

Opinion

What kind of network is the brain?

John D. Mollon ^{1,2,*} Chie Takahashi,¹ and Marina V. Danilova^{1,2}

The different areas of the cerebral cortex are linked by a network of white matter, comprising the myelinated axons of pyramidal cells. Is this network a neural net, in the sense that representations of the world are embodied in the structure of the net, its pattern of nodes, and connections? Or is it a communications network, where the same physical substrate carries different information from moment to moment? This question is part of the larger question of whether the brain is better modeled by connectionism or by symbolic artificial intelligence (AI), but we review it in the specific context of the psychophysics of stimulus comparison and the format and protocol of information transmission over the long-range tracts of the brain.

Two types of networks

Charles Bonnet lived a quiet life, and a comfortable one. Born in 1720 and enjoying private means, he is thought never to have left his native Geneva [1]. Although trained in law, he occupied himself with entomology and botany. While still in his twenties, he was elected a Fellow of the Royal Society of London and a Corresponding Member of the Académie Royale of Paris. However, his eyesight and his general health became weaker, and he turned – as many do – to theoretical cognitive neuroscience.

The first edition of Bonnet's *Essai de Psychologie* (1755) does not name the author on the title page and claims 'Londres' as the place of publication [2]: its materialist views were potentially dangerous even in Geneva. In the *Essai*, Bonnet explicitly raises the issue with which the present review is concerned. He writes (Figure 1A):

If all our ideas, even the most abstract, depend ultimately on motions that occur in the brain, it is appropriate to ask whether each idea has a specific fiber dedicated to produce it, or whether different motions of the same fiber produce different ideas.

The modern form of Bonnet's question is this – the brain is a network of neurons, but how far is it a **neural net** (see [Glossary](#)), and how far is it a **communications network**?

In a neural net, information is embodied in the structure of the net – its neural units (or 'nodes') and the signs and weightings of their connections (or 'edges') [3–5]. The term 'neural net' now has an independent life in computer science and engineering, but in biological modeling, the nodes of the net are usually individual cells, and the edges are synaptic connections. In typical models, individual cells – or subsets of cells – in the output layer may represent particular items or categories, such as words or faces. The responses of a given unit may change as learning proceeds, and what is represented by the response of a unit, especially in hidden layers, may have meaning only in the context of the activity in other units. At any one time, however, a given representation is embedded in the network of cells that extract it. There is no separation between storage and processing.

By contrast, in a communications network, the same physical substrate carries different information from moment to moment [6,7]. For example, in the case of the man-made Internet, a given

Highlights

Brain networks may be considered as neural nets in which the information is embodied in the structure of the net, or as communication networks in which the same physical substrate carries different information from moment to moment. Different types of network may be present at different levels.

Psychophysical results and the remarkable sparsity of long-range white matter connections suggest that the central connectome of the brain may be a communications network. The format of the data transmitted via this network is one of the most important unsolved questions of neuroscience.

If the white matter tracts do constitute a communications network, the network almost certainly requires analogs of the transmission control protocols (TCPs) used in man-made networks to control the flow of data and avoid overflow of buffers. A subset of axons within the tracts may carry the control signals.

¹Department of Psychology, University of Cambridge, Cambridge CB2 3EB, UK
²I.P. Pavlov Institute of Physiology, St. Petersburg, Russia

*Correspondence: jm123@cam.ac.uk (J.D. Mollon).



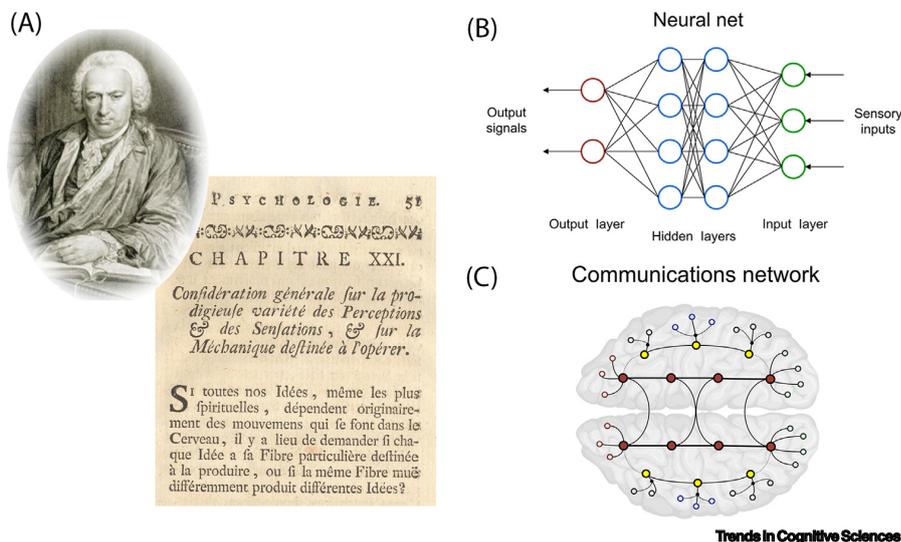


Figure 1. An unresolved issue from the 18th century. (A) Charles Bonnet and a passage from his *Essai de Psychologie* (our translation is given in the main text). In modern terms, his two alternative hypotheses correspond to (B) and (C) on the right. In a neural net (B), cells (or combinations of cells) in the output layer are fixed in their representations – at least until new learning occurs. Information is embodied in the connections of the net, and there is no separation between storage and processing. In a communications network (C), the information is in symbolic form, and the same physical substrate carries different information from moment to moment.

byte – ultimately a sequence of ones and zeros – may be part of the representation of a number, an alphanumeric character, or a pixel. Equally, it may carry information about the addresses of sender and receiver and about the format of the message.

