The Brain's Adaptive Response Process: Function and Anatomy

Ari Rappoport
The Hebrew University of Jerusalem, Israel
ari.rappoport@mail.huji.ac.il

August 2024

Abstract. I present a simple **adaptive response process (ARP)** through which the brain fulfills its main function, selection and release of responses. Each process stage is implemented by a different part of the brain's anatomy, providing a mechanistic explanation of the functional role of cortical layers, subcortical nuclei, excitatory vs. inhibitory neurons, and neuromodulators. The ARP explains how learning from past experience is reflected in the brain's anatomy, most notably in the basal ganglia, shows how hierarchical response sequences are generated without being explicitly represented, and provides a unified account of the notions of goals, predictions, and motor representations.

Keywords. Brain process; response selection; allow, restrain, amplify, respond (ARAR) architecture; motor control; goals; predictions; hierarchical sequences; automaticity (learning); inhibitory interneurons; language; consciousness; emotions;

1 Explaining the Brain

How the brain works is one of the most important open problems in science. The role of the brain is to manage the organism's interaction with the external environment, producing external responses to sensory inputs conveying survival threats and opportunities. It also controls internal organs to provide resources for external responses (e.g., increasing heart rate to increase available energy). Hence, the main task in explaining brain function is to explain how it selects and releases responses. The second task is to explain the roles of the various brain anatomical components (cortex, subcortical nuclei, etc.) and how they assist in the main task. Finally, a theory of brain function should show how the brain achieves its main adaptive temporal characteristic, the capacity of learning from experience.

Here I present a simple biological theory that fully addresses these tasks. The brain utilizes a single process for selecting and releasing responses, a process that I call the **adaptive response process (ARP)**. The ARP has two major subprocesses, one for selecting responses and one for releasing them. Each stage in the selection process utilizes different cortical layers and neural types, and is supported by different agents released by specialized nuclei. The release process relies on an architecture that I call "Allow, Restrain, Amplify, Respond (ARAR)" that includes the important subcortical nuclei such as the basal ganglia, habenula, and amygdala. Consequently, the ARP not only explains the brain's main task in abstract terms, but also ties each subfunction to specific anatomical structures, allowing a connection with empirical data and providing strong support for the theory. The anatomical components of each subprocess directly support learning from experience. Another attractive feature of the selection process is that it provides a unified view of the major notions of goals, predictions, and hierarchical sequences, showing how the latter emerge without explicit planning and representations.

2 The Adaptive Response Process (ARP)

2.1 Response selection

In the present paper, my discussion of response selection will focus on cortex. The same principles apply to responses generated in the spinal cord and other areas.

Selection stages. The selection part of the ARP has 4 stages: **sensory inputs**, generation of **competing response candidates (Cgen), competition resolution (Cres)**, and **execution of winning responses**, which are released via a set of Allow-Restrain-Amplify-Respond (ARAR) circuits.

The various ARP stages map into different neuronal networks distinguished by cortical layer, neuron type, and subcortical nucleus. Cortex is arranged in horizontal layers, L1 to L6. ARP selection stages map into these layers and define their functions. **L4 supports sensory inputs** (L2/3 supports inputs in areas with a negligible L4). **L2/3 supports Cgen (candidates)**, while **L5 supports responses.** These layers are referred to as the input network, the Cgen network, and the response network, respectively. L1 and L6 provide auxiliary functions.

Excitatory neurons (ENs) support inputs, Cgen, and response execution. **Inhibitory interneurons (IINs) support Cres and post-Cres execution**. Moreover, IIN subnetworks support different aspects of Cres and execution (i.e., different subnetworks for enhancing winners, suppressing losers, and supporting sustained responses). Projection (non-local) inhibitory neurons (INs) are mainly present in ARAR architecture nuceli and have a different role, see below.

The local minicolumn circuit. Most cortex is arranged in **minicolumns (MCLs)** [1] (for simplicity, I will not distinguish between MCLs, macrocolumns, and columns). These are tight neural sets spanning all cortical layers that are separated from each other by some distance. Each MCL represents a single focused entity (external and internal sensory objects, movements, etc).

The role of the ARP is to select the MCLs that represent the desired response to the current input, and execute that response.

Within a single MCL, there is a general well-defined connectivity pattern ("the canonical microcircuit") [2, 3]. Sensory inputs (mainly from the thalamus) arrive at L4 (or L2/3 when L4 is thin, and/or L5 when responses are largely automated/learned, see below). L5 is divided into two sublayers, L5a and L5b, the latter being the one driving midbrain and spinal responses. (L5a and L5b support novel/learned responses, see below.) L4 neurons excite L2/3 neurons (and L5b neurons in automaticity). L2/3 excites L5a, which excites L5b. Naturally, things are not that clean, and there are many other connections. Nonetheless, this is the main pattern. MCLs are strongly connected to each other across cortex, via L2/3/5 neurons.

There are other content representation areas besides cortex. For example, the basolateral nucleus of the amygdala (BLA) represents input and action valence, and the hippocampus represents events in space and time (see Table 1). As mentioned above, this paper focuses on cortex for clarity.

Bottom-up vs. top-down flow. The basic selection process goes as follows. Sensory inputs yield **bottom-up (feedforward)** flow across cortex. Since the input network excites the candidates network, this activates candidates (apart from state of complete automaticity). The current motivational state (needs, goals) of the organism is generally encoded in frontal areas, which generate **top-down (feedback)** flow to sculpt Cres and execution. The circuits participating in the response are modified by **plasticity** mechanisms, see automaticity/learning below.

2.2 Response release: ARAR circuits

Subcortex is arranged in a set of circuits whose goal is to provide specific input types and to release the response selected by the previous ARP stages. These circuits have a uniform architecture comprised of 4 components: **Allow, Restrain, Amplify, and Respond, at this order**. Each component is implemented by a neural nucleus (usually not by layers as in cortex). The first nucleus ('Allow') is the one receiving most inputs from the responses selected by the first ARP stage. It is comprised of projection INs that, when activated by the selected response, suppress the Restrain nucleus. The 'Restrain' nucleus contains INs that tonically suppress an 'Amplify' nucleus. The Amplify nucleus (mainly containing ENs) targets circuits that actually execute the response by driving muscles and conveying inputs to other circuits, and is capable of strong sustained firing to make sure that the selected and released 'Response' is continuously executed. Thus, activation of the Allow nucleus eventually disinhibits the Amplify nucleus to drive the response.

