

Attention: Multiple types, brain resonances, psychological functions, and conscious states

Stephen Grossberg^{1,*}

¹Center for Adaptive Systems, Graduate Program in Cognitive and Neural Systems, Departments of Mathematics & Statistics, Psychological & Brain Sciences, and Biomedical Engineering, Boston University, 677 Beacon Street, Boston, 02215 MA, USA

*Correspondence: steve@bu.edu (Stephen Grossberg)

DOI: [10.31083/j.jin.2021.01.406](https://doi.org/10.31083/j.jin.2021.01.406)

This is an open access article under the CC BY 4.0 license (<https://creativecommons.org/licenses/by/4.0/>).

Submitted: 06 December 2020 Revised: 24 December 2020 Accepted: 15 January 2021 Published: 30 March 2021

This article describes neural models of attention. Since attention is not a disembodied process, the article explains how brain processes of consciousness, learning, expectation, attention, resonance, and synchrony interact. These processes show how attention plays a critical role in dynamically stabilizing perceptual and cognitive learning throughout our lives. Classical concepts of object and spatial attention are replaced by mechanistically precise processes of prototype, boundary, and surface attention. Adaptive resonances trigger learning of bottom-up recognition categories and top-down expectations that help to classify our experiences, and focus prototype attention upon the patterns of critical features that predict behavioral success. These feature-category resonances also maintain the stability of these learned memories. Different types of resonances induce functionally distinct conscious experiences during seeing, hearing, feeling, and knowing that are described and explained, along with their different attentional and anatomical correlates within different parts of the cerebral cortex. All parts of the cerebral cortex are organized into layered circuits. Laminar computing models show how attention is embodied within a canonical laminar neocortical circuit design that integrates bottom-up filtering, horizontal grouping, and top-down attentive matching. Spatial and motor processes obey matching and learning laws that are computationally complementary to those obeyed by perceptual and cognitive processes. Their laws adapt to bodily changes throughout life, and do not support attention or conscious states.

Keywords

Attention; Learning; Adaptive resonance theory; Neural models; Cognitive processing; Neural networks

1. Introduction

1.1 Attention is an emergent property of interacting brain networks and systems

From our earliest years, parents and teachers may exhort us to “pay attention” to one or another important type of knowledge or event that we need to learn about, or action that we need to perform. Indeed, achieving success in life is quite unlikely unless one can pay attention to important tasks and valued goals for long periods of time. One just has to think about sports or the arts to realize this. The amount of sustained attention that is needed for a baseball player, bal-

let dancer, or virtuoso instrumentalist to achieve mastery is often spread over years, if not a lifetime, of effort. Even more mundane skills like learning to tie one's shoes or to drive a car take concerted attention over a period of days or weeks.

Many factors will determine how well we succeed in learning any of these skills. And understanding the central role of attention in achieving success is rendered difficult by the apparently intangible nature of this state of mind, seemingly so different from the vividness of seeing a friend's face, hearing a favorite piece of music, knowing our own name, or feeling a warm rush of feeling when we see someone we love.

One reason why understanding how we pay attention is so difficult is that the act of paying attention is not separable from multiple other processes that are going on at any time in our brains, such as seeing, hearing, knowing, or feeling. Indeed, attention is an emergent property of interactions among thousands, or even millions, of neurons within brain networks and systems. Moreover, as I will explain below, mechanistically distinct types of attention occur in different brain systems. Separating them is made even more difficult by the fact that they can all interact synchronously together to enable us to experience a unified sense of self.

This article is devoted to providing accessible explanations of how, where, and why attention works in our brains. Such explanations are based on the mostly highly developed neural models of how our brains make our minds, including how we become conscious, and how consciousness is linked to our ability to pay attention. These models have been getting incrementally developed over the past 40 years. A self-contained and non-technical summary of brain models and how they may be combined to make our minds can be found in [1].

1.2 From attention to the CLEARS processes and Adaptive Resonance Theory

Other reviews of attention can be found in Wikipedia (<https://en.wikipedia.org/wiki/Attention>) and Scholarpedia (<http://www.scholarpedia.org/article/Attention>). These reviews describe observable psychological properties of attention, but not the mechanisms that cause them or the functions that these mechanisms carry out during behavior.

The current article explains how attention is integrated within interacting psychological and brain processes of Consciousness, Learning, Expectation, Attention, Resonance, and Synchrony (CLEARS). These processes, and how they interact, are explained below. The CLEARS processes have been modeled and simulated on the computer as part of Adaptive Resonance Theory (ART), which is the currently most advanced cognitive and neural theory of how our brains learn to attend, recognize, and predict objects and events in a changing world that is filled with unexpected events self-contained and non-technical exposition of ART and other brain processes with which it interacts is found in the book [1]. Various articles include analyses of how normal and abnormal cognitive and emotional information processing, learning, recognition, memory, and consciousness interact [3–19, 28, 40]; how sequences of objects and events are temporarily stored in cognitive or motor working memories before being chunked, or unitized, into learned sequence categories [21, 23, 31, 36, 39, 41]; how adaptively-timed reinforcement learning interacts with cognitive processes [22, 29, 30, 32, 33]; how visual grouping, attention, perception, and search interact with ART learning processes [25, 34, 35, 37, 38]; how auditory, speech, and language processes are clarified by ART dynamics [2, 24, 26]; how social cognition processes like gaze following and joint attention interact with ART dynamics [27]; and how ART may be used to design autonomous adaptive algorithms and robots for technology [20].

1.3 ART uniquely follows from a thought experiment about correcting errors in a changing world

ART is not “just another” neural model. It has been derived from a Gedanken, or thought, experiment as the unique solution of the universal problem of how predictive errors can be autonomously corrected in a changing world [12]. The hypotheses from which the thought experiment is derived are, moreover, just a few familiar facts that we know from our daily lives. These facts are familiar because they are ubiquitous environmental constraints that have guided the evolution of our brains. When these hypotheses act together, as they regularly do in environments where individuals behave, they define a multiple constraint satisfaction problem that ART uniquely solves.

Grossberg ([12], p. 7) summarizes this evolutionary challenge as follows: “*The importance of this issue becomes clear when we realize that erroneous cues can accidentally be incorporated into a code when our interactions with the environment are simple and will only become evident when our environmental expectations become more demanding. Even if our code perfectly matched a given environment, we would certainly make errors as the environment itself fluctuates.*”

The thought experiment translates this purely logical inquiry about error correction into processes operating autonomously in real time with only locally computed quantities. The thought experiment thus shows how, when familiar environmental constraints on incremental knowl-

edge discovery are overcome in a self-organizing manner through evolutionary selection processes, ART circuits naturally emerge. As a consequence, ART architectures may, in some form, be expected to be embodied in all future truly *autonomous* adaptive intelligent devices, whether biological or artificial.

1.4 ART is a principled biological theory. Back propagation and Deep Learning are not

Perhaps because of the fact that ART uniquely follows from the hypotheses of the thought experiment, all of the basic neural mechanisms that ART has proposed have been supported by psychological and neurobiological data. ART has also provided a unified explanation of hundreds of other experiments, and has also made scores of predictions that have subsequently received experimental support, as the above cited articles about ART illustrate.

ART is thus a principled biological *theory* of how our brains learn to attend, recognize, and predict objects and events in a changing world. It is not just an *algorithm* defined by feedforward adaptive connections with no top-down attentional mechanism, as it the case with many popular neural learning algorithms, including competitive learning, simulated annealing, Boltzmann Machine, back propagation, and Deep Learning.

In particular, back propagation and Deep Learning lack a mechanism for paying attention to predictive data and for dynamically stabilizing learning of it. As a result, neither back propagation nor Deep Learning is trustworthy—because neither is *explainable*—nor *reliable*—because each can experience *catastrophic forgetting*. Explainability means that the basis for making a prediction can be explicitly derived from the state of the algorithm. Catastrophic forgetting means that an arbitrary part of an algorithm’s learned memory can unpredictably collapse. Life-or-death decisions, including medical and financial decisions, cannot confidently be made using an algorithm with these weaknesses. Grossberg [42] explains why back propagation and Deep Learning have these deficiencies.

Many learning algorithms can be trained as classifiers, but do not have the unique combination of properties that ART embodies, including how attention helps to realize ART’s ability to realize autonomous adaptive intelligence in response to a changing world. This article summarizes how attention does this.

Back propagation became popular in the 1980’s after a publication by Rumelhart, Hinton, and Williams [43] applied earlier discoveries of scientists like Amari [44], Werbos [45, 46], and Parker [47–49]. Schmidhuber [50] provides an extensive historical summary of various contributions to the development of back propagation. It was soon, however, realized that back propagation suffers from many serious problems. For example, Grossberg [51] summarized 17 problems of back propagation which Adaptive Resonance Theory had already overcome starting in the 1970s. They are, listed as follows:

- Real-time (on-line) learning vs. lab-time (off-line) learning
- Learning in nonstationary unexpected world vs. in stationary controlled world
- Self-organized unsupervised or supervised learning vs. supervised learning
- Dynamically self-stabilize learning to arbitrarily many inputs vs. catastrophic forgetting
- Maintain plasticity forever vs. externally shut off learning when database gets too large
- Effective learning of arbitrary databases vs. statistical restrictions on learnable data
- Learn internal expectations vs. impose external cost functions
- Actively focus attention to selectively learn critical features vs. passive weight change
- Closing vs. opening the feedback loop between fast signaling and slower learning
- Top-down priming and selective processing vs. activation of all memory resources
- Match learning vs. mismatch learning: Avoiding the noise catastrophe
- Fast and slow learning vs. only slow learning: Avoiding the oscillation catastrophe
- Learning guided by hypothesis testing and memory search vs. passive weight change
- Direct access to globally best match vs. local minima
- Asynchronous learning vs. fixed duration learning: A cost of unstable slow learning
- Autonomous vigilance control vs. unchanging sensitivity during learning
- General-purpose self-organizing production system vs. passive adaptive filter

Several of the most serious problems will be discussed below.

Due to such problems, back propagation was gradually supplanted by other neural network algorithms. Although Deep Learning shares these problems with back propagation, it has become popular lately, largely because of the advent in the intervening years of huge online databases—which make it easier to train the algorithm using lots of data—and much faster computers—which facilitate the multiple learning trials that are needed because of the algorithm’s slow learning rate. Many tend like to think of Deep Learning as “back propagation on steroids”, since it has not solved the core foundational problems of its back propagation learning algorithm.

1.5 Brain evolution needs to achieve behavioral success: A modeling method and cycle

It is important to understand ART in an evolutionary context. The concept of “survival of the fittest” is often used to describe how Charles Darwin’s proposed evolutionary mechanism of natural selection works [52]. When applied to our brains, natural selection requires that “brain evolution needs to achieve behavioral success”, because it is only through behaviors that the cumulative effects of a species’ evolutionary

specializations can be tested against the persistent challenges of changing environments.

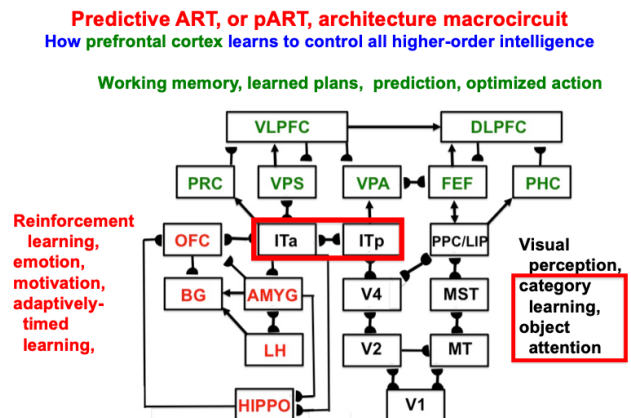


Fig. 1. The predictive ART, or pART, architecture unifies many of the brain processes that control biological intelligence. ART category learning and object (or prototype) attention processes take place in the model’s posterior inferotemporal cortex (ITp) and anterior inferotemporal cortex (ITa). The other brain regions and their processes are modelled by additional biological neural networks, whose abbreviations are printed with the same color (black, red, green) as the functions that they carry out. V1: striate, or primary, visual cortex; V2 and V4: areas of prestriate visual cortex; MT: middle temporal cortex; MST: medial superior temporal area; ITp: posterior inferotemporal cortex; ITa: anterior inferotemporal cortex; PPC/LIP: posterior parietal cortex; LIP: lateral intraparietal area; VPA: ventral prearcuate gyrus; FEF: frontal eye fields; PHC: parahippocampal cortex; DLPFC: dorsolateral hippocampal cortex; HIPPO: hippocampus; LH: lateral hypothalamus; BG: basal ganglia; AMGY: amygdala; OFC: orbitofrontal cortex; PRC: perirhinal cortex; VPS: ventral bank of the principal sulcus; VLPFC: ventrolateral prefrontal cortex. Output signals from the BG that regulate reinforcement learning and gating of multiple cortical areas are not shown. See Fig. 41 for some of these. Output signals from cortical areas to motor responses are also not shown. [Adapted with permission from [21] published in SAGE journals.]

1.6 pART: Increasingly comprehensive attentive brain architectures

Despite its status as the unique solution of the thought experiment about error correction using only locally computed quantities in a changing world, ART is just one of the biological neural network architectures that model how our brains make our minds. This is true because our brains need to solve many other problems than learning to attention, recognize, and predict objects and events in a changing world. The predictive ART, or pART, architecture embeds the core ART adaptive classification abilities within a more comprehensive brain macrocircuit of how human brains work (Fig. 1). Each of the processes in Fig. 1 have been developed as rigorous neural models, along with parametric simulations of many psychological and neurobiological data.

The two inferotemporal cortical areas in pART—posterior inferotemporal cortex (ITp) and anterior inferotemporal cortex (ITa)—carry out ART-like attentive

EMERGING UNIFIED THEORY OF VISUAL INTELLIGENCE

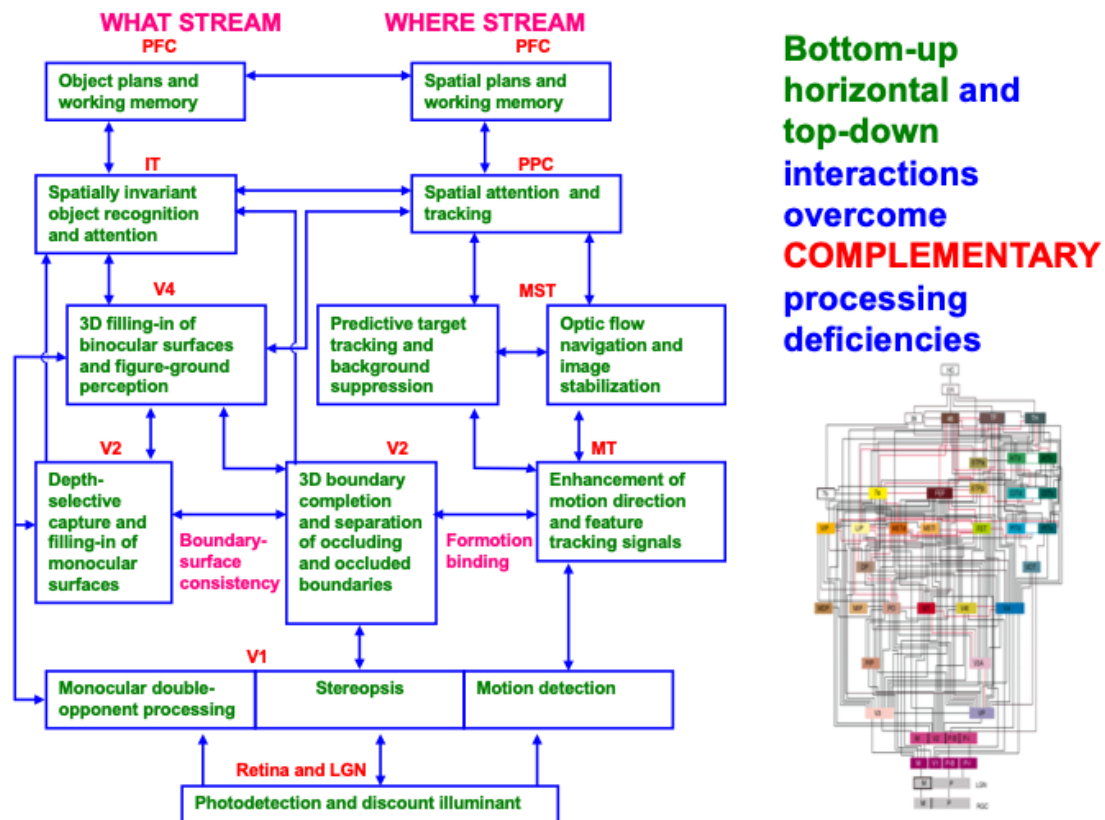


Fig. 2. A macrocircuit of some of the main brain regions and the processes that they carry out within and between the ventral, or What, cortical stream and dorsal, or Where cortical stream, that comprise an emerging neural model of biological vision. Bottom-up, horizontal, and top-down interactions among these regions overcome computationally complementary weaknesses that each process would exhibit if acted alone. [Reprinted with permission from [20] published in Elsevier.]

category learning. These cognitive networks receive pre-processed outputs from visual cortical areas in the lower right of the pART architecture—V1, V2, V4, MT, MST, PPC/LIP—that carry out the functionally distinct processes which together enable our brains to consciously see.

The macrocircuit in Fig. 2 summarizes key psychological processes that occur in different brain regions that interact within an emerging unified theory of visual intelligence. Note that these processes occur within both the ventral, or What, cortical processing stream and the dorsal, or Where, cortical processing stream (Fig. 3; [54–57]), and interact using a combination of bottom-up, horizontal, and top-down interactions. It will be explained below how these top-down interactions embody distinct object and spatial attentional processes, and how and *why* paying conscious spatial attention to an object enables us to look at it and reach for it.

These models have been derived incrementally over the years using a modeling method and cycle that reflects the fact that brain evolution needs to achieve behavioral success, as summarized in Fig. 4.

1.7 Fast learning without catastrophic forgetting and paying attention to critical feature patterns

Below it will be shown how attention plays a key role in the ability of ART to support our behavioral success, and thus survival. ART top-down expectations enable humans to learn how to attend to those combinations of *critical features* which control actions that have led to behavioral success in the past (Fig. 5). These top-down learned expectations enable ART, and ourselves, to learn quickly and to remember what we have learned, often for many years, without experiencing *catastrophic forgetting*, or the unexpected collapse of part of our learned memories. Despite the persistence of these learned memories, they can also be forgotten, or extinguished, when they lead to unexpected consequences that disconfirm them.

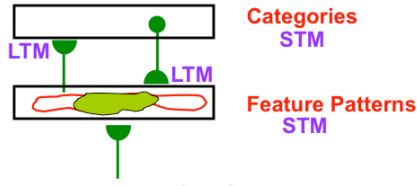
1.8 Feature-category resonance: Multiple resonances support attention, learning, and consciousness

As Fig. 5 illustrates, an attended critical feature pattern across a level of feature detectors reactivates the bottom-up adaptive filter pathways that activate a level of learned recognition categories. The activated category, in turn, reactivates

ADAPTIVE RESONANCE

Attended feature clusters reactivate bottom-up pathways

Activated categories reactivate their top-down pathways



Feature-Category resonance synchronizes amplifies prolongs system response

Resonance triggers learning in bottom-up and top-down adaptive weights: *adaptive resonance!*

Fig. 5. A *feature-category resonance* binds together an attended critical feature pattern with the category that represents it. See the text for details. [Adapted with permission from [18] published in Frontiers.]

In Eqn. 1, the automatic gain control terms $(B_i - x_i)$ and $(E_i x_i + F_i)$ keep the activities x_i within the finite bounds B_i/C_i and $-F_i/E_i$. The balance between the total excitatory term $(B_i - x_i) \left[\sum_j f_j(x_j) D_{ji} y_{ji} z_{ji} + I_i \right]$ and the total inhibitory term $(E_i x_i + F_i) \left[\sum_j g_j(x_j) G_{ji} Y_{ji} Z_{ji} + J_i \right]$ keep the activities from saturating at either their maximum or minimum values and, more generally, enables such a network to process spatial patterns of analog x_i values in response to external excitatory and inhibitory inputs, I_i and J_i , respectively, and excitatory feedback signals and inhibitory feedback signals, $\sum_j f_j(x_j) D_{ji} y_{ji} z_{ji}$ and $\sum_j g_j(x_j) G_{ji} Y_{ji} Z_{ji}$, respectively, that may vary greatly in size through time [51, 61].

The shunting dynamics in Eqn. 1 embody the membrane equations of neurophysiology operating in a recurrent on-center off-surround anatomy [62]. When the automatic gain control terms are removed, then the shunting STM equation reduces to the additive STM equation. The additive STM model cannot saturate, but it also does not have many of the valuable properties of the shunting model, including such essential properties for biological vision as the ability to *discount the illuminant*, or to compensate for huge changes in illumination that occur every day [63, 64].

The MTM Eqn. 2 describes how the chemical transmitter concentration at the ends of the axons, or pathways, between neurons balances between a process of accumulation $H(K - y_{ki})$ and one of gated release, or inactivation $-L f_k(x_k) y_{ki}$.

The LTM Eqn. 3 describes how learning switches on and off when the *stimulus sampling signal* $f_k(x_k)$ is positive or zero, respectively. When it is positive, then the adaptive weight, or LTM trace, z_{ki} tracks the sampled signal $h_i(x_i)$ by steepest descent.

With these STM, MTM, and LTM variables defined, it is possible to say that, when a feature-category resonance (Fig. 5) occurs between attended critical features and the

Table 1. A summary of the adaptive resonances that support conscious experiences of seeing, hearing, knowing, and feeling. [Reprinted with permission from [19] published in Taylor & Francis Group.]

CLASSIFICATION OF RESONANCES

Surface-shroud resonances support conscious seeing of visual qualia

Feature-category resonances support conscious recognition of visual objects and scenes

Stream-shroud resonances support conscious hearing of auditory qualia

Spectral-pitch-and-timbre resonances support conscious recognition of sources in auditory streams

Item-list resonances support conscious recognition of speech and language

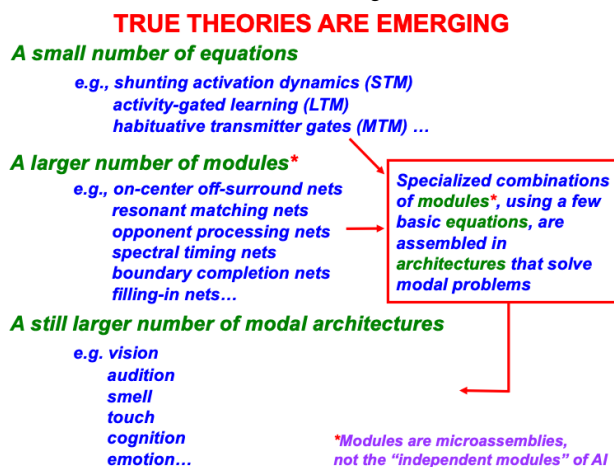
Cognitive-emotional resonances support conscious feelings and recognition of them

recognition category to which they are bound by bottom-up and top-down excitatory signals, then fast learning is triggered in the adaptive weights, or long-term memory (LTM) traces, in the synaptic knobs at the ends of the bottom-up adaptive filter and top-down expectation axons (Table 1). Learning regulates the size of the adaptive weights, which regulate the amount of chemical transmitter that is released from synaptic knobs to the abutting nerve cell body, or change the sensitivity of postsynaptic membranes to these presynaptic signals, or both [65–67].

