Contents lists available at ScienceDirect

Progress in Neurobiology

journal homepage: www.elsevier.com/locate/pneurobio

Cortical maps as a fundamental neural substrate for visual representation.

Bruss Lima [a], Maria M. Florentino [a], Mario Fiorani [a], Juliana G.M. Soares [a], Kerstin E. Schmidt [b],
Sergio Neuenschwander [b], Jerome Baron [c], Ricardo Gattass [a,*]

[a] Programa de Neurobiologia, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21941-902, Brazil
[b] Instituto do Cérebro, Universidade Federal do Rio Grande do Norte, Natal, RN 59056-450, Brazil
[c] Departamento de Fisiologia e Biofísica, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil

ARTICLE INFO

ABSTRACT

Visual perception is the product of serial hierarchical processing, parallel processing, and remapping on a dynamic network involving several topographically organized cortical visual areas. Here, we will focus on the topographical organization of cortical areas and the different kinds of visual maps found in the primate brain. We will interpret our findings in light of a broader representational framework for perception. Based on neurophysiological data, our results do not support the notion that vision can be explained by a strict representational model, where the objective visual world is faithfully represented in our brain. On the contrary, we find strong evidence that vision is an active and constructive process from the very initial stages taking place in the eye and from the very initial stages of our development. A constructive interplay between perceptual and motor systems (e.g., during saccadic eye movements) is actively learnt from early infancy and ultimately provides our fluid stable visual perception of the world.

1. Introduction

Only a limited portion of all the information available in the visual world is extracted by the retina. Indeed, our visual sampling is somewhat sparse in both the spatial and temporal domains. A meshwork of blood vessels and unmyelinated axons stands in the way between the visual field and the photoreceptors at the back of the retina. Additionally, axons and blood vessels bundle together into a structure known as the optic disk. The perceptual counterpart of the optic disk is an approximately 4–6°-diameter blind spot, completely oblivious to light stimulation. Despite this, our perception of the world is not disturbed by web-like filaments or by a large oval void, even when viewing the world monocularly. Notably, there is a clear disparity between image capture in the retina and our visual experience.

Other important aspects are worth considering when contemplating the reliability of our visual experience. First, there are large-scale inhomogeneities across the retinal surface. Retinal ganglion cells in primates are highly concentrated in the fovea, providing our center of gaze with a high spatial resolution of the scene. Therefore, in order to provide the focus of our attention with the highest spatial resolution possible, our eyes are constantly performing saccadic movements to precise and

specific points in space. Second, retinal ganglion cells are known to have small receptive fields (Hartline, 1938), which implies that the image is broken up into over a million pieces at the very first stage of visual processing. Finally, the temporal sampling of an image by the retinal ganglion cells is limited by its time constant and integration properties. Despite all of this, the visual world is not perceived as a sequence of pseudo-flashed images due to the saccadic eye movements, or as pixels with varying sizes as a function of retinal magnification, or as flickering frames based on the biophysical properties of the neuron membrane. On the contrary, our visual perception is a stable experience that can be subjectively described as an unwavering continuous stream of incoming information.

Evidence from patients that recover vision later in life through surgical intervention indicates that visuomotor learning plays a crucial role in our stable visual experience. Recovered patients often report discomfort when using vision to perform tasks that involve interpreting three-dimensional visual cues. In a case reported by (Fine et al., 2003), the subject preferred to close his eyes when crossing a street or when performing activities where vision is usually useful. Therefore, our sensory interaction with the environment relies fundamentally on the motor system. In the case of the visual system, the integration with eye,

head and body movements is crucial. Ultimately, sophisticated binding mechanisms (Lima et al., 2010) may come into play in order to link and stabilize visual features in time and space, and to allow for an efficient and fast communication between sensory and motor systems.

Here, we will discuss some of the neuronal substrates that subserve vision and that may enable stable perception. We argue that the concept of topographical organization, embodied by multiple visual maps established along several stages of early and intermediate perception, is essential to understanding vision. Here, we discuss how representational maps may contribute to a stable, coherent visual experience. We argue that the various transformational processes along the visual streams allow for multiple hierarchical representations of the visual world. These multifold maps cooperate in structuring large-scale network dynamics that underlie perception. Do maps inevitably lead us to the trap of a "master representation" equivalent to a Cartesian Theater? Other conceptual frameworks have been proposed that do not incorporate the idea of maps or, even more, avoid the notion of representation altogether (e.g., Thompson and Varela, 2001). As we argue below, maps alone cannot fully account for complex cognitive processes, such as perceptual stability. However, neuronal dynamics are likely constrained by representational maps. These hierarchical representations may allow multiple cooperative dynamic states to emerge via lateral, feedforward, and feedback connections.

An underlying assumption of visual maps, which has intrigued philosophers such as René Descartes (1596–1650; see also (Merleau-Ponty, 2013) and scientists (Lettvin et al., 1959), is that the world is represented inside our brain in a somewhat faithful manner (Fig. 1). Alternatively, we present evidence that vision is a highly constructive process, thereby undermining the primacy of a full representational model for vision. The relevance of this discussion is manifold. While it may seem straightforward that individual neurons code for simple features such as an oriented bar, what is the neuronal substrate for representing complex objects in the brain?

Charles Gross and collaborators provided the first physiological evidence that single neurons in the inferotemporal cortex were capable of coding for complex features, such as hands and faces (Gross et al., 1969, 1972). This experimental breakthrough provided support to the notion that abstract concepts and ideas could also be captured by the firing pattern of a single neuron. However, it contrasted with other prevalent concepts popular at the time based on distributed coding, as proposed by Walter Freeman. (Curiously, Gross spent the year of 1970 or so as a guest of Walter Freeman in the Department of Physiology and Anatomy in Berkeley before starting his career at Princeton University). This debate is still far from resolved. Would each newly acquired concept require its unique neuron? Alternatively, if some sort of cooperative coding is indeed in action between these classes of neurons, what is its nature and

mechanism? Notably, Charles Gross and collaborators made scientific contributions to a wide range of concepts that may underlie the way we represent the visual world, from visual maps to high concept coding neurons, and which are fiercely debated to this day. This discussion has direct implications for the way we probe neuronal activity: should we allocate resources to acquire reliable single unit activity (e.g., with tetrode recordings associated with spike sorting techniques) on the assumption that the necessary neuronal information is present at the single neuron level? Does multiunit activity provide a meaningful assessment of neuronal processing based on the assumption that local circuits and the columnar organization of the primate neocortex are basic processing units? Or does the key to understanding neuronal function rely on widespread multielectrode recordings? In an autobiographical assessment of his own work (Gross, 2002), Gross used the terms coarse coding, population coding, and ensemble to describe the potential mesoscopic level where neuronal coding might be taking place. The underlying techniques that we use to probe brain activity illustrate and underscore the unsolved issues regarding how the brain represents and processes information.

Next, we address the different streams of visual information processing (Ungerleider and Mishkin, 1982) and the impact of perceptual completion and filling-in on the visual system (Fiorani et al., 1992; De Weerd et al., 1995). We focus on visual representation in the brain and its effect on vision. We describe the visual representations as retinotopic, visuotopic, craniocentric or ambient. We propose that visual information be kept in a large and stable neural network composed of multiple areas and connected with the various processing pathways. These pathways extract the information necessary to identify objects (ventral stream), to understand language (lateral stream), to compute the trajectory of a target and/or to understand social interactions (dorsolateral stream), and to allow motor coordination for body movement, such as stepping and locomotion (dorsomedial stream). Below, we offer a broad assessment of the issues raised above. We intend to provide adequate detail throughout our text by focusing on work done by our group and on work performed by a limited number of groups working in related fields. We acknowledge, therefore, that several invaluable contributions made to the field of sensory visual neuroscience are not cited here, mainly due to space constraints.

2. Descartes's eyes

The French philosopher René Descartes (1596–1650) influenced several domains of human knowledge, and neuroscience was no exception. He had a particular mindset for spatial representation, as evidenced by the Cartesian coordinate system that he so successfully disseminated. Descartes provided one of the first comprehensive depictions of the visual system within the perspective of a representational framework (Fig. 1). His first important insight was to consider the role of both eyes in vision, and the potential integration of their input for further joint processing. There is no evidence that Descartes fully acknowledged binocular vision as crucial to three-dimensional perception, as we do today. However, his initial studies made clear that an object (depicted by an arrow in Fig. 1) could be independently projected onto both retinas and subsequently fused into a single and matched representation somewhere in the brain.

Second, Descartes made explicit the existence of multiple processing layers and offered an early hint to what we today define as a connectome. The concept of layers here can be interpreted in a more abstract way than the notion of cortical layers or maps. It can be viewed as having a direct parallel to the input, output and hidden layers found in artificial intelligence networks and machine learning algorithms. In Fig. 1, we can observe at least 3 layers, where corresponding points in visual space are depicted with distinct designations (i.e., letters or numbers) to imply that information is actually being processed and transformed as it travels along the system. Notably, a strict topological isomorphism is obeyed, at least up to layer 2. This implies that a
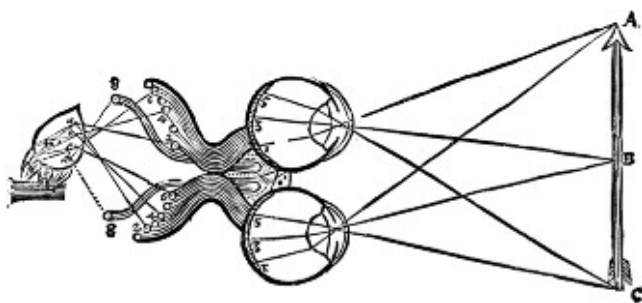
Fig. 1. Descartes's early depiction of the visual system embodied all essential concepts of a representational system. The diagram shows how distinct parts (letters A, B, C) of an object (arrow) within the external world can be orderly mapped, first onto the retina surface and then inside our brain, laying out the framework for an internal representational system of the physical world. The diagram is from Descartes's posthumously published *Treatise on Man*. For further details, see text.

topographical organization (i.e., a map) of the visual world is being explicitly created within the framework of a representational system.

On the interface between the second and last layers (i.e., the pineal gland) illustrated by Descartes in Fig. 1, he introduces the notion of input convergence, a vital concept in neuronal circuits and artificial networks alike. Input convergence allows for the coding of successively more complex feature as one progresses to layers 'higher' in the hierarchy. In Fig. 1, for example, points 4 and 4' (second layer) converge to point b in the third layer. Point b can, theoretically, code or represent a more elaborate joint feature of the visual stimulus arising from the coding of more primitive attributes by elements 4 and 4'. It is useful to interpret this model in light of Descartes' beliefs at the time. He considered the pineal gland to be the main seat of the soul and the location where thoughts are formed. Indeed, the complexity of human thoughts would require complex sets of associations and conjunctions made possible by input convergence. In the context of the specific topic discussed herein, input convergence is required to create the complex coding properties of inferotemporal neurons (Gross et al., 1969, 1972).

One final key concept that we could extract from Descartes' model is his acknowledgment of the sensory feedback to the motor system. In Fig. 1 (symbol '8'), the pineal gland, presumably after processing the sensory input to some extent, is able to feedback this information to the muscles that control eye position. This visuomotor integration would enable a continuous loop capable of fine-tuning perception, underscoring the importance of the motor system to sensation.

The underlying concept here is that distinct inputs from the external world can be orderly mapped inside our brain, laying out the framework for an internal representational system of the physical world. This concept is the foundation of the discussion below based on visual topography and maps.

3. Visual perception

Numerous studies from the fields of anatomy, neurophysiology and brain imaging have simultaneously addressed and challenged our current understanding of visual processing in the brain. Different works have shed light on the mechanisms of visual perception, each emphasizing distinctive attributes of vision. The early work of Daniel and Whitteridge (1961) emphasized the topographic organization of the primary visual cortex and its relation to visual acuity. They showed that the magnification factor of central vision was significantly higher than that of the visual periphery and suggested that this organization would account for the difference in visual acuity. Parallel pathways of visual processing are already observed within the retina with different classes of ganglion cells projecting to different subcortical structures, which in turn project differentially to cortical areas. The ganglion cells project to specific subdivisions of the dorsal lateral geniculate nucleus, which in turn project to different sublayers of layer IVc in V1. Separate compartments of V1 project to distinct cytochrome oxidase (CytOx)-rich and CytOx-poor stripes in V2 or to MT. We have been working with the notion that most visual processing is performed within cortical modules. Ascending and intrinsic circuits build these cortical modules to decode specific attributes of the sensory system. In the visual system, orientation modules were described first by Hubel and Wiesel (1968). They also proposed a hierarchical model for visual processing. This model implies that the concentric receptive fields of the dorsal lateral geniculate nucleus (dLGN) assemble orientation decoders or orientation selective cells arranged in columns in V1 to build simple cells, which would then produce complex and hypercomplex neurons. This notion suggests that these columns could construct edge detectors used in higher areas, such as the inferior temporal cortex, to assemble objects or form detectors, such as cells selective for complex objects described by Gross and collaborators (1972). While the hierarchical arrangement supports serial processing in the visual system (Hubel and Wiesel, 1968), the discovery of several areas with topographically organized maps (Daniel and Whitteridge, 1961; Allman and Kaas, 1973; Gattass and Gross, 1981;

Gattass et al., 1981; Gattass et al., 1987; Gattass et al., 1988), with modules selective for different attributes of the visual stimuli, such as motion or color (Zeki, 1974), creates the basis for parallel processing in the cortical visual areas. Sampling limitations of neural processing suggest that parallel processing is accountable for the efficient coding of an image. Several brain regions must work concurrently to allow for fast processing of the visual scene. The very nature of the neural signals and the connections between visual areas, with the action potential and the following refractory period lasting more than 1.2 ms, limits the propagation of the visual information to far less than 1 kHz. Thus, the interaction of cortical modules or synchronization between neurons is thought to be limited to an absolute maximum of 300 Hz. Nonetheless, when we open our eyes, we build a stable perception in a few hundreds of ms (Thorpe et al., 1996). To achieve this performance, parallel processing is crucial.