There are several ARAR circuits, all operating under the same principle. Virtually all of the important nuclei of the brain (besides nuclei whose role is to release specific agents such as dopamine and serotonin) participate in ARAR circuits (Table 1). ARAR circuits interact with the selection circuits. They receive response selection inputs, provide inputs to selection circuits, and interact with each other, e.g., by inhibiting other ARAR circuits to promote their own type of response.

Name	Function	Content	Allow	Restrain	Amplify	Respond
D1 BG path [4]	novel responses	all	STR D1s	GPi, SNr	thalamus	all
D2 BG path [4]	learned responses	all	STR D2s	GPe	STN	all
Extended amygdala [7]	valence, survival responses	BLA, thalamus	CeM	BNST	LHthalA	PAG, other
Cerebellum (CRB) [8]	quick or precise movements	granules	golgis	Purkinjes	deep crb nuclei	crb targets
Hippocampus [9, 10]	space time events	CA1-CA3 DG, subiculum	Lsep	Msep	Supm	nucleus incertus
Habenula [11]	switch to defend	many	STR	(some) ventral pallidum	MHb LHb	IPN, VTA
Superior colliculus [12]	saccades	all	STR	SNr	SCi	eye motor nuclei

Table 1: Summary of the main ARAR circuits. Note that many of the circuits are standalone in the sense that they have dedicated content representations in addition to the ARAR circuits. Many details (e.g., the role of the zona incerta in vigorous movements) are not shown, and some of the listing (e.g., the assigned functions of the Supm and nucleus incertus in the space circuit) should be refined. Other less central circuits having the same structure (e.g., the auditory dorsal cochlear nucleus) are not shown.
BG: basal ganglia. STR: striatum. D1s, D2s: D1, D2 dopamine receptor-expressing medium spiny neurons. GPi/GPe: internal/external glubus pallidus. STN: subthalamic nucleus. BLA/CeM: basolateral/centromedial nuclei of the amygdala. BNST: bed nucleus of the stria terminalis. LHthalA: lateral hypothalamic area. PAG: periaqueductal gray. crb: cerebellar. DG: dentate gyrus. CA1: cornu ammonis 1. Lsep/Msep: lateral/medial septum. Supm: supramammillary nucleus. LHb/MHb: lateral/medial habenula. IPN: interpeduncular nucleus. VTA: ventral tegmental area (via the rostromedial tegmental nucleus, RMTg). SCi: intermediate layers of the superior colliculus.

As an example, probably the most familiar ARAR circuit is the D1 ('direct') path of the basal ganglia (BG) [4]. It receives cortical, thalamic, amygdalar, hippocampal, and other inputs representing the selected response. Its Allow function is mediated by D1-expressing striatal medium spiny neurons (MSNs), which are projection INs. When activated, these suppress the Restrain nuclei, the internal globus pallidus (GPi) and substantia nigra (SNi). These tonically inhibit the thalamus, which is disinhibited when the selected response is released. The thalamus is a powerful Amplifier that sustains responses via repeated firing and widespread connections [5]. It targets both subcortical (via its CM/Pf nuclei [6]) and cortical circuits to drive the response.

2.3 Novel vs. Automated (Learned) Responses

Learning from past experience is important for the organism's survival. Following response selection and release, the circuits participating in response execution undergo plasticity processes, mainly generation of new synapses and modification of existing ones to make them more stable

and make their usage more efficient in terms of time and energy [13].

Technically, the difference between novel and familiar inputs is that learned states involve preceding partial excitation of neurons predicted to participate in the response. Such predictions are a natural consequence of the enhancement of synapses that have already been used under similar conditions. When neurons that are not partially excited receive novel firing-inducing inputs, they fire strong action potential bursts [14], such that the effect of a surprise is much larger than the effect of a predicted response.

In the ARP, novel inputs induce a stronger bottom-up flow and a larger number of candidates. Conversely, fully automated responses yield a direct activation of the response network by the input network. L2/3 excites L5a but not L5b, L5a is the one sustaining novel responses. The excitation pattern L2/3 -> L5a -> L5b may be needed because the alternative pattern, L2/3 -> L5b, would have caused immediate L5b firing for novel responses. The presence of L5a provides more time for top-down flow to guide inter-candidate competition (Cres).

Acute inputs generate flow in many MCLs (because no MCL is predicted, and they are strongly interconnected), exciting their L2/3 neurons. These MCLs are the response candidates. This stage is called candidate generation (Cgen).

The candidate set includes all MCLs that might be somehow relevant to the selected response, by being the ones that receive the current sensory input or being connected to them.

This means that the next time that similar inputs arrive, there would be fewer candidates activated (because IINs would suppress them faster, see the next section), ENs would be more synchronized, new synapses going directly from the thalamus to L5b might be generated, etc. Automated responses are more energy-efficient than acute ones, are quicker (due to skipping most Cres), and are more stable due to higher synchronization.

2.4 Agents

The brain utilizes a large number of agents to calibrate its processes. In the present context, it is imperative to mention 3 of them, which directly support the different ARP stages. **Norepinephrine** is released during sensory inputs to recruit energy and cognitive resources needed to deal with the new input [15]. **Dopamine** is released during novel responses, including 'charge' (sex, predation, 'reward') actions but also strong defensive fight (aversive situations) [16]. **Acetylcholine** [17] works with IINs to support both Cres and post-Cres execution.

Thus, the tight correspondence between the brain's function and anatomy also extends to agent-releasing nuclei, in addition to content and ARAR circuits. We do not have the space to deal with other important agents (e.g., orexin and oxytocin) here. I have detailed general theories of opioids, endocannabinoids and CRH in my papers presenting theories of depression (and bipolar disorder), schizophrenia, and autism, respectively [18, 19, 20]. I have also detailed a large number of plasticity-related agents in the context of Alzheimer's Disease [13]. Readers are welcome to consult these papers.