These equations are assembled within a somewhat larger number of *modules*, or microcircuits, that carry out different functions within each modality. They may be thought of as the “molecules” of biological intelligence. The modules include the following kinds of networks: shunting on-center off-surround networks, gated dipole opponent processing networks, associative learning networks, and adaptively timed spectral learning networks. Each type of module exhibits a rich set of useful computational properties, but are not general-purpose computers. Rather, each kind of module was shaped by evolution to carry out a range of different tasks that could be accomplished by specializations of its design.

For example, shunting on-center off-surround networks exhibit properties like contrast normalization, including discounting the illuminant during visual perception; contrast enhancement, noise suppression, and winner-take-all choice during the choice of a recognition category; short-term memory and working memory storage during the persistent short-term storage of individual events or sequences of events; attentive matching of bottom-up input patterns and top-down learned expectations, as occurs during a feature-category resonance; synchronous oscillations, as occurs during conscious resonances; and traveling waves that can occur during epileptic seizures.

Table 2. A small number of equations and a somewhat larger number of modules, or microcircuits, are specialized and assembled in modal architectures that carry out different modalities of intelligence.



These equations and modules are specialized and assembled into *modal* architectures. The term “modal” stands for different modalities of biological intelligence, including architectures for vision, audition, cognition, cognitive-emotional interactions, and sensory-motor control.

An integrated self is possible because it builds on a shared set of equations and modules within modal architectures that can interact seamlessly together.

Although they cannot compute everything, unlike a universal Turing machine or its hardware embodiment in a von Neumann computer (https://en.wikipedia.org/wiki/Von_Neumann_architecture), modal architectures are *general-purpose* in the sense that they can process all inputs to their modality, whether from the external world or from other modal architectures. Modal architectures are thus more general than a traditional AI algorithm. The types of resonances summarized in Table 3 form part of several different modal architectures, including modal architectures that enable conscious seeing, hearing, feeling, and knowing.

The exposition below will describe how the CLEARS processes interact within ART. Some of the psychological and neurobiological data for which ART has provided a unified explanation will be summarized, as well as ART predictions that not yet been tested experimentally.

1.10 Large-scale ART applications in engineering and technology

Before moving on to these explanations, it is useful to note that ART properties such as fast learning without catastrophic forgetting, and learned selection by attention of the critical features that control effective decisions and predictions, have encouraged the mathematical analysis, computer simulation and application of multiple ART algorithms towards the solution of large-scale problems in engineering and technology. See <http://techlab.bu.edu/resources/articles/C5.html>, [68] and [69] for a partial list of applications, and the following ar-

Table 3. Some pairs of computationally complementary processes and the brain regions where they occur. [Reprinted with permission from [19] published in Taylor & Francis Group.]

SOME COMPLEMENTARY PROCESSES

Visual Boundary Interbob Stream V1-V4	Visual Surface Blob Stream V1-V4
Visual Boundary Interbob Stream V1-V4	Visual Motion Magno Stream V1-MT
WHAT Steam Perception & Recognition Inferotemporal and Prefrontal areas	WHERE Stream Space & Action Parietal and Prefrontal areas
Object Tracking MT Interbands and MSTv	Optic Flow Navigation MT Bands and MSTd
Motor Target Position Motor and Parietal Cortex	Volitional Speed Basal Ganglia

ticles for the mathematical and computational development of various ART algorithms [70–84].

2. Object attention and spatial attention

2.1 Dynamics of exogenous and endogenous object and spatial attention shifts during visual search

For at least the past 40 years, the cognitive neuroscience literature has made the distinction between object attention [85] and spatial attention [86]. These Duncan and Posner articles included new experiments, but can be viewed as a further development of extensive earlier work by multiple authors. A small sample of early experiments includes those of Neiser [87], Lappin [88], and Treisman, Kahneman, and Burkeil [89] about object-based attention, often studied in displays that include multiple possible target and distractor objects. Early experiments probing spatial attention were carried out in alert monkeys [90–92], brain injured patients [93], and normal individuals [94].

The most compelling studies of object and spatial attention successfully dissociate paying attention to a single object at multiple positions, or to a single position inhabited by multiple objects (e.g., [95–99]). Although such studies represent an ideal that is possible to implement in the laboratory, in the real world, object attention and spatial attention often strongly interact, for example, when planning one’s escape from a predator in a forest. Such interactions have been the subject of hundreds of experiments which are often subsumed under the general rubric of *visual search*. Triesman and her colleagues interpret their data using their Feature Integration model (e.g., [100–104]), while Wolfe and his colleagues do so using variants of their Guided Search model (e.g., [105–110]).

Biological neural network models have provided unified mechanistic explanations of many of the most challenging data about visual search (e.g., [7, 22, 37, 111]). These models of perceptual and cognitive information processing had previously been used to explain and predict other kinds of psychological and neurobiological data. Visual search data

were hereby integrated within a broad landscape of experimental paradigms probing different aspects of how brains make minds.

For example, the Spatial Object Search, or SOS, model of Grossberg, Mingolla, and Ross [112] has the title “*A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations*”. The SOS model explains and simulates visual search data as emergent properties of interactions between visual perception processes of boundary completion and surface filling-in, object attention processes whereby ART categories are learned and recognized, and spatial attention processes whereby attention shifts to objects in different locations.

Interactions between object attention and spatial attention have been probed in the laboratory when, for example, they are sequentially primed by object or spatial cues during relatively brief time intervals. Theeuwes, Mathot, and Grainger [113] discuss such interactions in the context of “exogenously controlled object attention”, and review related studies by other authors. Exogenous control refers to bids for attention from the external world, whereas endogenous control refers to top-down attentional processes within our brains.

An exogenous attention shift to a location in space may be caused by the sudden appearance of an object in a scene [114, 115]. This automatic exogenous allocation of spatial attention can compete with endogenous top-down object attentional priming, thereby illustrating competition for attentional resources across the Where and What cortical streams. For example, if an observer is primed to look for a color singleton, then an abrupt onset of an object at a different position will cause an attentional shift that slows down search for the color singleton [114, 116].

Posner [86] emphasized this orienting process, as illustrated by the title of his article, “Orienting of attention”. The examples proposed by Posner [86] and Theeuwes *et al.* [113] describe an orienting response due to transient appearances of objects at different positions. As I will explain below, orienting can also be driven endogenously during an ART search, or hypothesis testing, for the internal representation of any event, whether or not it occurs along with a shift of spatial attention.

As noted by Theeuwes, Mathot, and Grainger [113], Posner and Cohen [117] studied exogenous attention shifts in an experimental setup where subjects fixate a central position surrounded by a regular array of outline boxes. Then one of two peripheral boxes is cued by brightening, before a target is presented inside a box. Participants detected the target faster when it appeared at the cued, relative to the uncued, box, thereby illustrating how a shift in spatial attention can facilitate object processing and attention at the cued position. It was also shown that this facilitation is coded in retinotopic coordinates.

Several labs have proposed that abrupt onsets capture attention by strongly activating transient cells that are designed to respond to rapid cue changes. These cells are abundant in

the Where cortical stream that is also often referred to as the magnocellular pathway (e.g., [118–121]) due to its abundance of retinal Y cells that are insensitive to object form and color blind, but highly sensitive to luminance transients and motion [122–126].

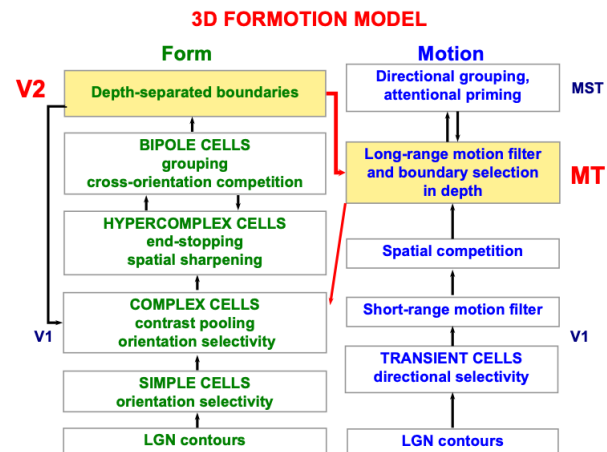


Fig. 6. Macro-circuit of the 3D FORMOTION model for form-to-motion, or FORMOTION, interactions from cortical area V2 in the What cortical stream to cortical area MT in the Where cortical stream. This interaction enables an observer to track a moving form in depth. See the text for details. [Adapted with permission from [127] published in Brill.]

2.2 From data about transient cells to the 3D FORMOTION model of visual form and motion perception

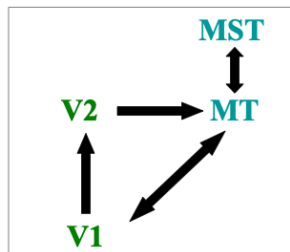
A detailed psychological, anatomical, and neurophysiological model has been incrementally developed to explain how transient cells, among many others, contribute to our brain’s ability to compute the direction and speed of objects that are moving within a cluttered environment that also contains many moving environmental distractors. This 3D FORMOTION model (Fig. 6) integrates form and motion information across multiple brain regions of the What and Where cortical streams, including cortical areas V1, V2, V4, MT, and MST (Fig. 7), to accomplish this feat.

The 3D FORMOTION does so by developing a rigorous computational solution of the *global aperture problem* whereby our brains convert the ambiguous hodgepodge of local motion signals arriving at our retinas from the environment into coherent cortical representations of object motion direction and speed (e.g., [128–135]).

As illustrated by Fig. 8, when an object moves under real world conditions, such as a leopard running across a grassy field, only a small subset of its image features, notably its bounding contours, may generate motion direction cues that accurately describe its direction-of-motion. The movements of the leopard’s limbs as it runs (red arrows) occur in multiple directions other than the direction that the leopard is moving (green arrows). The same is true of the contours of the spots on the leopard’s coat as they move with the limbs. Most lo-

cal motion signals differ from the direction of object motion because they are computed from local views of object motion within neurons' finite receptive fields.

MULTIPLE BRAIN REGIONS INTERACT TO SOLVE THE APERTURE PROBLEM



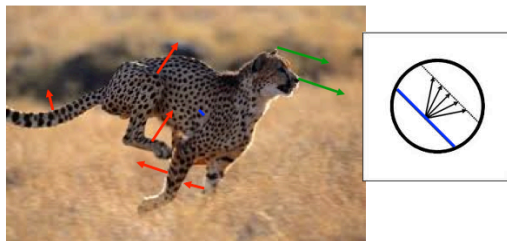
Form-Motion interactions are necessary to solve the problem

Fig. 7. Cortical regions and their bidirectional or cyclic interactions that carry out form-to-motion, or FORMOTION interactions. [Adapted with permission from [128] published in Elsevier.]

MOST MOTION SIGNALS MAY NOT POINT IN AN OBJECT'S DIRECTION OF MOTION

Aperture Problem

EVERY neuron's receptive field experiences an aperture problem



How does the brain use the small number of correct, unambiguous motion signals to compute an object's motion direction?

Fig. 8. The aperture problem arises because of two related limitations individual neuron's receptive fields to compute the true motion direction of an object. One limitation arises because a line moving in any direction within a circular aperture appears to move in the direction that is perpendicular to its orientation. Another limitation arises because the motion direction of a moving object often differs from the directions in which its parts move. The 3D FORMOTION model solves the aperture problem. See the text for details. [Reprinted with permission from [1] published in Oxford University Press.]

The right insert in Fig. 8 illustrates the insight of Hans Wallach in 1935 [136] that the motion direction of a line seen within a circular aperture is perceptually ambiguous. No matter what the line's real direction of motion may be, its perceived direction is perpendicular to its orientation. This phenomenon was called the *aperture problem* by Marr and Ullman

A CONFIRMED PREDICTION:

FORM and MOTION are COMPLEMENTARY

Prediction: Grossberg (1991, Perception & Psychophysics)
Data: Ponce, Lomber, and Born (2008, Nature Neuroscience)

Why separate cortical FORM and MOTION streams?



My 1991 prediction:

V2 - MT interaction achieves fine moving-form-in-depth which facilitates object tracking, etc.

Ponce et al. in 2008 provided strong evidence for it

Fig. 9. Form and motion computations are complementary: The form stream binocularly matches features that the two eyes receive from a nearby object to compute an estimate of the object's depth with respect to the observer. Each binocular match occurs between left eye and right eye representations of the same object feature in the world (see left image). These features thus represent the same object orientation in the world, thereby enabling a fine depth estimate to be computed. Restricting matches to the same orientation, however, enables them to compute only coarse direction estimates of object motion direction. The motion stream generates fine direction estimates of an object's motion by pooling over the object's differently oriented contours that are moving in the same direction (see right image). Pooling over orientation enables only coarse depth estimates of the object. FORMOTION interactions from V2-to-MT enable cells in MT to overcome these complementary weaknesses to compute fine moving-form-in-depth, which can be used to support accurate object tracking. If V2 is cooled, then only coarse depth estimates are recorded in MT, even though fine estimates of motion direction are unimpaired.

Note: Prediction: Grossberg [34]. Data: Ponce, Lomber, and Born [139].

[137]. The aperture problem is faced by any localized neural motion sensor, such as a neuron in the early visual pathway, that responds to a moving local contour through an aperture-like receptive field.

Until the aperture problem is solved, our brains cannot compute an object's direction and speed of motion, and thus cannot localize spatial attention to track the object. The microcircuit in Fig. 6 of the 3D FORMOTION model shows that the model's cortical area MST can control top-down attention upon the object motion direction and speed representation that is computed in cortical area MT. As described in [127] and [1], this attentional circuit obeys the same ART Matching Rule that is realized in all object attentional circuits, and which also supports learning and stable memory of the directionally tuned cells that enable leopards, and humans, to solve the aperture problem. Berzhanskaya, Grossberg, and Mingolla [127] also describe and explain many other data about object motion perception, including the coordinates in which it is computed. The ART Matching Rule will be explained in greater detail below.

2.3 Why FORMOTION is needed: Form and motion obey computationally complementary laws

Interactions between the form and motion streams are needed because the laws of form and motion processing are computationally complementary. As shown in Fig. 9, form processing within the What cortical stream areas of V1, V2, and V4 is *orientationally* sensitive, whereas motion processing within the Where cortical stream areas of V1, MT, and MST is *directionally* sensitive. Positionally and orientationally precise binocular matches between our two eyes (see left image) enable us to compute fine estimates of an object's depth. In contrast, to compute an object's motion direction, motion signals need to be pooled from multiple boundaries of the object with possibly different orientations (see right image). Such pooling cannot be done during a binocular estimate of depth. Hence *fine depth* estimates coexist with *coarse direction* estimates within the What stream. Because pooling across orientations must be done to derive a *fine direction* estimate, the Where stream can only compute a *coarse depth* estimate.

FORM AND MOTION AFTEREFFECTS

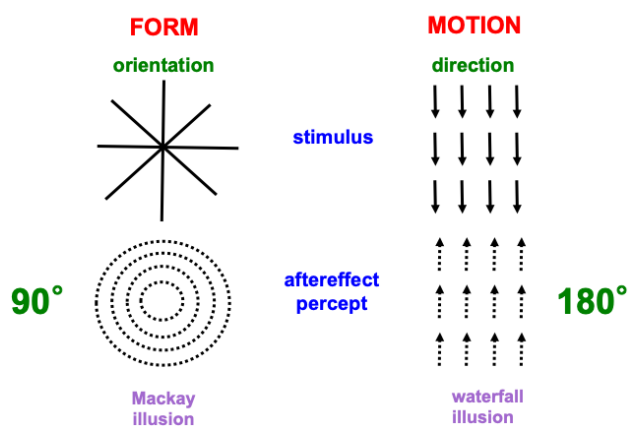


Fig. 10. Aftereffects of form and motion processing exhibit different asymmetries, with 90 degree differences between stimulus and aftereffect in the form stream, and 180 degree differences between stimulus and aftereffect in the motion stream. See the text for details. [Reprinted with permission from [1] published in Oxford University Press.]

Fig. 10 summarizes one of many demonstrations that computations of form and motion are, in fact, orientationally and directionally sensitive, respectively. The left half of the figure describes the Mackay illusion: Inspect intersecting lines before looking at a black wall or screen, where an afterimage of nested circles can be seen. The two images differ by 90 degrees. The right half of the figure describes the waterfall illusion: Inspect downward motion before looking at a black wall or screen, where an afterimage of upward motion can be seen. The two images differ by 180 degrees. These different symmetries for orientational vs. directional processing make clear that different brain systems support form and motion perception.

Form-to-motion interactions from V2-to-MT overcome the complementary weaknesses described in Fig. 9 to compute fine moving-form-in-depth estimates in MT (Fig. 11). This representation can then serve as a basis for attentively tracking a moving object, such as a predator or prey, in MST and beyond. I made this prediction in 1991 [138]. As Fig. 9 notes, Ponce, Lomber, and Born [139] confirmed my prediction in 2008 by reversibly cooling V2 and showing the predicted properties of fine direction and coarse depth estimates by MT cell responses, which returned to fine direction and fine depth estimates by MT responses after V2 recovered from cooling.

FORM and MOTION are COMPLEMENTARY

	Depth	Direction
Form	Fine	Coarse
Motion	Coarse	Fine
Formotion	Fine	Fine

Fig. 11. Summary of the complementary weaknesses of form and motion cortical computations that are overcome by FORMOTION interactions from V2-to-MT.

ViSTARS NAVIGATION MODEL

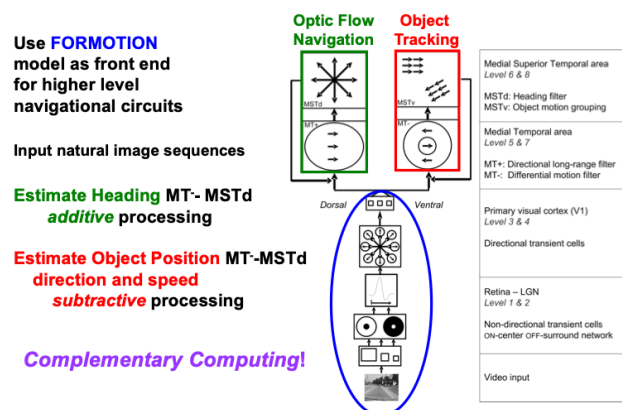


Fig. 12. The computation of heading during optic flow navigation, and of object position, direction, and speed during object tracking, obey computationally complementary laws. The ViSTARS model explains these different computations and the functions that they accomplish. See the text for details. [Adapted with permission from [140] published in Elsevier.]

2.4 Attentive object tracking and optic flow navigation: Complementary Computing

Once our brains overcome the aperture problem to compute reliable estimates of object motion, this information is used to support two different kinds of visually-guided navigational behaviors: *object tracking* and *navigation*. Object tracking enables us to track a target that is moving relative to us. A moving target can even be tracked behind multiple occluders, as could occur when a predator is tracking a moving prey that is intermittently occluded by bushes and trees in the forest. The predator can complete a continuous trajectory of the prey's motion, even behind the occluders, using long-range apparent motion [130], which also maintains spatial attention upon the location that is currently most active in the trajectory.

Navigation enables us to move relative to the world around us. Fig. 12 summarizes the macrocircuit of the Visual Steering, Tracking, And Route Selection, or ViSTARS, neural model of navigation [140–142]. The model shows that object tracking and optic flow navigation are carried out by two parallel cortical streams within cortical areas MT and MST (Fig. 7), with the ventral stream MT^- -to-MSTv carrying out object tracking, and the dorsal stream MT^+ -to-MSTd carrying out optic flow navigation. Optic flow is the pattern of motion that is generated in an observer's brain when objects in a visual scene move relative to the observer. See, for example, the video: <https://www.youtube.com/watch?v=tGYder2LHAA>.

Fig. 12 notes that these two parallel processing streams use computationally *complementary* processes: *Additive* processing enables the brain to determine the direction of *heading*, or a navigator's self-motion direction, whereas *subtractive* processing is used to determine the position, direction, and speed of a moving object. These complementary types of processing enable the computation of an observer's heading while moving relative to a scene, and of an object's movements relative to the observer. This latter information can, in turn, be used to avoid collisions with objects in a scene while moving through it. Both processes contribute to an observer's ability to lock attention onto a valued goal object and to maintain it while navigating towards that object.

2.5 Complementary matching and learning laws for recognition and action

The complementary laws of tracking and navigation are just one of many examples of the general principle of Complementary Computing that organizes how multiple pairs of interacting brain regions are specialized. Table 3 lists some pairs of psychological processes, and the cortical areas within which they occur, for which neural models have articulated computationally complementary properties.

Table 4 summarizes complementary properties of learning and matching within the What and Where cortical streams. The What stream learns invariant object recognition categories within the anterior inferotemporal cortex, or ITa (Fig. 1). This learning can occur quickly without caus-

Table 4. Learning and matching processes within the What and Where cortical streams obey complementary laws. See the text for details. [Adapted with permission from [17] published in Elsevier.]

WHAT and WHERE LEARNING and MATCHING are COMPLEMENTARY

<div> <div> Spatially-invariant object learning and recognition Fast learning without catastrophic forgetting IT </div> <div> Spatially-variant reaching and movement Continually update sensory-motor maps and gains PPC </div> </div>	
	<div> <div>WHAT</div> <div>WHERE</div> </div>
MATCHING	<div> <div>EXCITATORY</div> <div>INHIBITORY</div> </div>
LEARNING	<div> <div>MATCH</div> <div>MISMATCH</div> </div>
	<div> <div>ART</div> <div>VAM</div> </div>

ing catastrophic forgetting. A single invariant object category can recognize multiple views of an object at different positions and image sizes on the retina. The ability of an invariant category to recognize an object that is seen at different positions is the *spatially-invariant* property that is summarized in the figure, in order to contrast it with the *spatially-variant* allocation of spatial attention in the posterior parietal cortex, or PPC, to control reaching and other movements towards different positions in space. As our limbs change in size and strength over the years, the circuits that control them can continually update their motor maps and gains—that is, can experience “catastrophic forgetting”—to ensure skillful performance.

The bottom part of Table 4 contrasts the computationally complementary mechanisms of matching and learning in the What and Where streams that generate these distinct learning and behavioral properties. As in the case of object tracking and navigation, invariant category learning and movement control differ by obeying excitatory vs. inhibitory laws, in this case matching laws. The category learning properties have been modeled by ART, some of whose foundational properties will be reviewed below, including how ART learning occurs when there is a good enough *match* between bottom-up input patterns and top-down learned expectation signals that focus object attention upon the *critical features* that predict successful decisions and actions. Such a match is *excitatory* because it initiates an attentive resonance which triggers category learning. It is thus an *adaptive* resonance, hence the name of ART.

