UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

A BIOLOGICAL INTERPRETATION OF FLEXIBLE ARTIFICIAL NEURAL NETWORKS

TESIS

QUE PARA OBTENER EL TÍTULO DE:

FÍSICO

PRESENTA:

FERNANDO IGNACIO CALDERÓN DE LEÓN

DIRECTOR DE TESIS:
DR. ROMÁN ROSSI-POOL

CIUDAD UNIVERSITARIA CD.MX., 2024

Sinodales

Sinodal presidente: Dr. Luis Fernando Magaña Solís

Sinodal vocal: Dr. Luis Enrique Hernández-Lemus

Sinodal suplente 1: Dr. Juan Claudio Toledo Roy

Sinodal suplente 2: Dr. Alejandro Reyes Coronado

Agradecimientos

Académicos

Al Dr. Román, por abrirme la puerta al laboratorio y permitirme conocer todos estos temas, que encuentro increíblemente interesantes.

Al resto de los integrantes del laboratorio, especialmente a Jero y Gabriel, por su apoyo, paciencia y valiosos comentarios.

A todos los profesores, que, a lo largo de toda mi trayectoria académica, han contribuido a mi formación como científico y como persona.

Personales

A mi madre, a mi padre, y a mi hermana, por su apoyo, cariño, y guianza.

Al resto de mi familia, por siempre estar ahí.

A mis amigos, mis hermanos, por acompañarme en el camino.

Y finalmente, a Sandra, por todos los momentos.

Instituto de
Fisiología
Celular,
UNAM

A biological interpretation of flexible artificial neural networks

Undergraduate thesis by Fernando Ignacio Calderón de León

Abstract

Throughout this work the reported consistency between specific kinds of artificial neural networks (ANN) and biological neural systems (BNS) is examined from various perspectives. It is shown that flexible problem-solving abilities in a reference network rely on a set of functional features whose nature is consistent with the current understanding of the neurophysiological basis of biological cognition and perception. Such consistency is then analyzed from the perspective of complex systems, especially within the framework of physics and mathematics, concluding that the fact that both the ANN and BNS developed similar solutions to deal with parallel tasks is a direct consequence of both systems' inherent search for efficiency. Further, it is exposed how certain aspects of the phenomena treated here may also be observed in other kinds of naturally occurring complex systems, as a result of the interplay between a few physical principles. Finally, a general overview of the relationship between ANNs and BNSs is put forward, constituting an integral understanding of the topic as well as a brief proposal for future work, ultimately leading to this project's final conclusions. In sum, this work seeks to clarify the existing relationship between certain types of ANNs and BNSs in order to advance the understanding of cognition within a mathematical perspective.

Highlights

- Flexible problem-solving abilities in both ANN and BNS rely on a common set of functional features, suggesting that such properties may be fundamental for flexible cognition and perception.
- The interplay of such functional features generates compositional representations of tasks for both ANNs and BNSs, allowing the engagement of specialized functions in different patterns for varying computational requirements, conferring flexibility and efficiency in solving a variety of perceptual tasks.
- Some of said functional features are **excitability**, **mixed-selectivity**, **functional specialization**, and **compositionality**; these properties are purportedly fundamental for efficient cognitive abilities.
- The search for energetic efficiency generates such comparable functionality between artificial and biological neural systems, as it also the reason for their resemblance with other kinds of naturally occurring complex systems.
- Investigators found a reference network that had developed specialized functions resembling real-life neurophysiological processes, under the same task context.
- To bridge artificial and biological intelligence, the biological plausibility of the models' structure and functionality must be refined further.

Main objectives

- Synthetize relevant results from several lines of research, especially within the field of cognitive and computational neuroscience, into a general overview with the purpose of generating an accessible and consistent knowledge basis involving the subjects treated here.
- Clarify various important aspects, such as the nature of the perceptual tasks utilized, as well as the functional resemblance between ANNs and BNSs, that are not properly elucidated in the research work from which the current discussion takes its basis.
- Conceive original conclusions, as well as reinforce previously existing ones, regarding the computational modelling of biological cognition and perception.
- Implement various research skills, such as analysis, compilation, and communication of information, on a relevant subject in contemporary and multidisciplinary scientific research.

Index

I. Introduction

In the context of contemporary research in neuroscience, it is undeniable that the advancement of knowledge relies on the collective effort of multiple disciplines. In recent years, the fusion of two seemingly unrelated fields, computer science and neuroscience, has conceived remarkable results that contribute to a better understanding of both subjects. Regarding the matter of cognition and perception, a broad understanding has been formed, although there are still aspects that are being studied further for additional clarification. Thus, with the application of an extensive repertoire of computational techniques, valuable insights into the mathematical nature of these phenomena have been made possible.

Computational models known as **artificial neural networks** (ANNs) have been developed to emulate, in as close a manner as possible, specific aspects of some physiological functions identified in **biological neural systems** (BNSs) (Carpenter R. et al., 2013). In an effort to computationally simulate basic components of cognition and perception, several network models have been inspired by well-known processes performed by biological systems, conceiving remarkable results. However, most network models have been designed to solve only one or a few inter-related tasks, while BNSs are inherently capable of solving a great variety of cognitive tasks. Clearly, this striking difference has strong implications on the conclusions that can be drawn from studying the relationship between both systems, leading to the design of the networks discussed here.

Pursuant to this general outline, the current discussion is a conceptual analysis of various aspects of the existing relationship between ANNs and BNSs. Firstly, a synthesis of relevant results obtained from several lines of research in both cognitive and computational neuroscience will be put forward, allowing a comparison between certain properties of the neurophysiological basis of biological cognition and those of a reference computational model. This will attempt to establish a direct relationship between the features of both systems, leading to valuable conclusions concerning the mathematical modelling of cognitive and perceptive processes.

This way, the following discussion perspective will be based on the consideration of both ANNs and BNSs as complex systems. Finally, a general overview of the mathematical modelling of cognition by means of artificial neural networks will be integrated into the overall discussion of the current context within the field of neuroscience.

It is convenient to note that the models treated here have been worked on for a few decades now, yet the contributions that are of more interest to this work are quite recent. Their fundamental units were first proposed back in the 1940's as a prototype model of biological neurons, popularly known as the Perceptron, while inter-connected systems of said units were characterized in the following years (McCulloch W. et al., 1943) (Rosenblatt F., 1957). Nevertheless, disregarding their promising features, research on ANNs at the time was almost halted due to the unavailability of computational tools powerful enough to test their related hypothesis, generating in recent times a renewed interest in such systems and causing them to be present in remarkable applications in scientific research, as well as in private and public enterprises.

I.I. Cognition and perception

First of all, it is right to define in the broadest terms what is cognition and perception. Classically, the latter may be understood as the unconscious integration of sensory information in any biological system, consequently generating awareness and knowledge about its environment. Cognition, on the other hand, may be defined as the phenomenon responsible for consciousness, in which perceptive information is assessed in order to conceive conscious, or will-driven, actions. Crucially, this phenomenon occurs as a consequence of the complex interplay between diverse physiological structures, more specifically, the nervous system. Thus, it may be stated that cognition and perception is a pair of intrinsically inter-related processes inherent to many biological systems (Romo R. et al., 2013).

As an effort to study this natural phenomena, diverse experimental configurations can be implemented in the laboratory to analyze various aspects of cognitive and perceptive processes. Most times, these procedures will be based on a subject's performance of a series of actions based on perceptive information, a process that is commonly known as a perceptual task. This definition will prove to be quite useful for the present work, as the proposed relationship between ANNs and BNSs departs from their performance of multiple perceptual tasks.

I.II. Artificial neural networks

Essentially, the computational models treated throughout this work are relatively small recurrent neural networks (RNN) consisting of 256 LIF units as basic building blocks. Their general structure coincides with the canonical neural network configuration, as these characteristics are meant to resemble certain known properties of BNSs (Yang et al. 2020). To address the matter of emulating biological cognitive flexibility, said networks were trained over an entire set of tasks meant to be the computational analogs of well-studied perceptual tasks performed by animal subjects in laboratory studies, facilitating a qualitative comparison between certain aspects of such models and their biological analogs.

Before starting such discussion, it is convenient to keep in mind one of the fundamental principles of the mathematical modelling of any natural phenomenon, which states the impossibility of capturing in a model absolutely every one of its features, thus making such model's purpose to only focus on certain properties of interest. Evidently, this is the case of the current study, as it will mainly focus on a handful of functional features of neural systems, disregarding many aspects of their complex nature. This fact has strong implications on the following discussion, so it will be further discussed in later sections. Now, with these considerations, the main characteristics of the models treated here may be exposed.

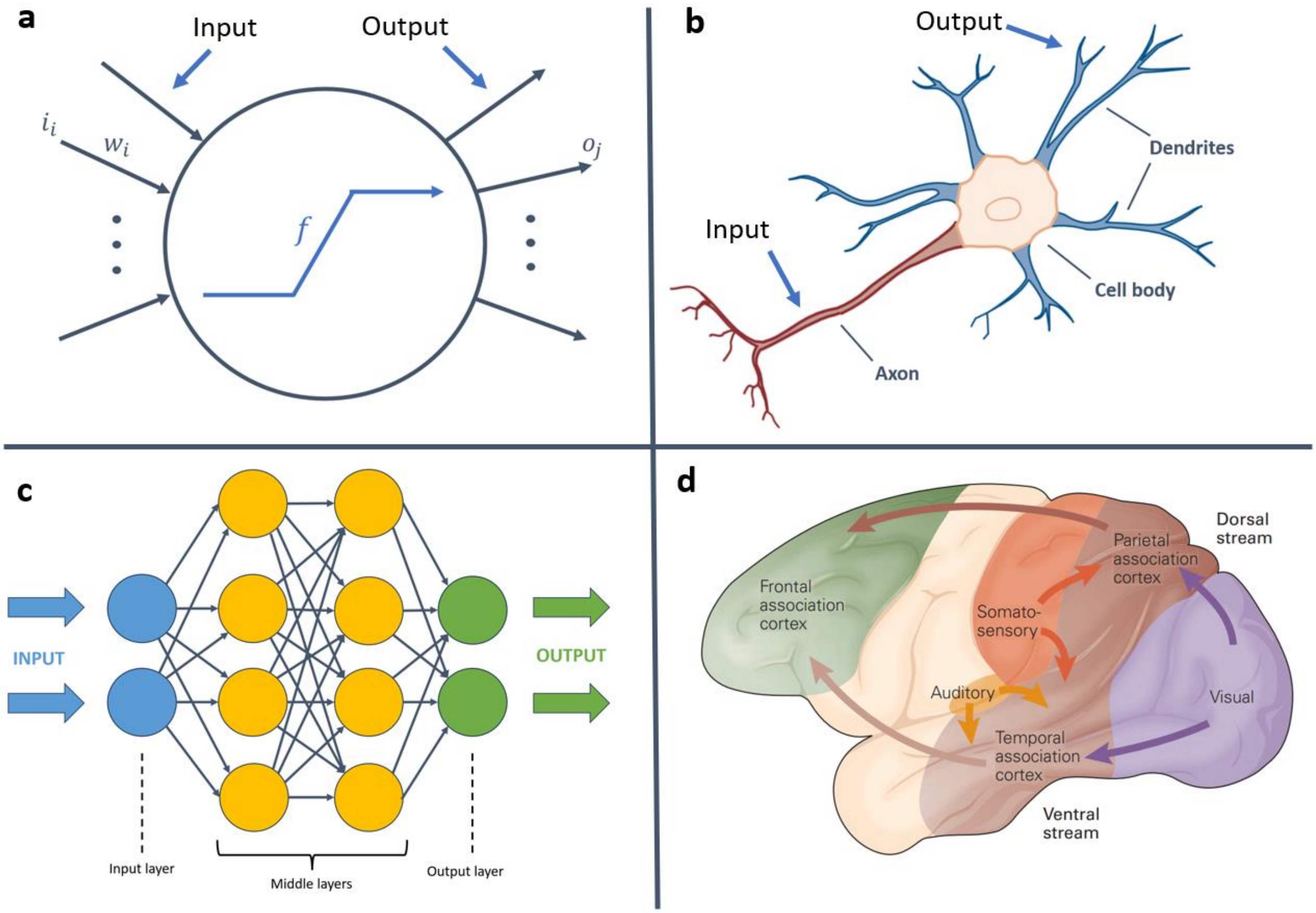


Figure 1. Schematic illustrations depicting the functional resemblance between artificial and biological neural systems at a single unit and a systemic perspective. Figures 1.a and 1.c (left column) represent the artificial case, while Figures 1.b and 1.d (right column) represent the biological case. a) LIF unit. Here i_i *stands for the i-th input the unit receives,* w_i *for its corresponding synaptic weigh, and* o_j *for the unit's j-th output. b) Basic morphology of a biological neuron. c) General structure of a RNN. d) Illustrative representation of a few regions of the primate brain in function of certain cognitive functions they are known to perform. Figures 1.b and 1.d have been readapted from Kandel R. et al., 2013.*

I.II.I. Single unit properties

Similarly to the way nervous cells may be understood as the fundamental functional units of biological nervous systems, ANNs also rely on the interconnections between single units whose characteristics have strong effects on the system's function. Regarding the networks treated here, these units are based on a widely used structure in machine-learning known as "leaky integrate-and-fire" (LIF), whose features are meant to resemble some of the most basic functional properties of biological neurons.

Starting with their characteristic firing function, LIF units integrate multiple inputs coming from previous units and apply a non-linear function to the result to generate an "all or nothing" output response that is transmitted to following units; a process that emulates a biological neuron's excitability (Equation 1). Further, such interconnections may be regarded as well as the cells' synapses, both incoming and outgoing. In the model, the scalar value of connections is mediated by the synaptic weight factor, a numerical value that enhances or decreases the strength of such interconnections, in a similar way to excitatory and inhibitory responses in nervous cells (Equation 2).

Importantly, this may lead to the functional specialization of units based on their input and output selectivity, a feature that has been extensively identified in BNSs and that is known as **functional selectivity**. In the case that a unit does not show preference for a given stimuli type, it is said that it shows mixed-selectivity. This functional resemblance between artificial and biological neural units is visually presented in Figures 1.a and 1.b, as the unit’s input (left) is associated with a neuron’s incoming signaling through its axon receptors, its output (right) with the cell’s outgoing signaling through its dendritic tree, and its non-linear operation (center) to a neuron’s critical firing function. It is right to note that regarding the models treated here, the applied non-linear function is the Rectified Linear Unit (ReLU), a two-step operation that is standardly used in this kind of network models.

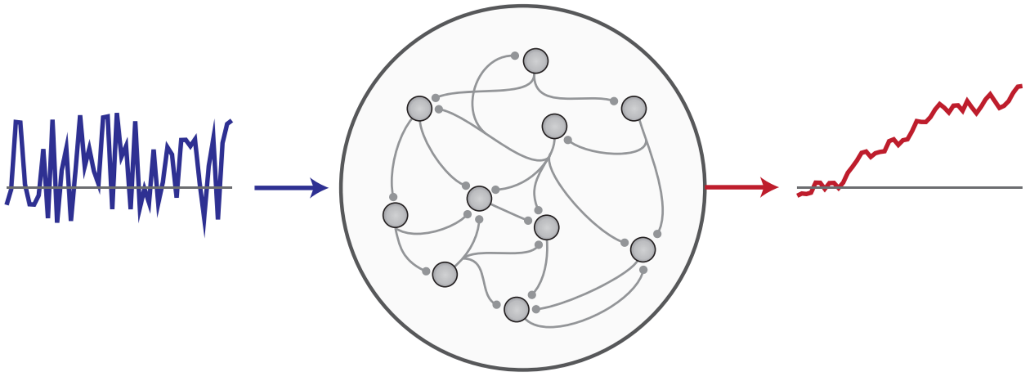


Figure 2. Schematic diagram of a generic RNN receiving a continuous input signal, as it processes in real time and consequently generates a continuous output signal. This functionality highly resembles that of the network models used here, as it has been utilized to model broad aspects of diverse neurophysiological functions. Figure readapted from Yang et al. 2020.