A fourth agent, **serotonin (5-HT)** is usually mentioned along with the other 3 agents. In the theory presented here, the basic role of serotonin is to deal with situations of excess (food,

warmth, hyperoxia), e.g., by promoting gut food absorption (about 90% of body serotonin is located in the gut) and glucose uptake by adipose and other tissue [21]. Since dealing with excess is in many ways opposite to the roles of the other 3 agents, which deal with deficiency-induced stress, serotonin can oppose them [22], a capacity that the brain has adapted for additional purposes. For example, in the habenular ARAR circuit, serotonin, via opposition to dopamine, can put a strong brake on vigorous charging actions, effectively switching from charge to strong passive defence.

3 IINs: Competition Resolution (Cres) & Execution

If the situation is not that urgent to handle, MCLs **compete** until a response emerges [23, 24]. Competition is implemented by IINs [25, 26, 24]. It is crucial to understand that although IINs are technically "inhibitory" (even this is not always the case, since they are initially excitatory in many areas), their overall role is NOT to inhibit responses. Their role is to synchronize the firing of their EN targets, which includes suppression of losing candidates but also enhancement of winning ones. When an IIN fires onto two ENs that have just fired, it does not inhibit their firing but ensures that they would start charging at the same time, and if they receive comparable inputs, this increases their synchronized firing. IINs inhibit ENs when they fire onto them when they are partially excited, preparing to fire. Thus, **IINs suppress most ENs, but enhance the coordinated firing of the ENs that are most temporally synchronized with the input**. Assuming that IINs are always inhibitory is a big mistake. A good theory-independent term for their role is **coordination**.

To allow their Cres function, IINs receive sensory inputs from the thalamus, from local MCL ENs, and from other IINs. Since IINs are small neurons, they are quick to fire, allowing quick Cres in automated situations. IINs keep their activity to support post-Cres execution, because they need to maintain the winning circuits.

In the MCLs that have won the competition, sustained input, sustained top-down flow from long-range connections, and/or sustained L4 firing (which may be the role of L4), sustain the firing of L2/3 neurons, which in turn excite L5a neurons, which excite L5b neurons to generate the response.

IINs can be classified into many subsets, based on location, connectivity, and agent expression [27]. For our purposes here, it is beneficial to highlight 3 such subsets. First, so-called Martinotti IINs are located in superficial layers and target apical dendrites. They are thus natural candidates to mediate the suppressing aspect of Cres. Second, VIP-expressing IINs located in L1 and L2/3 locally target the Martinotti IINs, thus supporting the winner-promotion aspect of Cres. Finally, parvalbumin-expressing IINs, located in most layers but mostly in the response layers, support post-Cres execution.

4 Goals and Predictions

Suppose that an MCL has won the competition, and its L5b neurons drive the current response. As a general rule, responses involve external movement. How does the brain know that the movement has attained its purpose so that the muscle contraction that drives it should be stopped? This is a crucial question for understanding motor control, and yet it is surprisingly not asked that often.

Our answer to this question is as follows. When L5b drive responses, L5a or L5b neurons in the same MCL also excite L5a or L5b neurons in other MCLs (L5a and L5b are used in novel and learned responses, respectively). **When a response is executed by an MCL, its L5 neurons excite the L5 neurons of the MCLs that represent the desired sensory consequences of the executed response (the goals)**. For example, if the executed response is to extend your arm in order to reach an apple, then the goal MCLs represent the state in which your fingers touch the apple (or get close to it). These kinds of physical configurations (a limb getting in contact with an external object) are generally represented in the parietal cortex [28].

Why is this technique useful? When the response goal has been met (reaching the apple), sensory inputs encoding this state excite the MCLs that represent it. They also excite other MCLs. However, the L2/3/5 neurons of the "correct" MCLs are already partially excited (by the response MCLs), which gives them a head advantage in winning the competition.

When the sensory inputs denoting the event arrive, the L5b neurons of the goal MCL are activated. This is a type of response. Unlike motor responses, which generate movement, this response indicates that a particular event has occurred in the real world. In other words, it indicates the fulfillment of a **prediction**. When this happens, the activated L5b neurons induce a response that suppresses the movement of the effector (arm) that participates in this MCL's representation.

Note that what the activation of motor MCLs represents are **target sensory configurations**, not just direct muscle contractions [29], plus the sum of external forces needed to sustain the desired configuration [30].

Note also that **the notions of goals and predictions are identical**. Instead of saying that the MCL's L5 neurons represent the goal of the movement, we can say that they represent a prediction for some sensory event (and vice versa).

5 The Basal Ganglia and the Cerebellum

BG D2 path. Above, I gave the BG D1 (direct) circuit as a well-known example of an ARAR circuit. Here I devote some space to the circuit known as the D2 (indirect) path of the BG, because its description in virtually all of the scientific literature is simply **wrong** [31]. Its standard description is that D2-expressing MSNs of the striatum suppress the external globus pallidus (GPe), which suppresses the subthalamic nucleus (STN), which excites the BG output nuclei GPi and SNi. The net result is thought to suppress the direct path and movement.

This is not an ARAR circuit, because STN projection neurons are ENs, and they supposedly project to Restrain nuclei, not to a Respond nucleus. **In the present theory, the correct arrangement of this circuit is an Allow nucleus (D2 MSNs) that targets a Restrain nucleus (GPe), which releases an Amplify nucleus (STN) that drives responses.** This conforms with the obvious anatomical data, with the STN taking the Amplify role of the thalamus in the D1 path, while GPe takes that of GPi. **Unlike the BG's D1 path, which supports vigorous "charging" movements (e.g., towards rewards, or for active defence), the role of the D2 circuit is to support learned movement**.

Indeed, the STN projects to brainstem locomotion circuits to produce movement [32], most STN neurons show locomotion-dependent activity increase, and brief optogenetic inhibition slows, dysregulates, and terminates locomotion [33]. Moreover, it is well known that the D2 path is the main learner of habits [34], and that it cooperates with the CRB in the execution of precise, learned movements [35]. Thus, my account here is completely opposite to the standard one, where the D2 path's role is 'stop'. As part of its role, this circuit indeed suppresses the direct BG circuit, but this is in order to prevent vigorous movements from competing with learned ones. Its main role is automated/learned movements. I will not expand upon this point any further in the present paper, deferring it to a detailed discussion of ARAR circuits.