The movement control properties have been modeled by Vector Associative Map, or VAM, dynamics [143, 144]. Unlike the *excitatory* matching and *match*-based learning of ART, a VAM model carries out *mismatch*-based learning that is used to calibrate its *inhibitory* matching computations, as when an arm's present position is subtracted from a desired target po-

sition to compute a difference vector that controls the direction and distance of a reaching movement to the target. VAM dynamics will not be further discussed herein, except to note that, because of its inhibitory matching dynamics, VAM models cannot pay attention or become conscious.

2.6 From object attention to prototype attention

The heuristic concepts of object and spatial attention can be refined in terms of the brain mechanisms that carry out these attentional processes. Object attention is replaced by the concept of *prototype attention* during recognition learning. As will be discussed more fully below, each active category in an ART architecture reads-out a top-down expectation which learns a prototype that encodes a time-average of the critical feature patterns that are attended when the category is active. This critical feature pattern also chooses the active category via its bottom-up adaptive filter signals to the category level. All of these operations occur within the What cortical stream.

2.7 From spatial attention to boundary and surface attention

The heuristic concept of spatial attention is replaced by the mechanistically more precise concepts of *boundary attention* and *surface attention* during visual perception. One reason that boundaries and surfaces are so important in spatial attention is that they are, when properly understood in terms of the processes of boundary completion and surface filling-in, the functional units of visual perception [112, 145–165].

Boundaries and surfaces are computed in the What cortical stream within visual cortical areas such as V1, V2, and V4 (Fig. 2). Boundaries and surfaces are another example of computationally complementary processes (Table 3). These complementary properties are summarized in Fig. 13, and are illustrated by the boundary completion and surface filling-in processes that occur during the visual illusion of neon color spreading.

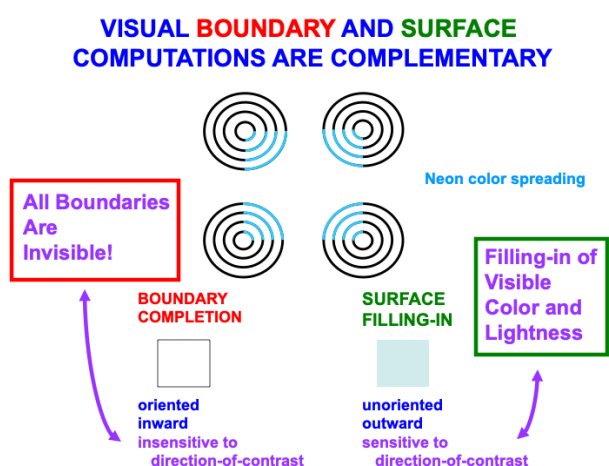


Fig. 13. Visual boundary and surface laws are complementary. [Adapted with permission from [166] published in Elsevier.]

2.8 Why are boundary completion and surface filling-in needed?

Boundary completion and surface filling-in compensate for gaps and uncertainties that exist in retinal images due to the existence of a blind spot, retinal veins, light scattering within the retina, and other image-degrading processes. The blind spot is as large as the region of maximal retinal sensitivity to light, the fovea, yet we are not aware of it. Partly this is because the retina jiggles rapidly in its orbit, thereby creating transient signals from external objects due to their motion relative to the retina. These transients refresh the signals from external objects at the retina. The blind spot and retinal veins do not create transients because they are attached to the retina, hence they fade from visibility.

The result is an incomplete representation of the external world where the blind spot and veins have occluded it (Fig. 14, top image). Boundary completion and surface filling-in complete visual representations over the occluded regions (Fig. 14, bottom three images).

EVERY LINE IS AN ILLUSION!



Fig. 14. After retinal occlusions such as the blind spot and retinal veins fade because they do not generate transient refresh signals on the retina, boundary completion and surface filling-in restore visual representations over the occluded retinal regions. [Reprinted with permission from [23] published in Springer.]

2.9 Hierarchical resolution of uncertainty and consciousness: Seeing to reach

It requires multiple processing stages in the visual cortex to create complete, context-sensitive, and stable cortical representations of visual boundaries and surfaces. I call this *process hierarchical resolution of uncertainty*. Hierarchical resolutions of uncertainty are needed in multiple brain processes to generate sufficiently complete representations of sensory data upon which to base successful actions.

The processing stage where perceptual representations are completed needs to be distinguished from the previous processing stages so that the complete representation can be used to control successful goal-oriented actions. Actions

Table 5. Multiple processing stages—that is, a *hierarchical resolution of uncertainty*—are needed to transform incomplete and ambiguous retinal representations, such as those occluded by the blind spot and retinal veins, into complete perceptual representations using processes like boundary completion and surface filling-in (Fig. 13). A conscious resonance “lights up” the processing stage in such a hierarchy at which a sufficiently complete, context-sensitive, and stable representation is computed with which to control actions. That is why, conscious seeing helps to ensure effective looking and reaching, conscious hearing helps to ensure effective auditory communication and speaking, and conscious feeling helps to ensure effect goal-oriented action.

CENTRAL CLAIM

Conscious states are part of larger adaptive behavioral capabilities that help us to adapt to a changing world

Resonances for conscious

seeing help to ensure effective reaching

hearing help to ensure effective speaking

feeling help to ensure effective goal-oriented action

based upon incomplete representations could cause serious problems. The selected processing stage resonates with the subsequent one in the cortical hierarchy to selectively “light up” the complete representation. The lighting-up process renders the complete representation conscious so it can be used to guide successful actions. In this sense, we consciously see in order to look and reach, hear to communicate and speak, and feel to control effective goal-oriented actions (Table 5).

2.10 From spatial attention to surface attention and conscious seeing by surface-shroud resonances

Spatial attention plays a major role in lighting up such a complete representation. It does so, in particular, by resonating with the object’s completed—notably, filled-in—surface representation. It is in cortical area V4 that such a completed surface representation is computed, thereby triggering a resonance with the posterior parietal cortex, or PPC.

This V4-to-PPC-to V4 resonating feedback loop is called a *surface-shroud resonance* (Fig. 15) because surface attention in PPC fits its shape to that of the surface with which it is resonating. Form-fitting spatial attention was called an *attentional shroud* by Tyler and Kontsevich [169].

Just as a feature-category resonance supports conscious *recognition* of a visual object or scene, a surface-shroud resonance supports conscious *seeing* of an object’s visual qualia (Table 1). When a feature-category resonance synchronizes with a surface-shroud resonance, our brain knows what a familiar object is as we see it, and are then also ready to reach it via the active surface representation in PPC (Fig. 16). The top-down feedback from PPC to V4 carries out surface *atten-*

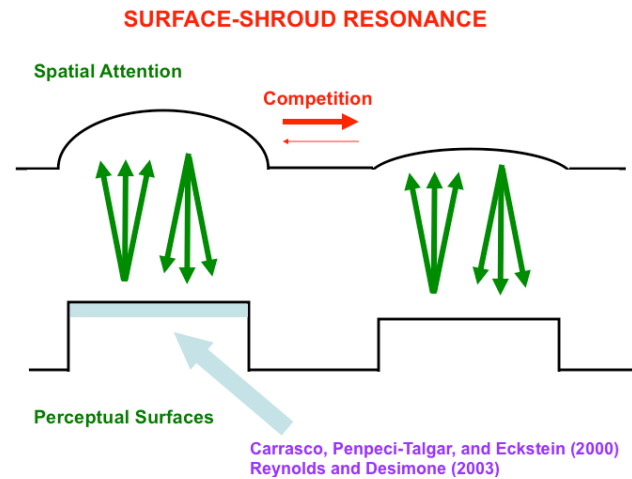
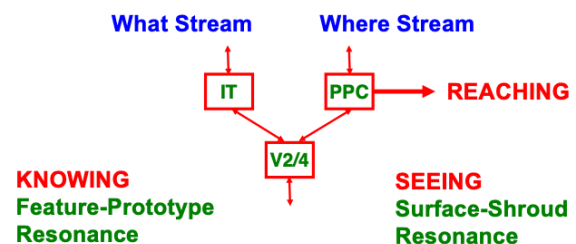


Fig. 15. A *surface-shroud resonance* between cortical areas V4 and PPC lights up the complete surface representation that can be used to control effective looking and reaching, as it focuses surface attention upon that surface representation, and thereby increases its effective contrast (light blue region). This enhancement has been reported in both psychophysical experiments (e.g., [167]) and neurophysiological experiments (e.g., [168]). [Reprinted with permission from [19] published in Springer.]

Note: Carrasco, Penpeci-Talgar, and Eckstein [167]. Reynolds and Desimone [168].

tion even while the bottom-up signals from PPC to downstream movement circuits embody the *intention* to move to the attended location (e.g., [170–174]). Boundary-shroud resonances can, in a similar way, become conscious and control reaches to an object’s contours.

WHAT KINDS OF RESONANCES SUPPORT KNOWING VS. SEEING?



Synchronous linkage between resonances enables us to consciously KNOW what the object is as we SEE it and REACH to it

Fig. 16. A *surface-shroud resonance* that supports conscious seeing can synchronize with a *feature-category resonance* that supports conscious knowing, or recognition (see Table 1), so that we can both see and know a familiar object when we attend it. The position of the attended object in PPC can then be used to look at or reach it. [Adapted with permission from [19] published in Springer.]

2.11 Using seeing and knowing resonances to explain psychophysical paradoxes and mental disorders

When a lesion in the feature-category resonance pathway occurs, then visual agnosia results during which humans, and the model, can reach to a target without knowing what it is, as in the patient DF reported by Goodale *et al.* [55]. Properties of surface-shroud resonances have been used to explain challenging data both about visual perception in normal humans and defects of consciousness in clinical patients. Here I will just note that normal properties such as limited capacity of attention, perceptual crowding, change blindness, and motion-induced blindness, and clinical properties of visual and auditory neglect, all get a unified explanation in [19], which also provides references to the relevant experimental literature.

For example, in response to a hemifield lesion of parietal cortex, the model explains many properties of the visual or auditory neglect that ensues. Explained clinical data go considerably beyond the familiar facts that such individuals may omit drawing parts of a scene, or neglect to dress the side of the body, that the lesioned hemifield would have processed. In all these cases, the visual cortex is intact but, without the parietal cortex to support a surface-shroud resonance, conscious seeing of, and spatial attention to, the afflicted hemifield does not occur.

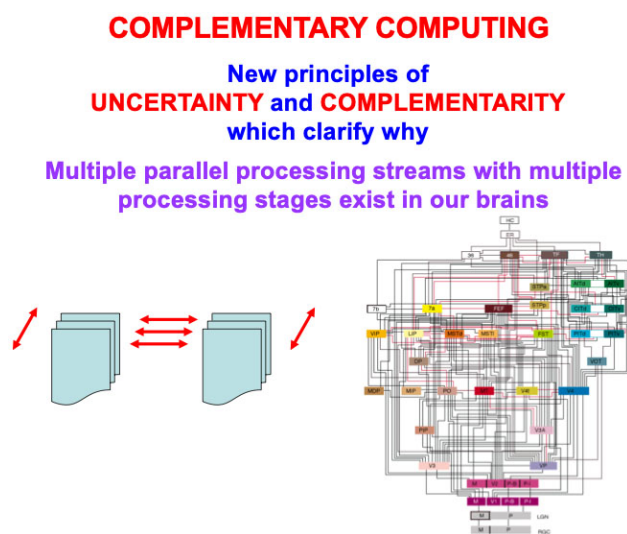


Fig. 17. Complementary computing requires new principles of uncertainty and complementarity, including the need for a hierarchical resolution of uncertainty, which clarify why multiple processing streams with multiple processing stages exist in our brains, as illustrated by the famous macrocircuit diagram of the visual system in the lower right corner that is adapted with permission from [175]. [Reprinted with permission from [20] published in Elsevier.]

2.12 Complementary Computing, hierarchical resolution of uncertainty, and consciousness

The above text has reviewed four examples of pairs of interacting processing streams in our brains that exhibit computationally complementary properties; namely, the first four

processing pairs that are listed in Table 1. The processes in Table 1 are only a partial list of known computationally complementary processes in our brains. All of these processes need multiple processing stages to carry out a hierarchical resolution of uncertainty, and thus also require that completed representations be “lit up” by conscious resonances. I will show in the next sections that ART circuits also exhibit complementary properties. Complementary Computing has hereby emerged, from results of multiple modeling studies over the years, as a basic principle of how brain systems are specialized, yet strongly interact (Fig. 17).

2.13 Shifting attention into and out of objects: Evaluating the processes underlying the object advantage

With this background about prototype, surface, and boundary attention in hand, it is easier to mechanistically explain data about interactions between them. The experiments of Brown and Denny [176] are particularly illuminating in this regard.

Explaining data of this kind is possible using the ARTSCAN model of Fazl, Grossberg, and Mingolla [7] and its extension to the ARTSCAN Search model of Chang, Grossberg, and Cao [5], whose macrocircuit is summarized in Figs. 18 and 19. Perhaps the most important computational property of these models is that they enable autonomous incremental learning of invariant object categories. Fig. 18 describes interactions that support learning and naming of invariant object categories. Interactions between surface and boundary attention in the Where cortical stream, and prototype attention in the What cortical stream, coordinate these learning, recognition, and naming processes. Fig. 19 describes the model processes that can direct a search for a previously learned, and currently desired, target object in a scene, thereby clarifying how our brains solve the Where’s Waldo problem. More about how invariant object categories are learned will be said in the following section.

The experiments of Brown and Denny [176] built upon experiments of Egly, Driver, and Rafal [177] who also studied how visual attention shifts between objects and locations, in both normal individuals and individuals with parietal lesions. Fig. 20 summarizes reaction time, or RT, data, as well as model computer simulations, from four different experimental conditions. These conditions are: (a) no shift of attention, (a) shift of attention to a different location on the same object, (c) shift of attention to a different location outside the object, and (d) shift of attention to a different location on a different object. In each condition, a cue precedes a target. Sometimes both are in the same object, sometimes in different objects, sometimes one outside an object, sometimes both outside an object.

Brown and Denney [176] showed that inter-object (Fig. 20d, left) and object-to-location (Fig. 20c, left) shifts of attention take longer than intra-object shifts (Fig. 20b, left). In all these cases, attention first needs to be engaged at the location of the cue. They also found that shifting attention from one object to another object, or from an object to an-

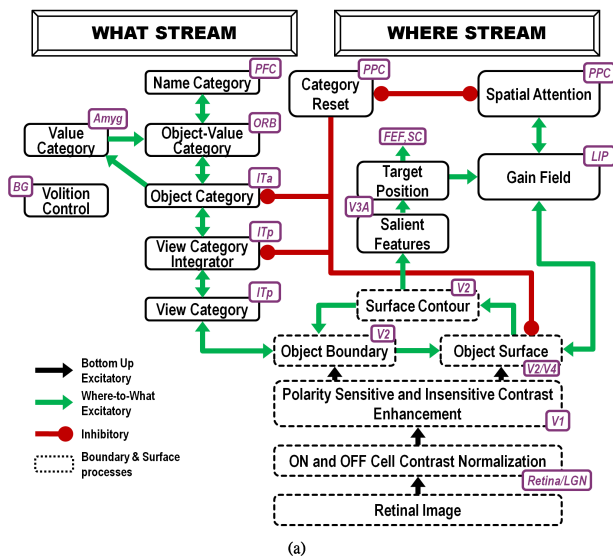


Fig. 18. Interactions of modelled brain regions within the ARTSCAN Search model enable them to learn to recognize and name invariant object categories. Invariant object category learning is modulated by Where-to-What stream interactions from spatial attention in the Where cortical stream—that is sustained by a surface-shroud resonance—and object attention in the What cortical stream—that obeys the ART Matching Rule. Dashed boxes indicate boundary and surface processes. Green arrows carry excitatory cortical signals from the Where stream to the What stream whereby invariant category learning and reinforcement learning occur. Red connections ending in circular disks indicate inhibitory connections. ITa: anterior inferotemporal cortex; ITp: posterior inferotemporal cortex; PPC: posterior parietal cortex; LIP: lateral intraparietal cortex; LGN: lateral geniculate nucleus; ORB: orbitofrontal cortex; Amyg: amygdala; BG: basal ganglia; PFC: prefrontal cortex; FEF: frontal eye fields; SC: superior colliculus; V1 and V2: primary and secondary visual areas; V3 and V4: visual areas 3 and 4. [Reprinted with permission from [5] published in Frontiers.]

other location, takes nearly the same amount of time (369 ± 10 msec versus 376 ± 9 msec, $P > 0.87$ in Fig. 20a, right). The main RT differences thus seem to be because attention does not need to be disengaged from the object as it moves inside it, and it takes longer to disengage attention when it needs to move to a different object or location outside the original object.

In every condition, the cue and target trigger a surface attention signal to their location, thereby leading to the formation of a surface-shroud resonance. The longer reaction times in the inter-object and object-to-location attention shifts compared to intra-object attention shifts are simulated in ARTSCAN by the time it takes for an attentional shroud caused by the cue to collapse, and with it the corresponding surface-shroud resonance, before any other location or object can form a new shroud in response to the target. Reaction time, or RT, in each trial was computed in the model as the time it takes for surface contour or eye movement activity at the target location to reach a prescribed threshold.

The shroud does not collapse in the intra-object case. In-

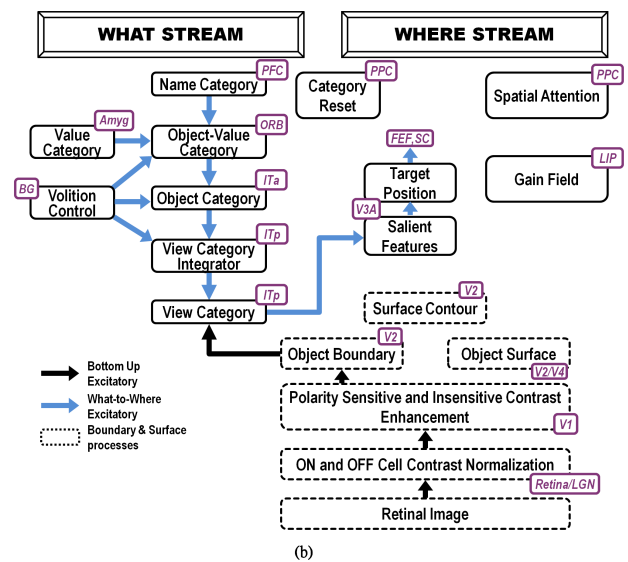


Fig. 19. ARTSCAN Search macrocircuit for Where's Waldo search. A cognitive search for a desired object can be initiated by an object's Name Category in the prefrontal cortex, or PFC. Search for a desired object can also be initiated by a Value Category in the amygdala, or Amyg. Either search can proceed via What-to-Where stream interactions. Black arrows represent bottom-up excitatory input signals. Blue arrows represent top-down excitatory search signals. Abbreviations are the same as in the caption of Figure 18. [Reprinted with permission from [5] published in Frontiers.]

ARTSCAN SIMULATIONS OF INTRA- VS. INTER-OBJECT ATTENTIONAL SHIFTS

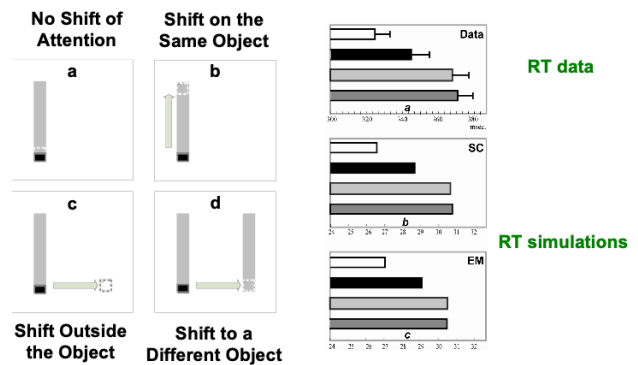


Fig. 20. Data from [176] about intra- vs. inter-object attentional shifts, and computer simulations of it by the ARTSCAN model [7]. [Adapted with permission from [7] published in Elsevier.]

stead, presenting a cue at one end of an object can cause a spread of both surface and boundary attention to the other end of the object. Such a spread of attention has been reported psychophysically by Roelfsema, Lamme, and Spekreijse [178] and simulated by Grossberg and Raizada [179]. Subsequent activation of surface attention by the target at the other end of the same object can add to the baseline of spreading attention, thereby reaching the RT threshold sooner than it can when the target lands outside the object.

Table 6. The ARTSCAN model family summarizes incremental developments of ARTSCAN to explain and predict increasingly large psychological and neurobiological databases. The 3D ARTSCAN Search architecture has not yet been developed, although all of its component architectures have. This synthesis would enable the Where's Waldo Problem to be solved in the natural world of objects in depth.

ARTSCAN MODEL FAMILY

ARTSCAN

View-invariant object category learning and recognition with freely moving eyes in 2D scene Fazl, Grossberg, and Mingolla (2009)

pARTSCAN

View-, position-, and size-invariant object category learning and recognition with freely moving eyes in 2D scene Cao, Grossberg, and Markowitz (2011)

dARTSCAN

ARTSCAN with transient and sustained distributed attention and prefrontal priming Foley, Grossberg, and Mingolla (2012)

ARTSCAN SEARCH

Solution of Where's Waldo Problem:
Directed search to find valued target object in a 2D scene
Chang, Grossberg, and Cao (2014)

3D ARTSCAN

ARTSCAN in 3D scene with 3D perceptual representations that remain stable as the eyes move Grossberg, Srinivasan, and Yazdanbakhsh (2014)

3D ARTSCAN SEARCH Unify all the above

Note: Fazl, Grossberg, and Mingolla [7]. Cao, Grossberg, and Markowitz [70]. Foley, Grossberg, and Mingolla [180]. Chang, Grossberg, and Cao [5]. Grossberg, Srinivasan, and Yazdanbakhsh [35].

2.14 Explaining how invariant categories are learned also explains all the Brown and Denny data

The original ARTSCAN model could simulate only four of the nine experimental conditions that were reported by Brown and Denny [176]. Simulating the entire data set became possible in a consistent extension of ARTSCAN called distributed ARTSCAN, or dARTSCAN (Table 6). dARTSCAN is one of several ARTSCAN developments over the years aimed at explaining ever larger databases about invariant category learning and search (Fig. 21).

The model circuits in Fig. 18 support learning of invariant object categories in the following way: Surface attention upon an object is maintained by a surface-shroud resonance between V2/V4 and PPC during learning of its view-invariant category. The attentional shroud in PPC that is maintained by the surface-shroud resonance inhibits a Category Reset stage that is also in PPC. Inhibition of Category Reset removes inhibition from ITa. The emerging view-invariant category in ITa can thus get associated with all the view-specific categories of the object that are learned in ITp. In this way, only views of the attended object can be incorporated into the view-invariant object category, thereby solving the *view-to-object binding problem*.