This set of facts suggests the reliance of both artificial and biological units on **criticality**, **excitability** and **mixed-selectivity**, functional features that appear to be fundamental for efficient neural processing (Kay M. et al., 2024). Thus, by fixing large arrays of these inter-connected units, it becomes possible to emulate some of the most general characteristics of certain neurophysiological processes of BNSs, allowing their study through this mathematical approach.

I.II.II. Systemic properties

As it has been exposed, inter-connected systems of operational units that functionally resemble biological neurons may be applied to emulate in a broad nature certain aspects of neurophysiological processes involved in cognition and perception. For instance, several remarkable lines of research have successfully developed ANN models that simulate diverse functional properties of the auditive, visual and sensorimotor pathways (Bi Z. et al., 2020) (Joglekar M. et al., 2018), from which a set of relevant features may be identified to facilitate the analysis of the networks treated here.

Starting with the direction of information flow within the systems, the networks discussed here rely on a bidirectional flow of information in which the operations that are directly associated to the network's inner task representation flow in one direction (Figure 1.c), and, as it will be exposed in the following section, information related to performance feedback flows in the opposite direction.

This configuration coincides with the canonical neural network structure (Figure 1.c), in which the purpose of the first layer of units is the recognition of stimuli and its transmission into further layers, where most information is transformed, as the last layer of units is specialized for decision report. Remarkably, this resembles to some extent the flow of neural activity in various sensory-cognitive pathways in BNSs, such as the auditory, visual and somatosensory hierarchies (Figure 1.d) (Figure 2) (Parra et al., 2022) (Campos A., 2022).

Further, it is also a fact that in order to deal with the tasks at hand in the most efficient way possible, neural networks tend to develop specialized functions that are performed through populations of selective units, a process that has been extensively observed in BNSs and that is commonly known as **functional specialization**. In addition, it has also been found that in both the artificial and biological case, neural systems are able to engage in different patterns said specialized functions to flexibly deal with multiple tasks involving varying computational requirements, a process that has also been proposed to be fundamental for efficient cognition and that is usually referred to as **compositionality** (Dela Fuente V. et al., 2006).

Again, it is evident that this mathematical modelling of neurophysiological functions does not contemplate many aspects of their essential nature, a fact that has strong implications on the current discussion. Now, the most relevant characteristics of the method by which the networks treated here were trained may be exposed.

I.II.III. Training

To this point, it has been shown that ANN models can be applied to broadly emulate certain aspects of cognitive and perceptive functions in biological systems. To achieve this, the systems must first be trained on specific tasks. Essentially, they are trained by means of stochastic gradient descent (SGD) algorithms, a family of numerical techniques widely used in machine-learning. They are fundamentally based on the search for ever-increasing task performance, for which the following cycle is continuously iterated: First, after the system completes one task trial, its performance is translated into a numerical parameter that is compared to the value of a perfect performance (Equation 6), generating a discrete gradient (Equation 5), that can be diminished by stochastically modifying the synaptic weights between the units engaged in such task (Equations 3 and 4), thus generating a stochastic gradient descent process that flows in an opposite direction to the network’s task representation. This way, by executing the task a high number of times, an increased performance is achieved (Murphy K., 2021).

Considering that the networks treated here were trained on 20 inter-related tasks, and in order to avoid a pair of well-known issues in automated learning known as catastrophic forgetting and overfitting, a continual learning technique was implemented. The first of these issues involves a network effect in which new tasks are learned by expense of forgetting previously learned ones, as the second one involves the over-specialization of functions, an effect that usually has negative effects on overall performance. Thus, the implemented continual learning technique, commonly known as FORCE algorithms, essentially consists of randomly interleaving different tasks on the network's training schedule, as well preserving through training relevant weighting factors, thus conceiving inner processes that bypass these issues (Sussillo D. et al., 2009).

$$Eq.(1) \qquad a_u(t) = \begin{cases} f * y_u(t) \quad if \quad y_u(t) > 0 \\ 0 \quad if \quad y_u(t) < 0 \end{cases}$$

$$Eq.(2) \qquad y_u(t) = \sum_{i}^{n} W_i(t) * i_i + b_i$$

$$Eq.(3) \qquad y_u(t+1) = \sum_{i}^{n} W_i(t+1) * i_i + b_i$$

$$Eq.(4) \qquad W_i(t+1) = C_u(t,i) * W_i(t)$$

$$Eq.(5) \qquad C_u(t,i) = \varepsilon_u(t,i) * \delta(t)$$

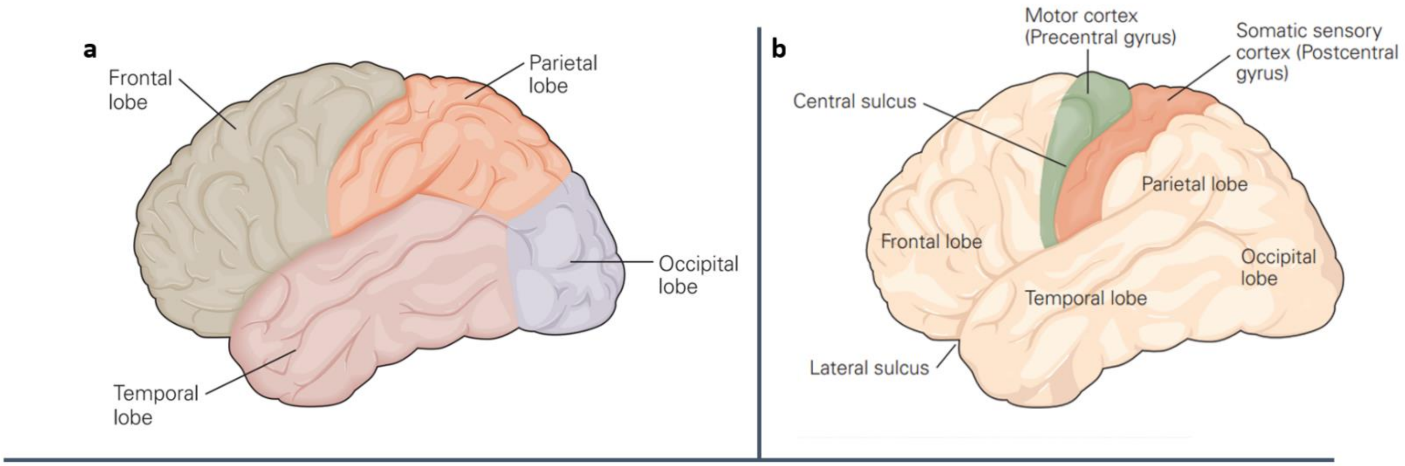
$$Eq.(6) \qquad \delta(t) = \frac{P_G}{P_F(t)}$$

Equations 1 – 6. This set of discrete equations show some of the most relevant mathematics on which the models treated here are based. They may also be found in literature in their differential and matrix forms. All the numerical factors shown here may be regarded as real numbers. Eq 1. Critical function of an ANN single unit. Here $a_u(t)$ stands for the activity at time t of unit u, $y_u(t)$ stands for its pre-critical activity, and f stands for the applied non-linear function. Eq 2. Pre-critical activity of a single unit at time t. Here $W_i(t)$ represents the i-th weighting factor at time t of unit u, i_i represents its i-th input, and b_i a modifiable scalar value. Eq 3. Pre-critical activity of a single unit at time $t+1$. Eq 4. Modified weight factor for the i-th input at time $t+1$. Here $C_u(t,i)$ stands for the correction factor at time t and i-th input of unit u. Eq 5. Correction factor. Here $\delta(t)$ represents the deviation factor, as $\varepsilon_u(t,i)$ represents the scaling factor for the i-th input at time t for unit u. Importantly, the latter is calculated by means of FORCE algorithms, by which they are also strategically implemented to the network. Eq 6. Deviation factor at time t. Here P_G stands for the performance goal of the network for any given task, while $P_F(t)$ represents the numerical value of the system's final performance on said task at time t.

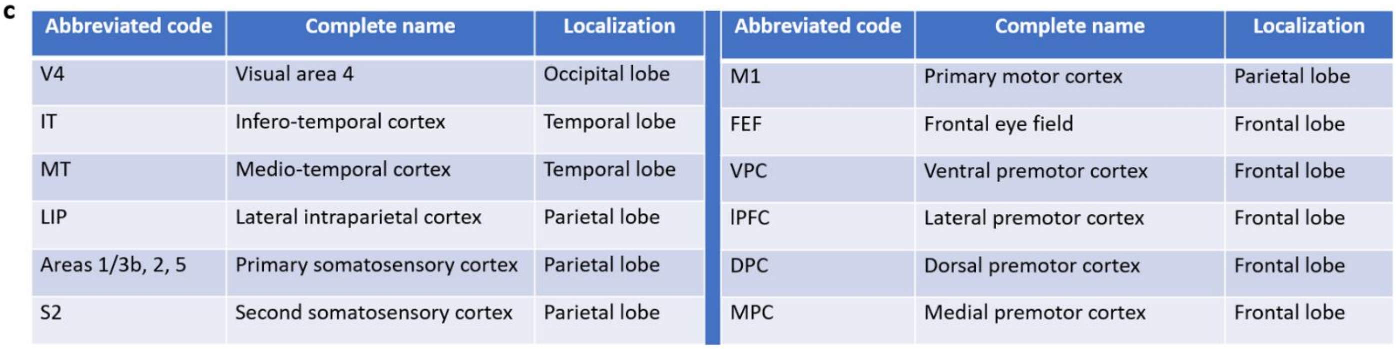
Clearly, these numerical techniques do not meet many important aspects of the learning process in BNSs, implying strong consequences on the current study. This fact will also be further discussed in following sections. Now, it is possible to expose the most general characteristics of the task set on which both ANNs and BNSs were trained, emphasizing on their general structure and neurophysiological results.

I.III. Task set

In recent years, a better understanding of many aspects of the neurophysiological basis of diverse cognitive and perceptive processes has been reached thanks to the contributions of multiple lines of research. Focusing on the auditory, visual, and somatosensory systems, an extensive repertoire of experimental and theoretical techniques has allowed significant findings regarding their underlying mechanisms to be made, both at a single-unit and a system-level perspective. To achieve this, most studies rely on specific task families that are performed by animal subjects in highly controlled environments, in which diverse physical measurements are made in the neural populations of interest. For instance, regarding the structural features of BNSs, magnetic resonances, or MRIs, are standardly used to identify in a non-invasive manner different anatomical formations of the encephalon, as well as certain aspects of their functioning, creating a good parting point for further studies (Figure 4.a). As for functionality, there exists another pair of techniques that are utilized to analyze the activity of neural populations: EEG, which stands for electroencephalography, is another widely used non-invasive method for measuring extracranial electrical activity (Figure 4.b), while *in-situ* electrophysiology analyses the electrical activity of single neurons (Figures 4.c and 4.d).



c



Abbreviated code	Complete name	Localization	Abbreviated code	Complete name	Localization
V4	Visual area 4	Occipital lobe	M1	Primary motor cortex	Parietal lobe
IT	Infero-temporal cortex	Temporal lobe	FEF	Frontal eye field	Frontal lobe
MT	Medio-temporal cortex	Temporal lobe	VPC	Ventral premotor cortex	Frontal lobe
LIP	Lateral intraparietal cortex	Parietal lobe	lPFC	Lateral premotor cortex	Frontal lobe
Areas 1/3b, 2, 5	Primary somatosensory cortex	Parietal lobe	DPC	Dorsal premotor cortex	Frontal lobe
S2	Second somatosensory cortex	Parietal lobe	MPC	Medial premotor cortex	Frontal lobe

Figure 3. a) Schematic diagram of the four lobes of the primate brain. It is currently understood that many neurophysiological processes involved in cognition and perception are widely distributed along different cortical and subcortical regions. b) Schematic diagram showing the location of the motor and somatosensory cortices in the primate brain. The respective functions of these regions are fundamental for the current study. c) Abbreviated nomenclature for some of the brain regions involved in the current discussion. Figures 3.a and 3.b have been readapted from Kandel R. et al., 2013.

Crucially, these techniques differ starkly on their resolution; while MRIs and EEG permit the analysis of neural activity at a systemic level, LFP (local field potential) measurements allow insights onto single-cell dynamics. It is convenient to note that most of the neurophysiological results utilized in this work come from measurements made in cortical regions of the primate brain (Figures 3. a - c) (Figures 5. a - h), as cognitive and perceptive processes are known to mostly take place in such outer brain regions. Inner brain structures and functionality (Figure 5.d), while being essential for cognitive function, may be regarded as not directly related to the neurocomputational processes discussed here.

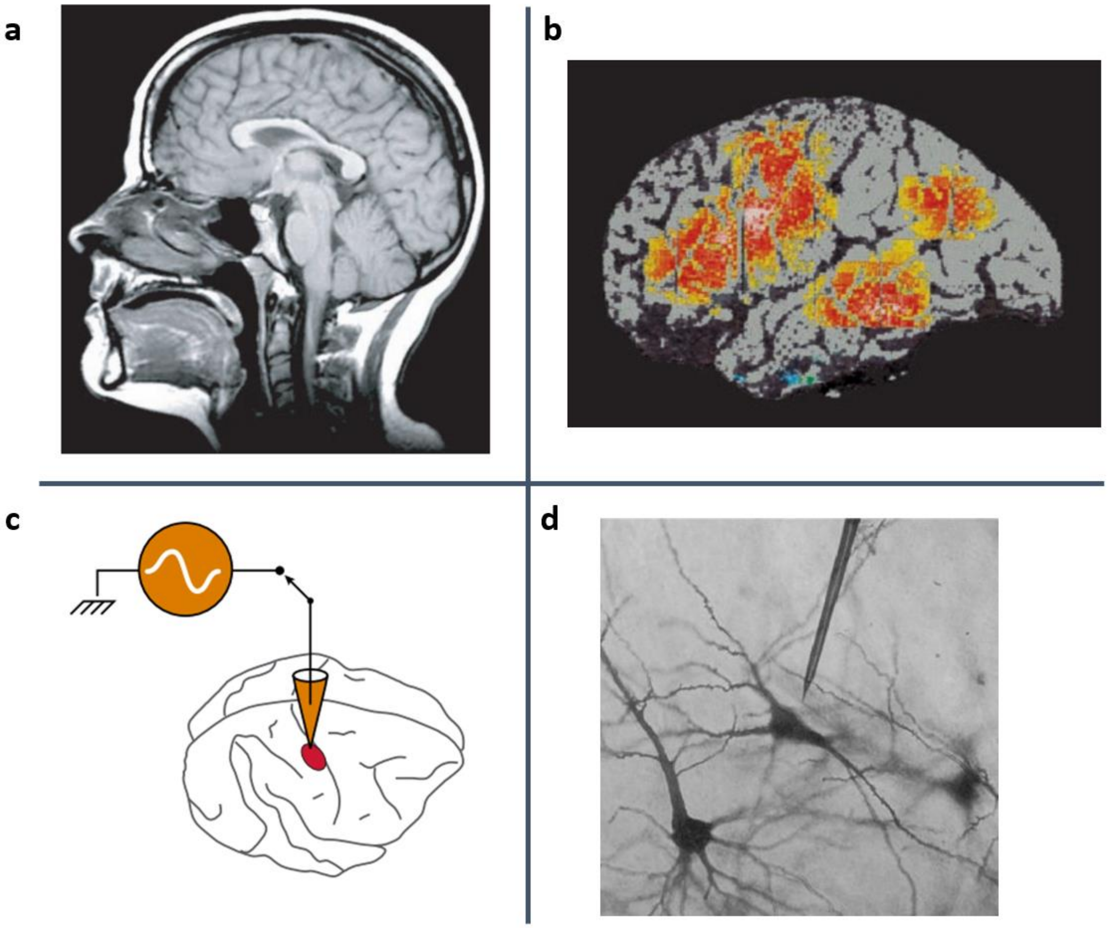


Figure 4. a) Single shot of a magnetic resonance of an adult human brain, on which diverse anatomical formations may be observed. b) Single shot of an EEG of an adult human brain showing the activation of various cortical regions during a cognitive task involving language. It may be observed how the somatosensory cortex is not predominantly engaged on such activity. c) Schematic diagram showing the location of an in-situ electrophysiological measurement on the primate brain. d) Photograph of an electrode measuring the extracellular electrical activity of single neurons located in the primate parietal cortex. Figures 4.a and 4.b have been readapted from Kandel R. et al., 2013. Figure 4.c has been readapted from Romo R. et at., 2020. Figure 4.d has been readapted from Carpenter R. et al., 2013.