Note that there is **redundancy** in the BG ARAR circuit (and possibly in all ARAR circuits). The number of inputs is larger than the number of neurons in the Allow nucleus, which could be larger than the number of neurons in the Restrain nucleus. This means that the Allowed muscles are not fully determined by the inputs. In my opinion, the ARAR circuit determines a superset of muscles that possibly participate in the response, but the actual muscles participating is selected by sensosry inputs. This is what lets the brain adapt so rapidly to external changes but still maintain its goals.

The BG is the ARAR circuit that controls the release of cortical goal-directed responses, novel and learned. The CRB provides both precision, enabled by its huge set of input content space (granules), and speed (vs. the BG). It is not specialized for kinematics and dynamics (when compared to the BG). The BG can do both, but less accurately and more slowly.

6 Hierarchical Sequences (HSEQs)

A major issue in motor control is how the brain represents action sequences, including hierarchical ones. The predominant notion is that of optimal control with explicit planning of hierarchical sequences [36]. However, the number of possible meaningful sequences is so large that it is not clear whether the brain has enough space to represent all of them explicitly. Moreover, the brain extremely rapidly recuperates from minor and major obstacles and changes during execution, and it is not clear how this can be done with explicit top-down representations.

Here I present an almost trivial account of HSEQs. There are four keys to this account. First, the fact that motor responses are phrased at the level of goal configurations, not as direct muscle contractions. Second, the fact that the ARP starts with and is guided by external and

proprioceptive sensory inputs. Third, the fact that frontal areas continuously receive interoceptive inputs denoting the internal needs of the organism. Finally, the relatively small number of ethologically meaningful high-level sequences, which are learned during early development.

Suppose that the organism needs to eat, and that there is an apple reachable by arm. The visual inputs denoting the apple trigger an ARP, whose winners include frontal nodes activated by the current hunger (because they are partially excited by it so they have an advantage in winning response selection). These frontal nodes receive valence input (based on previous experience) indicating that the apple can satisfy hunger. Thus, the winners include motivational frontal nodes, motor nodes whose responses drive the arm to reach the apple, and sensory nodes representing the fingers touching the apple (or just getting close to it). (Clearly, there should be additional nodes involved, whose identity is a good open question.) Now suppose that someone has moved the apple to the right, or has introduced an obstacle preventing the arm from reaching it. This yields new sensory inputs. Our point is that **there is no need for a renewed top-down plan to reach the apple**. The new inputs simply trigger a new ARP with the same motivation. This ARP yields Cgen, Cres, and probably a new goal winner (because the goal configuration of the arm could be different). This new ARP is resolved just like the previous one, so the brain fixes the problem on-the-fly without any difficulty.

The next response in the sequence is to grab the apple with the fingers, followed by taking it into the mouth. Why are these actions executed? A large part of development time is devoted to acquiring action plans that satisfy needs. In abstract terms, there aren't that many of them, and 'reach, grab, put in mouth' is a basic one. Given the hunger, the nodes that represent grabbing the apple, and those that represent taking it into the mouth (and chewing it), are partially excited by these schemes learned via previous experience, so they win the competition again (unless something more urgent comes up).

In summary, every new sensory input can trigger a new response. Selection of the response is guided by needs and opportunities. The actual movements done in order to execute the response are mainly a function of the existing network structure, which has been learned from past experience. Motor control is definitely not 'optimal', because past experience may have been non-optimal (e.g., early training may have been done with tied arms). It is only optimal from the point of view of the brain, in the sense that if a response has won, it had to win.

7 Additional Evidence

The evidence supporting the ARP is vast. Here I briefly mention some lines of evidence with some salient references.

Efference copy. A notion that is very close to goals and predictions is the old notion of efference copy (of action commands) [37]. It has been known for a long time that when an action executes, it sends links to sensory nodes. This has been thought to be for suppressing the effects of self-generated movements in order that they do not interfere with required responses to sensory inputs. In the present theory, **a main role of efference copy is to declare the sensory**

consequences of actions as the goals of the action and prepare the brain for their occurrence. The effect of this 'copy of an action command' is to partially excite the target, which causes its activation when the event occurs to be smaller than when it is not predicted (as in automaticity). Thus, the efference copy seems to suppress the action's consequences, while in fact it prepares the brain for their efficient activation. It has been shown that it is more likely that the efference copy specifies the sensory goal rather than a copy motor command [38].

It is also possible that some efference copies actually suppress competitors, supporting the full range of goals and predictions during the ARP [39].

Goals and posterior parietal cortex. Supporting our view that goal nodes suppress the current response when its goal has been attained, posterior parietal cortex (PPC) stimulation can generate finger and wrist movement (presumably of the antagonist muscles) [40]. Moreover, PPC firing rates rise before contact, peak at contact, and decline when grasp is secure [41]. It has also been shown that PPC encodes visual goals related to future movements [42].

Place cells. Place cells are neurons that fire in specific spatial locations that the rodent (usually) reaches [43]. They start firing a few cycles earlier, and exhibit interesting properties such as theta phase precession. Place cells provide strong evidence for activation of action goals during execution. The other famous spatial cell type, grid cells, might be goals expressed in internal-limb space.

Goal configurations. An extremely strong line of evidence supporting goals is given by the fact that motor cortex represents goal configurations rather than direct muscle contractions [29].

Event-related potentials (ERPs). The first sensory-induced ERPs are negative, and they then turn positive [44]. Negative ERPs indicate Cgen (activation of the superficial L2/3), while positive ERPs indicate Cres and responses.

BOLD. The central technique of showing brain activity via fMRI BOLD reflects the Cgen stage (BOLD is most salient in the superficial networks).

Sensory flow. There are different feedforward and feedback bidirectional streams, occurring in the superficial (Cgen) and deep (response) networks, respectively. The former flow is stronger and is more point focused [45].

Sensory input drives motor control. There is a wide line of evidence for sensory inputs being essential for all stages of response selection, execution, and learning, showing that top-down planning does not suffice (or does not exist). Sensory feedback is crucial for trained finger sequences [46]. Spinal stretch reflexes support efficient corrective responses, and these depend on goals, not just on muscle stretch [47]. Most S1 neurons induce movements that combine distinct body parts [48]. Rodent barrell cortex controls motor responses directly, like M1 [49]. S1 activity correlates with locomotion. Stimulating S1 increases speed, inhibition slows speed and terminates [50]. Supporting the ARP, new input is immediately reflected in movement strategies [51].