Processing of the different attributes of a scene relies on parallel processing. In V1, we have orientation columns, which carry out orientation selectivity, the main attribute for form perception (Hubel and Wiesel, 1969). Additionally, we have perceptual completion, an attribute for the binding of multiple features into a unified object (Fiorani et al., 1992; Azzi et al., 2015). In V1 and V2, we have the processing of retinal disparity, an attribute for three-dimensional vision (Barlow et al., 1967; Pettigrew et al., 1968). Area V4 shows color selectivity, an attribute for color vision (Zeki, 1974). Area MT shows selectivity for the axis of movement columns, an attribute for motion perception (Albright, 1984). In macaques, visual area V2 is the earliest site in the visual processing hierarchy in which neurons selective for relative disparity have been observed (Pettigrew et al., 1968; Thomas et al., 2002). By combining optical imaging, single unit electrophysiology and CytOx histology, Ts'o et al. (2001) revealed in greater detail the functional organization within the CytOx stripes of the visual area V2 of primates.

Based on the description of visual area PO in primates (Neuenschwander et al., 1994), we proposed a division of the dorsal pathway into dorsal medial and dorsal lateral streams (Gattass et al., 1990; Nascimento-Silva et al., 2003). The dorsal medial sub-stream is related to locomotion processing (Neuenschwander et al., 1994). The concept of visual information processing evolved in nonhuman primates toward three streams, with the first relay area in each corresponding stream receiving direct projections from the striate cortex (Zeki, 1974; Colby et al., 1988). Here, we extend this concept to four streams of information processing in humans, as illustrated in Fig. 2. Note that the different streams receive most of the feedforward connections from discrete portions of V1. The ventral stream receives projections from up to 30° of the central visual field representation of V1. The lateral stream is related to reading and receives projections from the central 5° representation. The dorsal lateral stream receives projections from the central 60°. Finally, the ventral medial stream receives projections from the peripheral field representation.

4. Streams of visual information processing in humans

Fig. 2 shows our proposal for four streams of visual information processing in humans related to different aspects of visual perception. The ventral pathway composed of projections representing the central visual field to inferior temporal areas is responsible for recognizing objects and faces. The lateral pathway composed of foveal projections to the angular gyrus, which subsequently projects to Wernicke's area, is responsible for language comprehension. The dorsolateral pathway composed of binocular field projections to areas of the superior temporal sulcus is responsible for the perception of movement, processing of target trajectories and social interactions. Finally, the dorsomedial pathway, consisting of projections representing the peripheral visual field to areas of the parieto-occipital sulcus and the parietal areas is responsible for motor coordination involved in walking and body movements. This figure emphasizes the extent of the visual field
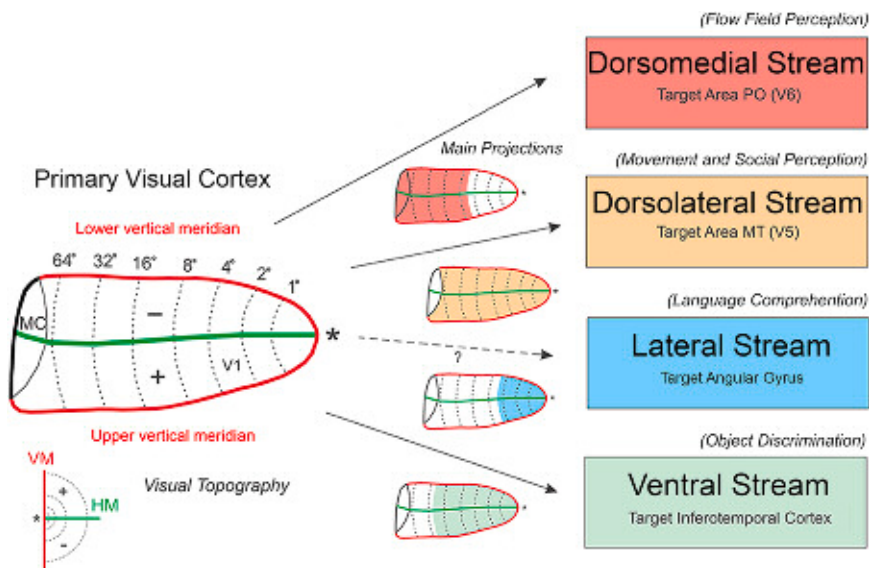
Fig. 2. **Early cortical visual processing can be sub-divided into four major pathways in humans**. The scheme illustrates the extent and topographic specificity of the portions of V1 that give rise to the four major cortical streams of visual information processing in humans. Right: flattened map of human V1, with the vertical meridian (VM) represented by the thick red line, the horizontal meridian by thick green line, the eccentricity lines by dashed lines, the visual field periphery by the thick black line, and the fovea by asterisk (*).

represented in V1, which mainly projects to these streams. Most, if not all streams receive direct projections from V1 to its initial target area. Studies in nonhuman primates confirmed a direct projection from V1 to the following areas: V2, V4, MT and PO (Colby et al., 1988; Cragg and Ainsworth, 1969; Kuypers et al., 1965; Nakamura et al., 1993; Yukie and Iwai, 1985; Zeki, 1978). Until now, there is no evidence based on neuronal tracer studies showing direct anatomical projections between V1 and the angular gyrus. However, there are results from diffusion-weighted magnetic resonance imaging (dMRI) showing fiber bundles interconnecting V1 and the angular gyrus in humans (Weiner et al., 2017). As a matter of fact, Weiner et al. (2017) argue that the emergence of language in anthropoid primates is paralleled by major evolutionary differences in the white matter bundles interconnecting the occipital pole and the temporal lobe. Regarding the connectivity between V1 and the angular gyrus, Roumazeilles et al. (2020) also provide evidence for direct projections using diffusion tensor imaging (DTI). More specific to our case, functional magnetic resonance imaging (fMRI) has shown sequential activation of the central region of V1, the angular gyrus, and Wernicke's area when subjects read a written word (Price, 2000).

The visual topography of V1 in humans is shown on the left of Fig. 2 and indicates a geometric decay in the magnification factor, with iso-eccentricity lines equally distributed at 1, 2, 4, 8 16, 32 and 64 degrees. Most of the available data from functional magnetic resonance imaging (fMRI) focus on the ventral stream of information processing, which is mainly comprised of V1, V2, V3, V4, TEO and TE. Areas in this stream located in the anterior temporal lobe have been shown to be responsive to faces, hands and houses (Kanwisher et al., 1997; Keizer et al., 2008; Pitcher et al., 2019). The fusiform face area in the temporal lobe is specialized for expert object recognition (Gauthier et al., 2000). As we probably spend more time looking at faces than at any other object, we therefore associate this stream with object discrimination.

fMRI data in humans show an activation of the central representation of V1, in addition to an activation of areas in the left occipital gyrus, when reading words. Activation of the central representation of V1 during reading was also reported by Bavelier et al. (1997). On the other hand, listening to words activates the auditory cortex and Wernicke's area, which is located in the superior temporal gyrus of the temporal lobe (Baars and Gage, 2010). High span readers show relatively greater activation of the left angular gyrus (Buchweitz et al., 2009). These results corroborate previous studies on listening and reading comprehension (Constable et al., 2004; Jobard et al., 2007; Michael et al., 2001).

The joint role of the visual cortex and the angular gyrus during reading was proposed by Geschwind (1965) in the first neurological model for language processing. Since his pioneering work, we have gained substantial knowledge regarding the intricate connectivity of the neural network involving the visual cortex, the angular gyrus and temporal regions such as Wernicke's area (Price, 2000). For example, the angular gyrus, whose function was thought to be restricted to the visual processing of written words, is now known to be a memory storage site for visual forms in general. In addition to attributing meaning to words during reading, it is also part of a wider semantic processing network; stimuli with shared meaning, such as visual objects, faces, and speech, can independently activate the angular gyrus. Accordingly, disconnections between the visual cortex and the left angular gyrus impair the ability to read letters, words, sentences, and symbols (i.e., alexia), with no significant impact on the writing capabilities of the individual (Damasio and Damasio, 1983). More recently, reading acquisition studies report a massive effect of reading in a specific brain site in left occipito-temporal cortex, which has been termed "visual word form area" (VWFA). During and after reading acquisition it starts to respond to orthographic stimuli in the learned script, and not to oral language, as attested by the lack of activation to spoken sentences (Dehaene, 2013). Reading effects are also sensed in a vast left hemisphere language network. This network is also activated by spoken language in all groups: left posterior, middle, and anterior superior temporal sulcus, left temporal pole, left and right premotor cortex, left inferior frontal gyrus, and left supplementary motor area. A significant left-hemispheric asymmetry of this effect is observed in all areas except for activation in the temporal pole and occipital cortex.

Behavioral data from patients with macular degeneration disease (McClure et al., 2000) suggest the existence of a new stream of visual information processing dealing with language comprehension. Here, we name this pathway associated with reading comprehension as the "lateral stream". Macular degeneration compromises mainly the foveal and parafoveal regions of the retina. In human and non-human primates, this central representation extends to approximately 5 degrees. Most of the acuity tests used in ophthalmological practice probe the properties of this lateral stream since the largest letter generally used in such tests (the letter "E") encompasses approximately 20 min of arc. Patients with advanced macular degeneration are unable to read with their peripheral vision (McClure et al., 2000). Hence, the dorsolateral stream, which contributes to visuomotor coordination used, for example, for driving, is not assessed by the standard acuity tests used in ophthalmology.

Recently, Eichert et al. (2020) investigated the extent to which

between-species alignment, based on cortical myelin, could predict changes in connectivity patterns across macaques, chimpanzees, and humans. Considering that the evolutionary adaptations of the temporoparietal cortex represent a critical specialization of the human brain, the authors were especially interested in how language evolved in primates. Their study showed that the difference between human and non-human primates could not be explained solely by changes in the positions of brain regions. Instead, it was found that the arcuate

fasciculus underwent additional changes in its course, which may have contributed to the evolution of language.

Most of the visual field representation, including its binocular overlapping region, projects to the dorsolateral stream of visual information processing, which includes areas MT and MST (Tootell et al., 1995) and areas in the intraparietal sulcus and parietal areas. These areas interact with sensory-motor areas and are responsible for the perception of object movement and visuomotor coordination (Gattass
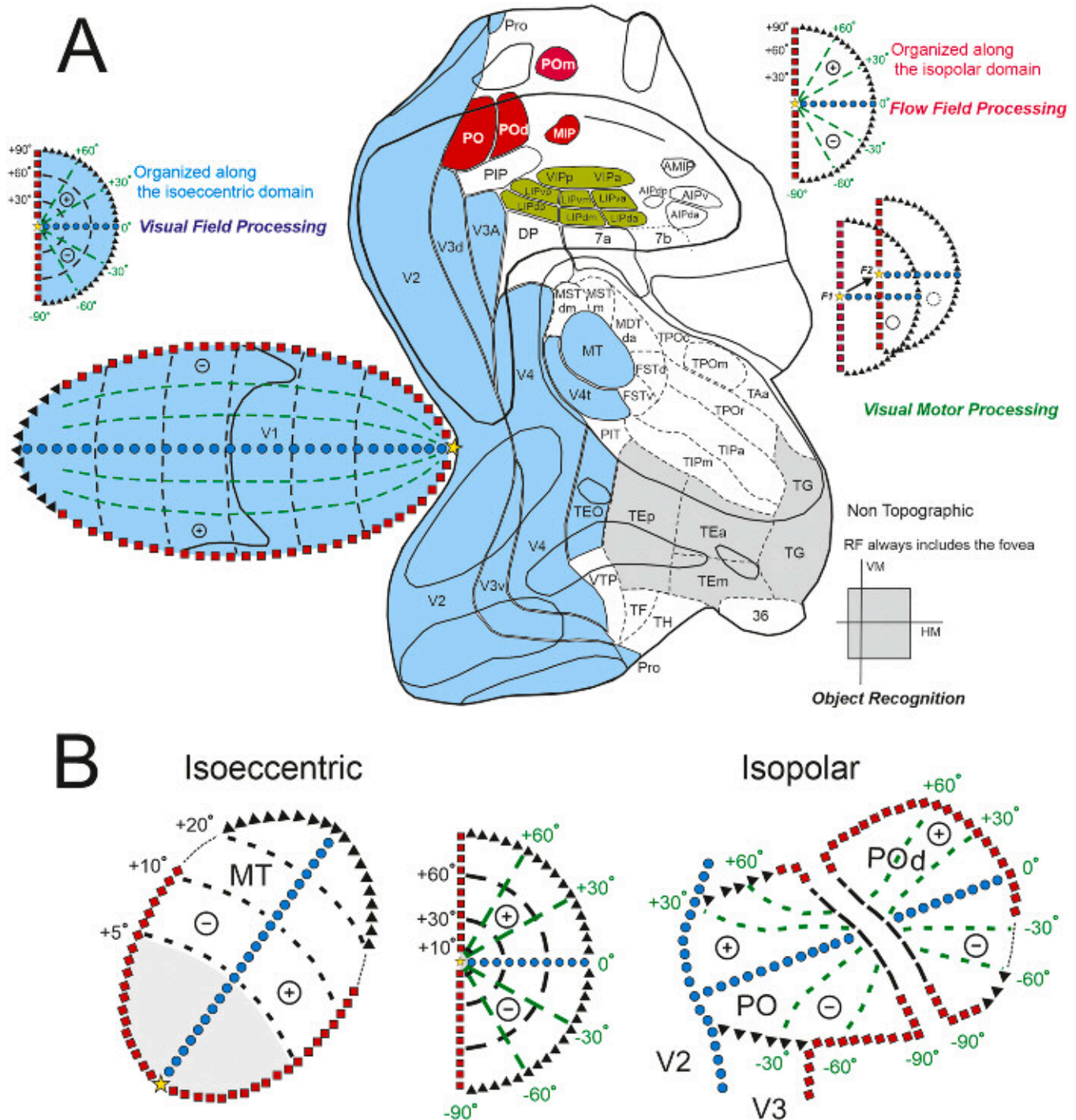
Fig. 3. The primate visual cortex is comprised of different types of topographic organizations. (A) A flattened map of the striate and extratriate cortices of a rhesus monkey brain. Colors illustrate the different types of topographic organizations attributed to the various visual cortical areas. Blue: areas with a formal visual topography. Red: areas with an organization based on the isopolar domain. Green: areas associated with visual motor processing. Gray: areas without a clear topographic organization. (B) Example of an area organized on the isoeccentric domain (area MT, left) and two examples of areas organized along the isopolar domain (areas PO and POd, right). A depiction of the corresponding visual field is shown in the center.

et al., 2020). This stream is used to aid most body movements and contributes to the ability to drive a motor vehicle. We thus consider the use of acuity tests to renew a driver's license to be insufficient. Visuo-motor tests in the binocular region are indeed more appropriate, even if stimulating only one eye.