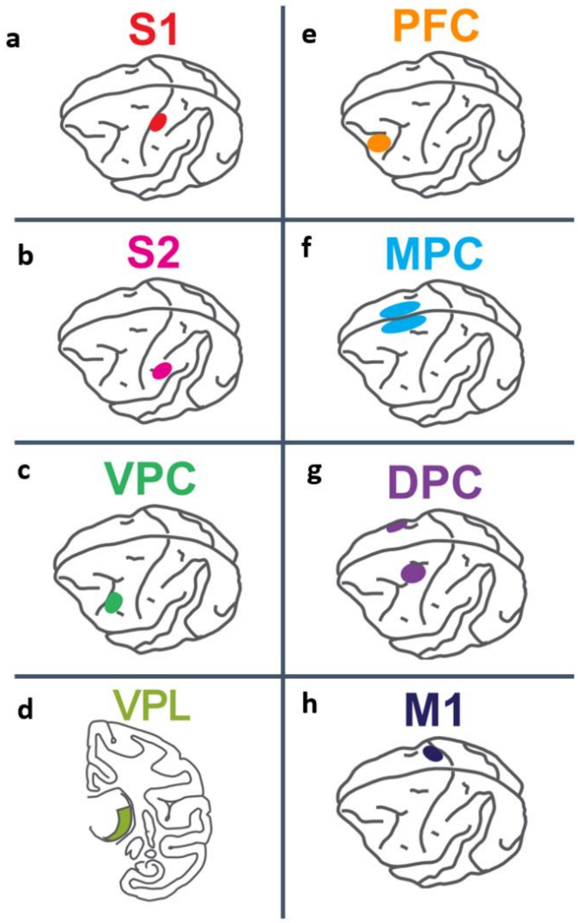


Figure 5. Schematic diagrams depicting various cortical regions of the primate brain involved in the current discussion. a) Primary somatosensory cortex. b) Secondary somatosensory cortex. c) Ventral premotor cortex. d) Ventral posterior lateral nucleus. e) Prefrontal cortex. f) Medial premotor cortex. g) Dorsal premotor cortex. h) Primary motor cortex. Figures 5 a – h have been readapted from Romo R. et al. 2020.

This way, the resulting evidence of these studies has conceived a consistent, yet still incomplete, understanding of the complex functionality of biological neural systems (Carpenter R. et al., 2013) (Kandel R. et al., 2013). Consequently, regarding this matter within a computational perspective, and with the purpose of studying how biologically-inspired ANN handle perceptual tasks classically performed by animal subjects, many tasks have been translated into a computational framework, thereby generating the task set on which the models treated here were trained.

Importantly, considering that this work is mostly focused on the cognitive and perceptive stages of the tasks, not to mention that the models do not contemplate movement execution, most findings regarding the tasks' decision report stage will not be exposed. All the results that will be now presented were obtained from studies performed on non-human primate subjects, more specifically on Rhesus monkeys (*macaca mulatta*). The task set consists of 20 interrelated tasks concerning mostly the auditive, visual and somatosensory systems, which allows the identification of 4 main task subsets in function of their general characteristics: The Go, Anti, Decision/making (DM) and Categorization families.

I.III.I. The Go subset

This first task family, which contains the basic Go task as well as two of its main modalities, RT Go and Dly Go, mainly concerns the neural mechanisms underlying relatively simple visuomotor tasks (Funahashi S. et al., 1989). In its basic modality, subjects must first fix their vision to a visual cue presented in the center of a screen; after a short fixation period the cue goes away, and a second visual cue appears somewhere in a ring around the center of the screen. Subjects must then perform a visual saccade into the direction of the second cue, finalizing the task (Figure 6.a). This way, by performing a thorough analysis of the neural activity in numerous brain regions known to be causally engaged in the performance of these tasks, it has been found that after literal stimuli information is conveyed through early visual processing stages, it is gradually transformed into an abstract representation interpretable by higher stages in the brain's cognitive processing hierarchy. Moreover, it has also been shown that highly selective PFC populations reliably encode all possible positions of the presented visual cues, supporting the idea that such brain region is crucially engaged in abstract cognitive processing as well with working memory functions. As for the other two modalities, which respectively add reaction time (RT) factors and delay periods (Dly) to this basic task structure, evidence further supports the reliance of biological visual processing on such neural dynamics.

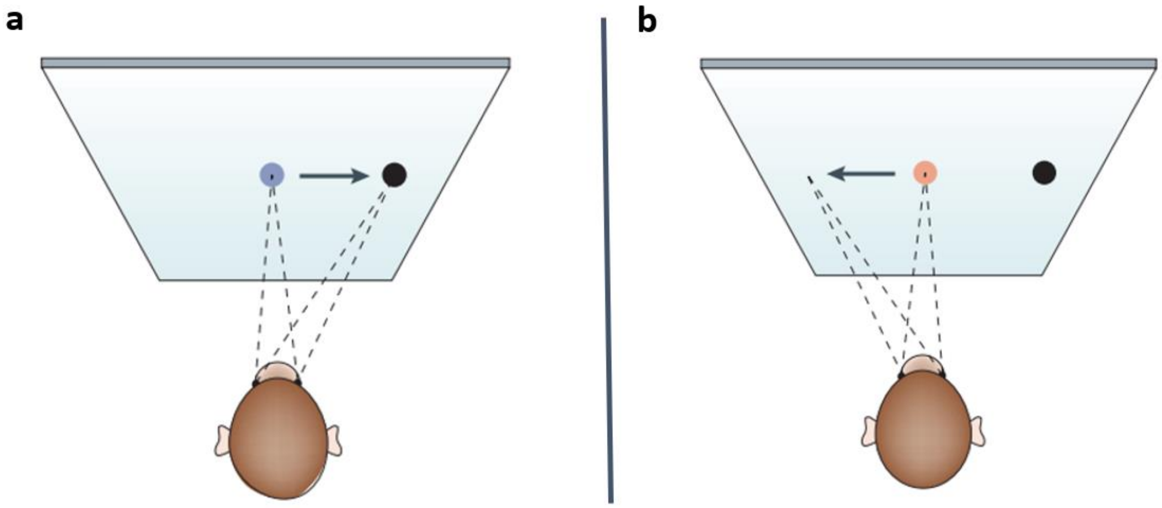


Figure 6. Schematic representation of the Go and Anti tasks. a) Go task. Subjects perform a saccade into the direction of the second visual cue (black dot). b) Anti task. Subjects perform the saccade into the opposite direction of the visual cue. Figures 6. a and b have been readapted from Munoz D.P. et al., 2004.

I.III.II. The Anti subset

This task family greatly resembles the previous subset, as it is composed by the same three task modalities, Anti, RT Anti, and Dly Anti, while adding its characteristic factor: In every Anti task, subjects must perform their report movement into exactly the opposite direction of the second visual cue, thereby generating an additional neural process (usually known as vector inversion) measurable by electrophysiological techniques (Munoz D.P. et al., 2004) (Figure 6.b). Thus, it has been found that for an adequate performance of these tasks, subjects must apply two crucial and consecutive stages: The voluntary inhibition of a reflexive eye movement into the direction of the visual cue, followed by the vector inversion. As for the first process, it is impossible to define a single neural population responsible for it, as it has been shown that it is rather performed throughout several inner and outer brain regions, and that it is fueled by important dopaminergic action. However, regarding the vector inversion process, neural activity associated to it has been traced to frontal brain regions, meaning that such function relies on the abstract computations known to be performed by PFC populations.

I.III.III. The DM subset

The decision-making (DM) task family is the most extensive of the entire task set, as it contains various modalities of at least 5 different tasks. Importantly, most of them require subjects to decide between two possible options, which is why most DM tasks will be represented as a pair.

Starting with the simplest modality, DM 1 and 2, subjects are presented with a screen showing a set of moving dots whose movement is restricted to a single orientation and must decide which is the preferential direction of the majority of dots. It has been found that given the bimodal nature of the presented stimuli (left vs. right, or up vs. down), as information flows through the brain's visual processing hierarchy, specific stimuli features, such as movement direction, are parametrically represented as a bimodal code in the activity of populations localized along the visual dorsal stream, a well-studied neurophysiological structure specialized for spatial recognition (Figure 9.b) (Figure 10). Further, it has been found that such encoding happens through the firing rates modes of the engaged populations, which is why two well-differentiated decisions may be reliably decoded from a subject's neural activity performing such operation. In these tasks decision report is usually done as a visual saccade or by activation of a button (Miller E.K. et al., 1996).

The next pair of tasks, Ctx DM 1 and 2, adds a contextual factor to the structure that has just been exposed. In this modality, besides from movement direction, dots also present another bimodal parameter: color (green or red). Thus, at the beginning of every trial, subjects are instructed on which feature they must report, color or movement direction. As a remarkable result, it has been found that several neural populations, particularly those of the brain's dorsal and ventral visual pathways, encode by means of their firing rates both task parameters of interest, suggesting the existence of a common neural code by which flexible computations regarding different types of information may be performed (Siegel M. et al., 2015).

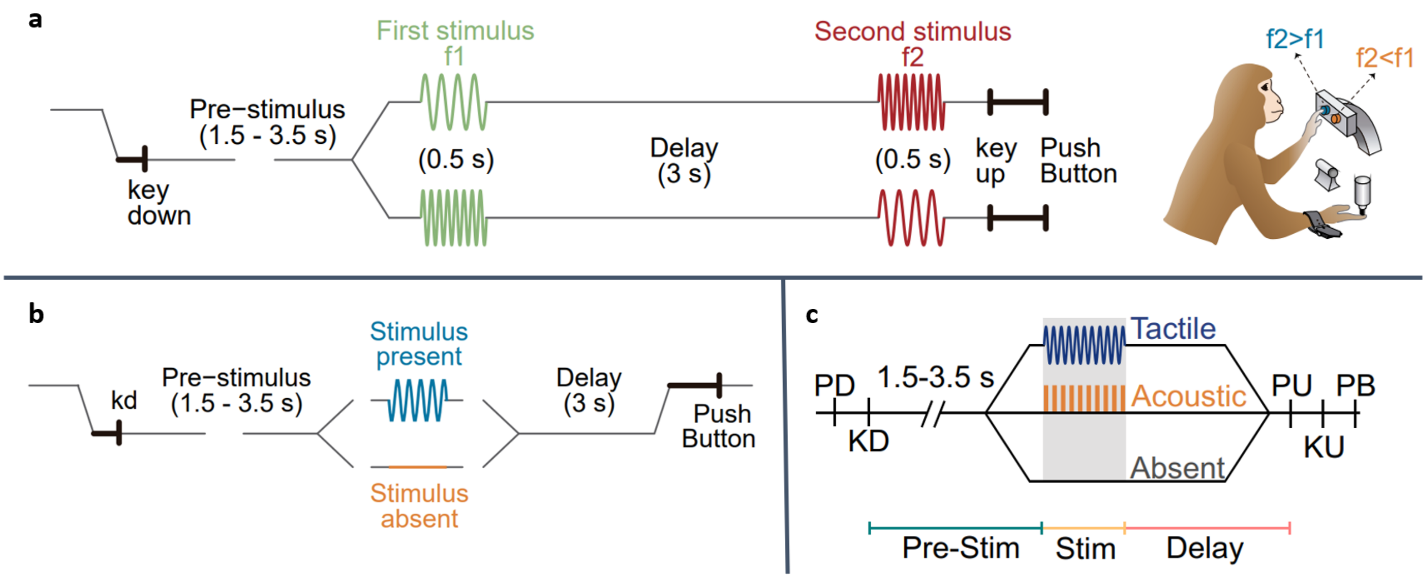


Figure 7. Schematic representation of the main structure of three cognitive tasks involved in the current discussion. a) Dly DM, also known as the vibrotactile frequency discrimination task (VFT). In this diagram, the upper line in which the frequency of the second stimulus is higher than first one, may be regarded as the Dly DM 1 task. The bottom line may be regarded as Dly DM 2. b) Vibrotactile stimulus detection task. c) Multisensory stimulus detection task. Figures 7.a and b have been readapted from Romo R. et al., 2020. Figure 7.c has been readapted from Parra S. et al., 2022.

The following pair of tasks, multisensory decision-making, MultSen DM and Dly MultSen DM, is the only one of its kind in the entire task set as it is the sole one that involves two different sensory modalities, visual and auditive. In its basic modality, subjects are presented with audiovisual periodic stimuli whose frequency may be higher or lower than a control frequency. For adequate task completion, subjects must report which is the given case. Importantly, stimuli may be presented as fully visual, fully auditive, or combined, since the purpose of the task is to study how multisensory integration happens in the brain. It has been found that populations of the subjects' posterior parietal (PPC) and medial temporal (MT) cortices encode, by means of their firing rates, relevant information about both task parameters, thus generating a flexible task representation that is integrated into an abstract code as neural activity flows to frontal brain regions.

It has also been shown that there exists a considerable enhancement in task performance if the stimulus is both auditive and visual, suggesting that a multisensory perceptive build-up might be performed through such populations. As for its Dly modality, a fully abstract representation of task parameters that is maintained through the delay period of the task has been identified in PFC populations, further supporting its involvement in abstract and executive cognitive functions (Raposo D. et al., 2014).

The next pair of inter-related tasks, delayed decision-making, Dly DM 1 and 2, may also be referred to as the vibrotactile frequency discrimination task (VFT), a well-studied paradigm that has also led to remarkable findings. Importantly, it is mostly focused on the activity of different populations to those that have been treated so far, mainly somatosensory, and frontal brain regions.