Proprioception (sensory input reporting limb and muscle state) is essential for locomotion

[52]. Movement is still possible with lesions impairing proprioception, but only with a lot of training in visually-defined goals [53].

MCL competition. Inter-MCL lateral inhibition has been demonstrated [23]. In addition, the well-known rich line of research of **surround suppression** supports inter-MCL competition [24]. Strong simultaneous stimulation of 4 pyramidal neurons suppressed all pyramides at a MCL distance [26].

Competition between selected and non-selected actions has been demonstrated to be constrained by competitor similarity, showing competition between relatively close alternatives [54]. More excitable neurons win the competition to be included in the response [55].

IIN competition. IINs clearly mediate competition. Martinotti IINs have been shown to mediate inhibitory competition [25]. High frequency (>70Hz) stimulation of a single pyramidal neuron suppresses others via Martinotti IINs, with suppression of neighboring MCLs [26]. VIP IINs 'poke holes' in the blanket of IIN inhibition, at roughly the radius between MCLs, to promote winners [56]. Supporting our view that they continue to support the winning response, a GABAA receptor antagonist in medial prefrontal cortex impaired effortful execution [57].

Mirror neurons (MNs). MNs are neurons that fire both when the animal performs a certain action and when it sees others perform it, and are claimed to identify an action's intention [58]. They are quite widespread. MNs support the ARP, which interprets them as candidates activated by sensory inputs resulting from the action. Due to the sensory similarity between the two states, activation of overlapping candidates is to be expected. Note that MNs respond to acting on geometric shapes without any high level intentions involved [59].

The attentional blink (AB). In the AB, during rapid presentation of sequences of sensory targets, subjects usually do not report seeing the second of a successive pair of pre-specified targets [60]. The sensory representations of pre-specified targets are activated in goal mode. According to the ARP, during the identification of the first target (i.e., goal attainment), the MCLs representing the goal are suppressed. If the second target arrives quickly, higher level task nodes may fail to reactivate the goal, in which case it would be completely missed when it occurs.

Repetition blindness is a related paradigm in which people cannot keep track of how many times a target appears in a rapid sequence.

8 Discussion

This paper presented a relatively simple mechanistic biological theory that explains how the brain fulfills its main function, response selection and execution. As such, it provides a large step forward towards a full understanding of the brain.

8.1 The explanatory power and biological logic of the theory

Given the mystery associated with the brain in all of the scientific, philosophical, and popular literature, it might be difficult to accept that it can be simply explained. However, the ARP as described in this paper provides a complete account of the major aspects of how the brain works.

The brain is simply another body tissue, its role being to select and release responses that interact with the external environment. It does so via a biological process that organisms use elsewhere as well. For example, the adaptive immune system also receives inputs (presented antigens), generates candidates (clonal expansion), its candidates compete for exit from lymph nodes [61], and the winners execute the response and establish long-term plasticity (memory B and T cells). Even more obviously, Cgen followed by Cres is the main principle of the theory of evolution.

8.2 Function and anatomy

Perhaps the main feature of the ARP that makes it a convincing biological theory is the elegant connection that it exposes between the brain's anatomy and its function. At a coarse level, all subcortical ARAR nuclei are arranged in the same architecture that has the same behavior. Each of them supports a different type of response or specific content, but their operating principles are identical. Similarly, the different stages of the response selection part of the ARP all use different cortical layers and different IIN networks, and are supported by different agents released by agent-releasing nuclei. Thus, the present theory provides a simple unified account of both the brain's function and of its anatomy. Naturally, this function-anatomy link enables the design and interpretation of empirical experiments to refine the theory.

8.3 Other theories; contribution

There is a very rich literature on brain theories. All of the notions we discussed have been previously discussed, including action-perception integration [62], sophisticated models for motor control [30], action goals, efference copy, cortical networks, subcortical nuclei, etc (see references above). Nonetheless, I am not aware of any theory that has both the explanatory power of the present one, and its clean description of the link between the brain's function and anatomy.

In the recently popular notion of 'predictive coding', the brain adapts to minimize the surprise conveyed by inputs by generating predictions [3]. This agrees with some aspects of this paper, but the present theory is much wider and complete in that it describes a mechanistic process, ties it to anatomy, and explains better the role of predictions and goals.

The initial idea of an ARAR-like architecture has been raised by Swanson, who focused on the extended amygdala [7]. Almost all of my analysis is different from his, sometimes contradictory (e.g., the BG D2 path). "CRB-like structures" such as the dorsal cochlear nucleus have been discussed [63], mainly in fish and mammals.

In summary, the contribution of the ARP over existing notions is as follows. First, the explicit phrasing of ARP stages (Cgen, Cres) unifies many important notions (feedforward vs. feedback, superficial vs. deep layers) using a very simple description. Second, the identification of the correspondence between anatomy and function provides a completely mechanistic account of brain function that is grounded by empirical data that can only be gathered (at least today) via anatomy-related means. 'Anatomy' refers here not just to the role of cortical areas, but to the function of cortical layers (e.g., the difference between L5a and L5b) and subnetworks of inhibitory interneurons. Third, an account of motor control that unifies the notions of goals, predictions and efference copy, and shows how hierarchical sequences are achieved without explicit representations. Finally, the identification of a detailed ARAR architecture, which unifies the anatomy and function of almost all subcortical nuclei.

8.4 Action execution, observation, imagery

The focus of this paper is on action execution. Action observation is a simple special case in which the animal does not generate movements, only watches what happens. During watching, motivation and sensory inputs keep arriving, which means that there is an on-going ARP. As explained above, goals and predictions are one and the same, and during observation these are better interpreted as predictions. Thus, it is not surprising that observation has large activity overlap with execution, and that most overlap occurs in relatively high-level areas [64] (see also mirror neurons above).

Imagery (which can be equated with 'thinking') is less immediate to explain. When we imagine a triangle, it has an abstract quality that is definitely devoid of the features of sensory inputs. We should first note that humans are not the only species capable of imagery, as rats simulate spatial paths before executing them [65]. Thus, imagery seems to be a universal feature of brains (at least in some species).