The model automatically controls surface attention shifts across the object so that it can inspect multiple object views to learn. These attention shifts use interactions between model cortical areas V2, V3A, V4, and LIP, among other brain regions. Details of these interactions, and data that support

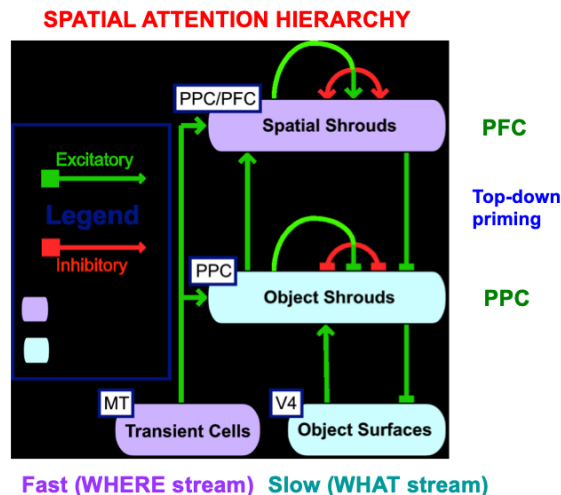


Fig. 21. The *distributed ARTSCAN*, or dARTSCAN, model adds spatial shrouds in the prefrontal cortex, or PFC, and both transient cell inputs from MT and sustained cell inputs from V4 to help form object shrouds and spatial shrouds in PFC and the posterior parietal cortex, or PPC. [Adapted with permission from [180] published in Elsevier.]

each of them, are found in [7].

When surface attention shifts to another object, the previously active surface-shroud resonance collapses, along with further category learning of the previously attended object. ARTSCAN can only learn view-invariant categories (Fig. 18). A consistent extension of ARTSCAN to the *positional ARTSCAN*, or pARTSCAN, model of Cao, Grossberg, and Markowitz [70] can learn view-, position-, and size-invariant object categories. pARTSCAN can, in turn, be extended to dARTSCAN in order to solve the Where's Waldo problem after invariant category learning ends.

COMPLETE BROWN AND DENNY (2007) SIMULATIONS

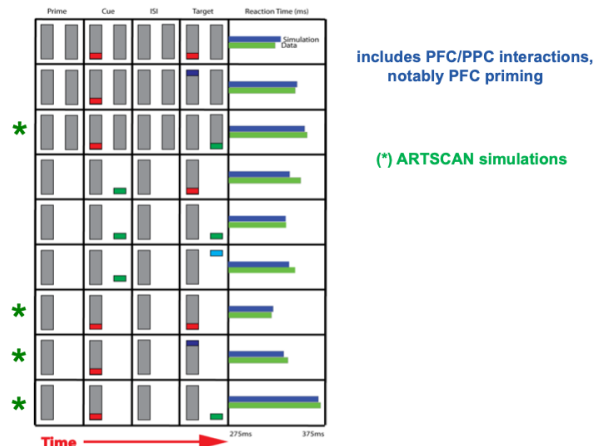


Fig. 22. dARTSCAN can simulate all nine connections from the experiments of Brown and Denny [176], not just the four conditions that ARTSCAN can simulate. [Adapted with permission from [180] published in Elsevier.]

dARTSCAN further developed ARTSCAN to incorporate prefrontal spatial shrouds and transient signals from the Where stream to both PPC and PFC (Fig. 21). These processes, and more (e.g., affective processing regions like amygdala, or Amyg, and volitional control regions like basal ganglia, or BG), are included in the ARTSCAN Search architecture of Figs. 18 and 19. Within both dARTSCAN and the more general ARTSCAN Search architecture, all nine conditions of the Brown and Denny [176] experiments can be simulated (Fig. 22), not just the four that ARTSCAN could simulate (marked by *), in addition to data from many other experiments. Details of these explanations are provided in [5].

3. Adaptive resonance theory: learning to attend, recognize, and predict a changing world

3.1 Stability-Plasticity Dilemma: Life-long learning without catastrophic forgetting

ART is a cognitive and neural theory of how our brains learn to attend, recognize, and predict objects and events in a changing world that may be filled with unexpected events. ART has the broadest explanatory and predictive range of current cognitive and neural theories. Its predictive power derives from its ability, shared with humans, to autonomously carry out fast, incremental, unsupervised and supervised learning and self-stabilizing memory in response to a changing world. ART hereby clarifies how humans can rapidly learn huge amounts of new information throughout life, and to integrate it into unified conscious experiences that support an emerging sense of self.

Fast learning in ART includes the possibility of learning an entire database on one learning trial [71, 73]. Our capacity for fast learning is often taken for granted, as when we see an exciting movie just once and then describe many details about it later to friends and family. In the laboratory, humans have been able to recognize thousands of pictures that they saw just once (e.g., [181–188]). Indeed, the ARTSCENE neural model [25] shows how ART can rapidly learn to classify natural scene photographs, and outperforms alternative models in the literature which use biologically implausible computations. The combination of fast learning and stable memory, upon which all human civilization builds, is an evolutionary achievement that many other popular learning algorithms, such as back propagation and Deep Learning, do not have [42].

ART mechanistically explains how humans can rapidly learn about wide range of novel and changing environments, even with no prior instruction about the statistics of these environments. ART also explains how humans can rapidly learn these things without just as rapidly forgetting them. Neither we, nor ART models, ever need to worry that, by learning to recognize a new friend's face within a few seconds, we will suddenly forget familiar faces of our family and friends. ART hereby avoids *catastrophic forgetting*. Most al-

ternative neural learning algorithms, including back propagation and Deep Learning, do experience catastrophic forgetting [189–191], whether they try fast learning, or slow learning of an environment whose statistics change through time.

Grossberg [12] has called the problem whereby our brains learn quickly without catastrophically forgetting its past knowledge the *stability-plasticity dilemma*. The stability-plasticity dilemma must be solved by every brain system that hopes to adaptively respond to the “blooming buzzing confusion” of signals that we experience each day. ART solves the stability-plasticity dilemma by specifying mechanistic links between processes of between CLEAR processes. Grossberg [11, 12, 192] predicted that all brain processes that solve the stability-plasticity dilemma use CLEAR mechanisms.

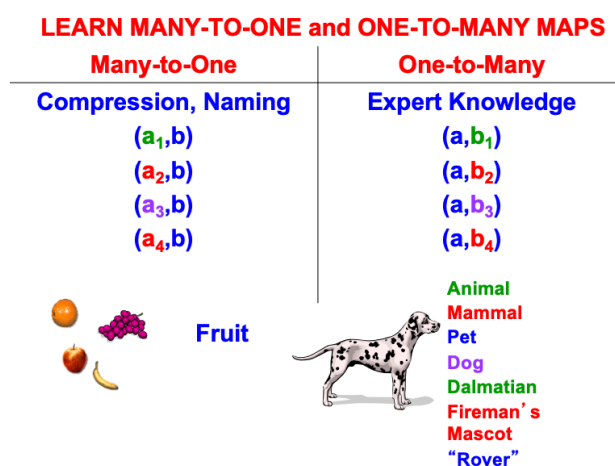


Fig. 23. A supervised ARTMAP system can learn both many-to-one maps and one-to-many maps. [Reprinted with permission from [1] published in Oxford University Press.]

ART uses CLEAR interactions to mechanistically explain why humans are intentional beings who pay attention to salient objects, why “all conscious states are resonant states”, and how we learn both *many-to-one maps* (representations whereby many object views, positions, and sizes all activate the same invariant object category) and *one-to-many maps* (representations that enable us to expertly know many things about individual objects and events). Fig. 23 summarizes what these concepts mean. How ART learns both kinds of maps will be explained after some more background is provided.

Before further explaining ART mechanisms, it is worth emphasizing that *all* of the foundational mechanisms of ART have received increasing support from subsequent psychological and neurobiological data since ART was introduced in [9, 10]. Since then, ART has undergone continual development to explain and predict increasingly large behavioral and neurobiological data bases, ranging from data about normal and abnormal human and animal perception and cognition, to the spiking and oscillatory dynamics of laminar thalamocortical networks in the visual and auditory modalities.

Some of these ART models unify explanations and predictions about behavioral, anatomical, neurophysiological, biophysical, and even biochemical data. ART currently provides (e.g., [17]) functional and mechanistic explanations of such diverse topics as:

- laminar cortical circuitry;
- invariant object and scenic gist learning and recognition;
- prototype, surface, and boundary attention;
- gamma and beta oscillations;
- learning of entorhinal grid cells and hippocampal place cells;
- computation of homologous spatial and temporal mechanisms in the entorhinal-hippocampal system;
- breakdowns of vigilance control during autism and medial temporal amnesia;
- cognitive-emotional interactions that focus motivated attention on valued objects in an adaptively timed way;
- planning and control of sequences of linguistic, spatial, and motor events using item-order-rank working memories and learned list chunks;
- influence of subsequent speech sounds on conscious speech percepts of previous sounds;
- segregation of multiple noisy sources by auditory streaming; and
- normalization of speech sounds to enable learning and imitation of speech from speakers of all ages.

The brain regions that ART models to explain such processes include visual and auditory neocortex; specific and nonspecific thalamic nuclei; inferotemporal, parietal, prefrontal, entorhinal, hippocampal, parahippocampal, perirhinal, and motor cortices; frontal and supplementary eye fields; cerebellum; amygdala; basal ganglia; and superior colliculus.

3.2 Experts learn prototypes of a changing world: Excitatory matching and match learning

As summarized in Table 4, our brains use different predictive mechanisms for perceptual/cognitive and spatial/motor learning that are carried out by computationally complementary cortical processing streams. Perceptual/cognitive processes in the What ventral cortical processing stream often use *excitatory matching* and *match-based learning* to create representations of objects and events in the world. Match-based learning in ART depends upon top-down learned expectations that focus prototype attention to solve the stability-plasticity dilemma. As noted above, this kind of learning can occur quickly without causing catastrophic forgetting. However, as Table 4 shows, match learning, and by extension ART, does not describe the only kind of learning that the brain needs to accomplish autonomous adaptation to a changing world. ART is thus not a “theory of everything”.

3.3 Resonance and learning of attended critical feature patterns

Excitatory matching and attentional focusing on bottom-up data using top-down expectations generates resonant brain states: When there is a good enough match between bottom-up and top-down signal patterns between two or

more levels of processing, their positive feedback signals amplify, synchronize, and prolong their mutual activation (Fig. 5), leading to a resonant state that focuses attention on a subset of features (the category prototype, or critical feature pattern) that can correctly classify the input pattern at the next processing level and lead to successful predictions and actions. Amplification, synchronization, and prolongation of activity triggers learning in the more slowly varying adaptive weights that control the signal flow along pathways between the attended features and the recognition category with which they resonate. Such a resonance embodies a global context-sensitive indicator that the system is processing data worthy of learning, hence the name *Adaptive Resonance Theory*.

ART hereby models a link between the mechanisms which enable us to learn quickly and stably about a changing world, and the mechanisms that enable us to learn expectations about such a world, test hypotheses about it, and focus attention upon information that may predict desired consequences. ART hereby explains how, in order to solve the stability-plasticity dilemma, only resonant states can drive fast new learning.

3.4 Prototype attention obeys the ART Matching Rule: Modulatory on-center, off-surround network

Carpenter and Grossberg [71] mathematically proved that the simplest attentional circuit that solves the stability-plasticity dilemma is a *top-down, modulatory on-center, off-surround network* (Fig. 24). This ART Matching Rule provides excitatory priming of critical features in the on-center, and driving inhibition of irrelevant features in the off-surround. The modulatory on-center emerges from a balance between top-down excitation and inhibition, driven by neurons that obey the membrane equations of neurophysiology.

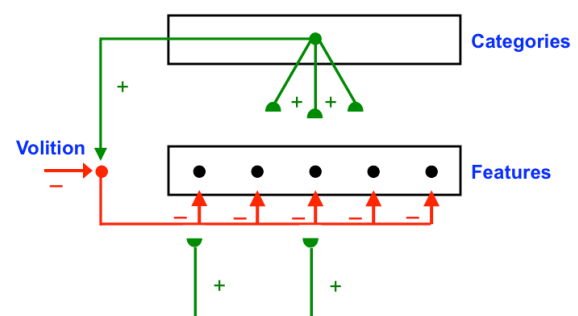


Fig. 24. The ART Matching Rule focuses prototype attention upon critical features that regulate category learning and prediction. It is realized by a top-down, modulatory on-center, off-surround network that also suppresses irrelevant features. See the text for details. [Adapted with permission from [193]. A closely related circuit appears in [71] published in Frontiers.]

3.5 Top-down attentional priming

Attentional priming by the modulatory on-center can be illustrated by the following example: Suppose you are asked to “find the yellow ball as quickly as possible, and you will win a \$100,000 prize”. Activating a top-down, modulatory expectation of a “yellow ball” can subliminally enhance, or “prime”, the activity of its critical feature pattern, without causing these cells to generate suprathreshold activities. If and when the ball does appear, the primed cells can fire more energetically and rapidly. Sensory and cognitive top-down expectations hereby support *excitatory matching* with consistent bottom-up data (Table 4). In contrast, a big enough mismatch between a top-down expectation and bottom-up data can suppress the mismatched part of the bottom-up data, while attention focuses upon the matched, or expected, part of the bottom-up data.

3.6 Data support for the ART Matching Rule

Many anatomical and neurophysiological experiments have provided support for the ART prediction of how attention works, including data about modulatory on-center, off-surround interactions; excitatory priming of features in the on-center; suppression of features in the off-surround; and gain amplification of matched data (e.g., [194–203]). The ART Matching Rule is often called the “biased competition” model of attention by experimental neurophysiologists [204, 205]. The ART Matching Rule property that bottom-up sensory activity may be enhanced when matched by top-down expectations is consistent with neurophysiological data showing facilitation by attentional feedback [178, 200, 206].

The experiments of Sillito *et al.* ([200], pp. 479–482) on attentional feedback from cortical area V1 to the Lateral Geniculate Nucleus (LGN) provided particularly strong support for an early prediction that the ART Matching Rule should act in this circuit [10]. Sillito *et al.* [200] wrote that “the cortico-thalamic input is only strong enough to exert an effect on those dLGN cells that are additionally polarized by their retinal input...the feedback circuit searches for correlations that support the ‘hypothesis’ represented by a particular pattern of cortical activity”. Moreover, the “cortically induced correlation of relay cell activity produces coherent firing in those groups of relay cells with receptive-field alignments appropriate to signal the particular orientation of the moving contour to the cortex...this increases the gain of the input for feature-linked events detected by the cortex”.

In other words, top-down priming, by itself, cannot fully activate LGN cells. Instead, it needs matched bottom-up retinal inputs to do so, and the LGN cells whose bottom-up signals support cortical activity get synchronized and amplified by this feedback.

Additional anatomical studies have shown that the V1-to-LGN pathway realizes a top-down on-center off-surround network [200, 207, 208]. This kind of circuit also occurs during auditory processing: Zhang *et al.* [209] reported that feedback from auditory cortex to the medial geniculate nucleus (MGN) and the inferior colliculus (IC) also has an on-center off-surround form. Temereanca and Simons [210]

have produced evidence for a similar feedback architecture in the rodent barrel system.

The appearance of the ART Matching Rule circuit in multiple modalities can be understood from the need to solve the stability-plasticity dilemma during development and learning in all of them.

3.7 Mathematical definition of the ART Matching Rule

The ART Matching Rule has been represented mathematically in a similar way by more than one lab. In particular, the “normalization model of attention” [211] simulates several types of experiments on attention using the same equation that the distributed ARTEXture (dARTEX) model ([145], equation (A5)) used to simulate human psychophysical data about Orientation-Based Texture Segmentation (OBTS) of Ben-Shahar and Zucker [212]. Reynolds and Heeger [211] defined an algebraic equation for attention with built-in divisive terms to model normalization. Bhatt *et al.* [145] modeled attention using the dynamics of an ART neural network model. When these dynamics reach steady state, the resulting algebraic equation includes a divisive term that is not built explicitly into ART dynamics. The divisive term arises from the automatic gain control, or shunting, terms that occur in Eqn. 1. When such an equation reaches equilibrium, its time derivative dx_i/dt equals zero. Solving for x_i puts the automatic gain control terms into the denominator. The automatic gain controls of a shunting STM equation are thus the cause of normalized attention at equilibrium.

3.8 Imagining, planning, and hallucinating

A balance between excitation and inhibition in the on-center of a top-down expectation is what makes it modulatory. This balance can be modified by volitional gain control signals from the basal ganglia (Fig. 24). For example, if volitional signals inhibit inhibitory interneurons in the on-center, then read-out of a top-down expectation from a recognition category can fire, not merely modulate, cells in the on-center prototype. Such volitional control may control mental imagery and the ability to think and plan ahead before choosing an appropriate action.

A similar modulatory circuit, again modulated by the basal ganglia, is predicted to control temporary storage in the prefrontal cortex of sequences of events in working memory [31] and a task-appropriate span of spatial attention (“useful-field-of-view”) in the parietal and prefrontal cortex [180].

ART predicts that all these properties arise from a circuit design which uses top-down expectations to dynamically stabilize fast learning throughout life. This ability to learn quickly without catastrophic forgetting can then be volitionally modulated by the basal ganglia to support imagination, internal thought, and planning. This modulation has brought huge evolutionary advantages to human civilization.

When basal ganglia modulation breaks down, various clinical disorders can occur. In particular, if these volitional signals become tonically hyperactive, then top-down expectations can fire without overt intention, leading to properties

Table 7. The ART attentional system and orienting system are computationally complementary. The attentional system is where learning and prototype attention occur. It involves brain regions like temporal and prefrontal cortex. The orienting system responds to unexpected and unfamiliar events to drive a memory search, or hypothesis testing, to discover and, if necessary, learn a better matching category. It involves brain regions like the hippocampal system and nonspecific thalamus.

ART COMPLEMENTARY Attentional and Orienting Systems

Attentional System	↔	Orienting System
Expected Events		Unexpected Events
Familiar Events		Unfamiliar Events
Resonance		Reset
Attention		Memory Search
Learning		Hypothesis Testing
Recognition		
Temporal cortex Prefrontal cortex		Hippocampal system Nonspecific Thalamus

like schizophrenic hallucinations [213]. Such mental disorders have probably persisted because the evolutionary advantages of the circuits that allow them are so momentous.

3.9 More Complementary Computing: Resonance and reset controls ART learning and search

Match learning occurs only if a good enough match occurs between bottom-up information and a learned top-down expectation that is read out by an active recognition category. A good enough match can trigger an adaptive resonance that learns a time-average of attended critical feature patterns. Such learning leads to stable memories of arbitrary events presented in any order.

Match learning, by itself, is insufficient for an obvious reason: If learning occurs only when a good enough match occurs between bottom-up data and learned top-down expectations, then how is anything really novel learned? ART solves this problem by using an interaction between complementary processes of *resonance* and *reset* that control attentive learning and memory search, respectively. These processes balance between processing the familiar and the unfamiliar, the expected and the unexpected. When an input is unexpected but familiar, memory search can discover a better matching, but known, category. When an input is unexpected and unfamiliar, search can discover an uncommitted cell population with which to learn a new category.

The resonance process is predicted to take place in an *attentional system* within the What cortical stream, in brain regions like the temporal and prefrontal cortices (Table 7). This system computes processes like resonance, prototype attention, learning and recognition. It is here that top-down ex-

pectations that obey the ART Matching Rule are matched against bottom-up inputs. When a top-down expectation achieves a good enough match with bottom-up data, this match process focuses prototype attention upon those critical features in the bottom-up input that are expected. If the expectation is close enough to the input pattern, then a state of resonance develops as the attentional focus takes hold, which is often realized by oscillatory dynamics that synchronize the firing properties of the resonant neurons [214].

A sufficiently bad mismatch between an active top-down expectation and a bottom-up input activates the complementary *orienting system*, which includes the hippocampal system and the nonspecific thalamus (Table 7). Output signals from the orienting system rapidly reset the recognition category that has been reading out the poorly matching top-down expectation (Fig. 25). The cause of the mismatch is hereby removed, thereby freeing the system to activate a different recognition category. To discover a better matching category, the reset event triggers memory search, or hypothesis testing.

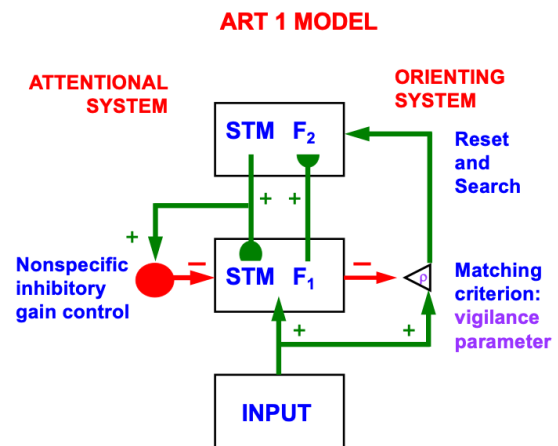


Fig. 25. Macrocircuit of the ART 1 model for learning to classify and stably remember arbitrarily many binary input patterns of an arbitrarily large finite dimension. [Adapted with permission from [193]. A closely related circuit appears in [71] published in Elsevier.]

3.10 The ART hypothesis testing and learning cycle

Fig. 26 describes how the ART hypothesis testing and learning cycle works that is triggered by a reset. In Fig. 26a, a bottom-up input feature pattern, or vector, I (represented by the upward facing green arrow to the first level F_1), activates a pattern X of activity across the feature detectors of F_1 . For example, a visual scene may be represented by its boundary and surface representations [111, 151]. This feature pattern (in yellow at F_1) represents the relative importance of different features in I . Pattern peaks represent more active feature detector cells. This feature pattern activates bottom-up signals S through an adaptive filter to the second level F_2 at which a recognition category Y (or compressed representation of the input) is activated in response to the

ART HYPOTHESIS TESTING AND LEARNING CYCLE

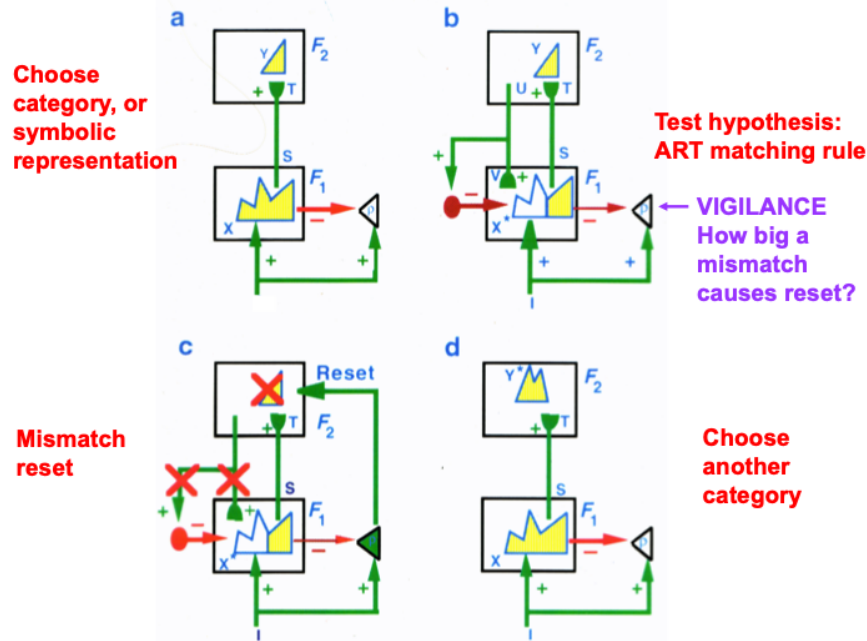


Fig. 26. The ART hypothesis testing and category learning cycle. [Adapted with permission from [193]. A closely related figure appears in [71] published in Elsevier.]

input pattern T . Input T is computed by multiplying the signal vector S by a matrix of adaptive weights, or long-term memory traces, that can be altered through learning to better activate familiar recognition categories. Recurrent competitive interactions across F_2 —notably, recurrent shunting lateral inhibitory signals—allow only a small subset of the most strongly activated cells to remain active in response to T . These active cells are the recognition category Y that represents the pattern of distributed features across level F_1 .