In its basic modality, two sequential periodic pulses, each with a different frequency, are applied to the tip of a subject's finger with a mechanical probe. Between both pulses there exists a delay period of variable duration, hence the term, delayed decision-making, that forces the subjects to retain some trace of the first pulse's frequency, as they must compare it to the frequency of the second pulse. Depending on whether it is higher or lower, subjects then report their decision mechanically, usually by activating a lever or pushing a button (Figure 7.a). Among many results, it has been found that literal stimuli information is first conveyed through primary and secondary somatosensory cortices, gradually transforming into an abstract and flexible representation as neural activity flows to various areas of frontal brain regions (Figure 9.a). Remarkably, it has been furtherly shown that relevant tasks parameters are reliably represented as a bimodal code based on the firing rate modulations of the engaged populations (Rossi-Pool R. et al., 2017).

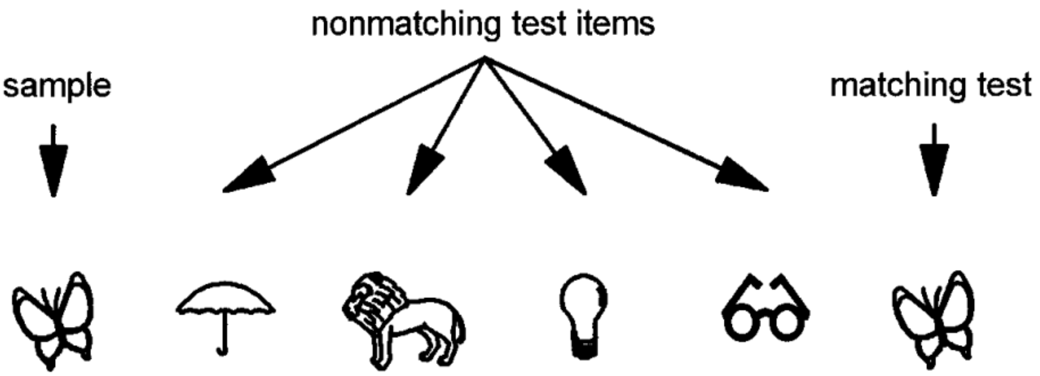


Figure 8. Schematic diagram of the structure of the DNMS task. Figure 8 has been readapted from Miller et al., 1996.

The last pair of decision-making tasks, Ctx Dly DM 1 and 2, may also be considered a special case of the task set, as this study is not aware of any real-life cognitive tasks that match its specifications. Nevertheless, there are indeed various lines of research whose findings rely on experimental configurations that highly resemble such characteristics, and whose results further support the proposals exposed here. (Figures 7.b and 7.c). They will be more appropriately presented in following sections.

I.III.IV. The Categorization subset

The last subgroup of the task set contains two main tasks with an extra modality each, adding up to 4 tasks in total. They are also mostly focused on the activity of different populations to those that have been exposed so far, as they are based on the categorization of complex static visual stimuli such as images from everyday objects. The first modality, delayed match-to-sample (DMS), presents subjects with a series of images, as they must indicate when one of them matches a sample image presented at the beginning of every trial. In the case that there is one or more images between match and sample, it is the delayed non-match-to-sample task (DNMS) that is being performed (Figure 8). Similarly, the delayed match-to-category task (DMC), as well as its non-matching modality (DNMC), require subjects to indicate when a sample image matches a specific category that has been specified at the beginning of every trial. Decision report is usually done mechanically, by means of activating a button or a lever (Freedman D.J. et al., 2016) (Gold J.I. et al., 2007).

This way, a thorough analysis of the activity of various populations pertaining to the visual ventral stream (Figure 10), an important neurophysiological structure specialized for the recognition of static visual stimuli, has conceived extensive evidence furtherly supporting the existence of a flexible neural code by which task parameters may be reliably encoded, as well as the role of PFC populations in higher cognitive functions such as decision-making and working memory.

I.III.V. Some additional tasks

As mentioned before, remarkable findings regarding these topics have been achieved by recent lines of research that rely on experimental configurations that highly resemble the tasks treated here, and whose results further support this work's conclusions. They are not included *per se* in the task set, yet some of their main findings, that will be taken from a pair of representative tasks, are undeniably a valuable contribution to this discussion. They are both decision-making tasks, and they mainly involve the auditory and somatosensory systems.

The first one is the vibrotactile stimulus detection task (Figure 7.b), in which subjects must report on the absence or presence of a vibrotactile mechanical stimulus that is applied to the tip of one of their fingers, in a very similar arrangement of the Dly DM tasks. It has been found that after literal stimuli information reaches the central nervous system, neural activity flows through a hierarchical structure involving both primary and secondary somatosensory cortices (Figures 5.a and 5.b), as well as several areas of parietal and frontal brain regions (Figures 5.c, 5.e, 5.f, and 5.g), gradually transforming from a literal representation of stimuli attributes into an abstract code that reliably represents task parameters of interest. Further, it has also been shown that such encoding is performed through the modulation of the firing rates of the engaged neural populations, generating a gradual perceptive build-up across cortical regions.

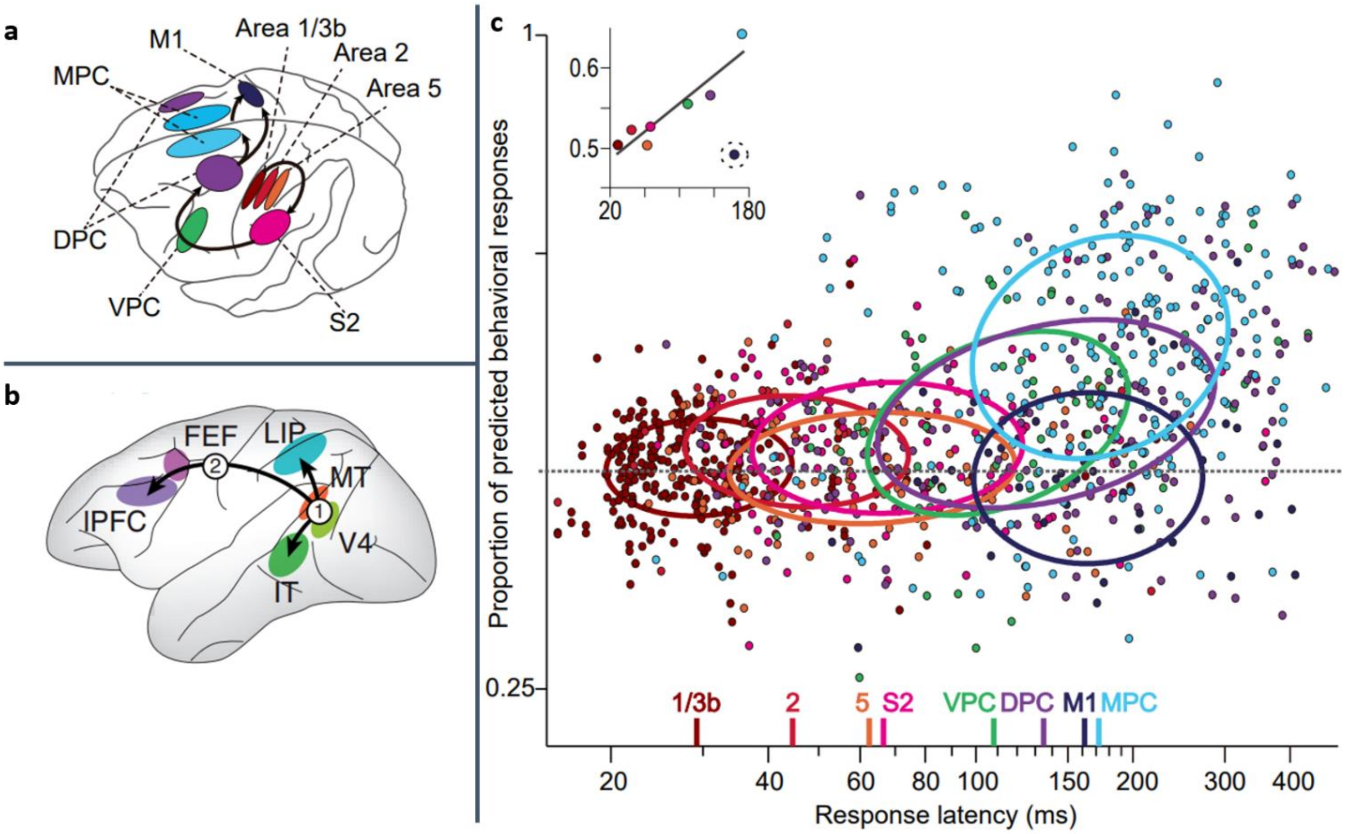


Figure 9. a) and b) Schematic diagrams depicting the flow of neural activity along various cortical regions of the primate brain during a vibrotactile task, and a visuomotor decision-making task, respectively. Their corresponding nomenclature is exposed in Figure 3.c. c) Visualization of the relationship between predictable choice probability (vertical axis) and the response latency of different cortical regions (horizontal axis) during a vibrotactile stimulus detection task. This graphic is directly related to the processing hierarchy depicted in Figure 9.a. Figures 9.a and 9.c have been readapted from Romo R. et al. 2020. Figure 9.b has been readapted from Siegel M. et al., 2015.

This process may be observed in Figure 9.c, in which the predicted behavioral response (vertical axis), which can be understood as a numerical factor associated with a neural population´s involvement in the abstract encoding of task parameters, as well as its commitment to a specific decision, gets higher as activity flows through various stages of the processing hierarchy (horizontal axis) (Romo R. et al. 2020).

Similarly, the other task may be regarded as a multisensory stimulus detection task (Figure 7.c), in which subjects must report on the absence or presence of periodical stimuli that can be either acoustic or vibrotactile. Importantly, although this task resembles to some extent the MultSen DM tasks, it mainly involves the auditory and somatosensory systems, while the task that does pertain to the set involves the auditory and visual systems. This way, it has been further found that early sensory stages of the brain, such as S1 and S2 in the case of the somatosensory system, and MT in the case of the auditive system, convey literal stimuli information into associative areas, such as various regions of the parietal and frontal cortices (Figures 5. a - h), where it is converted into an abstract and reliable representation of stimuli parameters. Interestingly, it has also been shown that several neural populations of the engaged structures unselectively respond to both sensory modalities, suggesting that a flexible neural encoding of task parameters is at play in the performance of these tasks (Vergara J. et al., 2016).

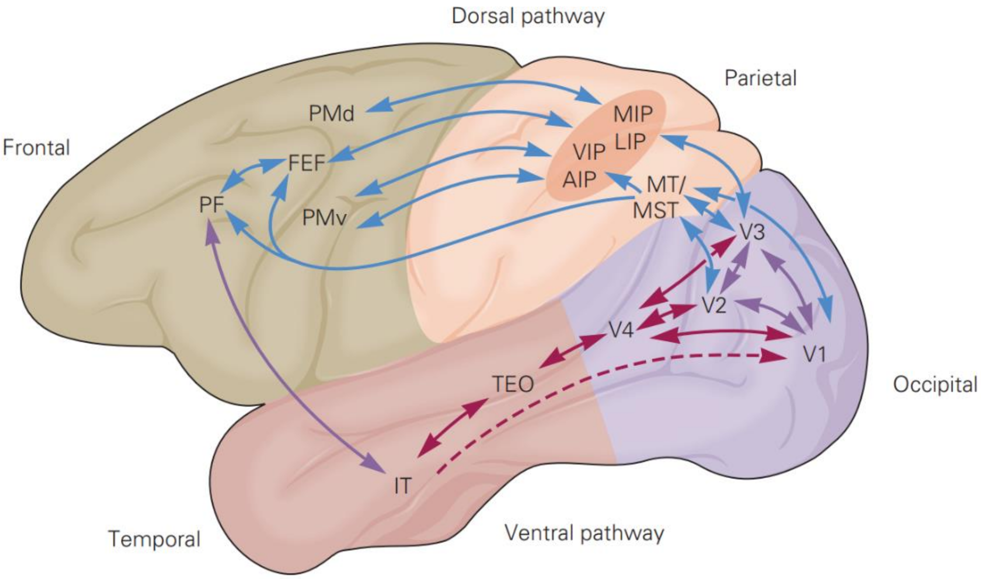


Figure 10. Schematic diagram showing various cortical regions of the primate brain that are engaged in the dorsal and ventral pathways. Figure 10 has been readapted from Kandel R. et al., 2013.

I.IV. Main takeaways I

Throughout this section, relevant results from cognitive and computational neuroscience from which the current discussion departs have been presented. Regarding the biological case, it has been shown that for the most part, sensory-cognitive processes are performed through hierarchical neurophysiological structures, distributed along cortical and subcortical brain regions, that transform literal stimuli information, conveyed by early sensory stages of different systems, into an abstract code interpretable by higher stages in such processing hierarchy, mostly located in frontal brain regions. To do this, the engaged neural populations encode by means of modulating their firing rates the task parameters of interest, generating a flexible and reliable neural representation that confers highly efficient cognitive abilities.

Remarkably, this set of facts may be evidence of the reliance of biological cognition and perception on a handful of functional features, **excitability** and **mixed-selectivity** at a single unit level, and **functional specialization** and **compositionality** at a systemic level. Further, this is exactly the basis for the modelling of this sort of processes by means of ANN models, as they too rely on such functional features.

II. Functional resemblance between ANNs and BNSs

Within the context that has been presented, stands out a remarkable line of research from which the current discussion takes its basis. As one of its main contributions, ANN models have been developed to perform the computational analogs of the cognitive tasks that have just been exposed, concluding that in order to efficiently perform the entire 20-task set, the models developed certain functional features that resemble those of their biological analogs (Yang G.R. et al., 2019). Disregarding this notable result, a finer explanation of the properties of said relationship has not been reported, becoming this work's main purpose to clarify such assertion. To do this, most of the model-related results that will be discussed in this section will be taken from a reference network, whose main characteristics will now be exposed.

II.I. The reference network

After the model had been successfully trained on the entire task set (Figure 12), various analytical and graphical methods were applied to visualize its functionality; some of which will be essential to the current discussion. Firstly, the functional selectivity of all the system's units was calculated for every task by means of the normalized task variance, a numerical factor ranging from 0 to 1, that corresponds to a unit's selectivity on a given task. Here, 0 would correspond to a completely non-selective unit, while 1 would correspond to a fully selective one. Thus, by assigning a color value to said factor and laying out the selectivity of every unit for every task, it is possible to obtain the raster plot shown in Figure 11.a. On this graphic, all 20 components of the task set are shown along the vertical axis, while the selectivity of all the network's units is shown along the horizontal axis. Importantly, it is possible to catalog units with similar selectivity profiles into differentiated groups that may be regarded as functional clusters, a categorization that will be central to this discussion. These clusters are also shown along the horizontal axis of the same figure.