More recently, Picher and Ungerleider (Pitcher and Ungerleider, 2021) proposed a new functionally relevant pathway projecting from the early visual cortex, via motion-selective areas, to the superior temporal sulcus (STS). Studies demonstrating that the STS computes the actions of moving faces and bodies (e.g., expressions, eye gaze, audio-visual integration, intention, and mood) show that this third visual pathway is specialized for the dynamic aspects of social perception (the dorsolateral pathway in Fig. 2).

Visual cortical areas have different types of topographic organizations (Fig. 3). We believe these different architectures reflect the particular type of visual processing that is taking place. Notably, they are intimately associated with the concept of visual maps that we employ here. The blue color in Fig. 3 depicts cortical areas with the traditional visual topography, where the isoeccentric and isopolar lines are clearly represented (e.g., areas V1, V2, V3, V4 and MT). It also shows areas that have their organization based on isopolar lines (e.g., areas PO and POd, in red), and areas with no clear topographic organization involved in object discrimination and recognition (e.g., area TE, in gray). Finally, there are cortical areas associated with visual motor processing (green). Fig. 3B shows detailed examples of two cortical areas organized along the isoeccentric or the isopolar domains (areas MT and PO/POd, respectively). For areas PO and POd, it is straightforward to define the representation of the isopolar lines, but not those of the isoeccentricity lines. This curious "disorganization" of the isoeccentricity lines in areas PO and POd is challenging to explain, especially when compared to the formal topographic organization of neighboring areas V2 and V3. We do not have definite data regarding the functional role of areas PO and POd, but we do put forward some hypothesis below.

In the dorsomedial stream, the peripheral field of V1 projects to areas PO and POd, areas in the intraparietal cortex, and areas in the parietal lobe (Colby et al., 1988). These regions are organized in the isopolar domain and are likely suitable for processing centrifugal and centripetal object movement. They project to areas of the intraparietal sulcus and areas of the parietal lobe (Gattass et al., 1985; Colby et al., 1988; Ungerleider et al., 2008; Mariani et al., 2019).

The segregation of the dorsal lateral and dorsal medial streams of visual information processing is based on several studies of the 1980´s and 1990´s. In the natural world, the flow of information as one moves through the environment (flow field processing) is the visual feedback generated by the observer's own locomotion, which provides cues for visuomotor coordination and the identification of features in the environment. Although related, we distinguish this self-motion visual feedback signal from optic flow. Areas VIP and MST are sensitive to optic flow (Cardin et al., 2012; Colby et al., 1993; Duffy and Wurtz, 1991), but there is no specific evidence that these areas code for the specific optic flow pattern produced as the individual moves through the environment. Here, we hypothesize that areas PO and POd, due to their isopolar-based organization, are better suited for this role (Neuenschwander et al., 1994). The segregation between the dorsolateral and dorsomedial streams is supported by neuronal connectivity data (Nascimento-Silva et al., 2003). On the other hand, the ability to code for the flow of information as one moves through the environment would characterize the dorsomedial stream. De Jong et al. (1994) have demonstrated that even the simplest form of optic flow is sufficient to activate two separate zones in the superior parietal and occipito-temporal cortices, zones which are commonly held to be part of the dorsal (spatial awareness) and ventral (object recognition) visual pathways, respectively (Mishkin et al., 1983; Newcombe et al., 1987). Therefore, the circuitry leading to these zones must involve both area MT (Gattass and Gross, 1981) and area PO (Gattass et al., 1985; Neuenschwander et al., 1994).

PO, also called V6, is a motion area of the dorsal medial visual stream in both macaque and human brains. Pitzalis and collaborators (Pitzalis et al., 2013) combined electrophysiological and neuroimaging methods to find the electrophysiological correlates of V6 and to define its temporal relationship with the neural activity observed in MT. They conclude that area PO plays a role in the perception of forward motion as previously suggested by Neuenschwander et al. (1994) based on the topographic organization of this area. Note that V6 is a small cortical area with no foveal magnification. For this reason, it is challenging to localize V6 in humans using retinotopic mapping with the traditional imaging techniques (Cardin et al., 2012).

Key differences between the ventral and dorsal streams of visual information processing can be traced back to anatomical projections originating in the early visual cortices that feed the corresponding pathways. The ventral stream inherits projections from the central 20° of the visuotopic map, with strong emphasis on the foveal representation. Likewise, characteristics of the dorsolateral and dorsomedial streams are rooted on early projection patterns. The foveal representation is deemphasized in both dorsolateral and dorsomedial streams (Gattass and Gross, 1981; Fiorani et al., 1989). The maps of the dorsolateral stream preserve the same isoeccentric organization as observed in early visual cortices (i.e., V1 and V2), with emphasis on peripheral vision. This latter characteristic is particularly relevant for the visuomotor transformations that take place in downstream areas of the dorsolateral pathway, such as LIP. As a direct consequence of the cortical magnification factor, peripheral vision is able to offer a relatively stable scene despite small eye movements. Cortical and subcortical regions with a visuotopic representation of the world, such as V1, V2, V3, V4, MT and the superior colliculus show clear representations of isopolar and isoeccentricity lines (Fig. 3). However, different from the ventral and dorsolateral streams, the dorsomedial pathway relinquished the cortical organization based on isoeccentric maps (Gattass et al., 2005) in favor of a cortical organization based on isopolar maps (Neuenschwander et al., 1994). This new isopolar coordinate system (Fig. 3B) is likely important for the implementation of our egocentric map and navigation in the world. We propose that the ventral and dorsal streams participate in two distinct dynamic neural networks. The ventral stream and its associated visuotopic maps, with emphasis on the foveal representation, are key to conscious and explicit perception. On the other hand, the dorsal pathway is key to visuomotor transformations, which may not be directly accessible to conscious perception (Goodale and Milner, 1992).

5. Visual pathways: the pervasiveness of parallel processing in the brain

That serial processing takes place in the brain is quite apparent and widely accepted. What may be less obvious are the numerous operations executed in parallel, making visual processing a highly distributed system. Parallel processing starts already in the retina, the obligatory entry point of all external visual information to the brain. The retina is not the equivalent of a single sensitive film from which neural images are formed. It is better compared to millions of tiny and overlapping neural films, each one of them breaking up the optical image formed by the eye and transmitting a separate filtered version for further visual processing (Shapley and Perry, 1986). Hartline (1938) was the first to establish that single ganglion cells respond to restricted portions of space. By dissecting and recording from individual optic nerve fibers in the frog, he observed discharges only when specific areas of the retina were illuminated. He defined this region as the "receptive field" of the neuron, being the first one to use this denomination for the visual system. However, Hartline used a single dot of light to stimulate the retina and, therefore, could not advance further into the detailed structure of ganglion cell receptive fields. Using two independently controlled light beams, Kuffler (Kuffler, 1953) was able to describe the center-surround structure of ganglion cell receptive fields in the cat retina. He reported that the receptive fields had a central region, which had, upon light

stimulation, a discharge pattern opposite to that found in the surround. The "ON" center ganglion cells required luminance increments for the center, and decrements for the surround, in order to discharge. The reverse was true for the "OFF"-center cells. Interestingly, retinal ganglion cells responded very poorly to homogenous illumination over their receptive fields. This property introduced an important concept in visual processing: relative contrast is far more relevant for information encoding than absolute luminance. In addition to the "ON" and "OFF"-center systems, the retina provides the essential neuronal substrate for the magno-, parvo- and koniocellular pathways that ascend to the primary visual cortex via to the lateral geniculate nucleus (reviewed in Merigan and Maunsell, 1993; Nassi and Callaway, 2009). Both pathways are considered to work in parallel performing complementary functions. The most striking difference between them is the sensitivity of the parvocellular pathway to color, which is absent in the magnocellular system. Parvocellular neurons also show a more tonic response to persistent light stimulation when compared to magnocellular neurons. The magnocellular pathway, however, is much more sensitive to contrast than the parvocellular system. For other dimensions, such as temporal and spatial frequency sensitivity, both systems cover a wide range of values despite significant differences in their mean response.

As shown in Fig. 4, the segregation between the magno- and parvocellular pathways is preserved in the projections from the retina to the lateral geniculate nucleus of the thalamus, as evidenced by physiological (Schiller and Malpeli, 1978) and anatomical (Conley and Fitzpatrick, 1989; Perry et al., 1984) studies. Projections of both pathways from the dLGN to the primary visual cortex are also believed to remain segregated. Hubel and Wiesel (Hubel and Wiesel, 1972) verified that projections of the magnocellular pathway terminated preferentially in layer 4Cα, while projections of the parvocellular pathway terminated preferentially in layer 4Cβ of the primary visual cortex. The intrinsic connectivity of V1 and its pattern of projections to downstream visual areas suggested that the segregation between the magno- and parvocellular system could pervade throughout the visual system. Early studies, for example, reported that layer 4B received inputs from layer 4Cα but not from layer 4Cβ (Lund, 1988), implying that the magnocellular system dominated neuronal processing in layer 4B. This dominance is particularly relevant because layer 4B provides major projections to MT, which is of crucial importance for visual motion processing. Based on the evidence that distinct channels of information arising in the retina remain segregated to the highest levels of the cortex, Livingstone and Hubel (1988) proposed that the visual system comprised two largely independent subsystems. One, mainly derived from the parvocellular pathway, would mediate visual behavior related to color and form. The other, dependent on the magnocellular pathway, would be responsible for the perception of movement and depth. A vast amount of anatomical and physiological evidence, however, suggested intense interactions between the magno- and parvocellular systems already in V1, and that their segregation was far from complete (Merigan and Maunsell, 1993; Nassi and Callaway, 2009). Sawatari and Callaway (Sawatari and Callaway, 1996), for example, reported that 4B layer neurons do receive, in fact, projections from layer 4Cβ and that the parvocellular system does indeed project to areas responsible for motion processing. Another influential model, which emphasized parallel processing in the visual system came from the work of Ungerleider and Mishkin (1982) and Mishkin et al. (1983). These authors also proposed the existence of two pathways for visual processing. However, different from the proposal of Livingstone and Hubel (1988), their model did not rely on the segregation of the magno- and parvocellular pathways. Mishkin and collaborators argued that the primary visual cortex and not the retina was the major source of two pathways. One of the pathways, the ventral stream, interconnected the primary visual cortex with the inferior temporal areas and enabled the identification of objects. The other cortical pathway, the dorsal stream, interconnected the primary visual cortex with parietal areas and allowed for the localization of objects. Experimental evidence for this "what/where" model relied heavily on lesion studies performed in nonhuman primates. Pohl (1973) demonstrated a dissociation of visual effects after inferior temporal and posterior parietal lesions in monkeys. Temporal but not parietal lesions produced severe impairment on object discrimination tasks, suggesting that the inferior temporal cortex participates mainly in acts of noticing and remembering the quality of objects. Parietal but not temporal lesions, on the other hand, produced deficits in localization tasks, suggesting that the posterior parietal cortex is concerned with spatial relations among objects and not their intrinsic qualities. A modification of the Ungerleider and Mishkin's model was subsequently proposed by Goodale and Milner (1992) on the basis of human case studies. According to this new model, "vision for perception" is processed in the ventral stream, while "vision for action" is processed in the dorsal stream. Thus, the what/where model and perception/action model attribute a fairly similar function to the ventral stream but not to the dorsal stream, and some results are consistent with one but not with the other (Sheth and Young, 2016). Notwithstanding, these two models have co-existed for about 30 years and have significantly contributed to our understanding of visual cortical function.

We usually experience our perceptions as single and unique events. Therefore, parallel processes being executed in our brains need to be eventually bound together into coherent percepts and actions. We still lack, however, a general mechanism for perceptual binding that is supported by solid experimental evidence. Mishkin and collaborators argued that the ventral and dorsal streams have both connections to the temporal and frontal lobe, and that each of these target areas constitute a potential site of convergence for object and space information. Single cells responding to the location of a specific object in space have been indeed described in the entorhinal cortex (Suzuki et al., 1997) and hippocampal formation (Rolls et al., 2005). This solution for perceptual binding, however, only exacerbates the problem of "combinatorial explosion" already discussed for the "grandmother cell" proposal. Additionally, a growing number of reports show appreciable anatomical crosstalk between the ventral and dorsal streams (DeYoe and Van Essen, 1988; Felleman and Van Essen, 1991), suggesting that binding may take place in a distributed manner rather than at highly convergent sites. Alternative mechanisms for perceptual binding, based on the precise timing of spiking responses, have been proposed (Singer, 1999). The advantage of such mechanisms is that they are flexible and can operate at various levels along the visual pathway. But they too face challenges posed by the architecture of the visual system. Consider, for example, the difference in conduction velocity for discharges traveling along fibers of the magno- and parvocellular pathways. The stouter magnocellular pathway axons conduct impulses more rapidly (Gouras, 1969). Therefore, retinal signals carrying information about color appear to be propagating some milliseconds slower than signals coding for other properties of the same object.