The active category cells, in turn, send top-down signals U back to F_1 (Fig. 26b). This signal vector U is converted into a top-down learned expectation V by being multiplied by another matrix of adaptive weights. Expectation V is delivered to F_1 via a circuit that obeys the ART Matching Rule.

The first time that a recognition category is activated in F_2 , it needs to resonate with its feature pattern in F_1 in order to be learned. But how does a new category in F_2 know what features are currently active in F_1 in order to be able to match and resonate with them? This is achieved by choosing *all initial* top-down adaptive weights to be large and uniformly distributed across the features with which they interact. These initial top-down weights can match *any* feature pattern. Learning *prunes* these initial weights to converge upon the correct critical feature pattern with which to code each category.

With this caveat in mind, it follows that, on every learning trial, matching between the input vector I and V selects that subset X^* of F_1 features that were “expected” by the active F_2 category Y . All of the selected features lie within the modulatory on-center of the ART Matching Rule circuit. These are

the critical features within the learned prototype of category Y .

3.11 Synchronous binding of features and categories during a conscious feature-category resonance

If the top-down expectation matches the bottom-up input pattern well enough, then the pattern X^* of attended features reactivates the category Y which, in turn, reactivates X^* . The network hereby locks into a resonant state through a positive feedback loop that dynamically links, or binds, the attended features across X^* with their category, or symbol, Y . A feature-category resonance (Table 1) results that synchronizes, gain amplifies, and prolongs the activities of the critical features and the category to which they are bound (Fig. 5), triggers learning of the critical features within both the bottom-up adaptive filter and the top-down learned expectation, and supports conscious recognition of the attended visual object or event.

3.12 Resonant bound states overcome two kinds of local ignorance

Individual features at F_1 have no meaning on their own, just as individual pixels in a picture are meaningless. The category Y in F_2 responds selectively to the global patterning of these features, but it does not know what they represent, because it is a compressed, or “symbolic,” representation. Resonance converts the *pattern* of attended features into a coherent, context-sensitive, state that is bound to its category through feedback. This bound state simultaneously represents both the critical features and the category that symbolizes them.

The next few sections describe the implications of the ART hypothesis testing and learning cycle in greater detail to clarify several basic issues about how ART benefits from Complementary Computing.

3.13 Complementary roles of the attentional and orienting systems

The attentional and orienting systems in an ART network play complementary roles during category learning and search for new categories. I will explain below that these complementary roles derive from computationally complementary *laws* that are obeyed by the ART attentional and orienting systems, in addition to the complementary laws that are summarized in Table 3, and the four of these that have already been discussed above.

The reason for this complementary relationship is as follows: When there is a good enough match in the attentional system, the bound state between critical features and their active category embodies information about the features that are currently being processed. However, at the moment when a predictive error occurs, the network does not know *why* the currently active category has failed to predict the correct outcome. Furthermore, when such a predictive mismatch in the attentional system activates the orienting system, the orienting system has no knowledge about *what* went wrong in the attentional system, just *that* something went wrong.

Thus, the attentional system has information about how inputs are categorized, but not whether the categorization is correct, whereas the orienting system has information about whether the categorization is correct, but not about what is being categorized. Given that neither system knows at the time of mismatch what went wrong, how does an ART memory search discover a more predictive outcome?

3.14 Novelty-sensitive nonspecific arousal: Novel events are arousing!

Because the orienting system does not know what category in the attentional system caused the predictive error, its activation needs to equally influence *all* cells in the category level. That is why a big enough mismatch in the orienting system triggers a burst of *nonspecific arousal* that equally activates all cells in the attentional system. In other words: Novel events are arousing! As shown in Fig. 26c, such an arousal burst selectively resets category cells that caused the mismatch, thereby initiating memory search, or hypothesis testing, to discover a more predictive category (Fig. 26d). Such a memory search shifts prototype attention from one category to another.

3.15 Medium-term memory: Habituated transmitters and limited capacity during hypothesis testing

How does a burst of *nonspecific arousal* selectively reset active category cells? Why do not the winning category cells remain winners after arousal acts?

Arousal-mediated category reset is selective because signals within the category level are multiplied, or gated, by habituated chemical transmitters, or MTM traces (see Eqn. 2).

Due to habituated gating, recently activate category cells are more habituated than inactive cells. The habituated category cells compete with other cells at the category level to determine which cells will be chosen next. This is a self-normalizing competition that is carried out by recurrent shunting on-center off-surround interactions across the category level (see Eqn. 1). Self-normalization interacts with habituation to suppress the habituated category cells, even though they were the most active cells when the arousal burst is received. This self-normalization property is also a computational basis of the *limited capacity* of attention (e.g., [87, 215–217]).

Mathematical proofs of how reset, or *antagonistic rebound*, happens can be found in [12] (Appendices A and E) and [14] (Appendix). Once the maximally activated cells are suppressed by a combination of habituation and competition during the search cycle, cells that initially got smaller inputs than the original winning cells can inherit the self-normalizing network activity that the winning cells no longer take up, and thereby become more active in the next time interval. This cycle of mismatch-arousal-reset continues until resonance can again occur.

Because the total activity of the category cell network is self-normalizing, the category level activations can be interpreted as a kind of “real-time probability distribution”, and the ART memory search cycle can be interpreted as a kind of probabilistic hypothesis testing and decision making that is competent in response to arbitrarily complicated non-stationary time series of input patterns. These properties are what makes ART a self-organizing *production system* in a changing world [42].

3.16 Processing Negativity in the attentional system is complementary to N200 in the orienting system

Psychological and neurobiological data have supported ART predictions about how the memory search process works. A particularly compelling kind of data have been derived from experiments using event-related potentials, or ERPs, that are measured from a human subject using arrays of scalp potentials.

For example, the predicted sequence of mismatch, arousal, and category reset operations during an ART search have been experimentally predicted, recorded, and modeled by Banquet and Grossberg [3] as sequences of P120, N200, and P300 ERPs during oddball experiments. During such an experiment, a human subject is primed to detect rare targets within sequences of more frequent distractors (Fig. 27).

In addition to ERP events that occur during search, the Processing Negativity, or PN, ERP has been recorded during sustained prototype attention [218, 219] while an ART feature-category resonance occurs.

The N200 ERP that is triggered by mismatch-mediated activation of the orienting system [220–222] has properties that are computationally complementary to those of PN, as summarized in Fig. 28. The complementary properties of PN (within the attentional system) and N200 (within the ori-

ERP SUPPORT FOR MISMATCH-MEDIATED RESET

Event-Related Potentials: Human Scalp Potentials

ART predicted correlated sequences of P120-N200-P300

Event Related Potentials during oddball learning

P120 - mismatch; N200 - arousal/novelty; P300 - STM reset

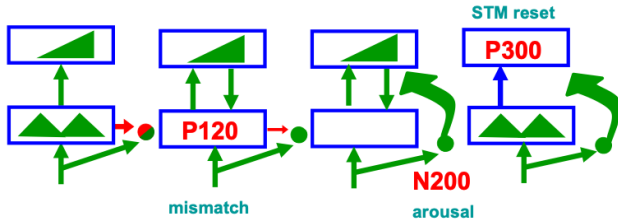


Fig. 27. Event Related Potential, or ERP, experimental support for the ART hypothesis testing cycle from [3]. [Reprinted with permission from [1] published in Oxford University Press.]

enting system) make precise the sense in which the ART attentional and orienting systems are computationally complementary.

COMPLEMENTARY COMPUTING IN ART

Complementary Event-Related Potentials

PN AND N200 ARE COMPLEMENTARY WAVES

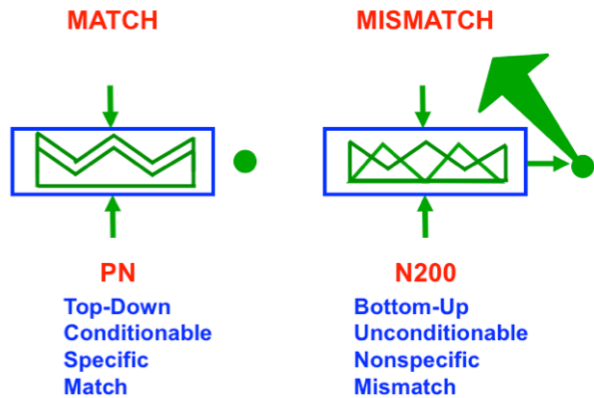


Fig. 28. The processing negativity (PN) and N200 ERPs illustrate the fact that the ART attentional and orienting systems are computationally complementary. [Adapted with permission from [19] published in Elsevier.]

3.17 Neurophysiological data about attentional and orienting systems from multielectrode experiments

ERP data have been supplemented by neurophysiological data about attentional and orienting system interactions. For example, Brincat and Miller ([223], p. 576) have reported neurophysiological data from prefrontal cortex (PFC) and hippocampus (HPC) in monkeys learning object-pair associations: "PFC spiking activity reflected learning in parallel with behavioral performance, while HPC neurons reflected feedback about whether trial-and-error guesses were correct or incorrect. Rapid ob-

ject associative learning may occur in PFC, while HPC may guide neocortical plasticity by signaling success or failure via oscillatory synchrony in different frequency bands".

PFC is a projection area of the inferotemporal, or IT, cortex where categories begin to be learned. These data thus add to a substantial experimental literature that have implicated the hippocampus as part of the orienting system for match/mismatch processing. See [19] for additional discussion of this literature.

3.18 Vigilance: A criterion for resonance vs. reset that determines category generality

How specific and concrete, or general and abstract, will learned ART categories be? On what combination of critical features will categories learn to focus prototype attention?

This is determined by how good a match is needed in order for resonance, attention, learning, and consciousness to occur. The matching criterion is determined by the size of a *vigilance parameter* ρ that is computed within the orienting system (Fig. 26b; [71]). If vigilance is chosen high, then even small mismatches can trigger search. As a result, only very good matches can resonate with a given category, so only a specific or concrete category prototype is learned, such as one that recognizes a frontal view of your mother's face. If vigilance is low, then even poor matches can resonate with a given category, so that multiple exemplars can be incorporated into the learned prototype, which hereby learns an abstract or general category, such as one that recognizes any face.

VIGILANCE CONTROL

$$\begin{aligned} \rho|I| - |X| &\leq 0 & \rho &\leq \frac{|X|}{|I|} & \text{resonate and learn} \\ \rho|I| - |X| &> 0 & \rho &> \frac{|X|}{|I|} & \text{reset and search} \end{aligned}$$

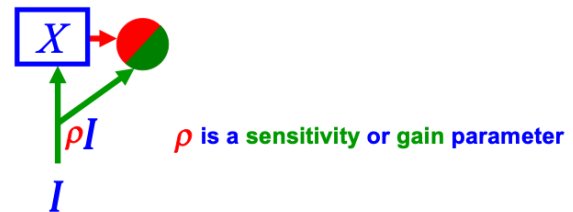


Fig. 29. How vigilance determines whether the attentional system will resonate and learn, or the orienting system will be activated to drive memory search and hypothesis testing. [Adapted with permission from [1] published in Oxford University Press.]

The vigilance parameter ρ multiplies the bottom-up excitatory inputs I to the orienting system. Vigilance is thus the *gain* of these excitatory inputs. The orienting system is activated when the total excitatory input ρI is greater than the total inhibition from the attended critical features X^* across

F_1 that survive top-down matching. This occurs when $\rho|I| - |X^*| > 0$ (Fig. 29), where $| \cdot |$ denotes the number of positive inputs $\rho|I|$ or matched critical features $|X^*|$. Rewriting this inequality as $\rho > |X^*||I|^{-1}$ shows that the orienting system is activated when ρ is larger than the ratio of the number $|X^*|$ of critical features in F_1 to the total number $|I|$ of features in I . If $\rho \leq |X^*||I|^{-1}$, then the orienting system remains quiet, so that resonance and category learning can proceed. Vigilance hereby controls how bad a match can be before reset of the current category, and search for a new category, is initiated.

Vigilance is typically chosen as low as possible to learn the most general possible categories, and thereby conserve memory resources, but without causing a reduction in predictive success. The baseline vigilance level is therefore set initially at the lowest level that has led to predictive success in the past. When a given task requires finer discriminations, vigilance is raised.

Extend UNSUPERVISED ART to SUPERVISED or UNSUPERVISED ARTMAP

FUZZY ARTMAP

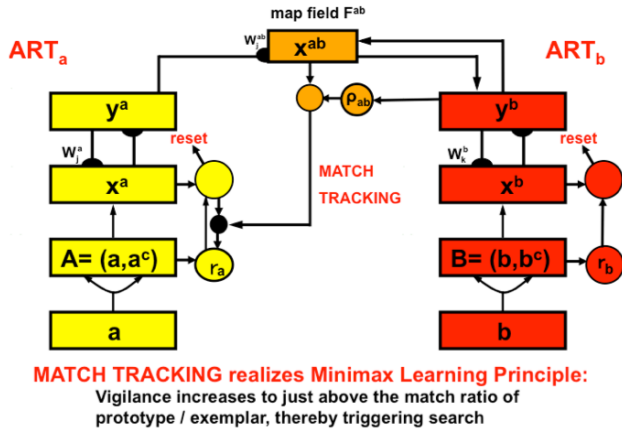


Fig. 30. ARTMAP systems can carry out arbitrary combinations of unsupervised or supervised learning during learning trials. During supervised learning, two unsupervised ART systems, ARTa and ARTb, are linked by a learned associative map which enables learning a map from arbitrarily large databases of m -dimensional vectors to n -dimensional vectors, where m and n can also be chosen arbitrarily large. [Adapted with permission from [75] published in Frontiers.]

3.19 From ART unsupervised category learning to ARTMAP supervised or unsupervised map learning

Predictive successes and failures can be computed in an ARTMAP system that is capable of learning to predict any number of arbitrarily large output vectors in response to any number of arbitrarily large input vectors ([75, 77]). An ARTMAP system is composed of two unsupervised ART systems, ARTa and ARTb, that are linked by associative learning in a map field (Fig. 30). By pairing an input vector \mathbf{a} with an output vector \mathbf{b} , an ARTMAP system can learn to associate input vector \mathbf{a} with output vector \mathbf{b} , thereby becoming a supervised mapping system. Input vector \mathbf{a} can learn to predict

MANY-TO-ONE MAP Two Stages of Compression Medical Database Prediction

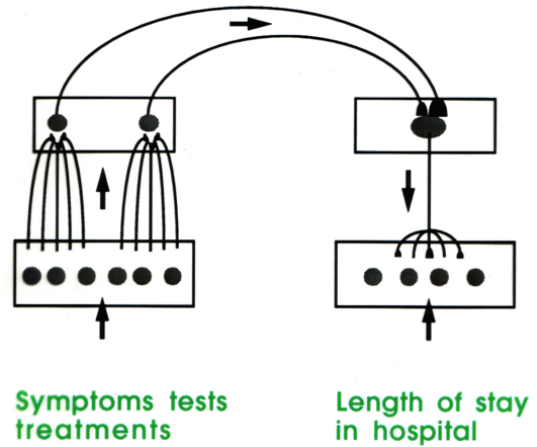


Fig. 31. ARTMAP systems can learn both many-to-one maps and one-to-many maps. Learning of a many-to-one map is illustrated in this figure. ARTa is trained to classify combinations of symptoms, tests, and treatments of medical disorders. ARTb is trained to classify properties of hospital care. Predictive errors in ARTb can drive memory searches in ARTa to focus attention upon, and learn to classify, predictive combinations of critical features about disease measurements with which to learn correct predictions about hospital care. [Reprinted with permission from [1] published in Oxford University Press.]

output vector \mathbf{b} because of the reciprocal bottom-up and top-down learned connections in both ARTa and ARTb.

ARTMAP clarifies how many-to-one maps and one-to-many maps are learned. Figs. 31 and 32 show one example of each. In the many-to-one example in Fig. 31, arbitrarily many, arbitrarily large vectors of medical symptoms, tests, and treatments can learn to predict multiple different medical outcomes of these combined constraints, including the estimated length of stay in the hospital. Fig. 32 illustrates how ARTMAP can learn a one-to-many map. For example, after learning to predict that the image on the bottom left is a dog, seeing that image leads to the response “dog”. However, if instead the answer “Rover” is given, then the mismatch between “dog” and “Rover” triggers an ART search in which prototype attention shifts to focus on that combination of features in the image that predicts the particular dog “Rover”.

3.20 Minimax learning by Match Tracking: Learning the most general predictive categories

Vigilance is typically raised in response to a predictive failure, or disconfirmation, because it must have been lower than the previously computed match value $|X^*||I|^{-1}$ in order to make the incorrect prediction in the first place (Fig. 29 and Fig. 33a). Increasing vigilance until it just exceeds the match value triggers a memory search for a better-matching category (Fig. 33b). Such an increase sacrifices that minimum amount of generalization in order to correct the predictive error. This operation is called *match tracking*.

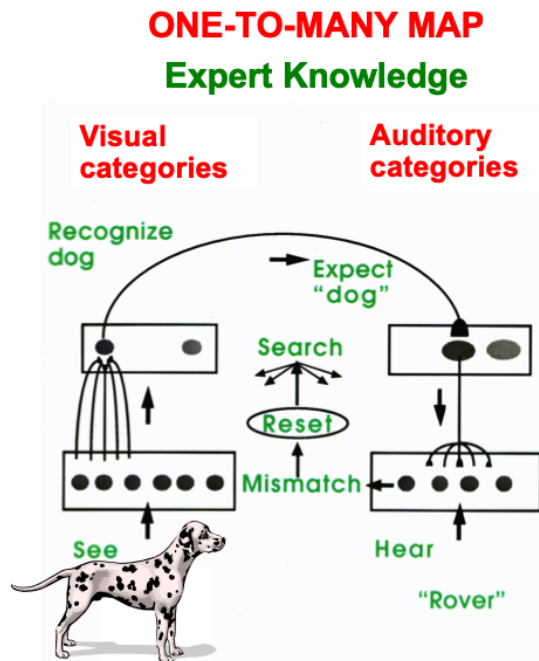


Fig. 32. An example of how a one-to-many map is learned by ARTMAP: After the association between the image and the name “dog” has been learned, showing the image on a subsequent trial with the teaching signal “Rover” causes a predictive mismatch between “dog” and “Rover” that drives a memory search which ends by focusing prototype attention upon a combination of critical features that is sufficient to recognize “Rover”. These critical features drive learning of a new category that is associated with the name “Rover”. An arbitrarily number of such associations can be learned in response to the image. These learned links can then be organized by higher categories that code all of this knowledge about the image. [Reprinted with permission from [1] published in Oxford University Press.]

MATCH TRACKING realizes MINIMAX LEARNING PRINCIPLE
Given a predictive error, vigilance increases just enough to trigger search and thus sacrifices the minimum generalization to correct the error

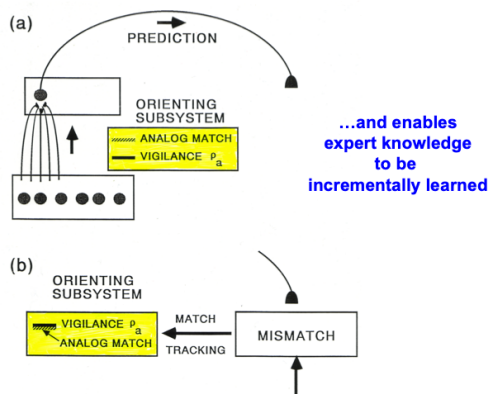


Fig. 33. Minimax learning by matching tracking minimizes predictive error while conjointly maximizing category generality. See the text for details. [Reprinted with permission from [42] published in Frontiers.]

Match tracking helps to explain how one-to-many maps are learned. It also helps to explain how predictive errors

are corrected. For example, a predictive error could occur if a viewer classifies an object as a dog, whereas it is really a wolf. Within ART, such a predictive disconfirmation causes an increase in vigilance that triggers a memory search. The memory search causes a shift in prototype attention to focus on a different combination of features that can successfully be used to recognize a wolf, and perhaps to recognize other similar looking wolves as well.

The vigilance increase in Fig. 33b is called *match tracking* because vigilance tracks the match $|X^*|/|I|^{-1}$ between the input pattern I and the attended feature pattern X^* . Match tracking is a form of *minimax learning*, or learning that *minimizes* predictive errors while it *maximizes* category generality. In this way, match tracking uses the minimum memory resources that can correct predictive errors. Because match tracking may change vigilance on each trial in order to maintain predictive success, ARTMAP can learn databases whose recognition categories differ greatly in their specificity or generality.

4. Laminar computing and the LAMINART model: Attention in laminar neocortical circuits

All the neurons in cerebral cortex are organized using a shared design of *layered* circuits whose specializations carry out all forms of higher-order biological intelligence [224, 225]. As noted in [192], this design realizes a revolutionary computational synthesis of the best properties of feed-forward and feedback processing, digital and analog processing, and data-driven bottom-up processing and hypothesis-driven top-down processing. Top-down processing including the focusing of attention.

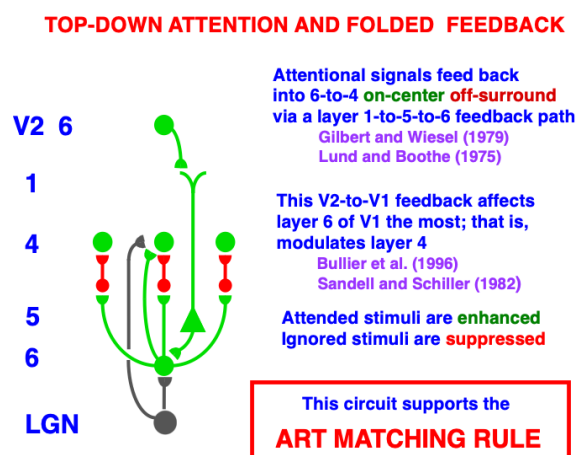


Fig. 34. Laminar circuit that carries prototype attention from layer 6 of cortical area V2 to layer 4 of cortical area V1 via a top-down, modulatory on-center, off-surround network that embodies the ART Matching Rule using a folded feedback circuit from layer 6 in V2 to layers 5-to-6-to-4 in V1. [Reprinted with permission from [1] published in Oxford University Press.]