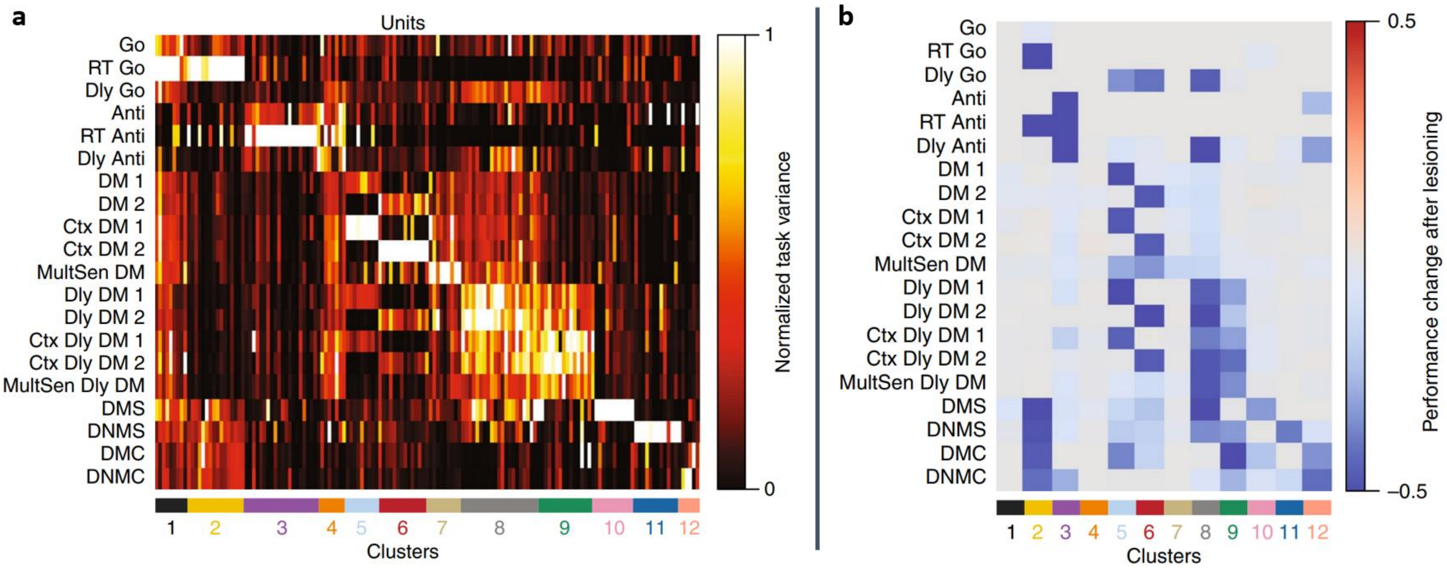


Figure 11. Emergence of functional selectivity in the reference network. The lighter each unit appears to be for a given task, the more selective it is on it. b) Emergence of specialized populations in the reference network. The darker each cluster appears to be for each task subset, the more it is causally involved in its performance. Figures 11.a and 11.b have been readapted from Yang G.R. et al., 2019.

Further, with the purpose of analyzing how said clusters are causally involved in task performance, a network lesioning technique can be applied. By “turning off” a given population and running the system over the task set, possible changes in the performance of specific tasks may be identified, indicating that such cluster is causally involved in their required processing. Thus, by repeating this process for every differentiated population in the reference network, it is possible to obtain the raster plot shown in Figure 11.b, where the reliance of specific task subsets on certain clusters can be observed.

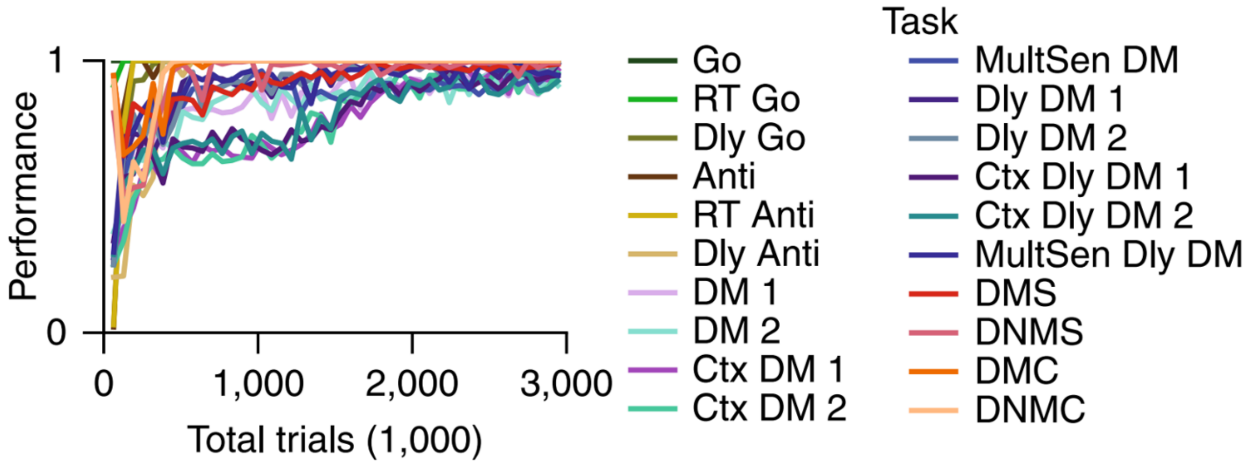


Figure 12. Reference network's training process. On this graphic, 1 corresponds to a perfect performance of any given task, while 0 corresponds to a fully negative outcome. Every line represents a different component of the task set, as the color code is shown on the right side. Figure 12 has been readapted from Yang G.R. et al., 2019.

On this graphic, the darker each task appears to be for a given cluster, the more dependent it is on it. This way, it can be seen that all Categorization tasks as well as the RT modalities of the Go and Anti subsets are associated with Cluster 2, or C2, as all components of the Anti tasks are related to Cluster 3. Analogously, it may be observed that the following pair of clusters, C5 and C6, are involved in each modality of all DM tasks, respectively. Finally, it can be easily noted that C8 is involved in the performance of all tasks that include a Dly factor.

Now, several aspects of this relationship may be exposed, but first, it is right to recall that the hereby proposed relationship between ANN and BNS is fundamentally based on the functional features that arise from the interplay of their given populations, not on their nature *per se*. It is quite important to keep this principle in mind during the following discussion, as it is the search for efficiency the fuel for the emergence of their resembling features.

As it has been exposed, to efficiently perform the entire task set, the reference network relies on a set of functional features both at a single unit and a systemic level, resembling to some extent those identified in its biological analogs. Further, it is the interplay between these features what confers the system highly flexible task-solving abilities. Interestingly, as it will be shown here, said features may be catalogued into a pair of categories, elemental and emergent, based on their causal origin in the system.

II.I.I. Unit-level features

Starting with general single unit characteristics, it may be stated that for both the artificial and biological case, **excitability** may be fundamental for their characteristic firing functions, as it has been shown that it is a very efficient process in regards of information processing. By conceiving a dual-state threshold activation function, an excitable system may adopt one out of two possible states as a consequence of information flow into it; thus, if the input complies with certain properties, it may cause the system to go into its excited state, while insufficient input will not cause any response (Figures 14.a and 14.b).

In logical terms, this process has proven to be quite useful, as it is true that all computable functions may be expressed as a combination of discrete values, such as the excited/basal states of neural units (Figure 14.c) (Leonard N.E. et al., 2023). Regarding the biological case, this fact may also be the basis for the abstract encoding known to be performed by the neural populations discussed here, as it has been shown that it relies on their firing rate modulations (Figure 14.d).

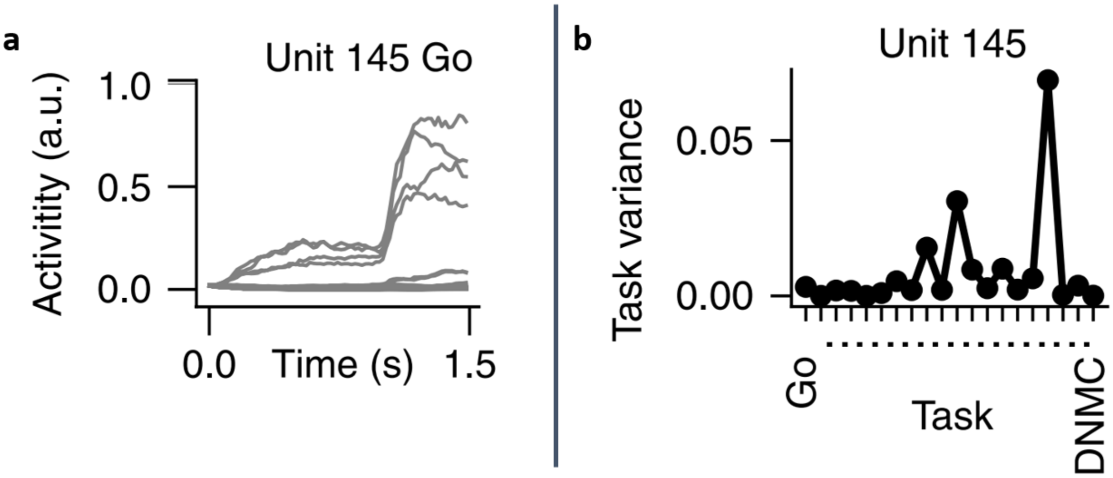


Figure 13. Activity in time of a single unit for the Go task. Different lines represent different task parameters. Here, it may be observed how this particular unit preferably responds to certain parameters. b) Normalized task variance of unit 145 for every component of the task set. Here, the selectivity of the unit for the DMS task can be easily noted. Figures 13. a and b have been readapted from Yang G.R. et al., 2019.

The next common feature between ANNs and BNSs may identified as **functional selectivity**, since, as it has been previously exposed, both systems count on neural populations tuned to respond to specific inputs. For instance, regarding visuomotor tasks, there is extensive evidence that supports the existence of highly selective units in early stages of the biological visual processing hierarchy (Figure 9.b) (Figure 10), as various populations are specialized for recognizing diverse visual features. Similarly, regarding vibrotactile tasks, selective units and populations have been identified along the sensory hierarchy involved in their performance (Figure 9.a). Further, another functional feature that has been proposed to be at play in efficient cognition is **mixed-selectivity**, which may be defined as the absence of response preferences in neural units, therefore allowing them to unselectively respond to inputs of variable nature. Remarkably, this feature may also be identified in both the artificial and biological systems treated here, noting that it is in fact fundamental for their functionality.

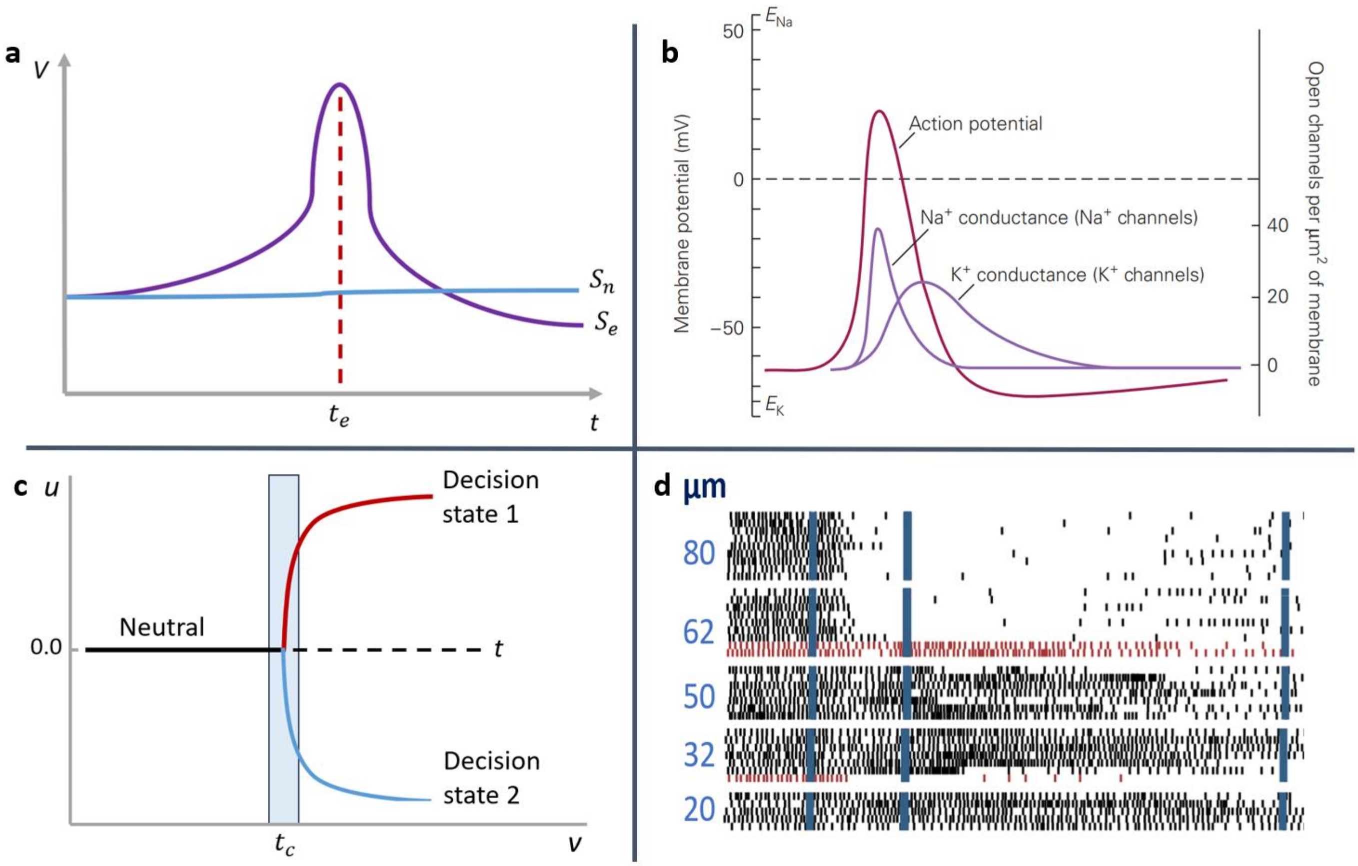


*Figure 14. a) Single-unit excitable dynamics. On this qualitative graphic, where potential energy (V, vertical axis) is a function of time (t, horizontal axis), the energy of two physical systems, an excitable and a non-excitable one, Se and Sn, respectively, are compared. It may be easily observed that for the case of the excitable system, a discrete point in time (te), as well as a reliable binary encoding of information (on/off), may be conceived by means of spiking dynamics, contrary to the non-excitable case, where it is evidently much harder to define discrete states and points in time. b) Presence of excitability in the fundamental firing function of biological neurons, as a consequence of their neurotransmitter-triggered electric depolarization. c) Qualitative graphic showing a decision-making process based on criticality. On this figure, a critical event that occurs at a time tc generates two possible well-differentiated system states. This process is at the basis of the neurocomputational systems discussed here. d) Raster plot of an S1 neuron responding to vibrotactile stimuli of varying amplitude. It may be observed that although this unit responds in all cases, it evidently shows a preference for low-amplitude stimuli. This set of facts suggests the presence of **excitability** and **functional selectivity** in the fundamental units of ANNs and BNSs. Figure 14.b has been readapted from Kandel R. et al., 2013. Figure 14.d has been readapted from Diaz-deLeon G. et al., 2022.*

As for the biological case, neural populations showing mixed-selectivity actually conform the majority of cases, as fully selective neurons are quite rare. For instance, regarding the cognitive tasks discussed here, most of the examined neurons show non-selective response profiles (except for the highly selective units of early processing stages), suggesting that they flexibly encode information of various stimulus parameters. This abstract representation is then conveyed to associative regions of the brain, where different neural mechanisms come at play. Interestingly, the nature of this process, in which early processing stages consisting mostly of highly selective units that conceive an abstract task representation interpretable by higher stages in its functional hierarchy, supports the hypothesis that efficient cognitive abilities may rely on a common and flexible neural code (Diaz-DeLeon G. et al. 2022) (Parra S. et al. 2022).