6. Beyond maps: coding complex multidimensional features

When asked by his students how neurons represent objects, neuroscientist Jerome Lettvin came up in 1969 with a humorous and fictitious story that has gradually become a serious theme in neuroscience circles (Gross, 2002). The main personage of his tale was his second cousin, Dr. Akakhi Akakhievitch, a neurosurgeon living in the distant Ural Mountains. Dr. Akakhievitch was convinced that single neurons were capable of representing ideas. Ever more obsessed, he decided to search for those neurons representing the most primitive of all objects: one's mother. He was able to find some 18,000 neurons in the human brain that responded exclusively to a mother, whichever way she was presented. Not only was Dr. Akakhievitch very pleased with his breakthrough, he also found immediate practical application for his discovery when a patient named Portney entered his office. Portney was psychologically tormented by problems involving his own mother, and Dr. Akakhievitch was keen to help. After being led to the operating table and having each one of his 18,000 mother neurons ablated, all his problems were finally solved: he
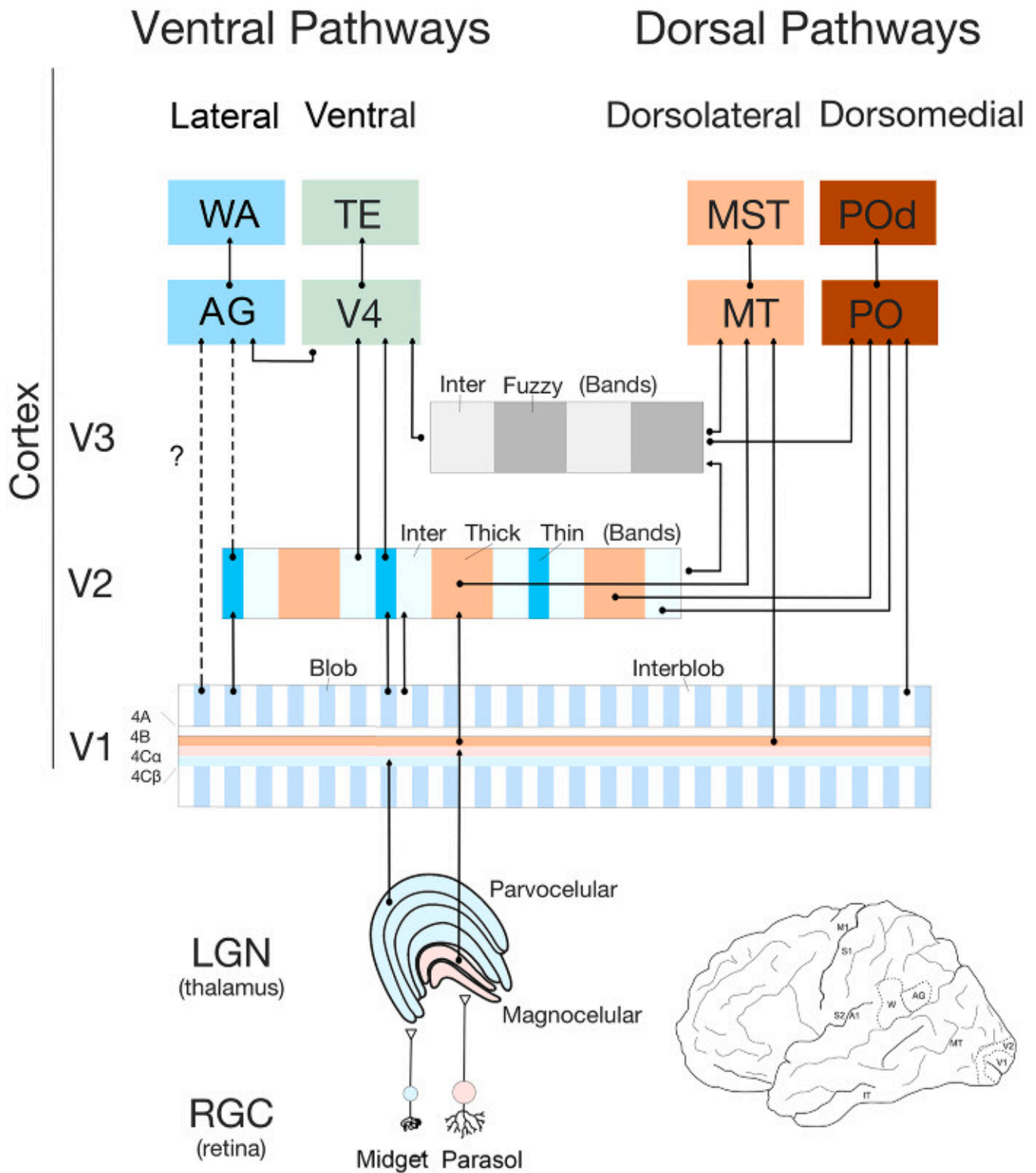
Fig. 4. Early cortical visual processing can be subdivided into four major pathways in humans: the ventral, lateral, dorsolateral and dorsomedial streams. The scheme illustrates the major streams of visual information processing starting at the retina and propagating through the thalamus, early and intermediate visual areas up to the inferotemporal and posterior parietal cortices. Parallel processing starts at the retina and thalamus with the well-established parvo- and magnocellular pathways. These two pathways provide inputs to V1, are intermingled to a considerable extent, and are processed through intricate circuits within well-defined compartments and modules. This gives rise to a qualitatively different type of information, which is then fed to new parallel streams of visual processing. Data suggest the existence of at least four distinct pathways, all partly originating in V1: the ventral, dorsolateral and dorsomedial streams of visual information processing. The latter two would be part of the larger dorsal stream. Abbreviations: AG, angular gyrus; LGN, lateral geniculate nucleus; MT, middle temporal area; RGC, retinal ganglion cells; TE, inferotemporal cortex, WA, Wernicke area. Data shown are based on Gattass et al. (1987), Colby et al. (1988), Gattass et al. (1988), (Rosa et al., 1988), Fiorani et al. (1989), Nakamura et al. (1993), Neuenschwander et al. (1994), (Gattass et al., 1997; Piñon et al., 1998), and Sousa et al. (1991), among others.

lost the concept of his mother. After solving the enigma of the mother cells, Dr. Akakhievitch then turned to his next great challenge: the grandmother cells.

Lettvin's interest on mother and grandmother cells probably originated from his own work in the frog (Lettvin et al., 1959). While recording in the frog's retina he was able to describe cells coding for various complex stimuli. One special type of ganglion cell, described as a "bug detector", responded uniquely when a small dark object entered its receptive field and moved about intermittently thereafter. Additional evidence coming from monkey recordings began to definitively convince the neuroscience community of the existence of cells capable of uniquely responding to specific complex objects (Gross et al., 1969; 1972). Serendipity led these researchers to verify that single neurons were capable of selectively responding to a hand (indeed, the first tool experimenters use to probe neuronal visual responsiveness are usually their own hands). In an effort to sort out which component of the hand (e.g., the high-frequency component of the fingers, or the low-frequency component of the palm) was preferentially driving the response, these authors ranked response-driven firing rates in six categories. Some neurons were particularly selective to the high spatial frequency components of the stimulus, which motivated testing other handy objects in the lab exhibiting high spatial-frequency components (personal accounts from these early experiments reported that the typical brush found in every toilet was particularly effective in driving the neuronal responses). Gross and collaborators performed a series of additional experiments after the two seminal works of 1969 and 1972, which addressed neuronal selectivity to other complex objects. They continued exploring neurons in the inferotemporal cortex, as well as neurons located in nearby regions, such as the superior temporal sulcus. Selectivity to human and monkey faces revealed to be of crucial importance. Neuronal responses turned out to be selective to faces, since the presentation of a scrambled monkey face elicited no response that was distinguishable from baseline. Several other works published thereafter not only confirmed these findings but extended them in describing cells responding specifically to other complex objects such as faces (Perrett et al., 1982; Desimone et al., 1984).

By the time the "grandmother cells" were discovered in the monkey temporal cortex, the hierarchical description of the visual system, as put forward by Hubel and Wiesel (1962, 1968), was already widely accepted. Based on the recording of single, well isolated units in the early visual cortex, they proposed that the projection of dLGN neurons having aligned receptive fields could entirely explain the origin of orientation selective cells in the primary visual cortex. These would initially consist in the simple cells, with their well determined ON and OFF sub-regions. The convergence of several simple cells would give rise to the complex cells capable of generalizing the coding of a specific orientation throughout its receptive field. Complex cells would give rise to the hypercomplex cells following this same line of thought (Hubel and Wiesel, 1968). In this framework, the "grandmother cells" constituted a natural unfolding of how the visual system, and ultimately the brain, would build a representation of complex stimuli in the visual world. Unit recordings in the human brain by Quiroga et al. (2005) corroborate the notion originally reported in monkeys that single neurons may embody the concept of a unique person or a unique object.

In 1972, Horace Barlow proposed the "neuron doctrine", which linked sensation directly to the activity of single units in the brain. In part, Barlow's proposal offered a significant paradigm shift. The activity of individual cells was considered too unreliable to be examined singly. Therefore, many turned to more macroscopic manifestations of neuronal activity, such as the electroencephalogram (EEG), in order to understand behavior. Barlow, on the other hand, was able to capture and synthesize a thinking that would dominate neurophysiology for many years to come: that single neurons are the prime movers leading to our perceptions and sensations (Hubel, 1982). Not every cortical neuron's activity had necessarily a simple perceptual correlate, but active high-level neurons directly and simply caused the elements of our perception.

The mechanism by which this was achieved consisted in the frequency of neural impulses (i.e., action potentials). Barlow argued that the rate of firing coded the certainty that the cause of a percept was present in the external world.

A pertinent criticism generally applied to the single unit doctrine is the one commonly referred to as "combinatorial problem" or "combinatorial explosion" (Singer and Gray, 1995). We are constantly confronted with unlimited combinations of elements and features, and yet we face no problem in recognizing patterns and objects presented in our visual scene. Additionally, we can easily cope with patterns not previously exposed to us, and we can effortlessly recognize objects encountered in unusual orientations and sizes, under different illumination conditions, or partially occluded by other objects (Logothetis and Sheinberg, 1996). Even though it may seem difficult to imagine that each one of our percepts is represented by a single high-level neuron, it has been suggested that there is a reasonable match between the number of possible percepts and the number of high-level neurons in the brain (Ghose and Maunsell, 1999). Contrary to what is usually assumed; however, Barlow did not believe in the existence of "grandmother cells" (Barlow, 1972). He acknowledged that there would not be enough "grandmother cells" or, as he called them, "pontifical cells" to account for the enormous variety of our percepts. He believed that the activity of a single neuron would not be able to convey the richness of our sensations, which consist of the combination of our various percepts. Barlow, therefore, proposed the term "cardinal cells" instead of "pontifical cells" to explain how the brain represented the world. The interaction of various "cardinal" cells, which are lower in the hierarchy and more numerous than "popes", would be responsible for the emergence of our perceptions and sensations.

The general mechanism used by the brain to build high level representations of the world is still largely unknown. We have, on the other hand, accumulated some knowledge regarding the functional organization of the visual system, which provides clues and constraints on how this representation might take place. Some aspects of neuronal pathways illustrate two basic points of early visual processing. First, that different features of a stimulus are processed by parallel pathways. Second, that the early visual system is organized in a modular and topographic fashion, such that a single object activates spatially distributed neurons with appropriate specificity. These two basic characteristics suggest that additional mechanisms are necessary in order to coordinate the neural activity processed in such a highly distributed manner.

## 7. Perceptual completion and fillingin as study cases for constructive vision

Perceptual completion is a phenomenon by which contours and shapes are perceived even though these features are not physically present in the retina. The study of completion is a case for perceptual constructivism since it violates the idea of visual representation. In general, retinas have a region naturally devoid of photoreceptors known as the blind spot. It corresponds to the head of the optic nerve. This discontinuity in the receptive surface is, under normal circumstances, not accompanied by abnormal perception, even in monocular conditions. Fiorani et al. (1992) showed that neurons within the cortical representation of the optic disk in V1 interpolate the receptive field position for the contralateral eye based on information of the stimulus present beyond the boundaries of the blind spot. Fig. 5 shows neuronal properties underlying perceptual completion (Azzi et al., 2015). It is possible to map receptive fields inside the blind spot representation of V1, where no corresponding photoreceptors are present in the retina. In Fig. 5A the multiple circular outlines depict automatically mapped receptive fields (Fiorani et al., 2014) for a multielectrode matrix inserted in a V1 region partially overlapping the blind spot representation (large oval dotted line). However, that is only possible when elongated stimuli, extending beyond the blind spot, are used for visual stimulation (global mapping). In Fig. 5A left, small light dots were used for visual
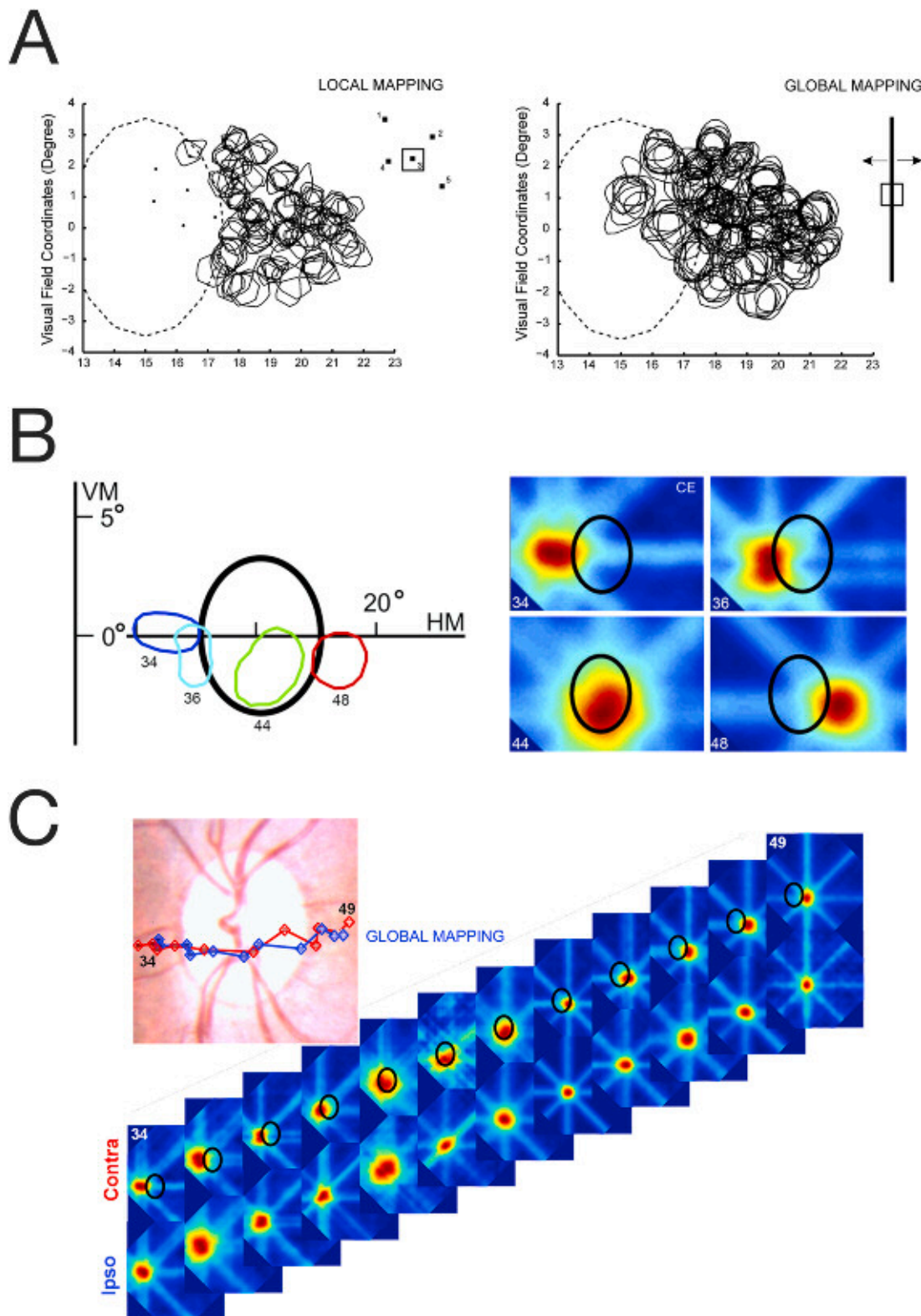
Fig. 5. V1 is best described as having a visuotopic map rather than a retinotopic map. (A) Receptive fields automatically mapped with a multielectrode matrix inserted in a V1 region partially overlapping the blind spot representation (large oval dotted line). Left: Flashing dots are used as visual stimulus (local mapping); Right: a moving bar was used as visual stimulus (global mapping). (B) Example cases of 4 linear single-electrode penetrations that cross the blind spot representation (black oval shape) in V1. The corresponding receptive field for each penetration (penetrations 34, 36, 44 and 48) was mapped using moving bars. The four subpanels on the right show the receptive field plots based on neuronal firing rate for the corresponding penetration. (C) Same data as in (B), but now including the entire set of penetrations that crossed the blind spot representation in V1. Abbreviations: VM, vertical meridian; HM, horizontal meridian. The insert in (C) is a photograph of the blind spot of the stimulated eye with the receptive fields centers mapped with an automated system.