The understanding of ART, and attention, incrementally developed until Grossberg [226] introduced the LAMINART model to show how bottom-up adaptive filtering, horizontal grouping, and top-down attentional interactions may be combined in the laminar circuits of neocortex. This article introduced the computational paradigm of Laminar Computing, which shows how predicted ART mechanisms may be embodied within known laminar cortical circuits, and solves a long-standing open problem, called the *attention-preattention interface problem*, thereby enabling the explanation and prediction of much more cognitive and brain data than before. More about the attention-preattention interface problem will be described in the next section.

This laminar synthesis combined ART as a theory of category learning and prediction—which emphasized bottom-up and top-down interactions within higher cortical areas such as V4, inferotemporal cortex, and prefrontal cortex—with the FACADE (Form-And-Color-And-DEpth) theory of 3D vision and figure-ground perception—which emphasized bottom-up and horizontal interactions for sensory filtering, followed by completion of boundaries and filling-in of surface brightness and color, within lower cortical processing areas such as V1, V2, and V4.

The unification of these two research streams in LAMINART proposed how *all* cortical areas combine bottom-up, horizontal, and top-down interactions, thereby beginning to functionally clarify why all granular neocortex has a characteristic architecture with six main cell layers [175]. In particular, this unification led to laminar neural models within variations of a shared canonical cortical design could be used to explain psychological and neurobiological data about vision, speech, and cognition. See [17] for a review.

4.1 Prototype attention, folded feedback, and multiplexed filtering, grouping, and attention

Fig. 34 shows the cortical circuit that realizes prototype attention within the visual cortex. Note that a top-down signal from layer 6 of a higher cortical area like V2 feeds topographically down to layer 6 of V1, either directly or via apical dendrites of layer 5 cells (as shown), before folding back up to layer 4 as a modulatory on-center, off-surround network. Prototype attention in visual cortex is thus realized by a *folded feedback* circuit.

Fig. 34 also shows that bottom-up filtering signals from the lateral geniculate nucleus, or LGN, to V1 and top-down attention from V2 to V1 share the modulatory on-center, off-surround decision circuit between layers 6 and 4. In this way, bottom-up input patterns can have critical features selected by top-down prototype attention in layer 4 of V1.

Fig. 35 shows that, in addition, long-range horizontal interactions in layer 2/3 (shown in black), which carry out boundary completion and perceptual grouping, also share the same decision circuit between layers 6 and 4. See, in particular, the excitatory pathway from layer 2/3 to layer 6.

Thus all the constraints of bottom-up adaptive filtering, horizontal grouping, and top-down attentive expectations

GROUPING AND ATTENTION SHARE THE SAME MODULATORY DECISION CIRCUIT
Layer 6-6-4-2/3 pathway shown; also a layer 6-1-2/3 path

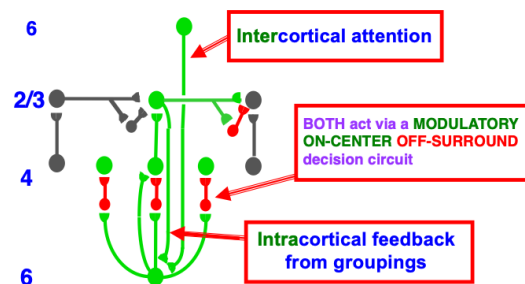


Fig. 35. Laminar circuit that joins the prototype attention circuit with a boundary completion circuit: Both boundary grouping and prototype attention share the same layer 6-to-4 decision circuit that uses modulatory on-center, off-surround interactions to choose the grouping that best balances between grouping and attentional constraints. The horizontal connections in layer 2/3 that complete boundary groupings use signals from layer 2/3-to-6-to-4, where they interact with prototype attentional signals. [Reprinted with permission from [1] published in Oxford University Press.]

share the same decision circuit between layers 6 to 4, thereby multiplexing all of their constraints before the best attentional focus and perceptual grouping is chosen in the current informational context. I therefore call this circuit the *attention-preattention interface*.

4.2 SMART: Resonance vs. reset lead to gamma vs. beta oscillations

The Synchronous Matching ART, or SMART, model [214] extends the LAMINART model to include hierarchies of laminar neocortical circuits that use spiking neurons to interact within and between each other, and with specific and nonspecific thalamic nuclei (Fig. 36). SMART explains and predicts the functional meaning of a larger body of anatomical data, of synchronized cortical interactions, and of the biochemical substrates of vigilance control. One anatomical refinement in Fig. 36 models different roles for cortical layers 6^I and 6^{II} in the laminar cortical circuit for prototype attention.

SMART simulations of synchronized cortical oscillations show how a good enough top-down attentive match may lead to fast gamma oscillations that support prototype attention and spike-timing dependent plasticity (STDP). In contrast, a big enough mismatch may generate slower beta oscillations. This match-mismatch gamma-beta story seems to occur in quite a few brain systems, with data supporting this prediction having recently been reported in V1, frontal eye fields, and hippocampus, as reviewed in [17].

4.3 Vigilance control by nucleus basalis acetylcholine release in health and disease

Of particular interest is the SMART prediction about how vigilance is controlled in our brains. As shown in Fig. 37, a big enough mismatch in the nonspecific thalamus activates

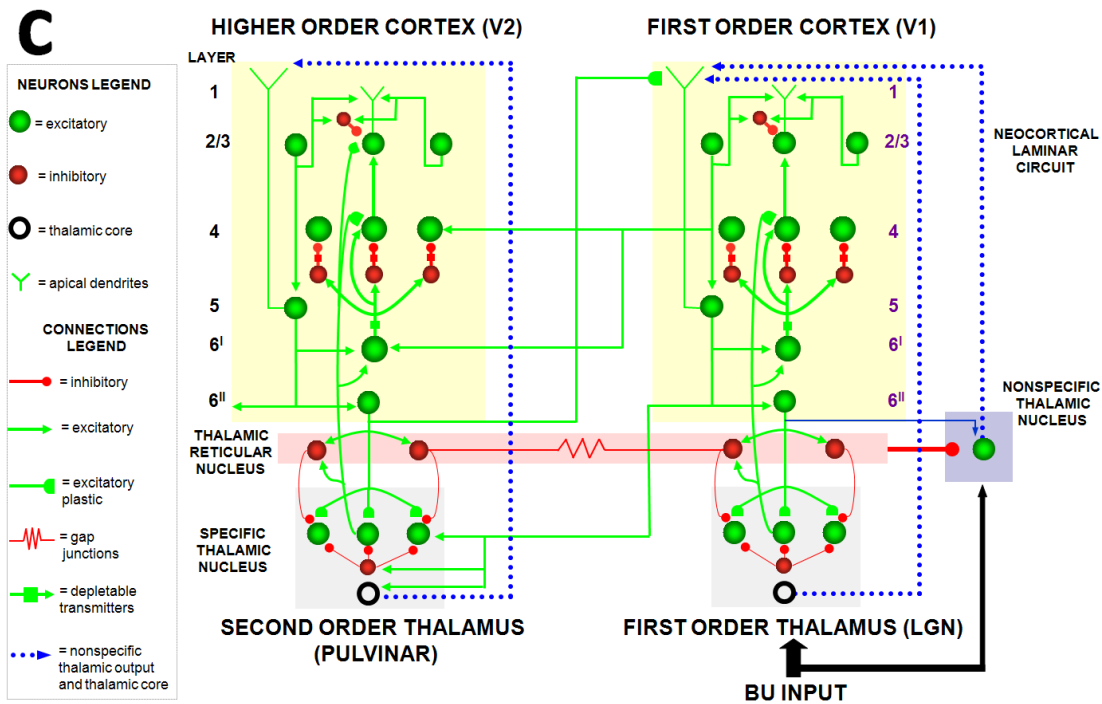


Fig. 36. Synchronous Matching ART, or SMART, microcircuit showing a hierarchy of laminar cortical areas interacting with both specific and nonspecific thalamic nuclei in order to incrementally learn a hierarchy of recognition categories. [Adapted with permission from [214] published in Elsevier.]

VIGILANCE CONTROL: MISMATCH-MEDIATED ACETYLCHOLINE RELEASE

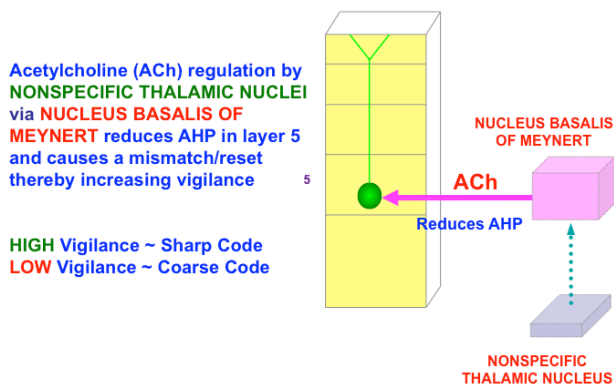


Fig. 37. Circuit showing how a big enough mismatch activates the nonspecific thalamus, which activates the nucleus basalis of Meynert, which releases acetylcholine, or ACh, in layer 5 of multiple cortical areas. The ACh reduces afterhyperpolarization, or AHP, currents and thereby causes an increase in vigilance. [Adapted with permission from [214] published in Elsevier.]

the nucleus basalis of Meynert [120], which, in turn, broadly releases acetylcholine, or ACh, in layer 5 of multiple cortical areas. ACh release reduces after-hyperpolarization (AHP) currents, which cause a transient increase in vigilance. If this increase is large enough (see Fig. 29), then it can disrupt an ongoing resonance to trigger a memory search, ending in se-

lection or learning of a category with a new focus of prototype attention. Many data about the role of ACh in vigilance control and reset are summarized and modeled in [227].

In normal individuals, this vigilance circuit controls processes like match tracking. In various mental disorders—such as Alzheimer’s disease, autism, amnesia, and disordered sleep—the circuit can break down in different ways, leading to behavioral symptoms of each disorder. Grossberg [18] describes what these symptoms are and how they may arise due to ACh malfunctions.

4.4 SACCART: Laminar circuits in SC choose spatial attentional targets for saccadic eye movements

After spatial attention orients to the position of a desired goal object, movements must follow to acquire it. One of the most important kinds of movements are saccadic, or ballistic, eye movements to foveate the target, followed by leg movements to navigate towards it, and arm movements to reach it.

The superior colliculus, or SC, contains a map of possible target positions. It is a multimodal map in which visual, auditory, and planned movement inputs compete for attention, leading to selection of a winning position that controls where the next saccadic eye movement will go. Visual and auditory signals are computed in different coordinates: Visual targets are registered on the retina, so are initially computed in retinotopic coordinates. The position in space that this retinal activation represents depends upon where the eyes are looking when they receive this retinal input. Both the posi-

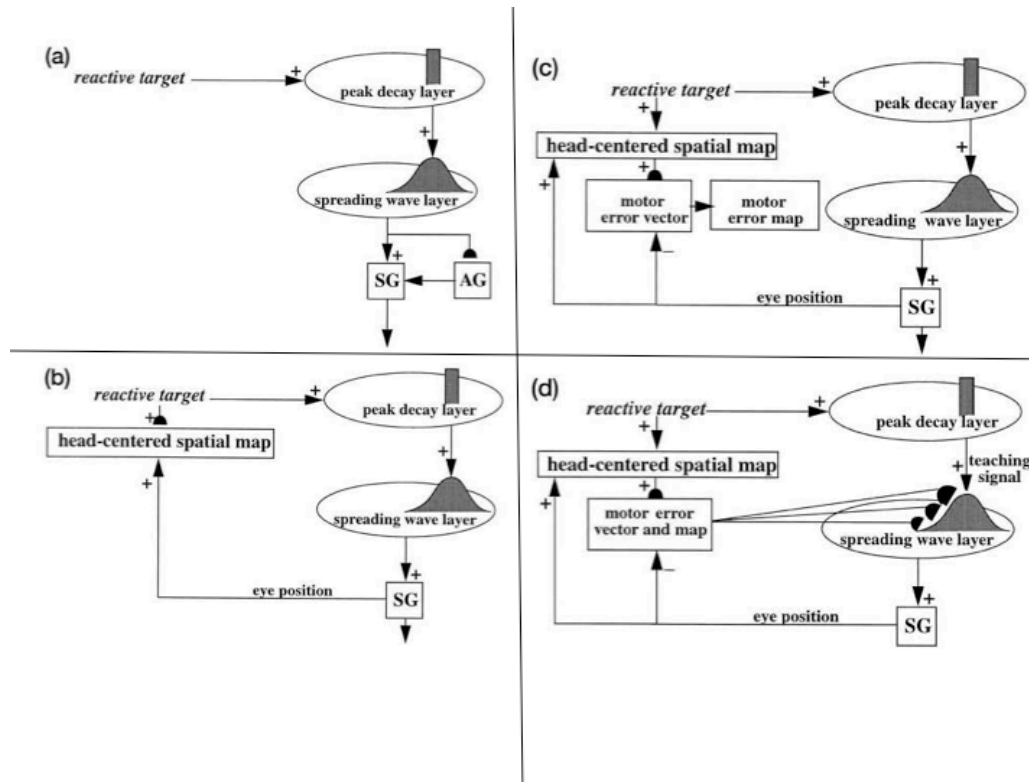


Fig. 38. (a) Retinal activation by a visual signal generates a reactive visual target signal to the peak decay, or burst, layer of the superior colliculus, or SC. This signal topographically activates the spreading wave, or buildup, layer of the SC. Other things being equal, the buildup cells activate the saccade generator, or SG, in the peripontine reticular formation to generate a saccadic eye movement. This movement is calibrated by learned adaptive gain control signals from the cerebellum, or adaptive gain (AG) stage. (b) When the SG fires, it generates both a signal to move the eye and a parallel corollary discharge signal that codes the current eye position in motor coordinates. The corollary discharge signal is added to the retinotopically-coded position of the target to define a head-centered spatial map of the target. (c) The head-centered representation of target position is converted into an error vector that is coded in motor coordinates when a corollary discharge of the current eye position, also coded in motor coordinates, is subtracted from it. This motor error vector codes the direction and distance that the eye needs to move to foveate the target. The hemi-disk denotes an adaptive synapse where head-centered and motor coordinates are rendered dimensionally consistent. This learning occurs when the eye is steadily looking at stationary target positions, so that the desired and current eye positions are the same. The motor error vector that is calibrated in this way is then converted into a motor error map. (d) The motor error map learns to activate the SC position of the target using inputs from the peak decay layer as teaching signals. The learned transformation from reactive target position to commanded SC saccadic eye movement is then complete. [Reprinted with permission from [228] published in Wiley.]

tion on the retina of the target, and the position of the eyes in their orbit, thus combine to compute target position. Auditory signals are computed in head-centered coordinates, since the ears move with the head.

Learning within the deeper layers of the SC converts these different coordinates into a shared coordinate system using visual inputs as a teaching signal. That is why both auditory and planned inputs get converted into a retinotopic coordinate system, that also codes a *motor error map* because each of its positions codes the direction and distance that a saccade must move in order to foveate a target at that position (Fig. 38). This learning process ensures that auditory, visual, and prefrontal planning inputs that represent the same position in space activate the same SC map position with which to select the target position of the next saccade. Multiple auditory, visual, and cognitive inputs to the SC compete through time to determine this winning target position. As in the case of cognitive category attention, learning, and choice, this

competition is realized by a recurrent on-center off-surround network.

The SACCART model [228] models multimodal map learning and choice within the deeper layers of the SC. Model interactions quantitatively simulate the temporal dynamics of SC burst and buildup cells under a variety of experimental conditions (Fig. 39). *Burst cells* respond with bursts that decay as the next saccadic position is chosen and performed. *Buildup cells* generate a spatially distributed pattern of activity that begins at the chosen position and then spreads towards the fovea as they command the eye to foveate.

This SC model is called SACCART because it is a specialized ART circuit whose map learning may be viewed as a kind of attentive motor category learning that takes place using bottom-up and top-down interactions between burst and buildup cells (Fig. 40). These ART dynamics also allow resonance to occur between burst and buildup cells when they choose the same target position (Fig. 40a), and reset within

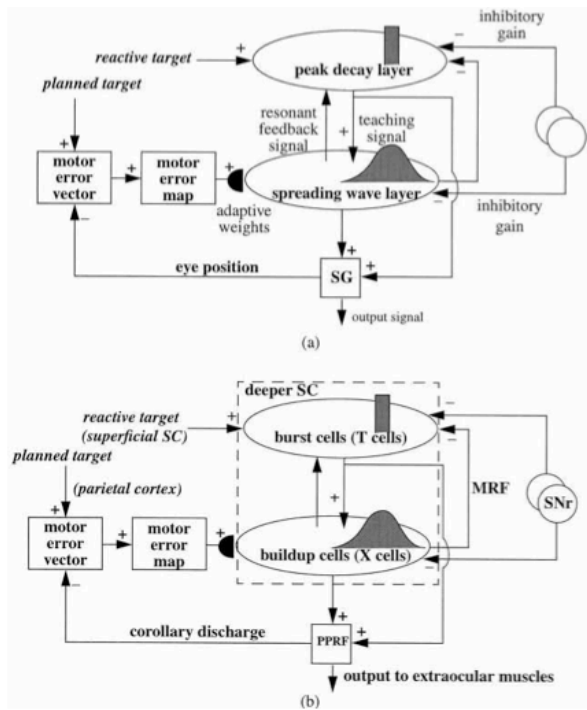


Fig. 39. Comparison of model processes with anatomical and neurophysiological correlates. (a) Schematic representation of model interactions. (b) Data correlates of the model depicted in (a). SNr: substantia nigra pars reticulata; PPRF: peripontine reticular formation; MRF: mesencephalic reticular formation; peak decay cells: burst, or T, cells; spreading wave cells: buildup, or X, cells. [Reprinted with permission from [228] published in Wiley.]

burst cells when a planned and reactive cue represent different possible saccadic targets (Fig. 40b).

This map enables visual, auditory, and planned movement commands to compete for attention, leading to selection of a winning position that controls where the next saccadic eye movement will go. Such map learning may be viewed as a kind of attentive motor category learning.

Grossberg, Palma, and Versace [227] extend this analysis to model how acetylcholine plays a role in SC map learning that is homologous to its role in regulating vigilance control during the learning of ART recognition categories, thereby establishing a mechanistic link between attention, learning, and cholinergic modulation during decision making within both cognitive and motor systems. The article also explains mechanistic homologs between the mammalian superior colliculus and the avian optic tectum [234–236], leading to predictions about how multimodal map learning may occur in the mammalian and avian brain and how such learning may be modulated by acetylcholine.

These results show how ART dynamics help to carry out self-stabilizing category learning, modulated by acetylcholine vigilance control, in all neocortical regions as well as in subcortical circuits devoted to movement control.

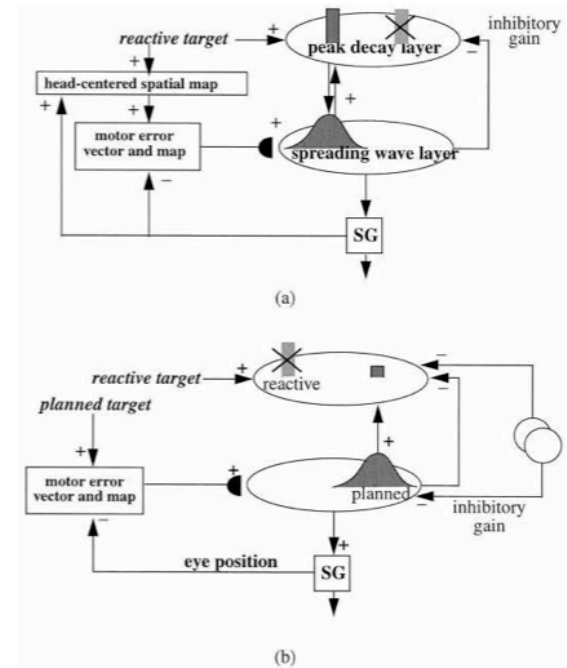


Fig. 40. (a) Topographically organized excitatory feedback between the spreading wave and peak decay layers triggers a resonance that drives the SC map learning process. A nonspecific inhibitory gain signal from the spreading wave layer reaches all target positions in the peak decay layer via the mesencephalic reticular formation [229–231]. Together these excitatory and inhibitory interactions embody the ART Matching Rule, thereby supporting the SACCART model name. An attended target position is chosen by competitive interactions within these layers, while also suppressing positions of irrelevant targets, thereby stabilizing map learning by preventing irrelevant targets from being learned. **(b)** Auditory, planned, and visual targets compete for attention in the deeper layers of the SC [232, 233]. Auditory or planned targets are received in the SC deeper layers, whereas reactive target signals are received by the superficial layers which then project them topographically to the deeper layers. The chosen target position blocks interruptions from other targets during its execution. When auditory or planned target positions and visually reactive target positions agree, then map learning is reinforced. When the auditory or planned target positions and visually reactive target positions disagree, then learning between these different representations is suppressed, and the distracting target is prevented from interrupting the saccade. Open circles are sources of inhibition from the substantia nigra pars reticulata, or SNr, of the basal ganglia which opens gates in the SC to enable targets at those positions to be performed. Rostral migration of activity in the spreading wave layer from its original position erodes feedback excitation to the burst cell layer at which visually reactive targets are stored, thereby leading to decay of activity at the peak decay layer, and justifying its name. [Reprinted with permission from [228] published in Wiley.]

