex brain dynamics and are related to cognitive processing. Nat. Hum. Behav. *7*, 1196–1215. https://doi.org/10.1038/s41562-023-01626-5.

139. Lundqvist, M., Brincat, S.L., Rose, J., Warden, M.R., Buschman, T.J., Miller, E.K., and Herman, P. (2023). Working memory control dynamics follow principles of spatial computing. Nat. Commun. *14*, 1429. https://doi.org/10.1038/s41467-023-36555-4.

140. Fröhlich, F., and McCormick, D.A. (2010). Endogenous Electric Fields May Guide Neocortical Network Activity. Neuron *67*, 129–143. https://doi.org/10.1016/j.neuron.2010.06.005.

141. Anastassiou, C.A., Perin, R., Markram, H., and Koch, C. (2011). Ephaptic coupling of cortical neurons. Nat. Neurosci. *14*, 217–223. https://doi.org/10.1038/nn.2727.

142. Anastassiou, C.A., and Koch, C. (2015). Ephaptic coupling to endogenous electric field activity: why bother? Curr. Opin. Neurobiol. *31*, 95–103. https://doi.org/10.1016/j.conb.2014.09.002.

143. Faber, D.S., and Pereda, A.E. (2018). Two Forms of Electrical Transmission Between Neurons. Front. Mol. Neurosci. *11*, 427. https://doi.org/10.3389/fnmol.2018.00427.

144. Schmidt, H., Hahn, G., Deco, G., and Knösche, T.R. (2021). Ephaptic coupling in white matter fibre bundles modulates axonal transmission delays. PLoS Comput. Biol. *17*, e1007858. https://doi.org/10.1371/journal.pcbi.1007858.

145. Pinotsis, D.A., and Miller, E.K. (2023). In vivo ephaptic coupling allows memory network formation. Cereb. Cortex *33*, 9877–9895. https://doi.org/10.1093/cercor/bhad251.

146. Chiang, C.-C., Shivacharan, R.S., Wei, X., Gonzalez-Reyes, L.E., and Durand, D.M. (2019). Slow periodic activity in the longitudinal hippocampal slice can self-propagate non-synaptically by a mechanism consistent with ephaptic coupling. J. Physiol. *597*, 249–269. https://doi.org/10.1113/JP276904.

147. Han, K.-S., Chen, C.H., Khan, M.M., Guo, C., and Regehr, W.G. (2020). Climbing fiber synapses rapidly and transiently inhibit neighboring Purkinje cells via ephaptic coupling. Nat. Neurosci. *23*, 1399–1409. https://doi.org/10.1038/s41593-020-0701-z.

148. Van Horn, J.D., Jacokes, Z., Newman, B., and Henry, T. (2023). Editorial: Is Now the Time for Foundational Theory of Brain Connectivity? Neuroinformatics *21*, 633–635. https://doi.org/10.1007/s12021-023-09641-7.

149. Shabel, S.J., Proulx, C.D., Piriz, J., and Malinow, R. (2014). Mood regulation. GABA/glutamate co-release controls habenula output and is modified by antidepressant treatment. Science *345*, 1494–1498. https://doi.org/10.1126/science.1250469.

150. Kim, S., Wallace, M.L., El-Rifai, M., Knudsen, A.R., and Sabatini, B.L. (2022). Co-packaging of opposing neurotransmitters in individual synaptic vesicles in the central nervous system. Neuron *110*, 1371–1384.e7. https://doi.org/10.1016/j.neuron.2022.01.007.

151. Brzosko, Z., Mierau, S.B., and Paulsen, O. (2019). Neuromodulation of Spike-Timing-Dependent Plasticity: Past, Present, and Future. Neuron *103*, 563–581. https://doi.org/10.1016/j.neuron.2019.05.041.

152. Lee, S.-H., and Dan, Y. (2012). Neuromodulation of Brain States. Neuron *76*, 209–222. https://doi.org/10.1016/j.neuron.2012.09.012.

153. Wren, A.M., Seal, L.J., Cohen, M.A., Brynes, A.E., Frost, G.S., Murphy, K.G., Dhillo, W.S., Ghatei, M.A., and Bloom, S.R. (2001). Ghrelin Enhances Appetite and Increases Food Intake in Humans. J. Clin. Endocrinol. Metab. *86*, 5992. https://doi.org/10.1210/jcem.86.12.8111.

154. Marsh, N., Marsh, A.A., Lee, M.R., and Hurlemann, R. (2021). Oxytocin and the Neurobiology of Prosocial Behavior. Neuroscientist *27*, 604–619. https://doi.org/10.1177/1073858420960111.

155. Calhoon, G.G., Sutton, A.K., Chang, C.-J., Libster, A.M., Glober, G.F., Lévêque, C.L., Murphy, G.D., Namburi, P., Leppla, C.A., Siciliano, C.A., et al. (2018). Acute Food Deprivation Rapidly Modifies Valence-Coding Microcircuits in the Amygdala. Preprint at bioRxiv. https://doi.org/10.1101/285189.

156. Weele, C.M.V., Siciliano, C.A., and Tye, K.M. (2019). Dopamine tunes prefrontal outputs to orchestrate aversive processing. Brain Res. *1713*, 16–31. https://doi.org/10.1016/j.brainres.2018.11.044.

157. Ginther, M.R., Walsh, D.F., and Ramus, S.J. (2011). Hippocampal Neurons Encode Different Episodes in an Overlapping Sequence of Odors Task. J. Neurosci. *31*, 2706–2711. https://doi.org/10.1523/JNEUROSCI.3413-10.2011.

158. Kira, S., Safaai, H., Morcos, A.S., Panzeri, S., and Harvey, C.D. (2023). A distributed and efficient population code of mixed selectivity neurons for flexible navigation decisions. Nat. Commun. *14*, 2121. https://doi.org/10.1038/s41467-023-37804-2.

159. DiCarlo, J.J., Zoccolan, D., and Rust, N.C. (2012). How does the brain solve visual object recognition? Neuron *73*, 415–434. https://doi.org/10.1016/j.neuron.2012.01.010.

160. Vaccari, F.E., Diomedi, S., Filippini, M., Hadjidimitrakis, K., and Fattori, P. (2022). New insights on single-neuron selectivity in the era of population-level approaches. Front. Integr. Neurosci. *16*, 929052. https://doi.org/10.3389/fnint.2022.929052.