stimulation (local mapping) and the receptive field did not extend into the blind spot. The 5 points plotted inside the blind spot correspond to the estimated receptive field centers. The neuronal activity elicited by the moving bar on the flanks of the contralateral blind spot can be interpolated by V1 neurons to generate activity inside the blind spot representation. We used bars with multiple orientations in order to map the center of the receptive fields. Notably, the resulting map of receptive fields follows an orderly topographical organization, continuous and comparable to the one found outside the blind spot representation. A retinotopic map for V1 would have predicted no neuronal activity inside the blind spot representation, in as much as no photoreceptors are present in the retina for this region. Therefore, V1's map is better described as visuotopic, since not only is it possible to map receptive fields inside the blind spot, but the mapped receptive fields are organized topographically relative to the visual field (i.e., visuotopically). Fig. 5B shows examples of 4 linear single-electrode penetrations that
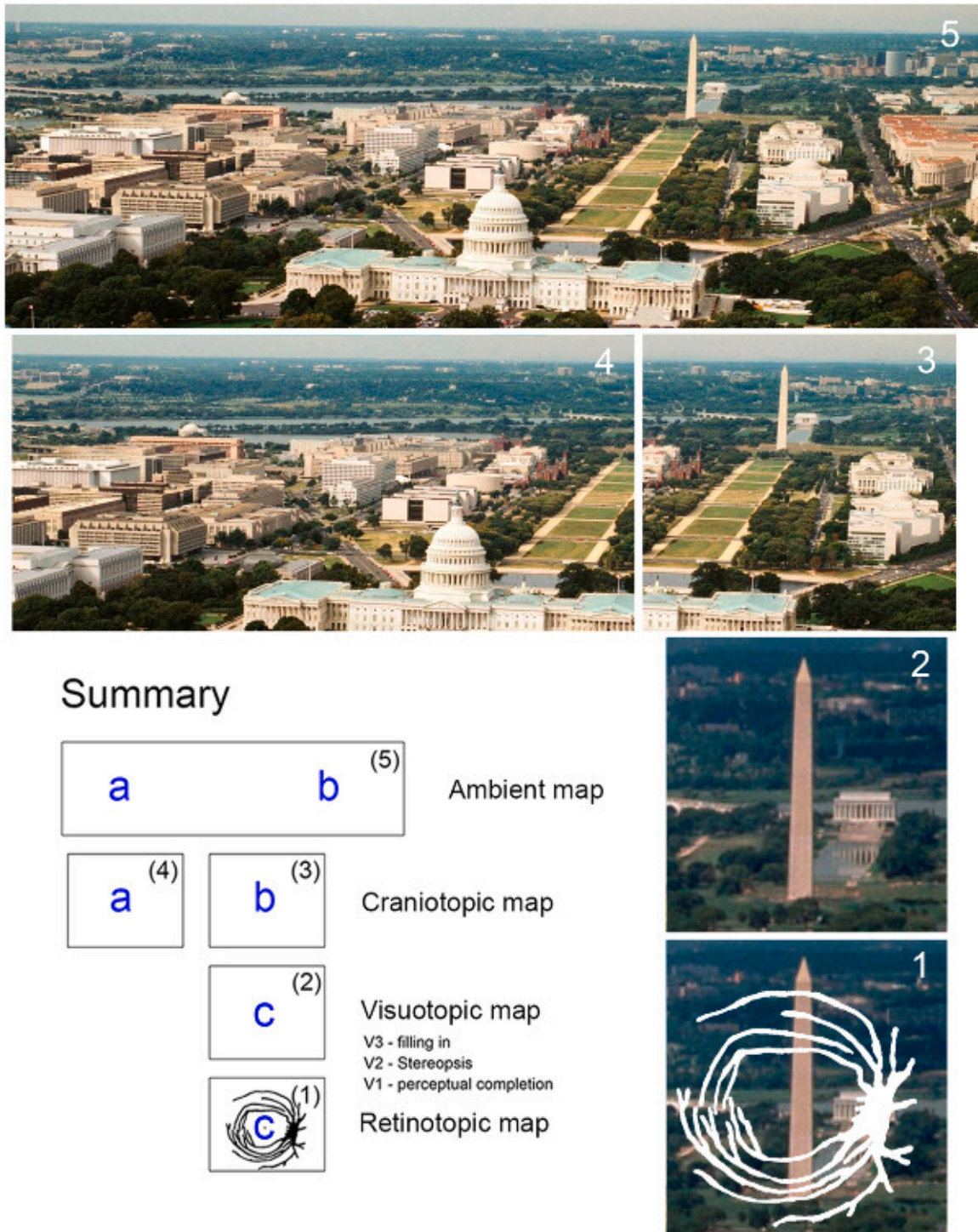
Fig. 6. A model for the several coordinate systems in vision - Neuronal representations. Representation of a retinotopic map of the obelisk in Washington, DC (1) and its corresponding visuotopic map (2). Different craniocentric maps depending on the position of the head (3,4). An ambient map (5) of an aerial view of Washington, DC assembled in the dynamic network of the cortex. (a-c, corresponding locations). Views 3 and 4 do not have the same magnification solely due to illustration purposes. For details, see text.

cross the blind spot representation (black oval shape) in V1. The corresponding receptive field for each penetration (receptive fields 34, 36, 44 and 48) was mapped using moving bars. The four sub-panels on the right show automatically mapped receptive fields (Fiorani et al., 2014) based on neuronal firing rate for the corresponding penetration. The colored ellipsoids on the left panel correspond to the receptive field outlines. Note that the position of the receptive fields follows the expected topography (i.e., linear progression), regardless of if its center is covered or not by photoreceptors in the retina. Fig. 5 C shows the same data as in Fig. 5B, but now including the entire set of penetrations that crossed the blind spot representation in V1. Here, we also present the receptive field plots for ipsilateral eye stimulation (blue and red for ipsilateral and contralateral eye stimulation, respectively). For ipsilateral eye stimulation, the contralateral eye was closed, and thereby no blind spot was present for this portion of visual field. Notably, receptive field positions for contralateral and ipsilateral eye stimulations match each other well, showing that the blind spot representation maintains a precise topographical map of the visual world. Most V1 neurons receive binocular input. Therefore, even though portions of the visual world are either permanently occluded by the natural scotoma produced by the constitution of the retina (i.e., optic disk and blood vessels) or occasionally occluded due to object/surface partial superposition, the topographically corresponding information reaching the contralateral eye in V1 can potentially compensate for this information loss. In other situations, however, there is no corresponding input from the contralateral eye. This phenomenon may be due to binocular occlusion of portions of the visual scene or because the individual has only one functional eye and must thereby deal with a permanent ~ 5º per 7º blind spot in his/her visual field. In either case, perceptual completion is an essential operation for "constructing" information that is not present in the retinal surface. This operation, based on intrinsic neural connections and contextual mechanisms, motivated us to define the V1 map as visuotopic, and not retinotopic. Arguably, only the dLGN shows a truly retinotopic map. Botelho and collaborators showed that the ability to interpolate receptive field position across substantial distances is present in neurons in other portions of V1 as well (Botelho et al., 2014).

8. Visual representations in the brain

Fig. 6 compares retinotopic (1) versus visuotopic representations (2). Different head positions generate distinct craniocentric maps (3 and 4). Regardless of head and body position, the ambient map (5) prevails. A retinotopic map refers to the orderly mapping of receptive field positions in retinotopic coordinates in the brain. A retinotopic map implies the existence of a neuronal representation organized in retinotopic coordinates. The significance of visual and brain maps originates with the 19th century debate concerning localization of brain function. Evidence for the existence of retinotopic maps and, by implication, for functional localization in the visual cortex arose from analyses of visual field scotomas resulting from partial injuries to the visual cortex caused by bullet wounds sustained by soldiers in combat (Holmes, 1918). These studies showed a predictable relationship between the region of damage in the striate cortex and the location of the area of blindness in the visual field (Baars and Gage, 2010). Curiously, we only found evidence for retinotopic maps only at the dLGN.

9. Active vision leads to visual perception

The map of V1 is visuotopic; that is, the area reconstructs the image representation based on predictable cues. When looking at a newspaper page printed with many imperfections, we automatically reconstruct the text, or the imperfections (partially interrupted letter fonts) based on local circuits or feedback connections to V1. Thus, the representation of the image in the primary visual cortex is visuotopic and not retinotopic. Visual perception is three-dimensional. It presents several proprieties described as perceptual completion in V1, filling-in in V3, stereoscopic

responses (due to retinal disparity) in V2, and color representation due to the processing in V4. Visual representations in the neocortex are based on extensive parallel, serial, and feedback circuit connections. It is stabilized due to feedback from the efferent copy of the control of the extraocular muscles (for review, see Sun and Goldberg, 2016). Remapping and perceptual inhibition are characteristics of image representation in the neocortex. If we keep our head fixed, the representation of the image (craniocentric) in the neocortex is stable, despite frequent eye movements. However, if one moves the head or, more generally, the position of the skull with respect to the ambient environment, the neural representation of the scene changes, and a new perspective of the image replaces the original one (Fig. 6). There is no evidence that the medulla, which controls neck muscles, sends efferent copy signals to the neocortex. Thus, compensatory mechanisms for stabilizing percepts are unlikely to occur in this case. Presumably, this is why we experience an afresh (craniocentric) percept of a scene shortly after each head shift. When the head assumes a fixed position and a subject opens hers/his eyes in a dark room with no light, no craniotopic neural representation is established. If the room is lit, however, about 100–300 ms after we open our eyes, a neural representation is formed, and we can compare it with our memory content to understand the visual scene. A sudden change in head position updates the neural representation and a new visual perception is formed. This neural representation is not retinotopic, like the maps of the layers of the lateral geniculate nucleus, or visuotopic, like the map of the striate cortex, V1. It also remains stable when we make eye movements. The neuronal representation of the scene is also modified when we move forward or backward in the ambient environment. Centrifugal and centripetal movement of objects in an ambient environment generate maps used for visual motor coordination. These visual representations of the ambient environment are based on an egocentric map.

10. Eye movement and remapping

When we move our eyes but maintain our head fixed, our perception of the visual world remains nevertheless stable. Different neuronal circuits and mechanisms likely subserve this perceptual phenomenon (Wurtz et al., 2011). We think two factors are essential for supporting this ability: the topographical organization of early visual areas and the low magnification exhibited by the peripheral visual field representation compared to the foveal representation (Gattass et al., 2005). As a result of the combined effect of these two factors, eye movements of small amplitude will only slightly disrupt image stability in the periphery. Take, for example, the topographical organization of area MT. Not only a decrease in magnification factor takes place as a function of eccentricity in this area, but also a deemphasis of the foveal representation relative to V1 and V2 (Gattass and Gross, 1981; Fiorani et al., 1989) . Considering that area MT is a major input source to VIP (Ungerleider and Desimone, 1986) and subsequently to LIP, its visual topography per se is likely to be determinant for the subsequent transformations leading to perception invariance in the face of eye movements.

Compared to the early visual cortex, associative cortical areas like LIP have far more complex and task-dependent spatial representations (Patel et al., 2014). A comprehensive understanding of how topographically organized projections from the early visual cortex influence visuomotor transformations in the associative cortex is still needed. Hardwired topography may coexist with dynamic maps, as suggested, for example, by the work of Ungerleider et al. (2008), which shows that the central representation of V4 is reciprocally connected with the anterior portion of LIP. Patel et al. (2014) offer an interesting discussion on the possible interplays between topographic projections, visual attention, and eye movement compensatory mapping in LIP.