5. All conscious states are attentive resonant states, but not conversely: the role of qualia

Many different behaviors have been linked to attentive resonances in specific parts of the brain, as reviewed in [17, 18]. Several of these resonances and their attentional substrates are summarized in Table 1, as part of the general prediction that “all conscious states are resonant states”. Given

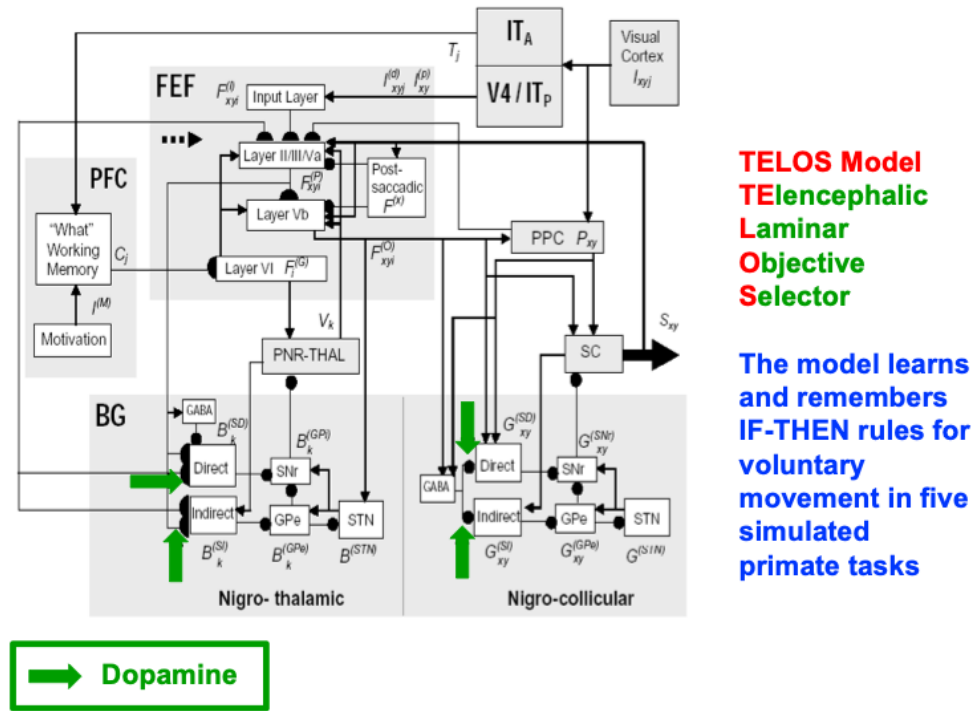


Fig. 41. The TELOS, or Telencephalic, Laminar, Objective, Selector, model is a laminar model that simulates interactions of the basal ganglia, or BG, with the frontal eye fields, or FEF, prefrontal cortex, or PFC, and superior colliculus, or SC. Separate gray-shaded blocks highlight the major anatomical regions whose roles in planned and reactive saccade generation are explained and simulated in the model. Excitatory links are shown as arrowheads, inhibitory as filled-in disks. Filled semi-circles terminate corticostriatal and corticocortical pathways that can adapt due to learning, which is modulated by reinforcement-related dopaminergic signals (green arrows). In the FEF block, Roman numerals I-VI label cortical layers; Va and Vb, respectively, are superficial and deep layer V. The BG decide which plan to execute and send a disinhibitory gating signal that allows thalamic activation V_k , which excites FEF layer Vb output cell activities $F^{(O)}_{xyi}$ to execute the plan. The model distinguishes a thalamus-controlling BG pathway, whose variables are symbolized by B , and a colliculus-controlling pathway, whose variables are symbolized by G [237]. See the original article for further details. [Adapted with permission from [238] published in Elsevier.]

that each kind of resonance activates some form of top-down attention and learning, a fundamental link is hereby shown between the CLEARs processes of Consciousness, Learning, Expectation, Attention, Resonance, and Synchrony.

The resonances in Table 1 may all be described using familiar words from daily language. That is true because each of them includes internal representations of external qualia, like brightness and color, or internal qualia, like hunger and pain, which become conscious during the resonance. Some functionally important resonances do not include internal representations of qualia, are not conscious, and thus do not have readily available names from day-to-day language.

For example, parietal-prefrontal resonances are predicted to trigger the selective opening of basal ganglia gates to enable the read out of context-appropriate actions [238, 240, 241]. Entorhinal-hippocampal resonances are predicted to dynamically stabilize the learning of entorhinal grid cells and hippocampal place cells that represent an animal's current position during spatial navigation [242–245].

For example, the TELOS model in Fig. 41 includes a parietal-prefrontal feedback loop between the frontal eye fields, or FEF, and the posterior parietal cortex, or PPC, that

focuses spatial attention within the PPC on target positions in space to which saccadic eye movements will be directed. Additional interactions between the FEF and the basal ganglia, or BG, regulate the opening of gates in brain regions like the superior colliculus, or SC, that code these positions, and thereby release saccadic eye movements to foveate them.

The GridPlaceMap model in Fig. 42 includes an entorhinal-hippocampal feedback loop that dynamically stabilizes the learning of entorhinal grid cells and hippocampal place cells, while focusing attention upon the stripe cells, grid cells, and place cells that represent an animal's current position in space. This loop provides a computational explanation of the hypothesis in the title of [243] that “Increased attention to spatial context increases both place field stability and spatial memory”.

Because each of these resonances does not include qualia, even though they do include resonant feedback loops that focus attention upon their respective representations, they do not support a reportable conscious state. That is why, although it is predicted that “all conscious states are resonant states”, it is not predicted that “all resonant states are conscious states”.

HIPPOCAMPAL PLACE CELLS AS ATTENTIVE SELF-STABILIZING SPATIAL CATEGORIES

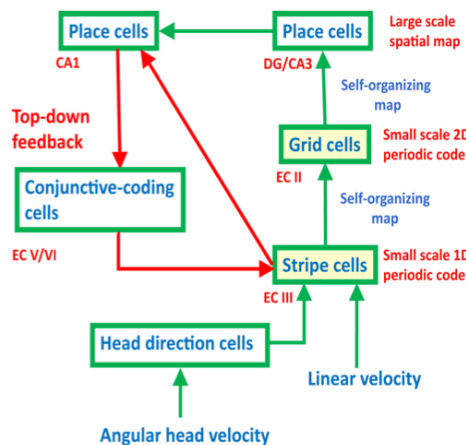


Fig. 42. Macrocircuit of GridPlaceMap model interactions within and between the entorhinal cortex (EC) and hippocampal cortex (DG/CA3 and CA1) to show how entorhinal grid cells and hippocampal place cells can be learned via a hierarchy of self-organizing maps in response to angular head velocity and linear velocity path integration signals while an animal navigates an environment. A resonant feedback loop between stripe cells, grid cells, and place cells is closed by top-down attentive feedback signals (red descending arrows) that dynamically stabilizes learning while focusing attention upon an animal's current positional representation. [Adapted from with permission from [239] published in Wiley.]

6. Concluding remarks

This article shows how multiple types of attention are incorporated into the CLEARS processes of Consciousness, Learning, Expectation, Attention, Resonance, and Synchrony that take place in perceptual, cognitive, affective, and motor systems throughout our brains. These processes enable us to continue learning about our unique and always changing experiences throughout life, while incorporating them into an emerging sense of self that defines our identities as human beings and supports all of our interactions with other individuals and the societies in which we live.

Author contributions

The author is the sole contributor to this article.

Acknowledgment

None.

Funding

There are no external funding sources for this article.

Conflict of interest

The author has no conflicts of interest.

References

- [1] Grossberg S. Conscious mind/resonant brain: How each brain makes a mind. New York: Oxford University Press. 2021.

- [2] Ames H, Grossberg S. Speaker normalization using cortical strip maps: a neural model for steady-state vowel categorization. *The Journal of the Acoustical Society of America*. 2008; 124: 3918-3936.
- [3] Banquet JP, Grossberg S. Probing cognitive processes through the structure of event-related potentials during learning: an experimental and theoretical analysis. *Applied Optics*. 2012; 26: 4931-4946.
- [4] Carpenter GA, Grossberg S. Normal and amnesic learning, recognition and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences*. 1993; 16: 131-137.
- [5] Chang HC, Grossberg S, Cao Y. Where's Waldo? How perceptual cognitive, and emotional brain processes cooperate during learning to categorize and find desired objects in a cluttered scene. *Frontiers in Integrative Neuroscience*. 2014; 8: 43.
- [6] Grossberg S. Competitive learning: from interactive activation to adaptive resonance. *Cognitive Science*. 1987; 11: 23-63.
- [7] Fazl A, Grossberg S, Mingolla E. View-invariant object category learning, recognition, and search: how spatial and object attention are coordinated using surface-based attentional shrouds. *Cognitive Psychology*. 2009; 58: 1-48.
- [8] Franklin DJ, Grossberg S. A neural model of normal and abnormal learning and memory consolidation: adaptively timed conditioning, hippocampus, amnesia, neurotrophins, and consciousness. *Cognitive, Affective, & Behavioral Neuroscience*. 2017; 17: 24-76.
- [9] Grossberg S. Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics*. 1976; 23: 121-134.
- [10] Grossberg S. Adaptive pattern classification and universal recoding, II: feedback, expectation, olfaction, and illusions. *Biological Cybernetics*. 1976; 23: 187-202.
- [11] Grossberg S. A theory of human memory: self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen and F. Snell (Eds.). *Progress in Theoretical Biology Volume 5* (pp. 233-374). New York: Academic Press. 1978.
- [12] Grossberg S. How does a brain build a cognitive code? *Psychological Review*. 1980; 87: 1-51.
- [13] Grossberg S. Processing of expected and unexpected events during conditioning and attention: a psychophysiological theory. *Psychological Review*. 1982; 89: 529-572.
- [14] Grossberg S. Some psychophysiological and pharmacological correlates of a developmental, cognitive and motivational theory. *Annals of the New York Academy of Sciences*. 1984; 425: 58-151.
- [15] Grossberg S. Unitization, automaticity, temporal order, and word recognition. *Cognition and Brain Theory*. 1984; 7: 263-283.
- [16] Grossberg S, Levine DS. Neural dynamics of attentionally modulated Pavlovian conditioning: blocking, interstimulus interval, and secondary reinforcement. *Applied Optics*. 1987; 26: 5015.
- [17] Grossberg S. Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks*. 2013; 37: 1-47.
- [18] Grossberg S. Acetylcholine neuromodulation in normal and abnormal learning and memory: vigilance control in waking, sleep, autism, amnesia, and Alzheimer's disease. *Frontiers in Neural Circuits*. 2017; 11: 82.
- [19] Grossberg S. Towards solving the hard problem of consciousness: the varieties of brain resonances and the conscious experiences that they support. *Neural Networks*. 2017; 87: 38-95.
- [20] Grossberg S. A half century of progress towards a unified neural theory of mind and brain with applications to autonomous adaptive agents and mental disorders. In R. Kozma, C. Alippi, Y. Choe, and F. C. Morabito (Eds.). *Artificial Intelligence in the Age of Neural Networks and Brain Computing*. Cambridge, MA: Academic Press. 2018.
- [21] Grossberg S. Desirability, availability, credit assignment, category learning, and attention: cognitive-emotional and working memory dynamics of orbitofrontal, ventrolateral, and dorsolateral

- eral prefrontal cortices. *Brain and Neuroscience Advances*. 2018; 2: 239821281877217.
- [22] Grossberg S. The embodied brain of SOVEREIGN2: from space-variant conscious percepts during visual search and navigation to learning invariant object categories and cognitive-emotional plans for acquiring valued goals. *Frontiers in Computational Neuroscience*. 2019; 13: 36.
 - [23] Grossberg S. The resonant brain: how attentive conscious seeing regulates action sequences that interact with attentive cognitive learning, recognition, and prediction. *Attention, Perception, & Psychophysics*. 2019; 81: 2237-2264.
 - [24] Grossberg S, Govindarajan KK, Wyse LL, Cohen MA. ART-STREAM: a neural network model of auditory scene analysis and source segregation. *Neural Networks*. 2004; 17: 511-536.
 - [25] Grossberg S, Huang TR. ARTSCENE: a neural system for natural scene classification. *Journal of Vision*. 2009; 9: 6.
 - [26] Grossberg S, Kazerounian S. Laminar cortical dynamics of conscious speech perception: a neural model of phonemic restoration using subsequent context in noise. *Journal of the Acoustical Society of America*. 2011; 130: 440-460.
 - [27] Grossberg S, Vladusich T. How do children learn to follow gaze, share joint attention, imitate their teachers, and use tools during social interactions? *Neural Networks*. 2010; 23: 940-965.
 - [28] Grossberg S, Markowitz J, Cao Y. On the road to invariant recognition: explaining tradeoff and morph properties of cells in inferotemporal cortex using multiple-scale task-sensitive attentive learning. *Neural Networks*. 2011; 24: 1036-1049.
 - [29] Grossberg S, Merrill JWL. A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. *Cognitive Brain Research*. 1992; 1: 3-38.
 - [30] Grossberg S, Merrill JW. The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. *Journal of Cognitive Neuroscience*. 2013; 8: 257-277.
 - [31] Grossberg S, Pearson L. Laminar cortical dynamics of cognitive and motor working memory, sequence learning and performance: toward a unified theory of how the cerebral cortex works. *Psychological Review*. 2008; 115: 677-732.
 - [32] Grossberg S, Schmajuk NA. Neural dynamics of attentionally-modulated Pavlovian conditioning: conditioned reinforcement, inhibition, and opponent processing. *Psychobiology*. 1987; 15: 195-240.
 - [33] Grossberg S, Seidman D. Neural dynamics of autistic behaviors: cognitive, emotional, and timing substrates. *Psychological Review*. 2006; 113: 483-525.
 - [34] Grossberg S, Somers D. Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*. 1991; 4: 453-466.
 - [35] Grossberg S, Srinivasan K, Yazdanbakhsh A. On the road to invariant object recognition: how cortical area V2 transforms absolute to relative disparity during 3D vision. *Neural Networks*. 2011; 24: 686-692.
 - [36] Grossberg S, Stone GO. Neural dynamics of word recognition and recall: attentional priming, learning, and resonance. *Psychological Review*. 1986; 93: 46-74.
 - [37] Huang T, Grossberg S. Cortical dynamics of contextually cued attentive visual learning and search: spatial and object evidence accumulation. *Psychological Review*. 2010; 117: 1080-1112.
 - [38] Raizada RDS, Grossberg S. Towards a theory of the laminar architecture of cerebral cortex: computational clues from the visual system. *Cerebral Cortex*. 2003; 13: 100-113.
 - [39] Silver MR, Grossberg S, Bullock D, Histed MH, Miller EK. A neural model of sequential movement planning and control of eye movements: Item-Order-Rank working memory and saccade selection by the supplementary eye fields. *Neural Networks*. 2012; 26: 29-58.
 - [40] Grossberg S. Associative and competitive principles of learning and development: The temporal unfolding and stability of STM and LTM patterns. In S. I. Amari and M. Arbib (Eds.). *Competition and Cooperation in Neural Networks* (pp. 295-341). New York: Springer Verlag. 1982.
 - [41] Cohen MA, Grossberg S. Masking fields: a massively parallel neural architecture for learning, recognizing, and predicting multiple groupings of patterned data. *Applied Optics*. 1987; 26: 1866.
 - [42] Grossberg S. A path towards Explainable AI and autonomous adaptive intelligence: deep learning, adaptive resonance, and models of perception, emotion, and action. *Frontiers in Neurobotics*. 2020; 14: 36.
 - [43] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986; 323: 533-536.
 - [44] Amari S. Characteristics of random nets of analog neuron-like elements. *IEEE Transactions on Systems, Man, and Cybernetics*. 1972; SMC-2: 643-657.
 - [45] Werbos P. Beyond regression: new tools for prediction and analysis in the behavioral sciences. Unpublished Doctoral Dissertation, Harvard University. 1974.
 - [46] Werbos P. The roots of backpropagation: from ordered derivatives to neural networks and political forecasting. New York: John Wiley & Sons, Inc. 1994.
 - [47] Parker DB. Learning-logic. Technical report TR-47, center for computational research in Economics and Management Science, MIT. 1985.
 - [48] Parker DB. A comparison of algorithms for neuron-like cells. In Denker J (Ed.). *Proceedings of the Second Annual Conference on Neural Networks for Computing*. Proceedings Vol. 151 (pp. 327-332). New York: American Institute of Physics. 1986.
 - [49] Parker DB. Optimal algorithms for adaptive networks: second order back propagation, second order direct propagation, and second order Hebbian learning. *Proceedings of the 1987 IEEE International Conference on Neural Networks, II* (pp. 593-600). New York: IEEE Press. 1987.
 - [50] Schmidhuber J. Critique of Honda Prize for Dr. Hinton. 2020. Available at: <http://people.idsia.ch/~juergen/critique-honda-prize-hinton.html#reply>.
 - [51] Grossberg S. Nonlinear neural networks: principles, mechanisms, and architectures. *Neural Networks*. 1988; 1: 17-61.
 - [52] Darwin C. *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life* (5th Ed.). London: John Murray. 1869.
 - [53] Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*. 2009; 12: 718-724.
 - [54] Goodale MA, Milner AD. Separate visual pathways for perception and action. *Trends in Neurosciences*. 1992; 15: 20-25.
 - [55] Goodale MA, Milner AD, Jakobson LS, Carey DP. A neurological dissociation between perceiving objects and grasping them. *Nature*. 1991; 349: 154-156.
 - [56] Mishkin M. A memory system in the monkey. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 1982; 298: 83-95.
 - [57] Mishkin M, Ungerleider LG, Macko KA. Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*. 1983; 6: 414-417.
 - [58] Grossberg S. Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences of the United States of America*. 1968; 60: 758-765.
 - [59] Grossberg S. On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*. 1969; 22: 325-364.
 - [60] Grossberg S. Recurrent neural networks. *Scholarpedia*. 2013; 8: 1888. Available at: http://www.scholarpedia.org/article/Recurrent_neural_networks.
 - [61] Grossberg S. Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*. 1973; 52: 213-257.
 - [62] Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*. 2003; 117: 500-544.
 - [63] Grossberg S, Hong S. A neural model of surface perception: light-

- ness, anchoring, and filling-in. *Spatial Vision*. 2006; 19: 263-321.
- [64] Grossberg S, Todorović D. Neural dynamics of 1-D and 2-D brightness perception: a unified model of classical and recent phenomena. *Perception & Psychophysics*. 1988; 43: 241-277.
- [65] Bi G, Poo M. Synaptic modification by correlated activity: Hebb's postulate revisited. *Annual Review of Neuroscience*. 2001; 24: 139-166.
- [66] Levy WB, Steward O. Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience*. 1983; 8: 791-797.
- [67] Markram H, Lübke J, Frotscher M, Sakmann B. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*. 1997; 275: 213-215.
- [68] Brito da Silva LE, Elnabarawy I, Wunsch DC. A survey of adaptive resonance theory neural network models for engineering applications. *Neural Networks*. 2019; 120: 167-203.
- [69] Wunsch II DC. Admiring the great mountain: a celebration special issue in honor of Stephen Grossberg's 80th birthday. *Neural Networks*. 2019; 120: 1-4.
- [70] Cao Y, Grossberg S, Markowitz J. How does the brain rapidly learn and reorganize view- and positionally-invariant object representations in inferior temporal cortex? *Neural Networks*. 2011; 24: 1050-1061.
- [71] Carpenter GA, Grossberg S. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*. 1987; 37: 54-115.
- [72] Carpenter GA, Grossberg S. ART 2: Stable self-organization of pattern recognition codes for analog input patterns. *Applied Optics*. 1987; 26: 4919-4930.
- [73] Carpenter GA, Grossberg S. The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*. 1988; 21: 77-88.
- [74] Carpenter GA, Grossberg S. ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks*. 1990; 3: 129-152.
- [75] Carpenter GA, Grossberg S, Markuzon N, Reynolds JH, Rosen DB. Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*. 1992; 3: 698-713.
- [76] Carpenter GA, Grossberg S, Mehanian C. Invariant recognition of cluttered scenes by a self-organizing ART architecture: CORT-X boundary segmentation. *Neural Networks*. 1989; 2: 169-181.
- [77] Carpenter GA, Grossberg S, Reynolds JH. ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*. 1991; 4: 565-588.
- [78] Carpenter GA, Grossberg S, Reynolds JH. A fuzzy ARTMAP non-parametric probability estimator for nonstationary pattern recognition problems. *IEEE Transactions on Neural Networks*. 1995; 6: 1330-1336.
- [79] Carpenter GA, Grossberg S, Rosen DB. ART 2A: an adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks*. 1991; 4: 493-504.
- [80] Carpenter GA, Grossberg S, Rosen DB. Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks*. 1991; 4: 759-771.
- [81] Carpenter GA, Milenova BL, Noeske BW. dARTMAP: a neural network for fast distributed supervised learning. *Neural Networks*. 1998; 11: 793-813.
- [82] Caudell TP, Smith SDG, Escobedo R, Anderson M. NIRS: large scale ART-1 neural architectures for engineering desiring retrieval. *Neural Networks*. 1994; 7: 1339-1350.
- [83] Gnat W, Grossberg S. SOVEREIGN: an autonomous neural system for incrementally learning planned action sequences to navigate towards a rewarded goal. *Neural Networks*. 2008; 21: 699-758.
- [84] Grossberg S, Wyse L. Invariant recognition of cluttered scenes by a self-organizing ART architecture: figure-ground separation. *Neural Networks*. 1991; 4: 723-742.
- [85] Duncan J. Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*. 1985; 113: 501-517.
- [86] Posner MI. Orienting of attention. *Quarterly Journal of Experimental Psychology*. 1980; 32: 3-25.
- [87] Neisser U. *Cognitive Psychology*. New York: Appleton-Century-Crofts. 1967.
- [88] Lappin JS. Attention in the identification of stimuli in complex displays. *Journal of Experimental Psychology*. 1967; 75: 321-328.
- [89] Treisman A, Kahneman D, Burkell J. Perceptual objects and the cost of filtering. *Perception & Psychophysics*. 1983; 33: 527-532.
- [90] Mountcastle VB. Brain mechanisms for directed attention. *Journal of the Royal Society of Medicine*. 1978; 71: 14-28.
- [91] Robinson DL, Goldberg ME, Stanton GB. Parietal association cortex in the primate: sensory mechanisms and behavioral modulations. *Journal of Neurophysiology*. 1978; 41: 910-932.
- [92] Wurtz RH, Mohler CW. Organization of monkey superior colliculus: enhanced visual response of superficial layer cells. *Journal of Neurophysiology*. 1976; 39: 745-763.
- [93] Weiskrantz L, Warrington EK, Sanders MD, Marshall J. Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain*. 1975; 97: 709-728.
- [94] Posner MI. *Chronometric explorations of mind*. Hillsdale, N.J.: Lawrence Erlbaum Associates. 1978.
- [95] Burgess N, Jeffery KJ, O'Keefe J (Eds.). *The hippocampal and parietal foundations of spatial cognition*. New York: Oxford University Press. 1999.
- [96] Chelazzi L, Duncan J, Miller EK, Desimone R. Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*. 1999; 80: 2918-2940.
- [97] Pohl W. Dissociation of spatial discrimination deficits following frontal and parietal lesions in monkeys. *Journal of Comparative and Physiological Psychology*. 1973; 82: 227-239.
- [98] Sommer T, Rose M, Weiller C, Büchel C. Contributions of occipital, parietal and parahippocampal cortex to encoding of object-location associations. *Neuropsychologia*. 2005; 43: 732-743.
- [99] Suzuki WA, Miller EK, Desimone R. Object and place memory in the macaque entorhinal cortex. *Journal of Neurophysiology*. 1997; 78: 1062-1081.
- [100] Treisman A. Perceptual grouping and attention in visual search for features and objects. *Journal of Experimental Psychology: Human Perception and Performance*. 1982; 8: 194-214.
- [101] Treisman A. Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*. 1985; 31: 156-177.
- [102] Treisman AM, Gelade G. A feature-integration theory of attention. *Cognitive Psychology*. 1980; 12: 97-136.
- [103] Treisman A, Gormican S. Feature analysis in early vision: evidence from search asymmetries. *Psychological Review*. 1988; 95: 15-48.
- [104] Treisman A, Sato S. Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*. 1990; 16: 459-478.
- [105] Cave KR, Wolfe JM. Modeling the role of parallel processing in visual search. *Cognitive Psychology*. 1990; 22: 225-271.
- [106] Wolfe JM. "Effortless" texture segmentation and "parallel" visual search are not the same thing. *Vision Research*. 1992; 32: 757-763.
- [107] Wolfe JM. Guided Search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*. 1994; 1: 202-238.
- [108] Wolfe JM. Guided Search 4.0: a guided search model that does not require memory for rejected distractors. *Journal of Vision*. 2010; 1: 349-349.
- [109] Wolfe JM, Cave KR, Franzel SL. Guided search: an alternative to the feature integration model of visual search. *Journal of Experimental Psychology: Human Perception and Performance*. 1989; 15: 419-433.
- [110] Wolfe JM, Gancarz G. Guided Search 3.0. in *basic and clinical applications of vision science* (pp. 189-192). V. Lakshminarayanan V (Ed.). Dordrecht, The Netherlands: Springer. 1997.
- [111] Grossberg S. 3-D vision and figure-ground separation by visual