As for the artificial case, and more specifically, that of the reference network, the existence of mixed-selectivity may also be identified from both a single-unit and a systemic perspective. As for the first one, it is possible to plot the activity in time of any random unit, visualizing its activation preferences for distinct task parameters (Figure 13.a). Analogously, its normalized task variance may be plotted for every component of the task set, conceiving its task selectivity profile (Figure 13.b). Regarding the specific unit and task that were chosen for said figure, unit 145 and the Go task, it may be noted that it definitely presents selectivity for certain task parameters, as well as it is preferably engaged in the DMS task, and to a lower degree, on the MultSen DM task. This way, by extrapolating this process to the entire network, it is possible to obtain the raster plots shown in Figure 11.a and 11.b, where the presence of mixed-selectivity at a systemic level may be easily observed. For instance, it can be seen that most units pertaining to Cluster 8 are engaged in all DM tasks, as well as some modalities of the Go and Anti tasks, which shows a high response variety. On the contrary, it is also easy to note the presence of highly selective responses in the system, as units pertaining to Cluster 3 are predominantly engaged in only the RT tasks. This set of facts consists direct evidence for the reliance of the reference network's functionality on mixed-selectivity.

II.I.II. System-level features

Within a systemic perspective, it is also true that the reference model shares some functional features with its biological analogs. As it has been previously shown, populations of units with similar response profiles conform functional clusters that are causally involved in various task subsets, which consists direct evidence for the emergence of **functional specialization** in the system. Thus, this functional feature in which whole populations are tuned to process specific aspects of certain tasks, may be qualitatively understood as an extension of **functional selectivity**.

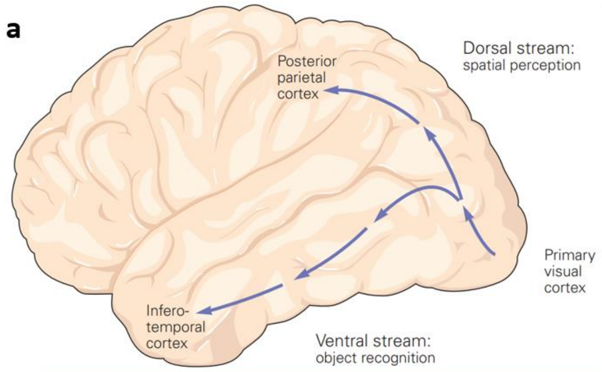
More specifically, it may be easily observed in Figure 11.b that the system did in fact develop at least five well-differentiated functions that subserve specific task subsets: C2 is causally involved in the performance of all Categorization and RT tasks, C3 is associated with the Anti tasks, C5 and C6 are engaged in each modality of the DM tasks, respectively, and C8 is involved in all tasks that include a Dly factor. Moreover, these functions may be consecutively engaged to conceive more elaborated processes, as is the case for Cluster 8 and all tasks that involve a working memory function, such as the Dly DM tasks (Figure 16.b). Once again, this is evidence for the reliance of the network's task representation on a functional feature shared with BNSs, **compositionality**, in which specialized functions can be consecutively applied to perform complex tasks. Biologically, this set of features has also been extensively identified to be at play in the cognitive processes discussed here.

Firstly, it is evident that as for auditive, visual, and vibrotactile tasks, there exist various specialized populations through which task information is first encoded and then relayed to associative areas of the brain. For instance, regarding auditive and vibrotactile tasks, such structures have been localized in parietal and temporal cortical regions (Figure 5), respectively, while early states of the visual processing hierarchy have been identified in occipital brain regions (Figure 10).

These populations gradually transform literal physical attributes of the stimuli into an abstract code that is conveyed to several areas of frontal and parietal regions, where higher cognitive functions, such as working memory and conscious decision-making, are known to take place. This set of facts further support that functional specialization and compositionality may be fundamental functional features of flexible cognition (Kanwisher N. et al. 2010).

II.I.III. Elemental and emergent features

To this point, it has been shown that the ANNs and BNSs discussed here share a common set of functional features that confers them highly flexible and efficient cognitive abilities. Interestingly, at least for the artificial case, these properties can be catalogued into a pair of categories depending on their causal origin in the system. Starting with elemental features, which can be defined as those present in the system even before it is trained on any task, it may be stated that **excitability** is evidently the sole component of the category. Since the networks are essentially blank before they start training, all the other functional properties that have been exposed, **mixed-selectivity**, **functional specialization**, and **compositionality**, emerge as a consequence of the system's development as it increases its performance, which is why they may be referred to as emergent features (Figure 15.b). Importantly, regarding the biological case, this categorization may not be valid, as most functional features are inherent to BNSs even before they are trained. This fact has strong implications on the current discussion, and it will be brought up again later.



b

ANNs	Elemental features	Emergent features
Unit level	• Excitability	• Mixed-selectivity
System level		• Functional specialization • Compositionality

c

Cluster (Task family)	Neurocomputational process	Relatable brain regions
2 (DMS / DMC / RT)	• Visual shape recognition • Comparison / Categorization	• **Visual ventral stream or "what" pathway** (shape recognition)
3 (Anti)	• Saccade inversion	• **PFC** (voluntary vector inversion)
5 & 6 (DM 1 and 2, respectively)	• Bimodal encoding of sensory information by parametric representation (audiovisual and tactile stimuli)	• **Somatosensory cortex, VPC, DPC, MPC** (vibrotactile discrimination) • **Visual dorsal stream or "where" pathway** (movement recognition)
8 (Dly)	• Abstract representation of stimuli • Working memory	• **PFC** (abstract and parametric representation of stimuli)

Figure 15. a) Schematic diagram depicting the dorsal and ventral visual streams of the primate brain, which are specialized for processing non-static visual stimuli and shape recognition, respectively. Remarkably, these functions are also present in the reference network. b) Elemental and emergent functional features of the ANNs discussed here. c) Common functions between the reference network and its biological analogs. Here, the specialized populations developed by the network are qualitatively associated to various brain regions through their shared processes. Figure 15.a has been readapted from Kandel R. et al., 2013.

II.I.IV. Common functions

Furthermore, since both the artificial and biological systems treated here were trained on analog tasks, they did not only develop resembling functional features, as their functionality also relies on common processes (Figure 15.c). For instance, regarding Cluster 2, which is selectively involved in all the Categorization tasks and whose function can be traced to visual shape recognition, may be associated to the highly selective neural populations involved in the real-life version of the tasks, which have been localized along various areas of the occipital and temporal cortices (Figure 10), as well as populations distributed through several inner brain regions. As a whole, this neurophysiological structure specialized for the recognition of static visual stimuli is known as the ventral stream, or rather informally, the brain's "what" pathway (Figure 15.a). Similarly, the function of Cluster 3, which is causally involved in the performance of all the Anti tasks, may be related to the executive processes carried out by frontal brain regions, known to be predominantly engaged in conscious decision-making. Further, neural activity directly associated with the vector inversion required for this task's adequate performance has been identified in various regions of the primate prefrontal cortex.

The function of the following pair of specialized populations developed by the reference network, Clusters 5 and 6, also coincides with a few interesting neural processes. Given the multisensory stimuli they involve (auditive, visual and vibrotactile), the bimodal nature of the decision-making (DM) tasks in which they are engaged, and that each population is causally applied in the performance of one of their modalities, their main function can be traced to the parametric representation of stimuli attributes, as there is extensive evidence that this process underlies the biological version of these tasks (Figure 16).

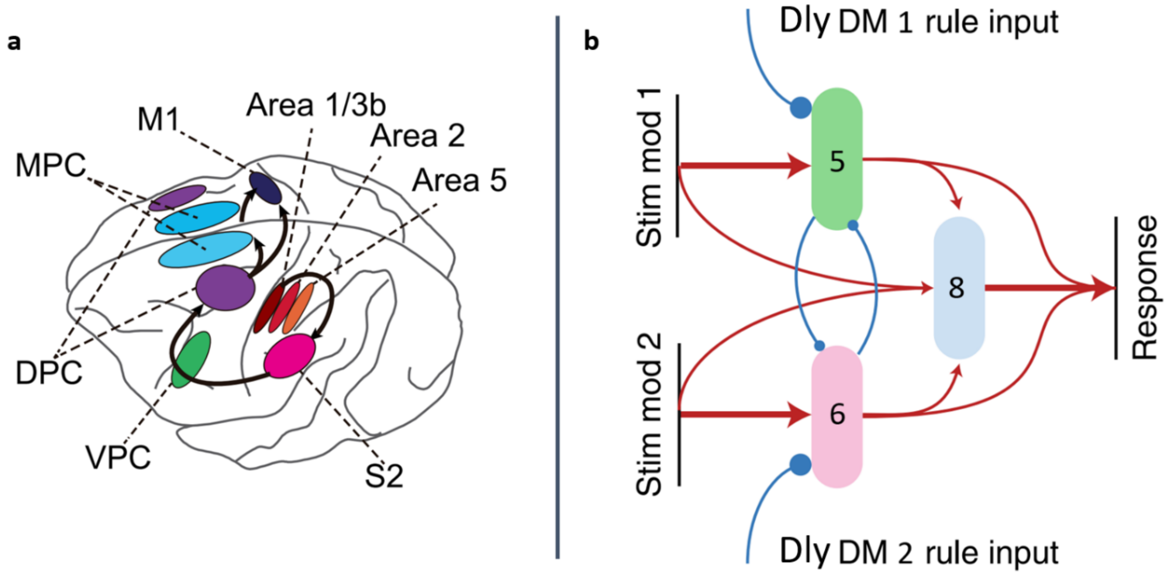


Figure 16. Schematic diagrams depicting the interplay of ***functional specialization*** *and* ***compositionality*** *in both ANNs and BNSs. a) Processing hierarchy involved in vibrotactile tasks, in which different neural populations abstractly encode relevant task parameters. This representation may then be composed with higher cognitive functions, such as working memory and decision-making, which are mostly localized in frontal cortices (not shown here). b) Compositional functionality in the reference network for the Dly DM tasks. Here, information corresponding to the first modality of the task is mostly processed through Cluster 5, as information corresponding to the second modality is processed through Cluster 6. As such task requires the maintenance of a parameter during a delay period, this representation is then composed with the function of Cluster 8, working memory, to finally conceive the system's response. This functionality is remarkably similar to the biological case, considering the same task context. Figure 16.a has been readapted from Romo R. et al. 2020. Figure 16.b has been readapted from Yang G.R. et al., 2019.*

Essentially, by analyzing the activity of different stages of their associated processing hierarchies, it has been found that a gradual encoding of literal stimuli features into an abstract representation is fundamental for effective decision-making (Figure 9.c). For instance, regarding vibrotactile tasks, such processing has been identified along the hierarchical structure depicted in Figure 16.a, which involves several regions of the parietal and frontal cortices.

Remarkably, as for the visual DM tasks, similar parametric representations of stimuli parameters, from which a reliable bimodal code needed for decision-making is conveyed, have also been identified along the visual dorsal stream or the brain's "where" pathway, a hierarchical structure distributed along occipital and parietal cortices and that is specialized for the recognition of non-static visual stimuli (Figure 10) (Figure 15.a).

Finally, the function of Cluster 8, which is causally engaged in the performance of all tasks that include a Dly factor, may be traced to working memory processes as such tasks require the system to maintain some trace of task parameters during a delay period. Biologically, this process has been localized almost exclusively in frontal brain regions, and, since by that point of the perceptive hierarchy the neural representation at play is almost purely abstract, said populations may flexibly process information coming from different sensory modalities. Notably, this sort of working-memory functions has been identified in the real-life version of many of the tasks discussed here, such as Dly Go (visuomotor task), Dly DM (vibrotactile tasks) and MultSen DM (multisensory decision-making tasks).

II.I.V. Main takeaways II

Throughout this section, resembling functional features between a reference ANN model and the BNSs on which it has been inspired were presented. Firstly, it was shown that flexible problem-solving abilities in both systems rely on a set on functional features both at a single unit and a systemic level; **excitability** and **mixed-selectivity** for the first one, and **functional specialization** and **compositionality** for the latter. Remarkably, regarding the biological case, extensive evidence from multiple lines of research supports the presence of these features in cognition and perception. Then, it was shown how both systems even developed resembling specific processes, given they were trained on analog tasks. Now, with these considerations, the next major perspective from which the subject will be discussed may be introduced.

II.II. Complex systems and neural networks

Although there exist inconsistencies in the definition of “complexity”, it may be unmistakably stated that the subjects treated here qualify as complex systems. Therefore, this perspective can yield valuable insights regarding the relationship between ANNs and BNSs. To this end, the following discussion will be based on three main subjects: a mathematical, a physical, and a biological context.

II.II.I. A mathematical perspective

Naturally, one of the most viable ways to study the subject of complexity is by means of mathematics. As discussed previously, the models treated in this work are essentially interconnected systems of individual units (the basic mathematical definition of a network), that are meant to emulate certain known properties of BNSs (Figure 1). This approximation allows their analysis through almost purely mathematical terms. Thus, given the hierarchical structure of the considered systems, an abstract topological space where all possible network configurations reside exists theoretically. This permits the testing of hypotheses related to structural and physiological properties across a diversity of models, allowing the conception of the following scenarios: Considering the fundamental properties of the models treated here, and given an appropriate selection of parameters, it is true that any sufficiently large ANN is capable of solving out any mathematically well-behaved function by means of the composition of specialized processes. Such task representation would fully rely on functionally selective units, generating well-differentiated populations performing separate, specialized functions (Figure 17, upper left quadrant). Evidently, this functionality would entirely rely on **functional specialization** and **compositionality**. Hence, if such a system was applied to solve diverse tasks, their adequate performance would require the existence of many subpopulations of selectively tuned units, *i.e.,* a greater number of tasks would exponentially increase the necessary population size of the network, as well as its computing resource requirements.

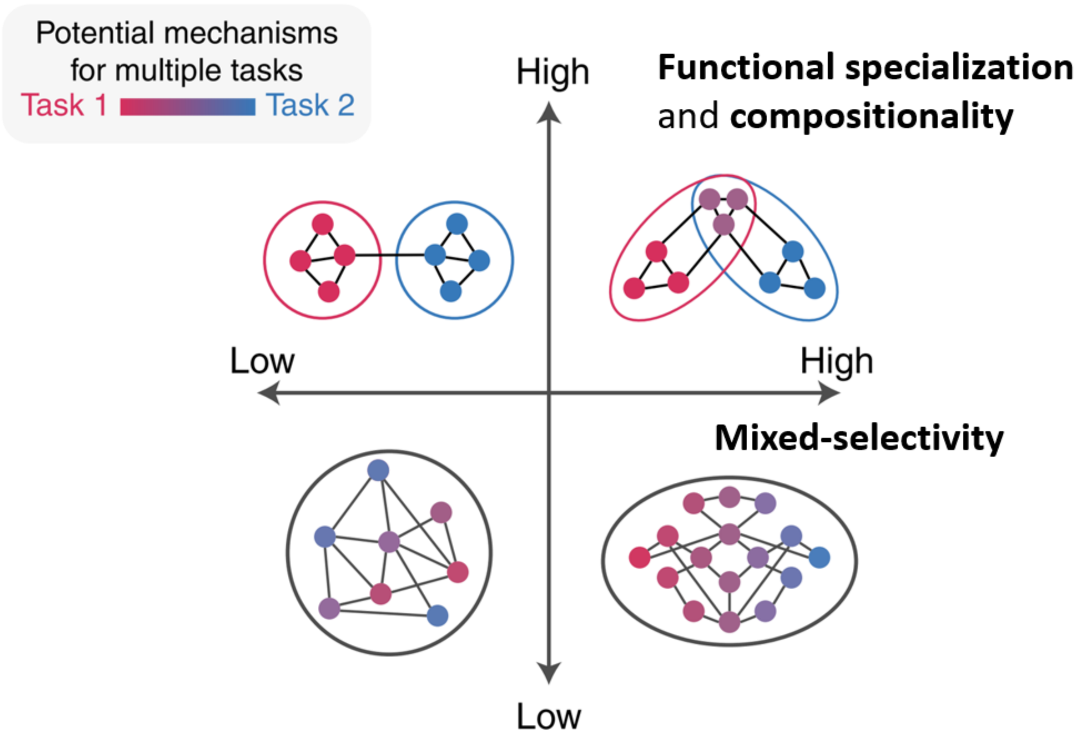


Figure 17. Two fundamentally different methods for multiple task-solving. On the upper left quadrant, a fully differentiated and compositional representation, and on the lower right one, a fully distributed and non-selective one. On the upper right quadrant, a combination of both methods that confers the system a highly flexible and efficient task representation may be observed. Figure 17 has been readapted from Yang G.R. et al., 2019.