Fig. 7 compares the representations of a craniocentric map (Fig. 7A) with those of visuotopic maps (Fig. 7B-D). When we look at the waterfall based on the three fixation points (colored plus signs) the craniocentric perception is stable (Fig. 7A) regardless of the changes in the retinal

Fig. 7. Perceptual spatial constancy in the presence of eye movements and visual perception with natural and artificial eye movement. Different parts of the scene shown in A can be scanned with eye movements without altering the perception of the image. This perceptual constancy occurs even though a different image is represented within the retina when the eye shifts its gaze to a new location (+ in B, C, and D). (E and F). The brain has developed intricate circuits to stabilize our perception of the visual world. Visual capture of the external world is inherently unstable. Eye movements are a prime source of this instability, as are body movements and movements in the environment. Nevertheless, our subjective perception of the corresponding scene is incredibly stable. Note that our subjective perception of the visual world is much closer to a stabilized continuous frame (E) as compared to (F), the image predicted by the movement of our eyes. Notable, our brain is efficient in stabilizing vision for inherently generated perturbations, such as saccadic eye movements. However, if we tap one of our eyeballs with our finger, we perceive the unstable jittered image in (F).

image and the corresponding V1 representations. Notably, the representation within the network of visual areas remains stable. However, if we move our head, perception changes, and a new perspective of the visual scene emerges. This difference is related to the nature of the integration process taking place at areas controlling eye movements, areas controlling head movements and cortical visual areas (Duhamel et al., 1992; Goldberg et al., 2002). Motor nuclei receiving inputs from extrinsic eye muscles send their efferent copies to areas located at the intraparietal cortex. Proprioceptor information is well integrated in the cerebellum and is responsible for harmonious motor coordination, resulting in precise control of head movements. Thus, stabilization by feedback of the efferent copy of eye movement signals contribute to a stable representation of the visual scene (Fig. 7E), while the perception of the same scene during an extrinsic movement of the eye (e.g., by externally tapping one's eyeball with your finger) causes a destabilized (fuzzy) perception (Fig. 7 F).

Any map or coordinate system implemented by the brain is dependent on and built upon a precise retinotopic coordinate system. As a matter of fact, receptive field remapping requires the interplay of precise visuotopic coordinates and precise coordinates of eye position. Neurons in areas such as V1 depend entirely on retinal position for their coding properties, while neurons in downstream areas, such as LIP and FEF, can remap their receptive field positions based on intended eye movements. There is a profound change in the encoding properties of single neurons from the latter cortical areas compared to V1. The impact this has on vision is debatable, but a reasonable assumption is that it allows for the emergence of a more stable perception that is less dependent on eye position. How to characterize the spatial coordinate system that results from this distinct transformation? We think that "craniotopic" or "craniocentric" are suitable designations here. Moreover, we believe that receptive field remapping is a key component for generating a craniotopic map. To our knowledge, experimental data has not yet revealed any single downstream cortical area capable of entirely and unequivocally encoding the visual scene in craniocentric coordinates. It is true that the coding properties of hippocampal place cells take into account the visual environment, regardless of gaze or body orientation (O'Keefe and Dostrovsky, 1971). However, place cells do not respond to the visual scene per se; they respond to the position of the animal within the context of a visual scene. If these neurons were selective to a specific scene or object, they would silence their response to changes in the visual environment. This is not usually the case. What happens is that the neuron acquires a new selectivity for an animal position, dependent on the new visual context (O'Keefe, 1979). Therefore, we believe any property resembling a craniotopic map likely emerges across the network of cortical and subcortical areas.

Bisiach and Luzzatti (1978) reported on two patients with left unilateral neglect following brain injuries to their right parietal cortex. The patients were asked to describe a familiar place from memory (the cathedral square of Milan, their native city) based on two vantage points selected by the authors. The two vantage points were opposite each other so that buildings on the right side of the patient from one vantage point were transposed to their left side from the perspective of the other vantage point. Notably, the patients could easily describe the buildings on their right side. However, when the vantage point was switched, they could no longer describe those same buildings now situated on their left neglected side. The findings by Bisiach and Luzzatti (1978) could offer valuable insights regarding the spatial coordinate system employed by higher cognitive functions. The fact that the left side of the retrieved images is consistently not reported is intriguing. It could be used as an argument that the explicit memory system relies on a coordinate system that is inherently dependent on eye position (i.e., a visuotopic map). Here, we argue that this notion would require some unlikely assumptions. First, the patients stood at that particular vantage point at some moment in their lives and had their eyes looking relatively straight and still when acquiring the reported memorized views. A second assumption is that the patients disregarded all other memorized views of the

cathedral square, acquired from several other vantage points during their lives, that could potentially help in their description. We believe it is more likely that the patients used all memorized views at their disposal to depict the cathedral square. These views were acquired across a variety of eye and head positions. The brain, nevertheless, consolidated this diverse collection of images into a unified mental picture, which requires a sequence of coordinate system transformations, starting from the retinotopically-based representation in the retina and dLGN to the visuotopic representation in V1, followed by the emergence of a craniotopic map, which we believe is distributed across several downstream areas. This information is ultimately fed to the explicit memory encoding system in the ventral stream. In Bisiach and Luzzatti (1978), the patients generated a mental image that was oriented relative to the left or the right side of their body. Once the mental image was available, it is fair to postulate that the unilateral neglect patients did not equally highlight the representations of the left and right sides. Therefore, we believe that the behavioral deficit reported in Bisiach and Luzzatti (1978) cannot be explained by retinotopic encoding or retrieval of familiar visual scenes but by a biased attentional system impaired by lesions in the right parietal cortex.

11. Propagation in the neural network

The visual representation at the conscious level can be related to visual maps in the cortex (Vernet et al., 2020). Intraub (2012) and Shioiri et al. (2018) proposed a new multistage model for visual scene representation in terms of an egocentric spatial framework that integrates visual sensory input, amodal perception, expectations and constraints derived from rapid-scene classification and object-to-context associations.

Topographic maps that replicate the retinal receptor surface have been described in at least 12 visual areas located in the occipital, parietal and temporal cortices. The connections between these areas are topographically organized and reciprocal, enabling the simultaneous processing of various attributes of the visual information. The preferential distribution of certain portions of the visual field in different visual information processing pathways creates a substrate for the construction of additive neural networks, while feedback connections create a recurrent network with storage properties.

After reviewing the connectivity of the cortical visual areas, Felleman and Van Essen (1991) presented the concept of a wide network of connections among them, suggesting that all areas are virtually connected to each other. This view suggests that these areas can also interact together within a dynamic network, enabling the aware and conscious access to a visual scene by areas of the occipital, temporal and parietal cortices by means of feedforward and feedback connections (Fig. 8). It also suggests that the selectivity of individual cells or the activity of neurons in one (or more) streams of visual information processing within this network can be accessed and contributes to vision at the conscious level. We propose that feedforward and feedback connections play an important role in determining the activity of each module in a wider network (Gattass et al., 1990; Jansen-Amorim et al., 2011; Correia et al., 2021). For example, the activity of a locus in V2 may depend on the activity of several loci of extrastriate areas located anteriorly. During GABA inactivation of areas MT and V4, we observed prominent changes in feature tuning and neural excitability of V2 neurons, suggesting that feedback projections play a more profound role in neuronal firing properties than previously acknowledged (Galuske et al., 2002; Jansen-Amorim et al., 2011; Schmidt et al., 2011; see Correia et al., 2021 for a recent review regarding the roles of feedback projections in feature tuning and neuronal excitability in the early primate visual system). Fig. 8 shows a topographically organized network representing one point in the upper visual field. This network is composed of subcortical and cortical areas, starting with different types of ganglion cells in the retina (Gattass et al., 1990). Axons from ganglion cells of the eye project to the superior colliculus (SC) and the dorsal lateral
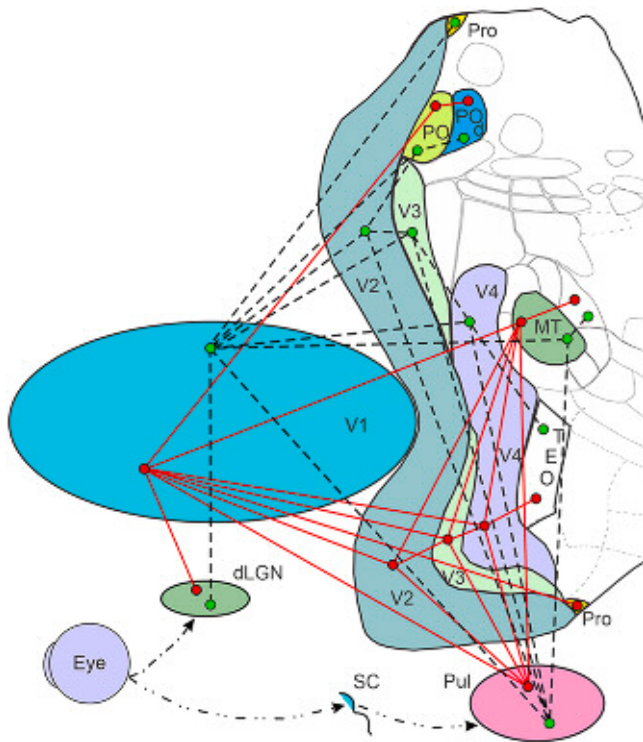
Fig. 8. Visual perceptual representations are generated on a wide dynamic network. The scheme illustrates two wide dynamic networks, where red and green dots depict, respectively, visual loci belonging to the upper and lower visual fields. Dashed and continuous red lines depict the corresponding functional connectivity. Information flow starts at the retina, propagates through the thalamus, and reaches early and intermediate visual areas. Information extracted from these retinotopically and visuotopically organized networks are used to generate craniocentric and ambient representations of the scene. Such neural representations enable access to the information content of perceptual moments associated with conscious visual experience.

geniculate nucleus (dLGN). Cells of the dLGN project mainly to the primary visual cortex (V1), while cells from the SC project to the pulvinar, which in turn projects to several cortical areas. Retinal ganglion cells project to visual areas and form an extensive topographically organized network of feedforward and feedback connections (Sousa et al., 1991). The visuotopic map of V1 provides the information of the central 5 degrees with high acuity, while the map of V4 provides the color information of the central 40 degrees of vision. The scanning of the visual field by eye movements gradually and successively accrues more information that contributes to the craniocentric map.

The craniocentric perceptual map is a neural representation of a scene emerging from a network of cortical visual areas; it accesses visual properties of cortical modules on topographic maps of all visual areas. This representation does not correspond to an individual map of a visual area, but it corresponds to the topographical combination of maps of all visual areas. Retinotopic and visuotopic maps change with eye movement, the master craniocentric perceptual map does not.

In search for models that could account for stable visual perception based on the premises described above, we came across concepts based on deep learning algorithms. Accordingly, the integration of several visual maps can eventually result in a "global space" compatible with visual perception. Deep learning (a subfield of machine learning) consists of algorithms somewhat inspired by brain architecture and function. More specifically, it is interesting to compare the process we propose for dynamic cortical network with the Attention in DeepMind Learning literature. Attention in the deep learning literature is a mechanism by which a network can weigh features by level of importance to a specific task and use this weighting to help achieve the task

(Mnih et al., 2015). These authors focus on the idea that convolution neural network architectures utilize a sliding window approach (iterating convolution filters over the extent of the image) whereas humans only process areas of an image most relevant to a given task. Deep learning has steadily improved the state of the art in artificial intelligence for single, well-defined neuroscience tasks in recent years. Novel advanced neural network architectures are required to create more general-purpose artificial intelligence systems with flexible and robust capabilities.

Indeed, recent advances in deep learning have allowed artificial intelligence to reach near-human-level performance in several sensory, perceptual, linguistic, and cognitive tasks. According to VanRullen and Kanai (2021), using deep-learning techniques to revisit brain-inspired models established in cognitive science could allow further progress. More specifically, the authors examine the potential advantages of this approach in relation to the Global Workspace Theory (GWT), a large-scale cognitive architecture integrating and distributing information among networks of specialized modules to create higher-level forms of cognition and awareness (Baars, 2005). GWT was primarily designed to account for conscious information processing in the human brain, but in principle, its associated functional advantages could generalize to artificial systems. In turn, considering an artificial global workspace can help constrain neuroscientific investigations of brain function and consciousness. VanRullen and Kanai (2021) propose a roadmap based on unsupervised neural translation between multiple latent spaces to create a unique, amodal global latent workspace. This theoretical framework appears to be well suited to address the issue of how to create a unique visual percept from multiple visual maps.

12. Conclusion

Vision captures information from the external world by means of an inhomogeneous retinal surface, which is impeded within an ocular structure going through discrete eye fixations. In spite of this, and despite the fact that the feedforward information flow is punctuated by saccadic suppression, scenes are perceived as coherent, continuous, and meaningful. The search for an appropriate model that can account for all these phenomena has led us to the notion that perception is the product of an emergent neuroanatomical network capable of integrating all visual areas by means of their joint topographical connectivity. The activity of this dynamic network creates a distributed dynamic map accessible when the subject is awake, allowing the extraction of properties which are decoded within modules such as cortical columns. No single cortical area alone is able to encompass this unified map. Rather, information would be held in the neural network made up of all visual areas. One emergent property of this unified map is perceptual stability, immune to disturbances resulting from saccadic eye movements. In the cascade of events spanning from area V1 all the way to the inferotemporal cortex, each neuron's receptive field would gradually evolve, allowing the integration of an increasingly larger area of visual space (Gattass et al., 1985). In this process, successive neuronal ensembles within the processing chain would gradually change from a topographic visuotopic system to a holistic representational system, allowing for the detection of ever more complex objects.