The explanation for imagery is strongly tied to the issue of representation. Cortex has a vast space for representing objects, with strong evidence for visual objects such as faces and letters [66] representation in ventral temporal cortex, and integrated object representations at the anterior temporal lobe (ATL) [67] (my view is that the ATL represents knowledge involving language). These representations provide an abstraction over specific objects conveyed by sensory inputs. Assume for simplicity that 'triangle' is represented by a single MCL (an open issue). To explain imagery, we can make the leap to assume that frontal circuits can activate these object representations without supporting sensory input. This does not contradict the ARP, because these object nodes still receive other inputs. If this account is correct, then external sensory inputs may occur if and only if there are external responses, while thinking/imagery use an ARP without external sensory inputs. It would definitely use interoceptive sensory inputs, because these are the ones that provide the motivation for the whole process.

Humans most likely have a larger imagery capacity than other animals. Although the differences between the human and non-human primate brain are a matter of debate [68], humans clearly have a larger temporal lobe with faster bundles of axons leading to it from frontal circuits

[69, 70, 71], and they have more connections in the frontal candidate network (L2/3), which can generate powerful flow throughout the brain [72].

A very interesting question related to imagery is how we humans know that we are imagining, as opposed to perceiving reality. We do not have the space to discuss this here.

8.5 Language

Language is yet another topic widely considered as mysterious. I think that language is very easy to explain. The role of the brain is to manage the organism's interaction with the external environment. This can be done using your own muscles, or via others' muscles. Language (and communication in general, using hands, eyes, etc) is a way of generating responses that use others' muscles. Since our brains are not connected to theirs, a way of overcoming the neural gap has emerged, and this is language.

What allows human language is 'hardware' for producing and perceiving sounds (with humans at least, sound is the quickest medium), and the evolution of a cortical content space to store linguistic representations (object-like representations for words in Wernicke's area in the temporal lobe, the lateral inferior frontal cortex, with higher level action representations for word sequences, and lateral motor areas for control over speech organs). Everything else about language is managed by the ARP as described here. The often cited capacity for an 'infinite number of sentences' simply parallels the capacity for an infinite number of movement sequences. Obviously, the existence of language dramatically enhances the human capacity for internal movement-less thoughts and enables the rich human culture, so there are endless levels of richness built upon language. However, the biological basis of language is covered by the ARP.

8.6 Fully explaining the brain

There are several topics that have not been touched upon here and that are required for a full account of the brain. These include memory recall vs. imagining things that never happened (people normally distinguish between the two, and the mechanism for this needs to be explained), familiarity vs. precise recollection, the precise nature of representations (e.g., is 'me' simply yet another notion that we learn through language and represent as other language-acquired notions, or it is represented by animals as a single notion even without language?), the nature of the representations in the various cortical areas (e.g., the claustrum), the precise representation of time (assuming that space is quite well-understood by now, although there are clearly a host of less understood details), the function of many agents (i.e., neurotensin, CCK), volition, the detailed structure of medullar, pontine and spinal circuits (these operate according to the ARP, but are not layered), the nature of the communication between the two hemispheres, striatum matrix vs. striosomes, oscillations, music, working memory, layer 6, visual stability, etc.

No body tissue is fully understood. In my opinion, the present theory provides the important basis of how the brain works, which puts its on par with other tissue and allows decades of more

detailed work.

8.7 Consciousness and emotions

Consciousness is yet another mysterious and attractive topic that has generated some highly creative accounts [73]. In my view, consciousness is not difficult to explain. What we call consciousness is an internal report that we make of some sensory state (an idea that has already been raised [73].) Being conscious of being alive is most likely simply a report that heart beat inputs continue reaching the insula, as attested by reports of people with lesions in this area (e.g., with Cotard's syndrome) that they feel dead. Awareness of external objects is a report that a response to perceiving that object has been made. In addition to the mechanisms already described above, I hypothesize that such awareness may involve a switch from the sympathetic nervous system (which conveys alerts to new inputs) to the parasympathetic nervous system (and/or the serotonergic system).

As to the report mechanisms, we humans have language, and language areas receive continuous inputs from all areas of the brain, allowing them to actively generate reports. Whether other animals have a 'language' that allows them reports that yield consciousness that is similar in some sense to ours is an open question. However, note that the fact that animals can produce responses that look like ours is not an indication. We, and dogs, can make very similar faces in response to similar external situations, but these are the responses themselves, not reports. What generates our internal feeling of consciousness are reports, not responses.

In the same vein, emotions are reports that we make to ourselves about our basic survival state. For example, in my theory of depression, feeling dejected and sad occurs when we report that the survival response 'disengage, withdraw' has been executed in the face of stress that we do not know how to resolve [18].

Author's Comment

Most of the ideas in this paper have been described in a previous paper I wrote [74, 75] (long and short versions respectively). However, these papers were not so accessible. I wrote the present paper completely from scratch, and it should be viewed as the currently definitive version of my brain theory.

References

[1] Mountcastle VB. The columnar organization of the neocortex. Brain. 1997;120(4):701-22.

[2] Morishima M, Kawaguchi Y. Recurrent connection patterns of corticostriatal pyramidal cells in frontal cortex. Journal of Neuroscience. 2006;26(16):4394-405.

[3] Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. Canonical microcircuits for predictive coding. Neuron. 2012;76(4):695-711.

[4] Graybiel A, Mink J. The basal ganglia and cognition. In: Gazzaniga M, editor. The Cognitive Neurosciences IV. MIT Press, Cambridge, MA; 2009. p. 565-85.

[5] Cruikshank SJ, Ahmed OJ, Stevens TR, Patrick SL, Gonzalez AN, Elmaleh M, et al. Thalamic control of layer 1 circuits in prefrontal cortex. J Neurosci. 2012;32(49):17813-23.

[6] Smith Y, Galvan A, Ellender TJ, Doig N, Villalba RM, Huerta-Ocampo I, et al. The thalamostriatal system in normal and diseased states. Frontiers in systems neuroscience. 2014;8.

[7] Swanson LW. Cerebral hemisphere regulation of motivated behavior. Brain research. 2000;886(1-2):113-64.

[8] Prati JM, Pontes-Silva A, Gianlorenço ACL. The cerebellum and its connections to other brain structures involved in motor and non-motor functions: a comprehensive review. Behavioural Brain Research. 2024:114933.