In contrast, it is also theoretically plausible for an ANN, with appropriate parameters and population size, to be able to solve many different tasks by means of a fully distributed representation relying only on non-selective units; a dependence that would not generate differentiated subpopulations nor specialized functions (Figure 17, lower right quadrant). Clearly, such method is based on **mixed-selectivity**, and it would also considerably increase the system's total number of units as well as its resource needs, rendering it inefficient along with the fully compositional representation that was exposed before. In sum, these hypothetical scenarios mean to illustrate how the sort of networks treated here may develop two fundamentally different methods to achieve flexible task-solving; either a fully specialized and composable representation, or an entirely distributed and nonselective one.

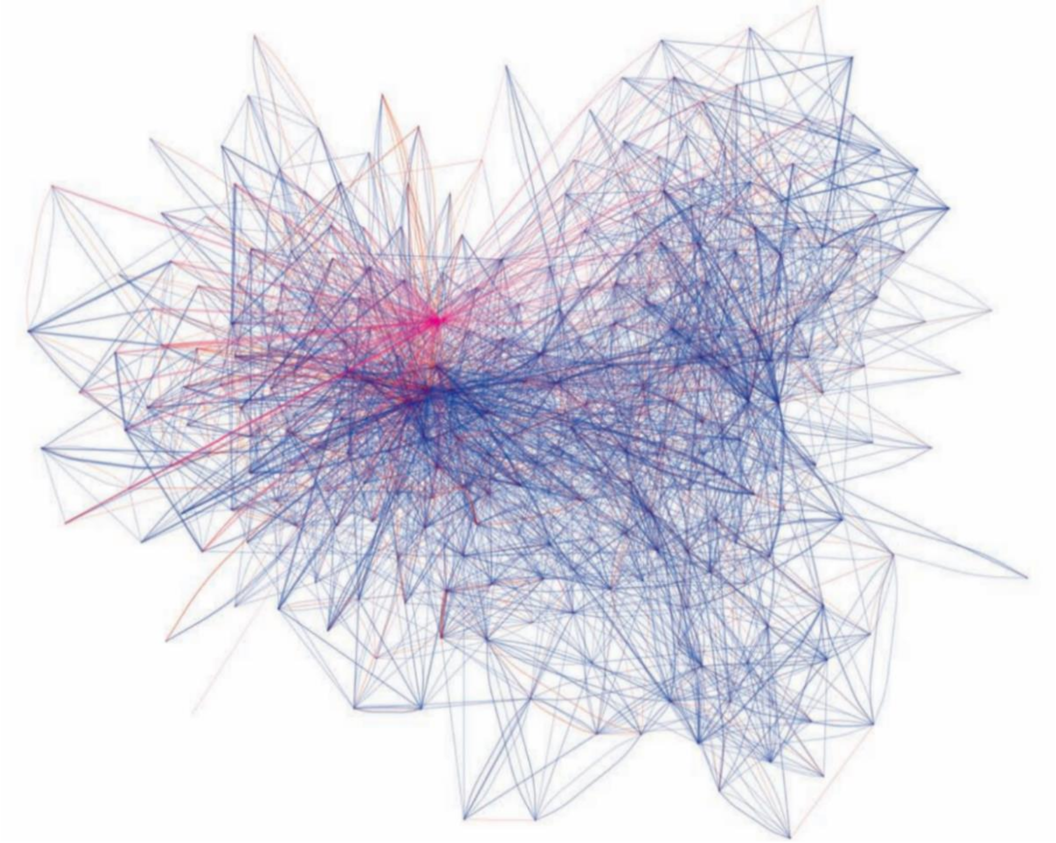


Figure 18. Partial connectome of the central nervous system of a C. elegans specimen. Here, each nodule represents a neuron, as different colors correspond to varying types of intercellular connections. Clearly, within a network perspective, this rather "simple" BNS shows a functional and structural complexity that is still unaccounted for in most biologically-inspired ANNs. Figure 18 has been readapted from Kandel R. et al., 2013.

Crucially, this functionality is fueled by the systems' inherent search for energetic efficiency, as they must conceive the best method possible to achieve their purpose while administrating available resources. Importantly, generating these exclusive representations may only be possible if the systems have access to unlimited resources, which is evidently not the case for ANNs and BNSs. Therefore, it follows that resource-restrained networks as those treated here must conceive more efficient solutions than those that have just been exposed if they are to optimally achieve their purpose, making possible for these systems to develop task representations subserved by neural dynamics situated halfway between both possibilities (Bassett D.S. et al., 2017).

Within a computational perspective, valuable insights into this matter may also be reached. Firstly, the emergence of system-level features such as **functional specialization** and **compositionality** are a direct consequence of the tasks' structure since they clearly represent mechanisms of greater efficiency; neural populations tuned to specific purposes will generally outperform nonselective populations. However, **mixed-selectivity**, a fundamental unit-level feature, may optimize the readout from non-selective populations; the combination of activity from single-units can achieve similarly efficient performance. Thus, a mixed application of these features may conceive a highly flexible and efficient method for multiple task solving, given limited resources (Figure 17, upper right quadrant). Remarkably, out of such functionality emerges one of the main resemblances between the artificial and biological systems treated here, as they both rely on neural dynamics that coincide with these properties. This suggests that ever-increasing energetic efficiency may be at the basis of their development and subsequent characteristics (Newman M. et al., 2006).

II.II.II. A physical perspective

The search for energetic efficiency is purported to be the primary motive for the emergence of similar functional features between different natural complex systems. This search also explains the similarity between BNSs and other physical systems. In this context, greater efficiency means greater stability. For example, water can be thought of as a natural and complex system: While flowing downstream, a river takes the path of least resistance. Less resistance can be directly equated to greater efficiency. Over time, the path formed by the river becomes more stable as the water erodes the environment around it. Interestingly, these dynamics eventually generate main pathways, or streams, that facilitate the system's "purpose" of bringing water to its lowest potential, resembling the way specialized neurophysiological functions and their underlying structures facilitate the "purpose" of cognitive function.

Ideally, this analogy pretends to illustrate how certain functional features of both the neural systems treated here, such as **functional specialization**, **compositionality**, and even **mixed-selectivity**, are relatable to those of purely physical complex systems, flowing rivers, in the way that both dynamical systems adopt, as a direct consequence of a few physical principles, their most possibly efficient configuration (Barab’asi A.L. et al., 2002).

II.II.III. A biological perspective

Shifting perspectives, the inherent search for efficiency of naturally occurring complex systems may also be identified within a purely biological context; both ANNs and BNSs seek optimal efficiency in structure and function. Regarding the artificial case, this fact may be easily observed as the mathematical algorithms used to train the networks are essentially efficiency-guided (Equations 1 - 6). However, the search for efficiency is more difficult to isolate in biological systems. For survival, all organisms evolve to improve the efficiency of an individual’s fitness. In this manner, all organisms have developed the most efficient method of survival given the long-term conditions of their environment. Fundamentally, this fact may be extrapolated to the ANNs and BNSs treated here, in the way that both systems have developed the most efficient method to achieve their purpose. This is furtherly supported by yet another well-studied biological phenomenon, evolutionary convergence, in which unrelated species come up with similar features (which are usually also the most efficient) to make their way through resembling ecosystemic pressures (Gould et al., 2002). Applied to the current discussion, this assertion is also valid for ANNs and BNSs, as they rely on resembling functional features to optimally deal with analogous tasks, suggesting that such functionality is the most efficient both systems have been able to come up with.

II.II.IV. Main takeaways III

In sum, throughout this section various aspects of the existing relationship between naturally occurring complex systems and neural networks have been discussed. Essentially, it has been found that the search for energetic efficiency fuels the functional resemblance between ANNs, BNSs, and other complex systems, in the way that they all tend to adopt their most stable and efficient configuration by means of the interplay between a handful of physical principles. This is, that their resembling functionality may also be the most efficient one.

II.III. Further considerations

The similarities between the observable features in ANNs and BNSs have been described and justified with a variety of perspectives. At this point, it should be relatively evident that these parallel processing mechanisms are inherent parts of the systems or emerge due to their training process, and, since the purpose of the current work is to elucidate and consider the relationship between the two types of systems, there is another crucial consideration that must be addressed: the biological underpinnings of the ANNs, oftentimes referred to as their biological plausibility. Understanding this aspect of the mathematical models treated here will prove to be a fruitful avenue for further investigation.

II.III.I. Biological plausibility of ANNs

One detraction of the use of mathematical modeling of natural phenomena is the fact that capturing all variables or features of any physical system is essentially impossible. Due to this, investigators isolate characteristics of particular interest, a procedure that inevitably ignores features that occur outside of the model's parameters, rendering it a manageable, yet incomplete, description of said phenomenon. It is necessary to keep this fact in mind for the following discussing of the biological plausibility of the models we have observed previously. In the first place, the network's unit-level operators are dynamic systems that model the electrochemical process known as an action potential for individual cells; in this manner, a single unit or operator can be thought of as a single neuron (Figures 1.a and b). Focusing on the action potential and its resultant extracellular signal transmission, a breadth of models exist that emulate the spiking dynamics observed experimentally. However, a practical model implementation of intricate extracellular and intracellular chemical interactions (Figure 20), a vast variety of nervous and non-nervous cells at play in cognition (Figures 19 a - d), as well as a representation of their profound interconnected nature has yet eluded investigators. Further, at a systemic level, stark differences between ANNs and BNSs may also be noted. In the previously described models, 256 interconnected LIFs were separated into input, middle, and output layers (Figure 1.c). This is a very simplified network compared to an active BNS, which is composed of millions of cells that are interconnected to a much higher degree (Figure 18). An effective model for representing the various functions of differing cell types and temporal dynamics of modular brain areas has not been proposed. A final distinction to discuss is the likely disparity between the ANN machine-learning algorithms and the computations performed within BNSs, as there are many fundamental processes of biological learning, such as attention and reward-based motivation, that have been largely ignored.

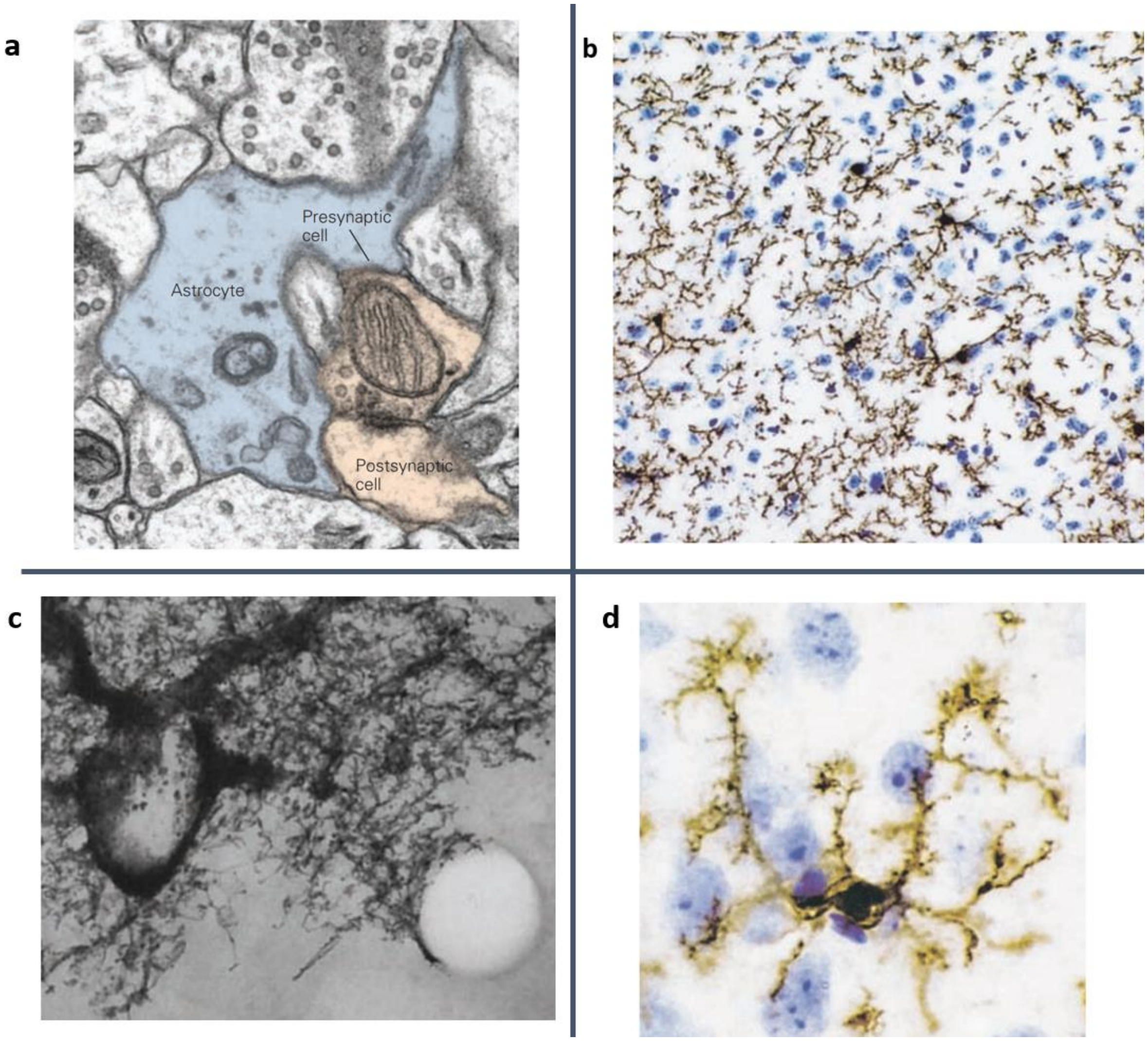


Figure 19. Examples of diverse factors that are not considered by the network models treated here. For instance, it is known that astrocytes and glial cells have important roles in neuronal metabolism, synapsis and inter-neuron communication. a) Micrograph of an astrocytic cell engulfing a presynaptic and a postsynaptic motor neuron. b) Photograph of glial (brown) and non-glial cells (blue) scattered along a layer of the motor cortex of a rat brain. c) Electron micrograph of an astrocyte enveloping a blood vessel. d) Photograph showing the connections between one glial cell and several motor neurons. Figures 14 a, b, c and d have been readapted from Kandel R. et al., 2013.