In this review, we discussed the difference between retinotopic, visuotopic, craniocentric (or cyclopic) and ambient maps. The retinotopic map can be observed in subcortical structures, while the others above mentioned maps can be found in the neocortex. The ambient or the egocentric map enables the interaction between the visual map and the individual's motor map. It is important to translate the location of the skull-centered map to a location on the map of the nearby extracorporeal space. It is also important to correlate visual space in the nearby extracorporeal space during ambulation, as in the case of ambulation in an immersive bubble. The three-dimensional map is a craniocentric, skull-centric or cyclopic map. It is perceptually stable regardless of eye movements. Keeping the head in one position and

scanning the field of view with the foveal region of the eyes can reconstruct a scene with high spatial resolution. High acuity is dependent upon the density of retinal ganglion cells and its projection to the primary visual cortex, V1. This representation is reconstructed in the neocortex within a set of visual areas. These areas together constitute a very efficient network capable of generating a percept with both color information and high spatial resolution that spreads out toward the visual mid-periphery up to approximately 40°. We propose that both the high resolution and the color information propagate within the network toward the representation of the periphery. Our conscious perception of the scene for each position of the head is both stable and uniform. The ambient or egocentric map is a conscious three-dimensional reconstruction of the scene, optimized for spatial resolution, color, and contrast across the entire field of view. It is built from several contiguous craniocentric representations. For each head position, the oculomotor system scans the scene using the eyes, specifically its foveal region, to construct a high-resolution color scenario that is updated based on an expected visual model generated by the network. Thus, a large scenario is "constructed" piece-by-piece in the network to allow for the reconstruction of the ambient scene. It is useful to compute object trajectories in space, to estimate the location of static and moving objects, and to estimate distances in order to compute direction and changes in trajectories. This representation is crucial to navigate in three-dimensional space, to determine road trajectories and to estimate spatial relations across landmarks.

It is essential to understand how representational maps link to current concepts of network dynamics if we are to avoid the classic Cartesian Theater fallacy. It has been proposed that synchronous neuronal oscillations work as a flexible mechanism that integrates activity across space and time in distributed networks relevant to feature binding and attention (Fries, 2015; Singer, 2021). This mechanism also seems to coordinate neuronal activity across eye saccades (Ito et al., 2011). Therefore, oscillatory dynamics may constitute an essential substrate for perceptual stability alongside receptive field remapping. More generally, various spatiotemporal dynamics can result from different network topologies (lateral, feedforward, and recurrent) within and across maps. An example of how maps can shape neuronal dynamics can be found in a study of gamma synchronization in monkey V1 (Lima et al., 2010). In this study, it was shown that gamma oscillation frequency depends on receptive field eccentricity. Thus, neuronal synchronization is constrained by cortical architecture.

Higher cognitive functions like attention and expectation also recruit and organize network dynamics during goal-directed behavior and may contribute to perception stability. In accordance with this view, it has been shown that selective spatial attention (Fries et al., 2001) and temporal expectation (Lima et al., 2011) are associated with oscillatory dynamics. In a nutshell, hierarchical maps do not exclude but contribute to our understanding of how dynamics emerge from cortical networks, dissipating the need for a master map and the fallacy of a Cartesian Theater.

Our visual experience is generally anchored to objects in a structured world determined by physical laws. In this respect, it makes sense to think that our internal representations veridically reflect the outside world. On the other hand, in many instances, our subjective experience contradicts this premise. As we discussed, filling-in is one example out of many such illusions. The lithographic work of M. C. Escher exquisitely illustrates how we can build mental images of physically impossible worlds (Fig. 9). Thus, our internal representations, even though coupled with the physical world, are built according to their own set of rules.

A tribute to leslie ungerleider and mortimer mishkin

The motivation of this review is to pay a tribute to Leslie Ungerleider

Fig. 9. Neural Representations and the outside world. Ongoing, high-order neural representations that emerge from processes taking place in the visual cortex provide us with consciously accessible, abstract models of the external world. These models may conflict, at times, with other brain 'constructs' of the outside world. In order to depict these conscious or unconscious mental conflicts, many visual illusions, whether produced by art or science, have been created. The famous lithograph print, Waterfall, made by the Dutch artist Maurits Cornelis Escher in 1961 features perplexing elements: the water goes up against gravity and finally goes down in the waterfall.

and Mortimer Mishkin, both of whom died recently. They were extremely influential to the work that has been carried out in our laboratory for the last 40 years. Their influence started when one of us (Ricardo Gattass) joined Charles Gross' group at Princeton University at the beginning of the 1980s. Their influence persisted during the multiple visits to the National Institutes of Mental Health (NIMH, NIH, USA) in the following decades. Mortimer Mishkin and Leslie Ungerleider made fundamental contributions to the understanding of cortical organization, visual perception and memory processing in primates. Their enthusiasm, scientific expertise and rigorous application of anatomical techniques have shaped various fields of research. Mortimer Mishkin's incredible surgical abilities enabled him to induce precise brain lesions in non-human primates. His double dissociation research design involving cortical and callosal lesions proved that elegant experimental paradigms should always be our highest standard. His work on semantic and episodic memory, on anatomically defined neural systems, and on visual streams of information processing has left us with the gift of greater understanding in key areas of neuroscience.

Significance statement

Electrophysiological recording tools associated with histochemical and immunohistochemical techniques have the potential to revolutionize our understanding of brain functional organization and, thereby, of visual perception. Early recordings in the striate and extrastriate areas were overwhelmingly performed using single electrodes. The advance of multi-unit recordings should allow us to delve into the dynamic

underpinnings of brain function. Here, we discuss the basis of neuronal properties and neuronal representations that allow us to construct our internal representation of stabilized scenes.

Funding information

Data Availability

Data will be made available on request.

References

Albright, T.D., 1984. Direction and orientation selectivity of neurons in visual area MT of the macaque. J. Neurophysiol. 52, 1106–1130.

Allman, J.M., Kaas, J.H., 1973. The organization of the second visual area (VII) in the owl monkey: a second order transformation of the visual hemifield. Brain Res. 76, 247–265.

Azzi, J.C., Gattass, R., Lima, B., Soares, J.G., Fiorani, M., 2015. Precise visuotopic organization of the blind spot representation in primate V1. J. Neurophysiol. 113, 3588–3599.

Baars, B.J., 2005. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. Prog. Brain Res. 150, 45–53.

Baars, B.J., Gage, N.M., 2010. Vision. In cognition, brain, and consciousness. Introd. Cogn. Neurosci. 156–193.

Barlow, H.B., 1972. Single units and sensation. A neuron doctrine for perceptual psychology? Perception 1, 371–394.

Barlow, H.B., Blakemore, C., Pettigrew, J.D., 1967. The neural mechanism of binocular depth discrimination. J. Physiol. 193, 327–342.

Bavelier, D., Corina, D., Jezzard, P., Padmanabhan, S., Clark, V.P., Karni, A., Prinster, A., Braun, A., Lalwani, A., Rauschecker, J.P., Turner, R., Neville, H., 1997. Sentence reading: a functional MRI study at 4 tesla. J. Cogn. Neurosci. 9, 664–686.

Bisiach, E., Luzzatti, C., 1978. Unilateral neglect of representational space. Cortex 14, 129–133.

Botelho, E.P., Ceriatte, C., Soares, J.G.M., Gattass, R., Fiorani, M., 2014. Quantification of early stages of cortical reorganization of the topographic map of V1 following retinal lesions in monkeys. Cereb. Cortex 24, 1–16. https://doi.org/10.1093/cercor/bhs208.

Buchweitz, A., Mason, R.A., Tomitch, L.M.B., Just, M.A., 2009. Brain activation for reading and listening comprehension: an fMRI study of modality effects and individual differences in language comprehension. Psychol., Neurosci. 2, 111–123.

Cardin, V., Sherrington, R., Hemsworth, L., Smith, A.T., 2012. Human V6: functional characterisation and localisation. PLoS One 7 (10), e47685.

Colby, C.L., Duhamel, J.R., Goldberg, M.E., 1993. Ventral intraparietal area of the macaque: anatomical location and visual response properties. J. Neurophysiol. 69, 902–914.

Colby, C.L., Gattass, R., Olson, C.R., Gross, C.G., 1988. Topographic organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. J. Comp. Neurol. 269, 392–413.

Conley, M., Fitzpatrick, D., 1989. Morphology of retinogeniculate axons in the macaque. Vis. Neurosci 2, 287–296.

Constable, R.T., Pugh, K.R., Berroya, E., Mencl, W.E., Westerveld, M., Ni, W., Shankweiler, D., 2004. Sentence complexity and input modality effects in sentence comprehension: an fMRI study. NeuroImage 22, 11–22.

Correia, A.R.A., Amorim, A.K.J., Soares, J.G.M., Lima, B., Fiorani, M., Gattass, R., 2021. The role of feedback projections in feature tuning and neuronal excitability in the early primate visual system. Brain Struct. Funct. 226, 2881–2895. https://doi.org/10.1007/s00429-021-02311-x.

Cragg, B.G., Ainsworth, A., 1969. The topography of the afferent projections in the circumstriate cortex of the monkey studied by the Nauta method. Vis. Res. 9, 733–747.

Damasio, A., Damasio, H., 1983. The anatomic basis of pure alexia. Neurology 33, 1573–1583.

Daniel, P.M., Whitteridge, D., 1961. The representation of the visual field on the cerebral cortex in monkeys. J. Physiol. 159, 203–221.

De Jong, B.M., Shipp, S., Skidmore, B., Frackowiak, R.S.J., Zeki, S., 1994. The cerebral activity related to the visual perception of forward motion in depth. Brain 117, 1039–1054.

De Weerd, P., Gattass, R., Desimone, R., Ungerleider, L.G., 1995. Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. Nature 377, 731–734.

Dehaene S., 2013. Inside the letterbox: how literacy transforms the human brain. Cerebrum. 2013:7. PMID: 23847714; PMCID: PMC3704307.

Desimone, R., Albright, T.D., Gross, C.G., Bruce, C., 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. J. Neurosci. 4, 2051–2062.

DeYoe, E.A., Van Essen, D.C., 1988. Concurrent processing streams in monkey visual cortex. Trends Neurosci. 11, 219–226.

Duffy, C.J., Wurtz, R.H., 1991. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. J. Neurophysiol. 65, 1329–1345.

Duhamel, J.R., Colby, C.L., Goldberg, M.E., 1992. The updating of the representation of visual space in parietal cortex by intended eye movements. Science 255, 90–92.

Eichert, N., Robinson, E.C., Bryant, K.L., Jbabdi, S., Jenkinson, M., Li, L., Krug, K., Watkins, K.E., Mars, R.B., 2020. Cross-species cortical alignment identifies different types of anatomical reorganization in the primate temporal lobe. Elife 9, e53232. https://doi.org/10.7554/eLife.53232.

Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–4.

Fine, I., Wade, A.R., Brewer, A.A., May, M.G., Goodman, D.F., Boynton, G.M., Wandell, B.A., MacLeod, D.I.A., 2003. Long-term deprivation affects visual perception and cortex. Nature Neuroscience 6, 1–2.

Fiorani Jr., M., Gattass, R., Rosa, M.G.P., Sousa, A.P.B., 1989. Visual area MT in the cebus monkey: location, visuotopic organization, and variability. J. Comp. Neurol. 287, 98–118.

Fiorani Jr., M., Rosa, M.G.P., Gattass, R., Rocha-Miranda, C.E., 1992. Dynamic surrounds of receptive fields in primate striate cortex: a physiological basis for perceptual completion? Proc. Natl. Acad. Sci. 89, 8547–8551.

Fiorani, M., Azzi, J.C.B., Soares, J.G.M., Gattass, R., 2014. Automatic mapping of visual cortex receptive fields: a fast and precise algorithm. J. Neurosci. Methods 221, 112–126.

Fries, P., 2015. Rhythms for cognition: communication through coherence. Neuron 88, 220–235.

Fries, P., Reynolds, J.H., Rorie, A.E., Desimone, R., 2001. Modulation of oscillatory neuronal synchronization by selective visual attention. Science 291, 1560–1563.

Galuske, R.A.W., Schmidt, K.E., Goebel, R., Lomber, S.G., Payne, B.R., 2002. The role of feedback in shaping neural representations in cat visual cortex. Proc. Natl. Acad. Sci. 99, 17083–17088.

Gattass, R., Gross, C.G., 1981. Visual topography of the striate projection zone in the posterior superior temporal sulcus (MT) of the macaque. J. Neurophysiol. 46, 521–638.

Gattass, R., Gross, C.G., Sandel, J.H., 1981. Visual topography of V2 in the macaque. J. Comp. Neurol. 201, 519–539.

Gattass, R., Sousa, A.P.B., Covey, E., 1985. Cortical visual areas of the macaque: possible substrates for pattern recognition mechanisms. In: Chagas, C., Gattass, R., Gross, C.G. (Eds.), Pattern Recognition Mechanisms. Pontifical Academy of Sciences, Vatican City, pp. 1–20.

Gattass, R., Sousa, A.P.B., Rosa, M.G.P., 1987. Visual topography of V1 in the Cebus monkey. J. Comp. Neurol. 259, 529–548.

Gattass, R., Sousa, A.P.B., Gross, C.G., 1988. Visuotopic organization and extent of V3 and V4 of the macaque. J. Neurosci. 8, 1831–1845.

Gattass, R., Soares, J.G.M., Lima, B., 2020. Effects of MT lesions on visuomotor performance in macaques. Prog. Neurobiol. 195, 101931 https://doi.org/10.1016/j.pneurobio.2020.101931.

Gattass, R., Sousa, A.P., Mishkin, M., Ungerleider, L.G., 1997. Cortical projections of area V2 in the macaque. Cerebral Cortex 7, 110–129.

Gattass, R., Rosa, M.G.P., Sousa, A.P.B., Piñon, M.C.G.P., Fiorani, M., Neuenschwander, S., 1990. Cortical streams of visual information processing in primates. Brazilian Journal of Medical and Biological Research 23, 375–393.

Gattass, R., Nascimento-Silva, S., Soares, J.G., Lima, B., Jansen, A.K., Diogo, A.C., Farias, M.F., Botelho, M.M., Mariani, O.S., Azzi, J., Fiorani, M., 2005. Cortical visual areas in monkeys: location, topography, connections, columns, plasticity and cortical dynamics. Philos. Transect R. Soc., Lond., B Biol. Sci. 360, 709–731.

Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C., Anderson, A.W., 2000. The fusiform "face area" is part of a network that processes faces at the individual level. J. Cogn. Neurosci. 12, 495–504.

Geschwind, N., 1965. Disconnection syndromes in animals and man: part I. Neuropsychol. Rev. 20, 128–157.