- cortex. *Perception & Psychophysics*. 1994; 55: 48-121.
- [112] Grossberg S, Mingolla E, Ross WD. A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations. *Psychological Review*. 1994; 101: 470-489.
- [113] Theeuwes J, Mathôt S, Grainger J. Exogenous object-centered attention. *Attention, Perception & Psychophysics*. 2013; 75: 812-818.
- [114] Theeuwes J. Stimulus-driven capture and attentional set: selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance*. 1994; 20: 799-806.
- [115] Theeuwes J. Top-down and bottom-up control of visual selection. *Acta Psychologica*. 2010; 135: 77-99.
- [116] Schreij D, Owens C, Theeuwes J. Abrupt onsets capture attention independent of top-down control settings. *Perception & Psychophysics*. 2008; 70: 208-218.
- [117] Posner MI, Cohen Y. Components of visual orienting. In Bourma H, Bowhuis D (Eds.). *Attention and Performance X* (pp. 531-556.). Hillsdale NJ: Lawrence Erlbaum Associates. 1984.
- [118] Breitmeyer BG, Ganz L. Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*. 1976; 83: 1-36.
- [119] Mathôt S, Cristino F, Gilchrist ID, Theeuwes J. A simple way to estimate similarity between pairs of eye movement sequences. *Journal of Eye Movement Research*. 2012; 5: 1-15.
- [120] Van der Werf YD, Witter MP, Groenewegen HJ. The intralaminar and midline nuclei of the thalamus. Anatomical and functional evidence for participation in processes of arousal and awareness. *Brain Research*. 2003; 39: 107-140.
- [121] Yantis S, Jonides J. Abrupt visual onsets and selective attention: evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*. 1984; 10: 601-621.
- [122] Enroth-Cugell C, Robson JG. The contrast sensitivity of retinal ganglion cells of the cat. *The Journal of Physiology*. 2006; 187: 517-552.
- [123] Hochstein S, Shapley RM. Linear and nonlinear spatial subunits in Y cat retinal ganglion cells. *The Journal of Physiology*. 1977; 262: 265-284.
- [124] Hochstein S, Shapley RM. Quantitative analysis of retinal ganglion cell classifications. *The Journal of Physiology*. 1977; 262: 237-264.
- [125] Theeuwes J. Abrupt luminance change pops out; abrupt color change does not. *Perception & Psychophysics*. 1995; 57: 637-644.
- [126] Todd JT, Van Gelder P. Implications of a transient-sustained dichotomy for the measurement of human performance. *Journal of Experimental Psychology: Human Perception and Performance*. 1979; 5: 625-638.
- [127] Berzhanskaya J, Grossberg S, Mingolla E. Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision*. 2007; 20: 337-395.
- [128] Grossberg S, Mingolla E, Viswanathan L. Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*. 2001; 41: 2521-2553.
- [129] Baloch AA, Grossberg S. A neural model of high-level motion processing: line motion and formotion dynamics. *Vision Research*. 1997; 37: 3037-3059.
- [130] Grossberg S. How is a moving target continuously tracked behind occluding cover? In T. Watanabe (Ed.). *High Level Motion Processing: Computational, Neurobiological, and Psychophysical Perspectives* (pp. 3-52). Cambridge, MA: MIT Press. 1998.
- [131] Baloch AA, Grossberg S, Mingolla E, Nogueira CA. Neural model of first-order and second-order motion perception and magnocellular dynamics. *Journal of the Optical Society of America. a, Optics, Image Science, and Vision*. 1999; 16: 953-978.
- [132] Chey J, Grossberg S, Mingolla E. Neural dynamics of motion processing and speed discrimination. *Vision Research*. 1998; 38: 2769-2786.
- [133] Francis G, Grossberg S. Cortical dynamics of form and motion integration: persistence, apparent motion, and illusory contours. *Vision Research*. 1996; 36: 149-173.
- [134] Grossberg S, Rudd ME. A neural architecture for visual motion perception: group and element apparent motion. *Neural Networks*. 1989; 2: 421-450.
- [135] Grossberg S, Rudd ME. Cortical dynamics of visual motion perception: short-range and long-range apparent motion. *Psychological Review*. 1992; 99: 78-121.
- [136] Wuerger S, Shapley R, Rubin N. "On the visually perceived direction of motion" by Hans Wallach: 60 Years Later. *Perception*. 1996; 25: 1317-1367.
- [137] Marr D, Ullman S. Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society, London, B*. 1981; 211: 151-180.
- [138] Grossberg S. Why do parallel cortical systems exist for the perception of static form and moving form? *Perception & Psychophysics*. 1991; 49: 117-141.
- [139] Ponce CR, Lomber SG, Born RT. Integrating motion and depth via parallel pathways. *Nature Neuroscience*. 2008; 11: 216-223.
- [140] Browning N, Grossberg S, Mingolla E. Cortical dynamics of navigation and steering in natural scenes: motion-based object segmentation, heading, and obstacle avoidance. *Neural Networks*. 2009; 22: 1383-1398.
- [141] Browning NA, Grossberg S, Mingolla E. A neural model of how the brain computes heading from optic flow in realistic scenes. *Cognitive Psychology*. 2009; 59: 320-356.
- [142] Elder D, Grossberg S, Mingolla E. A neural model of visually guided steering, obstacle avoidance, and route selection. *Journal of Experimental Psychology: Human Perception & Performance*. 2009; 35: 1501-1531.
- [143] Gaudiano P, Grossberg S. Vector associative maps: Unsupervised real-time error-based learning and control of movement trajectories. *Neural Networks*. 1991; 4: 147-183.
- [144] Gaudiano P, Grossberg S. Adaptive vector integration to endpoint: Self-organizing neural circuits for control of planned movement trajectories. *Human Movement Science*. 1992; 11: 141-155.
- [145] Bhatt R, Carpenter GA, Grossberg S. Texture segregation by visual cortex: perceptual grouping, attention, and learning. *Vision Research*. 2008; 47: 3173-3211.
- [146] Cao Y, Grossberg S. Stereopsis and 3D surface perception by spiking neurons in laminar cortical circuits: a method of converting neural rate models into spiking models. *Neural Networks*. 2012; 26: 75-98.
- [147] Cao Y, Grossberg S. A laminar cortical model for 3D boundary and surface representations of complex natural scenes. from Parallel to Emergent Computing. In A. Adamatzky S. Akl, G. Sirakoulis (Eds.). *From Parallel to Emergent Computing*. Taylor and Francis/CRC. 2018.
- [148] Dresch B, Grossberg S. Spatial facilitation by color and luminance edges: boundary, surface, and attentional factors. *Vision Research*. 1999; 39: 3431-3443.
- [149] Dresch-Langley B, Grossberg S. Neural computation of surface border ownership and relative surface depth from ambiguous contrast inputs. *Frontiers in Psychology*. 2016; 7: 1102.
- [150] Fang L, Grossberg S. From stereogram to surface: how the brain sees the world in depth. *Spatial Vision*. 2009; 22: 45-82.
- [151] Grossberg S. Cortical dynamics of three-dimensional figure-ground perception of two-dimensional figures. *Psychological Review*. 1997; 104: 618-658.
- [152] Grossberg S. Outline of a theory of brightness, color, and form perception. In E. Degreef and J. van Buggenhout (Eds.). *Trends in Mathematical Psychology* (pp. 59-85). Amsterdam: North-Holland. 1984.
- [153] Grossberg S. Filling-in the forms: Surface and boundary interactions in visual cortex. In L. Pessoa and P. DeWeerd (Eds.). *Filling-in: From Perceptual Completion to Skill Learning* (pp.13-37). New York: Oxford University Press. 2003.
- [154] Grossberg S, Kuhlmann L, Mingolla E. A neural model of 3D

- shape-from-texture: multiple-scale filtering, boundary grouping, and surface filling-in. *Vision Research*. 2007; 47: 634-672.
- [155] Grossberg S, Hwang S, Mingolla E. Thalamocortical dynamics of the McCollough effect: boundary-surface alignment through perceptual learning. *Vision Research*. 2002; 42: 1259-1286.
- [156] Grossberg S, Mingolla E. Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychological Review*. 1985; 92: 173-211.
- [157] Grossberg S, Mingolla E. Neural dynamics of perceptual grouping: textures, boundaries, and emergent segmentations. *Perception & Psychophysics*. 1985; 38: 141-171.
- [158] Grossberg S, Mingolla E. Neural dynamics of surface perception: boundary webs, illuminants, and shape-from-shading. *Computer Vision, Graphics, and Image Processing*. 1987; 37: 116-165.
- [159] Grossberg S, McLoughlin N. Cortical dynamics of 3-D surface perception: binocular and half-occluded scenic images. *Neural Networks*. 1997; 10: 1583-1605.
- [160] Grossberg S, Pessoa L. Texture segregation, surface representation and figure-ground separation. *Vision Research*. 2002; 38: 2657-2684.
- [161] Grossberg S, Swaminathan G. A laminar cortical model for 3D perception of slanted and curved surfaces and of 2D images: development, attention, and bistability. *Vision Research*. 2004; 44: 1147-1187.
- [162] Grossberg S, Mingolla E, Williamson J. Synthetic aperture radar processing by a multiple scale neural system for boundary and surface representation. *Neural Networks*. 1995; 8: 1005-1028.
- [163] Grossberg S, Yazdanbakhsh A. Laminar cortical dynamics of 3D surface perception: stratification, transparency, and neon color spreading. *Vision Research*. 2005; 45: 1725-1743.
- [164] Kelly F, Grossberg S. Neural dynamics of 3-D surface perception: figure-ground separation and lightness perception. *Perception & Psychophysics*. 2000; 62: 1596-1618.
- [165] Mingolla E, Ross W, Grossberg S. A neural network for enhancing boundaries and surfaces in synthetic aperture radar images. *Neural Networks*. 2019; 12: 499-511.
- [166] Grossberg S. The complementary brain: unifying brain dynamics and modularity. *Trends in Cognitive Sciences*. 2000; 4: 233-246.
- [167] Carrasco M, Penpeci-Talgar C, Eckstein M. Spatial covert attention increases contrast sensitivity across the CSF: support for signal enhancement. *Vision Research*. 2000; 40: 1203-1215.
- [168] Reynolds JH, Desimone R. Interacting roles of attention and visual salience in V4. *Neuron*. 2003; 37: 853-863.
- [169] Tyler CW, Kontsevich LL. Mechanisms of stereoscopic processing: stereoattention and surface perception in depth reconstruction. *Perception*. 1995; 24: 127-153.
- [170] Bushnell MC, Goldberg ME, Robinson DL. Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention. *Journal of Neurophysiology*. 1981; 46: 755-772.
- [171] Colby CL, Goldberg ME. Space and attention in parietal cortex. *Annual Review of Neuroscience*. 1999; 22: 319-349.
- [172] Gnadt JW, Andersen RA. Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental Brain Research*. 1988; 70: 216-220.
- [173] Snyder LH, Batista AP, Andersen RA. Coding of intention in the posterior parietal cortex. *Nature*. 1997; 386: 167-170.
- [174] Snyder LH, Batista AP, Andersen RA. Change in motor plan, without a change in the spatial locus of attention, modulates activity in posterior parietal cortex. *Journal of Neurophysiology*. 1998; 79: 2814-2819.
- [175] Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*. 1991; 1: 1-47.
- [176] Brown JM, Denney HI. Shifting attention into and out of objects: evaluating the processes underlying the object advantage. *Perception & Psychophysics*. 2007; 69: 606-618.
- [177] Egly R, Driver J, Rafal RD. Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology General*. 1994; 123: 161-177.
- [178] Roelfsema PR, Lamme VA, Spekreijse H. Object-based attention in the primary visual cortex of the macaque monkey. *Nature*. 1998; 395: 376-381.
- [179] Grossberg S, Raizada RDS. Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research*. 2000; 40: 1413-1432.
- [180] Foley NC, Grossberg S, Mingolla E. Neural dynamics of object-based multifocal visual spatial attention and priming: object cueing, useful-field-of-view, and crowding. *Cognitive Psychology*. 2012; 65: 77-117.
- [181] Brady TF, Konkle T, Alvarez GA, Oliva A. Remembering thousands of objects with high fidelity. *Proceedings of the National Academy of Sciences U.S.A.* 2008; 105: 14325-14329.
- [182] Oliva A. Gist of the scene. In Itti L, Rees G, Tsotsos JK (Eds.). *Neurobiology of Attention*, Chapter 41 (pp. 251-256). 2005.
- [183] Potter MC. Recognition and memory for briefly presented scenes. *Frontiers in Psychology*. 2012; 3: 32.
- [184] Potter MC, Levy EI. Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*. 1969; 81: 10-15.
- [185] Shepard RN. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*. 1967; 6: 156-163.
- [186] Standing L. Learning 10,000 pictures. *The Quarterly Journal of Experimental Psychology*. 1973; 25: 207-222.
- [187] Standing L, Conezio J, Haber RN. Perception and memory for pictures: single-trial learning of 2500 visual stimuli. *Psychonomic Science*. 1970; 19: 73-74.
- [188] Wolfe JM. Visual memory: what do you know about what you saw? *Current Biology*. 1998; 8: R303-R304.
- [189] French. Catastrophic forgetting in connectionist networks. *Trends in Cognitive Sciences*. 2019; 3: 128-135.
- [190] McCloskey M, Cohen NJ. In G. H. Bower (Ed.). *The Psychology of Learning and Motivation*, Vol. 24 (pp. 109-164). 1989.
- [191] Ratcliff R. Connectionist models of recognition memory: constraints imposed by learning and forgetting functions. *Psychological Review*. 1990; 97: 285-308.
- [192] Grossberg S. Consciousness CLEARs the mind. *Neural Networks*. 2007; 20: 1040-1053.
- [193] Grossberg S. The attentive brain. *American Scientist*. 1995; 83: 438-449.
- [194] Bullier J, Hupé JM, James A, Girard P. Functional interactions between areas V1 and V2 in the monkey. *Journal of Physiology, Paris*. 1997; 90: 217-220.
- [195] Caputo G, Guerra S. Attentional selection by distractor suppression. *Vision Research*. 1998; 38: 669-689.
- [196] Downing CJ. Expectancy and visual-spatial attention: effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance*. 1988; 14: 188-202.
- [197] Hupé JM, James AC, Girard DC, Bullier J. Feedback connections from V2 modulate intrinsic connectivity within V1. *Society for Neuroscience Abstracts*. 1997; 406.15: 1031.
- [198] Mounts JRW. Evidence for suppressive mechanisms in attentional selection: feature singletons produce inhibitory surrounds. *Perception & Psychophysics*. 2000; 62: 969-983.
- [199] Reynolds JH, Chelazzi L, Desimone R. Competitive mechanisms subserve attention in macaque areas V2 and V4. *The Journal of Neuroscience*. 1999; 19: 1736-1753.
- [200] Sillito AM, Jones HE, Gerstein GL, West DC. Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature*. 1994; 369: 479-482.
- [201] Somers DC, Dale AM, Seiffert AE, Tootell RBH. Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences*. 1999; 96: 1663-1668.
- [202] Steinman BA, Steinman SB, Lehmkuhle S. Visual attention mechanisms show a center-surround organization. *Vision Research*. 1995; 35: 1859-1869.
- [203] Vanduffel W. Attention-dependent suppression of metabolic ac-

- tivity in the early stages of the macaque visual system. *Cerebral Cortex*. 2000; 10: 109-126.
- [204] Desimone R. Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 1998; 353: 1245-1255.
- [205] Kastner S, Ungerleider LG. The neural basis of biased competition in human visual cortex. *Neuropsychologia*. 2002; 39: 1263-1276.
- [206] Luck SJ, Chelazzi L, Hillyard SA, Desimone R. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*. 1997; 77: 24-42.
- [207] Dubin MW, Cleland BG. Organization of visual inputs to interneurons of lateral geniculate nucleus of the cat. *Journal of Neurophysiology*. 1977; 40: 410-427.
- [208] Weber AJ, Kalil RE, Behan M. Synaptic connections between corticogeniculate axons and interneurons in the dorsal lateral geniculate nucleus of the cat. *The Journal of Comparative Neurology*. 1989; 289: 156-164.
- [209] Zhang Y, Suga N, Yan J. Corticofugal modulation of frequency processing in bat auditory system. *Nature*. 1997; 387: 900-903.
- [210] Temereanca S, Simons DJ. Topographic specificity in the functional effects of corticofugal feedback in the whisker/barrel system. *ScienceDirect*. 2004; 41: 639-651.
- [211] Reynolds JH, Heeger DJ. The Normalization Model of Attention. *Neuron*. 2009; 61: 168-185.
- [212] Ben-Shahar O, Zucker SW. Sensitivity to curvatures in orientation-based texture segmentation. *Vision Research*. 2004; 44: 257-277.
- [213] Grossberg S. How hallucinations may arise from brain mechanisms of learning, attention, and volition. *Journal of the International Neuropsychological Society*. 2000; 6: 579-588.
- [214] Grossberg S, Versace M. Spikes, synchrony, and attentive learning by laminar thalamocortical circuits. *Brain Research*. 2008; 1218: 278-312.
- [215] Broadbent DE. *Perception and communication*. London: Pergamon Press. 1958.
- [216] Moray N. Where is capacity limited? A survey and a model. *Acta Psychologica*. 1968; 27: 84-92.
- [217] Norman DA. Toward a theory of memory and attention. *Psychological Review*. 1968; 75: 522-536.
- [218] Hansen JC, Hillyard SA. Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology*. 1980; 49: 277-290.
- [219] Näätänen R, Gaillard AW, Mäntysalo S. Early selective attention effect on evoked potential reinterpreted. *Acta Psychologica*. 1978; 42: 313-329.
- [220] Näätänen R, Picton TW. N2 and automatic versus controlled processes. *Electroencephalography and Clinical Neurophysiology*. 1986; 38: 169-186.
- [221] Pritchard WS, Shappell SA, Brandt ME. Psychophysiology of N200/N400: a review and classification scheme. *Advances in Psychophysiology*. 1991; 4: 43-106.
- [222] Snyder E, Hillyard SA. Long-latency evoked potentials to irrelevant, deviant stimuli. *Behavioral Biology*. 1976; 16: 319-331.
- [223] Brincat SL, Miller EK. Frequency-specific hippocampal-prefrontal interactions during associative learning. *Nature Neuroscience*. 2015; 18: 576-581.
- [224] Brodmann K. *Vergleichende Lokalisationslehre der Grosshirnrinde in Ihren Prinzipien Dargestellt auf Grund des Zellenbaues*. Leipzig: Barth. 1909.
- [225] Martin JH. *Neuroanatomy: text and atlas*. Norwalk, CT: Appleton and Lange. 1989.
- [226] Grossberg S. How does the cerebral cortex work? Learning, attention and grouping by the laminar circuits of visual cortex. *Spatial Vision*. 1999; 12: 163-186.
- [227] Grossberg S, Palma J, Versace M. Resonant cholinergic dynamics in cognitive and motor decision-making: attention, category learning, and choice in neocortex, superior colliculus, and optic tectum. *Frontiers in Neuroscience*. 2016; 9: 501.
- [228] Grossberg S, Roberts K, Aguilar M, Bullock D. A neural model of multimodal adaptive saccadic eye movement control by superior colliculus. *The Journal of Neuroscience*. 1997; 17: 9706-9725.
- [229] Edwards S, de Olmos J. Autoradiographic studies of projections of midbrain reticular-formation-ascending projections of nucleus cuneiformis. *Journal of Comparative Neurology*. 1976; 165: 417-432.
- [230] Edwards SB, Ginsburgh CL, Henkel CK, Stein BE. Sources of subcortical projections to the superior colliculus in the cat. *The Journal of Comparative Neurology*. 1979; 184: 309-329.
- [231] Sparks D, Hartwich-Young R. The deep layers of the superior colliculus. In R. Wurtz and M. Goldberg (Eds.). *The Neurobiology of Saccadic Eye Movements*. Chapter 3 (pp. 213-255). New York: Elsevier. 1989.
- [232] Kowler E, Anderson E, Doshier B, Blaser E. The role of attention in the programming of saccades. *Vision Research*. 1995; 35: 1897-1916.
- [233] Deubel H, Schneider WX. Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision Research*. 1996; 36: 1827-1837.
- [234] Knudsen EI. Instructed learning in the auditory localization pathway of the barn owl. *Nature*. 2002; 417: 322-328.
- [235] Knudsen EI. Control from below: the role of a midbrain network in spatial attention. *The European Journal of Neuroscience*. 2011; 33: 1961-1972.
- [236] Knudsen EI, Brainard MS. Creating a unified representation of visual and auditory space in the brain. *Annual Review of Neuroscience*. 1995; 18: 19-43.
- [237] Kemel ML, Desban M, Gauchy C, Glowinski J, Besson MJ. Topographical organization of efferent projections from the cat substantia nigra pars reticulata. *Brain Research*. 1988; 455: 307-323.
- [238] Brown JW, Bullock D, Grossberg S. How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks*. 2004; 17: 471-510.
- [239] Mhatre H, Gorchetnikov A, Grossberg S. Grid cell hexagonal patterns formed by fast self-organized learning within entorhinal cortex. *Hippocampus*. 2012; 22: 320-334.
- [240] Buschman TJ, Miller EK. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*. 2007; 315: 1860-1862.
- [241] Grossberg S. Neural dynamics of the basal ganglia during perceptual, cognitive, and motor learning and gating. In J.J. Sghomnian (Ed.). *The basal ganglia: Novel Perspectives on Motor and Cognitive Functions* (pp. 457-512). Berlin, Germany: Springer. 2016.
- [242] Grossberg S, Pilly PK. Coordinated learning of grid cell and place cell spatial and temporal properties: multiple scales, attention and oscillations. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2014; 369: 20120524.
- [243] Kentros CG, Agnihotri NT, Streater S, Hawkins RD, Kandel ER. Increased attention to spatial context increases both place field stability and spatial memory. *Neuron*. 2004; 42: 283-295.
- [244] Morris RGM, Frey U. Hippocampal synaptic plasticity: role in spatial learning or the automatic recording of attended experience? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. 1997; 352: 1489-1503.
- [245] Pilly PK, Grossberg S. How do spatial learning and memory occur in the brain? Coordinated learning of entorhinal grid cells and hippocampal place cells. *Journal of Cognitive Neuroscience*. 2012; 24: 1031-1054.