[9] Tsanov M. Differential and complementary roles of medial and lateral septum in the orchestration of limbic oscillations and signal integration. European Journal of Neuroscience. 2018;48(8):2783-94.

[10] Martínez-Bellver S, Cervera-Ferri A, Martínez-Ricós J, Ruiz-Torner A, Luque-García A, Blasco-Serra A, et al. Regular theta-firing neurons in the nucleus incertus during sustained hippocampal activation. European Journal of Neuroscience. 2015;41(8):1049-67.

[11] Benarroch EE. Habenula Recently recognized functions and potential clinical relevance. Neurology. 2015;85(11):992-1000.

[12] White BJ, Munoz DP. The superior colliculus. The Oxford Handbook of Eye Movements. 2011;5(4):195.

[13] Rappoport A. A Lipid-Raft Theory of Alzheimer's Disease. Ann Rev Biochem. 2025;94.

[14] Cheong E, Shin HS. T-type Ca2+ channels in normal and abnormal brain functions. Physiol Rev. 2013;93(3):961-92.

[15] Szabadi E. Functional neuroanatomy of the central noradrenergic system. Journal of psychopharmacology. 2013;27(8):659-93.

[16] Klein MO, Battagello DS, Cardoso AR, Hauser DN, Bittencourt JC, Correa RG. Dopamine: functions, signaling, and association with neurological diseases. Cellular and molecular neurobiology. 2019;39(1):31-59.

[17] Muñoz W, Rudy B. Spatiotemporal specificity in cholinergic control of neocortical function. Curr Opin Neurobiol. 2014;26:149-60.

[18] Rappoport A. A Dynorphin Theory of Depression and Bipolar Disorder. arXiv preprint arXiv:240806763. 2024.

[19] Rappoport A. A Polyunsaturated Fatty Acid (PUFA) Theory of Schizophrenia. arXiv preprint arXiv:240806794. 2024.

[20] Rappoport A. A CRH Theory of Autism Spectrum Disorder. arXiv preprint arXiv:240806750. 2024.

[21] Yabut JM, Crane JD, Green AE, Keating DJ, Khan WI, Steinberg GR. Emerging roles for serotonin in regulating metabolism: new implications for an ancient molecule. Endocrine reviews. 2019;40(4):1092-107.

[22] Fink KB, Göthert M. 5-HT receptor regulation of neurotransmitter release. Pharmacol Rev. 2007;59(4):360-417.

[23] Hirata Y, Sawaguchi T. Functional columns in the primate prefrontal cortex revealed by optical imaging in vitro. Neuroscience research. 2008;61(1):1-10.

[24] Angelucci A, Bijanzadeh M, Nurminen L, Federer F, Merlin S, Bressloff PC. Circuits and mechanisms for surround modulation in visual cortex. Annual review of neuroscience. 2017;40(1):425-51.

[25] Silberberg G, Markram H. Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. Neuron. 2007;53(5):735-46.

[26] Berger TK, Perin R, Silberberg G, Markram H. Frequency-dependent disynaptic inhibition in the pyramidal network: a ubiquitous pathway in the developing rat neocortex. The Journal of physiology. 2009;587(22):5411-25.

[27] Tremblay R, Lee S, Rudy B. GABAergic interneurons in the neocortex: from cellular properties to circuits. Neuron. 2016;91(2):260-92.

[28] Borra E, Gerbella M, Rozzi S, Luppino G. The macaque lateral grasping network: a neural substrate for generating purposeful hand actions. Neuroscience & Biobehavioral Reviews. 2017;75:65-90.

[29] Graziano MS. Ethological action maps: a paradigm shift for the motor cortex. Trends Cogn Sci. 2016;20(2):121-32.

[30] Feldman AG, et al. Referent control of action and perception. Springer; 2015.

[31] Rappoport A. Basal Ganglia Paths Support Acute vs. Automated Execution, Not Movement vs. Stopping. doi: 10.20944/preprints201808.0157. preprtingsorg. 2019.

[32] Roseberry TK, Lee AM, Lalive AL, Wilbrecht L, Bonci A, Kreitzer AC. Cell-Type-Specific Control of Brainstem Locomotor Circuits by Basal Ganglia. Cell. 2016;164(3):526-37.

[33] Callahan JW, Morales JC, Atherton JF, Wang D, Kostic S, Bevan MD. Movement-related increases in subthalamic activity optimize locomotion. Cell reports. 2024;43(8).

[34] Shan Q, Christie MJ, Balleine BW. Plasticity in striatopallidal projection neurons mediates the acquisition of habitual actions. Eur J Neurosci. 2015;42(4):2097-104.

[35] Bostan AC, Dum RP, Strick PL. Cerebellar networks with the cerebral cortex and basal ganglia. Trends Cogn Sci. 2013;17(5):241-54.

[36] Wong AL, Haith AM, Krakauer JW. Motor planning. The Neuroscientist. 2015;21(4):385-98.

[37] Straka H, Simmers J, Chagnaud BP. A new perspective on predictive motor signaling. Current Biology. 2018;28(5):R232-43.

[38] Niziolek CA, Nagarajan SS, Houde JF. What does motor efference copy represent? Evidence from speech production. J Neurosci. 2013;33(41):16110-6.

[39] Greenlee JD, Jackson AW, Chen F, Larson CR, Oya H, Kawasaki H, et al. Human auditory cortical activation during self-vocalization. PloS one. 2011;6(3):e14744.

[40] Rathelot JA, Dum RP, Strick PL. Posterior parietal cortex contains a command apparatus for hand movements. Proc Natl Acad Sci U S A. 2017;114(16):4255-60.

[41] Gardner EP, Babu KS, Reitzen SD, Ghosh S, Brown AS, Chen J, et al. Neurophysiology of prehension. I. Posterior parietal cortex and object-oriented hand behaviors. Journal of neurophysiology. 2007;97(1):387-406.

[42] Kuang S, Morel P, Gail A. Planning movements in visual and physical space in monkey posterior parietal cortex. Cerebral cortex. 2016;26(2):731-47.

[43] Sosa M, Gillespie A, Frank L. Neural Activity Patterns Underlying Spatial Coding in the Hippocampus. Current Topics in Behavioral Neurosciences. 2018;37:43-100.

[44] Luck SJ, Kappenman ES. The Oxford handbook of event-related potential components. Oxford university press; 2012.