Ghose, G.M., Maunsell, J., 1999. Specialized representations in visual cortex: a role for binding? Neuron 24, 79–85.

Goldberg, M.E., Bisley, J., Powell, K.D., Gottlieb, J., Kusunoki, M., 2002. The role of the lateral intraparietal area of the monkey in the generation of saccades and visuospatial attention. Ann. N. Y. Acad. Sci. 956, 205–215.

Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. Trends Neurosci. 15, 20–25.

Gouras, P., 1969. Antidromic responses of orthodromically identified ganglion cells in monkey retina. J. Physiol. 204 (2), 407–419.

Gross, C.G., 2002. Genealogy of the "Grandmother Cell". Neuroscientist 8, 512–518.

Gross, C.G., Bender, D.B., Rocha-Miranda, C.E., 1969. Visual receptive fields of neurons in inferotemporal cortex of the monkey. Science 166, 1303–1306.

Gross, C.G., Rocha-Miranda, C.E., Bender, D.B., 1972. Visual properties of neurons in inferotemporal cortex of the Macaque. J. Neurophysiol. 35, 96–111.

Hartline, H.K., 1938. The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. Am. J. Physiol. 121, 400–415.

Holmes, G., 1918. Disturbances of vision by cerebral lesions. Brit. J. Ophthal. 2, 353–384.

Hubel, D.H., 1982. Exploration of the primary visual cortex, 1955–78. Nature 299, 515–524.

Hubel, D.H., Wiesel, T.N., 1968. Receptive fields and functional architecture of monkey striate cortex. J. Physiol. 195, 215–243.

Hubel, D.H., Wiesel, T.N., 1969. Anatomical demonstration of columns in the monkey striate cortex. Nature 221, 747–750.

Hubel, D.H., Wiesel, T.N., 1972. Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. J. Comp. Neurol. 146, 421–450.

Intraub, H., 2012. Rethinking visual scene perception. Wiley Interdisciplinary Reviews. Cogn. Sci. 3, 117–127.

Ito, J., Maldonado, P., Singer, W., Grün, S., 2011. Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. Cereb. Cortex 21, 2482–2497.

Jansen-Amorim, A.K., Lima, B., Fiorani, M., Gattass, R., 2011. GABA inactivation of visual area MT modifies the responsiveness and direction selectivity of V2 neurons in *Cebus* monkeys. Vis. Neurosci. 28, 513–527.

Jobard, G., Vigneau, M., Mazoyer, B., Tzourio-Mazoyer, N., 2007. Impact of modality and linguistic complexity during reading and listening tasks. Neuroimage 34, 784–800.

Kanwisher, N., Josh McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.

Keizer, A.W., Nieuwenhuis, S., Colzato, L.S., Teeuwisse, W., Rombouts, S.A., Hommel, B., 2008. When moving faces activate the house area: an fMRI study of object-file retrieval. Behav. Brain Funct. 4, 50–56.

Kuffler, S.W., 1953. Discharge patterns and functional organization of mammalian retina. J. Neurophysiol. 16, 37–68.

Kuypers, H.G.J.M., Szwarcbart, M.K., Mishkin, M., Rosvold, H.E., 1965. Occipitotemporal corticocortical connections in the rhesus monkey. Exp. Neurol. 11, 245–262.

Lettvin, J.Y., Maturana, H.R., McCulloch, W.S., Pitts, W.H., 1959. What the frog's eye tells the frog's brain. Proc. Inst. Radio Eng. 47, 1940–1951.

Lima, B., Singer, W., Neuenschwander, S., 2011. Gamma responses correlate with temporal expectation in the monkey primary visual cortex. J. Neurosci. 31, 15919–15931.

Lima, B., Singer, W., Chen, N.H., Neuenschwander, S., 2010. Synchronization dynamics in response to plaid stimuli in monkey V1. Cereb. Cortex 20, 1556–1573. https://doi. org/10.1093/cercor/bhp218.

Livingstone, M.S., Hubel, D.H., 1988. Segregation of form color movement and depth: anatomy physiology and perception. Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater. 240, 740–749.

Logothetis, N.K., Sheinberg, D.L., 1996. Visual object recognition. Annu. Rev. Neurosci. 19, 577–621.

Lund, J.S., 1988. Anatomical organization of macaque monkey striate visual cortex. Annual Review of Neuroscience 11, 253–288.

Mariani, O.S.C., Lima, B., Soares, J.G.M., Mayer, A., Franca, J.G., Gattass, R., 2019. Partitioning of the primate intraparietal cortex based on connectivity pattern and immunohistochemistry for Cat-301 and SMI-32. J. Comp. Neurol. 527, 694–717.

McClure, M.E., Hart, P.M., Jackson, A.J., Stevenson, M.R., Chakravarthy, U., 2000. Macular degeneration: do conventional measurements of impaired visual function equate with visual disability? Br. J. Ophthalmol. 84, 244–250. https://doi.org/ 10.1136/bjo.84.3.244.

Merigan, W.H., Maunsell, J.H.R., 1993. How parallel are the primate visual pathways? Annu Rev. Neurosci. 16, 369–402.

Merleau-Ponty, M., 2013. Phenomenology of Perception (Landes, R., Trans.), Routledge, Oxfordshire (Original work published 1945).

Michael, E.B., Keller, T.A., Carpenter, P.A., Just, M.A., 2001. fMRI investigation of sentence comprehension by eye and by ear: modality fingerprints on cognitive processes. Hum. Brain Mapp. 13, 239–252.

Mishkin, M., Ungerleider, L.G., Macko, K.A., 1983. Object vision and spatial vision: two cortical pathways. Trends Neurosci. 6, 414–417.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D., 2015. Human-level control through deep reinforcement learning. Nature 518, 529–533. https://doi.org/10.1038/nature14236.

Nakamura, H., Gattass, R., Desimone, R., Ungerleider, L.G., 1993. The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. J. Neurosci. 13, 3681–3691.

Nascimento-Silva, S., Gattass, R., Fiorani Jr., M., Sousa, A.P.B., 2003. Three streams of visual information processing in V2 of Cebus monkey. J. Comp. Neurol. 466, 104–118.

Nassi, J.J., Callaway, E.M., 2009. Parallel processing strategies of the primate visual system. Nat. Rev. Neurosci. 10, 360–372.

Neuenschwander, S., Gattass, R., Sousa, A.P.B., Piñon, M.C., 1994. Identification and visuotopic organization of areas PO and POd in Cebus monkey. J. Comp. Neurol. 340, 65–86.

Newcombe, F., Ratcliff, G., Damasio, H., 1987. Dissociable visual and spatial impairments following right posterior cerebral lesions: clinical, neuropsychological and anatomical evidence. Neuropsychologia 25 (IB), 149–161.

O'Keefe, J., 1979. A Review of the hippocampal place cells. Prog. Neurobiol. 13, 419–439.

O'Keefe, J., Dostrovsky, J., 1971. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. 34, 171–175.

Patel, G.H., Michael, D.K., Snyder, L.H., 2014. Topographic organization in the brain: searching for general principles. Trends Cogn. Sci. 18, 351–363.

Perrett, D.I., Rolls, E.T., Caan, W., 1982. Visual neurones responsive to faces in the monkey temporal cortex. Experimental Brain Research 47, 329–342.

Perry, V.H., Oehler, R., Cowey, A., 1984. Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. Neurosci 12, 1101–1123.

Pettigrew, J.D., Nikara, T., Bishop, P.O., 1968. Responses to moving slits by single units in cat striate cortex. Exp. Brain Res. 6, 373–390.

Piñon, M.C., Gattass, R., Sousa, A.P.B., 1998. Area V4 in Cebus Monkey: Extent and Visuotopic Organization. Cerebral Cortex 8, 685–701.

Pitcher, D., Ungerleider, L.G., 2021. Evidence for a third visual pathway specialized for social perception. Trends Cogn. Sci. 25, 100–110.

Pitcher, D., Ianni, G., Ungerleider, L.G., 2019. A functional dissociation of face-, body- and scene-selective brain areas based on their response to moving and static stimuli. Sci. Rep. 9, 8242–8291.

Pitzalis, S., Bozzacchi, C., Bultrini, A., Fattori, P., Galletti, C., Di Russo, F., 2013. Parallel motion signals to the medial and lateral motion areas V6 and MT+. NeuroImage 67 (2013), 89–100. https://doi.org/10.1016/j.neuroimage.2012.11.022.

Pohl, W., 1973. Dissociation of spatial discrimination deficits following frontal and parietal lesions in monkeys. J. Comp. Physiol. Psychol. 82, 227–239.

Price, C.J., 2000. The anatomy of language: contributions from functional neuroimaging. J. Anat. 197, 335–359.

Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., Fried, I., 2005. Invariant visual representation by single neurons in the human brain. Nature 435, 1102–1107.

Rolls, E.T., Xiang, J., Franco, L., 2005. Object, Space, and Object-Space Representations in the Primate Hippocampus. Journal of Neurophysiology 94, 833–844.

Rosa, M.G.P., Sousa, A.P.B, Gattass, R., 1988. Representation of the visual field in the second visual area in the Cebus monkey. Journal of Comparative Neurology 275, 326–345.

Roumazeilles, L., Eichert, N., Bryant, K.L., Folloni, D., Sallet, J., Vijayakumar, S., Foxley, S., Tendler, B.C., Jbabdi, S., Reveley, C., Verhagen, L., Dershowitz, L.B., Guthrie, M., Flach, E., Miller, K.L., Mars, R.B., 2020. Longitudinal connections and the organization of the temporal cortex in macaques, great apes, and humans. PloS Biol. 18 (7), e3000810.

Sawatari, A., Callaway, E.M., 1996. Convergence of magno- and parvocellular pathways in layer 4B of macaque primary visual cortex. Nature 380, 442–446.

Schiller, P.H., Malpeli, J.G., 1978. Functional specificity of lateral geniculate nucleus laminae of the rhesus monkey. J. Neurophysiol. 41, 788–797.

Schmidt, K.E., Lomber, S.G., Payne, B.R., Galuske, R.A.W., 2011. Pattern motion representation in primary visual cortex is mediated by transcortical feedback. Neuroimage 54, 474–484.

Shapley, R., Perry, H., 1986. Cat and monkey retinal ganglion cells and their visual functional roles. Trends Neurosci. 9, 229–235.

Sheth, B.R., Young, R., 2016. Two visual pathways in primates based on sampling of space: exploitation and exploration of visual information. Front. Integr. Neurosci. 10, 37.

Shioiri, S., Kobayashi, M., Matsumiya, K., Kuriki, I., 2018. Spatial representations of the viewer's surroundings. Sci. Rep. 8, 7171.

Singer, W., 1999. Neuronal synchrony: a versatile code for the definition of relations? Neuron 24, 49–65.

Singer, W., 2021. Recurrent dynamics in the cerebral cortex: Integration of sensory evidence with stored knowledge. Proc. Natl. Acad. Sci. USA 118 (33), e2101043118.

Singer, W., Gray, C.M., 1995. Visual feature integration and the temporal correlation hypothesis. Annu. Rev. Neurosci. 18, 555–586.

Sousa, A.P., Piñón, M.C., Gattass, R., Rosa, M.G., 1991. Topographic organization of cortical input to striate cortex in the Cebus monkey: a fluorescent tracer study. J. Comp. Neurol. 308, 665–682.

Sun, L., Goldberg, M.E., 2016. Corollary discharge and oculomotor proprioception: cortical mechanisms for spatially accurate vision. Annu. Rev. Vis. Sci. 2, 61–84.

Suzuki, W.A., Miller, E.K., Desimone, R., 1997. Object and Place Memory in the Macaque Entorhinal Cortex. Journal of Neurophysiology 78, 1062–1081.

Thomas, O.M., Cumming, B.G., Parker, A.J., 2002. A specialization for relative disparity in V2. Nat. Neurosci. 5, 472–478.

Thompson, E., Varela, F.J., 2001. Radical Embodiment: neural dynamics and consciousness. Trends Cogn. Sci. 5, 418–425.

Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. Nature 381, 520–522.

Tootell, R.B., Reppas, J.B., Kwong, K.K., Malach, R., Born, R.T., Brady, T.J., Rosen, B.R., Belliveau, J.W., 1995. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. J. Neurosci. 15, 3215–3230.

Ts'o, D.Y., Roe, A.W., Gilbert, C.D., 2001. A hierarchy of the functional organization for color, form and disparity in primate visual area V2. Vis. Res. 41, 1333–1349.

Ungerleider, L.G., Mishkin, M., 1982. Two cortical visual systems. In: Ingle, D.J., Goodale, M.A., Mansfield, R.J.W. (Eds.), Analysis of Visual Behavior. MIT Press, Cambridge, MA.

Ungerleider, L.G., Desimone, R., 1986. Cortical connections of visual area MT in the macaque. J. Comp. Neurol. 248, 190–222.

Ungerleider, L.G., Galkin, T.W., Desimone, R., Gattass, R., 2008. Cortical connections of area V4 in the macaque. Cereb. Cortex 18, 477–499.

VanRullen, R., Kanai, R., 2021. Deep learning and the global workspace theory. Trends Neurosci. Vol. 44 (No. 9) https://doi.org/10.1016/j.tins.2021.04.00.

Vernet, M., Quentin, R., Japee, S., Ungerleider, L.G., 2020. From visual awareness to consciousness without sensory input: the role of spontaneous brain activity. Cogn. Neuropsychol. 25, 1–4.

Weiner, K.S., Yeatman, J.D., Wandell, B.A., 2017. The posterior arcuate fasciculus and the vertical occipital fasciculus. Cortex 97, 274–276.

Wurtz, R.H., Joiner, W.M., Berman, R.A., 2011. Neuronal mechanisms for visual stability: progress and problems. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 366, 492–503.

Yukie, M., Iwai, E., 1985. Laminar origin of direct projection from cortex area V1 to V4 in the rhesus monkey. Brain Res 346, 383–386.

Zeki, S.M., 1974. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. J. Physiol. 236, 549–573.

Zeki, S.M., 1978. Functional specialisation in the visual cortex of the rhesus monkey. Nature 274, 423–428.