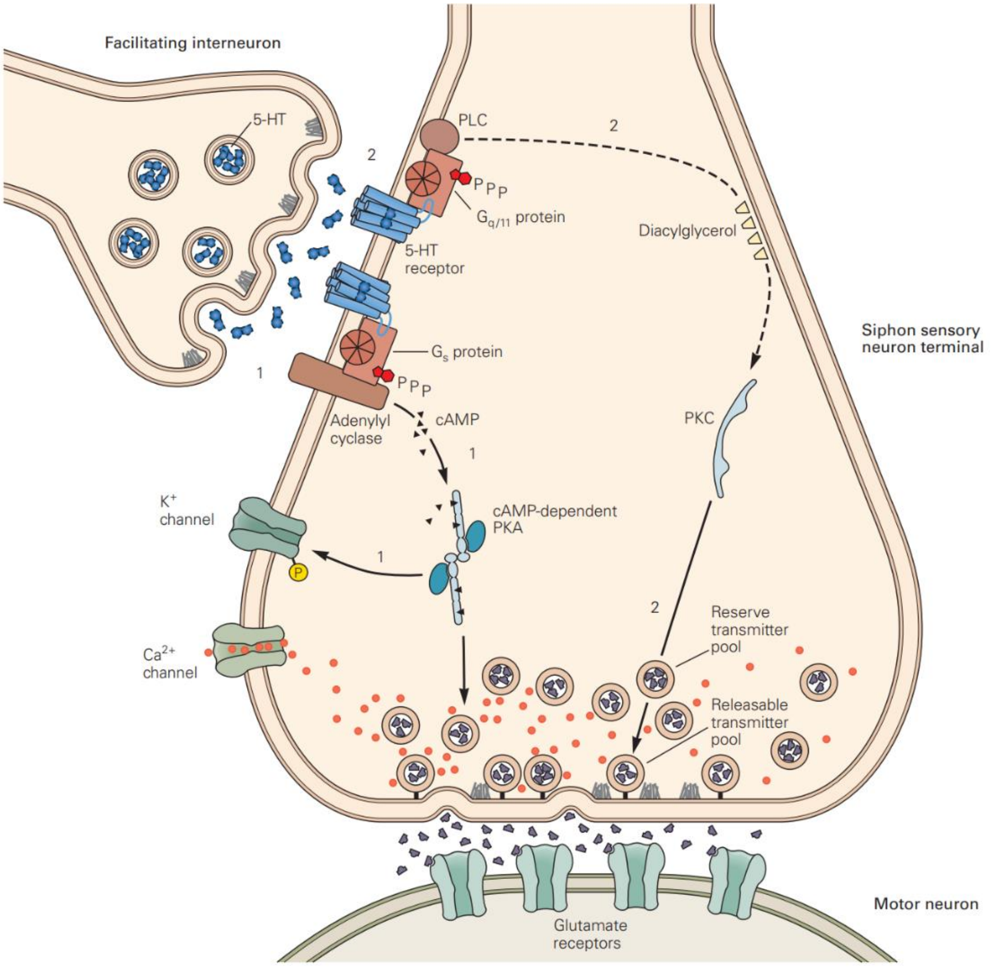


Figure 20. Schematic diagram depicting several complex chemical interactions, both intracellular and extracellular, at play in the synaptic region between a presynaptic sensory neuron, a postsynaptic motor neuron, and a glial cell. Most of these intricate processes, fundamental for any biological neural system's physiology, are not considered by the models discussed here. Figure 20 has been readapted from Kandel R. et al., 2013.

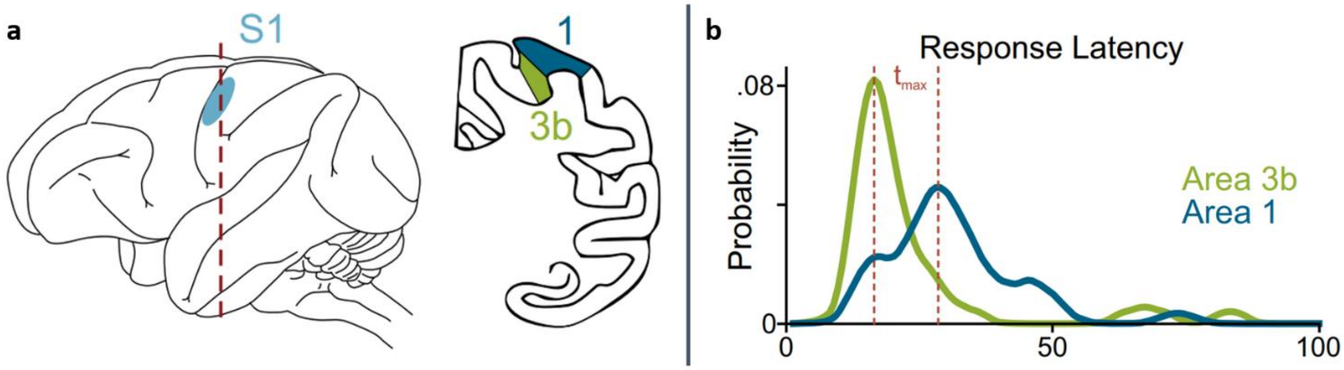


Figure 21. Schematic diagrams of the temporal dynamics underlying a MultSen DM task. On this simple example, the sequential engagement in time of different neural populations engaged in the task may be easily observed. These dynamics are not considered in the ANN models discussed here. a) Cortical regions known to be involved in the encoding of vibrotactile stimuli; areas 1 and 3b of the primary somatosensory cortex. b) Quantitative graphic showing how both areas are sequentially engaged in the encoding of vibrotactile stimuli. The horizontal axis represents time, starting from the beginning of the task, while the vertical axis represents the overall activation probability of such populations. t_{max} *stands for the maximum amount of time between both peaks of activity. Figures 21 a and b have been readapted from Parra S. et al., 2022.*

II.III.II. Neural circuitry and manifolds

The set of functional features discussed in ANN models falls under the umbrella of neural circuitry, a major perspective for studying the nervous system that refers to how diverse functions are associated to different areas in BNSs. Importantly, this approach has been conformed over a long history of research, composing a rather classical way to study said systems. Nevertheless, modern methods for analyzing said circuitry have been thoroughly applied, leading to stellar advances in the field. These mathematical techniques are mostly based on topology, the concept of manifolds (Sussillo D. et al., 2013). One manner to study the activity of both artificial and biological neural systems in a comparable framework is by studying the manifolds that emerge from multi-dimensional state spaces. Frequently, the dimensionality of said manifolds is reduced for the simplification of the possible interpretations derived from both kinds of systems; a process that facilitates the identification of parallel neural codes (Pujalte S., 2022).

Crucially, although these methods are not explicitly applied in this work, an integral perspective that considers both approaches has conceived the actual understanding of the neurophysiological basis of cognition from which most of the results exposed here have been taken. Thus, it follows that the conclusions resultant from the current discussion are consistent with both the neural circuitry and the manifolds perspectives (Langdon et al., 2023).

II.III.III. Future work

To further advance the understanding of the relationship between ANNs and BNSs, neural models can be refined to address various aspects both at a single-unit and a systemic level. Regarding the first one, a broader set of biophysical variables, such as a diverse family of neurotransmitters and their functional properties, as well different cell types, could be implemented (Figure 20). This would allow the study of agent-triggered responses across diverse neural populations within a computational perspective, a process that is not considered here. On the other hand, at a systemic level, several features could be investigated: increasing network size and applying modular functionality to the systems would allow the examination of more biologically accurate ANNs (Figure 19) (Figure 21). In addition, recalling the lesioning method by with the specialized populations of the reference network were identified, it would be interesting to study how the system deals in real-time with impaired functionality; this is, with deactivated populations. This would allow further insights into the network's flexibility, even serving as a computational model to study neural plasticity. Finally, the functional features that have been identified throughout this work could be implemented into the network's basic structure, aiming to conceive even more biologically plausible ANNs. The basis for this proposal is quite simple, as such features emerge in the artificial systems only after they had been thoroughly trained on the tasks, while the same properties are inherent to their biological analogs. Ideally, implementing these characteristics would help to bridge artificial and biological intelligence. These proposals are in no way exhaustive in terms of possible future work.

III. Concluding remarks

In the current work, the relationship between a certain kind of ANNs and BNSs has been examined from various perspectives. When training both systems on comparable perceptual tasks, the functional features that emerge are unquestionably similar. Crucially, this similar functionality for flexibly solving multiple tasks does not emerge directly from the interplay between individual units in either type of system; it is rather the combination of distinct functional properties in neural populations that plays a greater role in producing the necessary features.

Essentially, this functionality relies on the gradual encoding of perceptual stimuli as information is conveyed through specialized functional structures, generating a flexible representation that reliably encodes diverse task parameters. This representation may then be composed with other functions to conceive more complex processes. Remarkably, these system-level phenomena, which may be referred to as **functional specialization** and **compositionality**, respectively, are valid for both artificial and biological neural systems. Moreover, given that both systems were trained on resembling tasks, it was also found that the very functions they perform are relatable. As for unit-level features, it was also found that efficient functionality for both ANNs and BNSs rely on **excitability** and **mixed-selectivity**.

Regarding the artificial case, said functional features may be catalogued into a pair of categories, emergent and fundamental, depending on their causal origin in the systems. Fundamental features, such as excitability, are those present in the networks even before they are trained on tasks, while emergent features, mixed-selectivity, functional specialization, and compositionality, are conceived through training. This constitutes a major difference to the biological case, as said features are naturally inherent to BNSs.

Further, it was also shown that the emergence of this flexible functionality is a direct consequence of both artificial and biological systems search for energetic efficiency; they both seek optimal performance while administrating available resources. This is also the reason for the resemblance between neural networks and other kinds of naturally-occurring complex systems; they all tend to adopt their most efficient functionality based on the interplay between a handful of physical principles.

Finally, several differences between ANN models and their biological analogs were discussed, most of which concern the biological plausibility of the artificial systems. This has paved the way for a simple proposal for promising future work on this line of research.

References

- Andrade B. *Uniendo el aprendizaje de máquina con lo biológico: corteza ventral prefrontal y redes neuronales artificiales en tarea de detección bimodal.* (2022) Universidad Nacional Autónoma de México

- Barab'asi A.L. and Albert R. *Statistical mechanics of complex networks.* (2002) Reviews of modern physics

- Barack D. and Krakauer J. *Two views on the cognitive brain.* (2021) Nature Reviews

- Bassett D.S. and Sporns O. *Network neuroscience.* (2017) Nature Neuroscience

- Bi Z. and Zhou C. *Understanding the computation of time using neural network models.* (2020) PNAS

- Campos A. *Codificación bimodal en una población neuronal de la corteza premotora dorsal durante la memoria de trabajo.* (2022) Universidad Nacional Autónoma de México

- Carpenter R. and Reddi B. *Neurophysiology. A conceptual approach. Fifth edition.* (2013) Hodder Arnold

- Chaisangmongkon W., Swaminathan S., Freedman D., and Wang X. *Computing by robust transience: How the fronto-parietal network performs sequential, category-based decisions.* (2017) Cell Press

- DelaFuente V., and Romo R. *Neural correlate of subjective sensory experience gradually builds up across cortical areas.* (2006) PNAS

- Diaz-deLeon G., Alvarez M., Bayones L., ... and de Lafuente V. *An abstract categorical decision code in dorsal premotor cortex.* (2022) PNAS

- Douglas R., and Martin K. *Mapping the matrix: The ways of neocortex.* (2007) Neuron

- Freedman D.J. and Assad J.A. *Neuronal mechanisms of visual categorization: an abstract view on decision making.* (2016) Annu. Rev. Neuroscience

- Funahashi S., Bruce C.J., and Goldman-Rakic P.S.J. *Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex.* (1989) Neurophysiology

- Gold J.I. and Shadlen M.N. *The neural basis of decision making.* (2007) Annu. Rev. Neuroscience

- Gould and Stephen J. *The structure of evolutionary theory.* (2002) Harvard University Press

- Joglekar M., Mejias J., Yang G., and Wang X. *Interareal balanced amplification enhances signal propagation in a large-scale circuit model of the primate cortex.* (2018) Neuron

- Kay M., Miller E., Taschbach F., ... and Fusi S. Mixed selectivity: Cellular computations for complexity. (2024) Neuron

- Kandel R., Shwartz J., Jessell T., Siegelbaum S., and Hudspeth A. *Principles of neural science. Fifth edition.* (2013) McGraw Hill Medical

- Kanwisher N. *Functional specificity in the human brain: A window into the functional architecture of the mind.* (2010) PNAS

- Langdon C., Genkin M., and Engel T. *A unifying perspective on neural manifolds and circuits for cognition.* (2023) Nature Reviews Neuroscience

- Leonard N.E., Bizyaeva A., and Franci A. *Fast and flexible multi-agent decision-making.* (2023) Annual review of control, robotics and autonomous systems.

- Mante V., Sussillo D., Shenoy K., and Newsome W. *Context-dependent computation by recurrent dynamics in prefrontal cortex.* (2013) Nature

- McCulloch W. and Pitts W. *A logical calculus of ideas immanent in nervous activity.* (1943) Bulletin of Mathematical Biophysics

- Miller E.K., Erickson C.A., and Desimone R. J. *Neural mechanisms of visual working memory in prefrontal cortex of the macaque.* (1996) Neuroscience

- Munoz D.P. and Everling S. *Look away: the anti-saccade task and the voluntary control of eye movement.* (2004) Nat. Rev. Neuroscience

- Murphy K. *Probabilistic machine-learning: An Introduction* (2021) MIT Press

- Newman M., Barabasi A.L., and Watts D.J. *The structure and dynamics of networks.* (2006) Princeton University Press

- Parra S., Díaz H., Zainos A., … and Rossi-Pool R. *Hierarchical unimodal processing within the primary somatosensory cortex during a bimodal detection task.* (2022) PNAS

- Pujalte S. *Codificación de potenciales de acción utilizando reducción de dimensión no lineal.* (2022) Universidad Nacional Autónoma de México

- Raposo D., Kaufman M.T., and Churchland A.K. *A category-free neural population supports evolving demands during decision-making.* (2014) Nat. Neuroscience

- Romo R., Brody, C.D., Hernández A., and Lemus L. *Neuronal correlates of parametric working memory in the prefrontal cortex.* (1999) Nature

- Romo R. and DelaFuente V. *Conversion of sensory signals into perceptual decisions.* (2013) Progress in neurobiology

- Romo R. and Rossi-Pool R. *Turning touch into perception.* (2020) Neuron

- Rosenblatt F. *The Perceptron – a perceiving and recognizing automaton.* (1957) Cornell Aeronautical Laboratory

- Rossi-Pool R., Zainos A., Alvarez M., Zizumbo J., Vergara J., and Romo R. *Decoding a decision process in the neuronal population of dorsal premotor cortex.* (2017) Neuron

- Rossi-Pool R., Zainos A., Alvarez M., ... and Romo R. *Invariant timescale hierarchy across the cortical somatosensory network.* (2021) PNAS

- Siegel M., Buschman T.J., and Miller E.K. *Cortical information flow during flexible sensorimotor decisions.* (2015) Science

- Sussillo D., and Abbott L.F. *Generating coherent patterns of activity from chaotic neural networks.* (2009) Neuron

- Sussillo D., and Barak O. *Opening the black box: Low-dimensional dynamics in high dimensional recurrent neural networks.* (2013) Neural computation

- Trautmann E., Stavisky S., Lahiri S., ... and Shenoy K. *Accurate estimation of neural population dynamics without spike sorting.* (2019) Neuron

- Vergara J., Rivera N., Rossi-Pool R., And Romo R. *A neural parametric code for storing information of more than one sensory modality in working memory.* (2016) Neuron

- Yang G.R., Joglekar M.R., Song H.F., Newsome W.T., and Wang X.J. *Task representations in neural networks trained to perform many cognitive tasks.* (2019) Nature Neuroscience

- Yang G.R. and Wang X.J. *Artificial Neural Networks for Neuroscientists: A Primer.* (2020) Neuron