[45] Markov NT, Vezoli J, Chameau P, Falchier A, Quilodran R, Huissoud C, et al. Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. Journal of Comparative Neurology. 2014;522(1):225-59.

[46] Popp NJ, Hernandez-Castillo CR, Gribble PL, Diedrichsen J. The role of feedback in the production of skilled finger sequences. Journal of Neurophysiology. 2022;127(4):829-39.

[47] Weiler J, Gribble PL, Pruszynski JA. Spinal stretch reflexes support efficient control of reaching. Journal of Neurophysiology. 2021.

[48] Halley AC, Baldwin MK, Cooke DF, Englund M, Krubitzer L. Distributed motor control of limb movements in rat motor and somatosensory cortex: the sensorimotor amalgam revisited. Cerebral Cortex. 2020;30(12):6296-312.

[49] Matyas F, Sreenivasan V, Marbach F, Wacongne C, Barsy B, Mateo C, et al. Motor control by sensory cortex. Science. 2010;330(6008):1240-3.

[50] Karadimas SK, Satkunendrarajah K, Laliberte AM, Ringuette D, Weisspapir I, Li L, et al. Sensory cortical control of movement. Nature neuroscience. 2020;23(1):75-84.

[51] Vargas-Irwin CE, Franquemont L, Black MJ, Donoghue JP. Linking objects to actions: encoding of target object and grasping strategy in primate ventral premotor cortex. Journal of Neuroscience. 2015;35(30):10888-97.

[52] Formento E, Minassian K, Wagner F, Mignardot JB, Le Goff-Mignardot CG, Rowald A, et al. Electrical spinal cord stimulation must preserve proprioception to enable locomotion in humans with spinal cord injury. Nat Neurosci. 2018;21(12):1728.

[53] Cole J, Paillard J. Living without touch and peripheral information about body position and movement: Studies with deafferented subjects. The body and the self. 1995:245-66.

[54] Labruna L, Lebon F, Duque J, Klein PA, Cazares C, Ivry RB. Generic inhibition of the selected movement and constrained inhibition of nonselected movements during response preparation. Journal of cognitive neuroscience. 2014;26(2):269-78.

[55] Josselyn SA, Frankland PW. Memory allocation: mechanisms and function. Annu Rev Neurosci. 2018;41:389-413.

[56] Karnani MM, Jackson J, Ayzenshtat I, Sichani AH, Manoocheri K, Kim S, et al. Opening holes in the blanket of inhibition: localized lateral disinhibition by VIP interneurons. J Neurosci. 2016;36(12):3471-80.

[57] Piantadosi PT, Khayambashi S, Schluter MG, Kutarna A, Floresco SB. Perturbations in reward-related decision-making induced by reduced prefrontal cortical GABA transmission: relevance for psychiatric disorders. Neuropharmacology. 2016;101:279-90.

[58] Rizzolatti G, Sinigaglia C. Mirrors in the brain: How our minds share actions and emotions. Oxford University Press, USA; 2008.

[59] Press C, Catmur C, Cook R, Widmann H, Heyes C, Bird G. fMRI evidence of 'mirror'responses to geometric shapes. PloS one. 2012;7(12):e51934.

[60] Kawahara JI, Kumada T, Di Lollo V. The attentional blink is governed by a temporary loss of control. Psychonomic Bulletin & Review. 2006;13(5):886-90.

[61] Victora GD, Schwickert TA, Fooksman DR, Kamphorst AO, Meyer-Hermann M, Dustin ML, et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. Cell. 2010;143(4):592-605.

[62] Ridderinkhof KR. Neurocognitive mechanisms of perception–action coordination: A review and theoretical integration. Neuroscience & Biobehavioral Reviews. 2014;46:3-29.

[63] Bell CC, Han V, Sawtell NB. Cerebellum-like structures and their implications for cerebellar function. Annu Rev Neurosci. 2008;31(1):1-24.

[64] Rueschemeyer SA, Ekman M, van Ackeren M, Kilner J. Observing, performing, and understanding actions: Revisiting the role of cortical motor areas in processing of action words. Journal of Cognitive Neuroscience. 2014;26(8):1644-53.

[65] Pfeiffer BE, Foster DJ. Hippocampal place-cell sequences depict future paths to remembered goals. Nature. 2013;497(7447):74-9.

[66] Grill-Spector K, Weiner KS. The functional architecture of the ventral temporal cortex and its role in categorization. Nature Reviews Neuroscience. 2014;15(8):536-48.

[67] Wang Y, Collins JA, Koski J, Nugiel T, Metoki A, Olson IR. Dynamic neural architecture for social knowledge retrieval. Proc Natl Acad Sci U S A. 2017;114(16):E3305-14.

[68] Sousa AM, Meyer KA, Santpere G, Gulden FO, Sestan N. Evolution of the human nervous system function, structure, and development. Cell. 2017;170(2):226-47.

[69] Rilling JK, Glasser MF, Preuss TM, Ma X, Zhao T, Hu X, et al. The evolution of the arcuate fasciculus revealed with comparative DTI. Nature neuroscience. 2008;11(4):426.

[70] Chance SA, Sawyer EK, Clover LM, Wicinski B, Hof PR, Crow TJ. Hemispheric asymmetry in the fusiform gyrus distinguishes Homo sapiens from chimpanzees. Brain Structure and Function. 2013;218:1391-405.

[71] Patel GH, Yang D, Jamerson EC, Snyder LH, Corbetta M, Ferrera VP. Functional evolution of new and expanded attention networks in humans. Proceedings of the National Academy of Sciences. 2015;112(30):9454-9.

[72] Elston GN, Benavides-Piccione R, Elston A, Zietsch B, Defelipe J, Manger P, et al. Specializations of the granular prefrontal cortex of primates: implications for cognitive processing. The Anatomical Record. 2006;288(1):26-35.

[73] Dehaene S. Consciousness and the brain: Deciphering how the brain codes our thoughts. Penguin; 2014.

[74] Rappoport A. The Brain, Explained: A Comprehensive Theory of Brain Function. doi: 10.20944/preprints201805.0137. preprtingsorg. 2018.

[75] Rappoport A. Short version of 'the brain, explained': a response process theory of brain function. doi: 10.20944/preprints201806.0092. preprtingsorg. 2018.