Undoubtedly, the early stages of our perceptual systems do resemble neural nets [3]. But is the entire brain a neural net? Are the axons that comprise the great **white matter** tracts of the brain no more than extended edges in a confederation of neural nets, or is the brain centrally a communications network? What is the format of the information being transmitted between local modules over the long-range tracts of the brain, and what are the **transmission control protocols (TCPs)** that govern the network? Curiously, these questions are seldom discussed.

Our questions lie within the domain of a grander debate between **connectionism** and **symbolic AI** [5,8], but we limit ourselves here to asking what is being transmitted over the long-range tracts of the brain. We also limit ourselves to cortico-cortical connections, although we would readily acknowledge the importance of cortico-thalamic connections in the overall operation of the brain.

The modern connectome

The present era offers unprecedented knowledge of the network of networks that is the brain [9]. It is generally agreed that the brain consists of dense, well-connected, clusters or modules, with sparser long-range connections between the modules [4,10,11]. The latter corresponds to the bundles of myelinated axons that constitute the white matter of the major **fasciculi**.

Graph theory can be applied to a cerebral network without knowing whether it is a neural net or a communications network [11–13]. The network can be represented formally in terms of its nodes

Glossary

- Affective illness:** an illness that primarily alters emotions.
- Bus:** a shared channel that can deliver signals from multiple devices to multiple recipients.
- Communications network:** a network that supports communication between devices such as computers or mobile phones.
- Comparator unit:** a neuron that compares the signals of two or more cells at an earlier level of analysis.
- Connectionism:** the general doctrine that cognitive processes depend on neural nets/networks (*q.v.*) and that learning depends on the selective strengthening or weakening of connections between units in the net.
- Discriminanda:** in psychophysical tasks, the items that the participant is asked to distinguish between.
- Engram:** the memory trace, the hypothetical physical substrate that underlies memory.
- Fasciculi:** tracts of white matter (Latin, 'little bundles').
- Gnostic unit:** a neuron that responds only to a member of a particular category or only to a unique entity.
- Grandmother cell:** a hypothetical neuron that only responds to a unique entity such as an individual person.
- Handshaking protocols:** the procedures for establishing a communication session between two devices on a network.
- Hypercolumn:** a functional unit within the visual cortex (~1 mm in width) that corresponds to a local region of the visual field and comprises a set of component columns each of which is specific for a different stimulus orientation.
- Ipsilateral:** on the same side.
- Isomorphism:** doctrine that physiological representations of external objects have the same structure as the concurrent mental events.
- Neural net:** a network of nerve cells (or simulated cells) linked by excitatory and inhibitory connections and typically organized into a succession of layers. The pattern of the output of the net can be modified by changes in the weighting of the connections.
- Optic flow:** the flow of texture across the retina when there is relative motion between the eye and the world.
- Parity bit:** an extra bit added to each unit of a binary message for the purpose of detecting errors in transmission.

and edges. 'Hubs' can be identified as nodes that are especially rich in their connections. The network as a whole can be assessed for **small-worldness** and average path length. Subgraphs can be distinguished. The plasticity of the network can be monitored over time [14]. Extensive data on the long-range connections of the brain have been drawn from diffusion tensor imaging (DTI) and functional imaging.

In functional studies, networks have predominantly been inferred from temporal covariance between cortical regions in electrophysiological signals (e.g., [15,16]) or in blood oxygen level-dependent (BOLD) fMRI signals (e.g., [9,11]). Such studies have produced many interesting results. For example, a single measure of functional connectivity, based on 1 h of resting-state fMRI, has been reported to account for 20% of the variance in general intelligence (g), where g was derived by factor analysis of ten cognitive tests, including tests of working memory, processing speed, vocabulary, and reasoning [17]. In many psychiatric conditions (including both **psychotic illness** and **affective illness** without psychosis), connectivity is reported to be reduced and the fronto-parietal network is especially disrupted [18,19]. Based on fMRI data acquired while participants performed a test of divergent thinking, creativity has been related to functional connections between the hubs of three networks: 'default mode' (posterior cingulate cortex), 'executive' (right dorsolateral prefrontal cortex), and 'salience' (left anterior insula [20]).

These examples are a tiny sample from a large field of research. Such studies have their own validity and – very reasonably – their authors seldom comment on what is being transmitted along the edges of the graph; however, this is an attractive question to explore. Surely it is very detailed information that is being transmitted (rather than gross excitation or inhibition), but what is the nature and the format of the information? Are the individual axons of the fasciculi simply point-to-point connections dedicated in their representations, or do they constitute a communications network in which the same substrate carries different information from moment to moment?

Connectionist models, gnostic units, and ensembles

Charles Bonnet himself favored a connectionist model of the brain, in which, for example, words and concepts each had their dedicated 'fiber' [2]. In the 20th century, an early and influential model was that of Anne Treisman, who envisaged a network of 'dictionary units' that represented particular words – the 'units' were linked by excitatory and inhibitory connections that reflected, for example, the **transitional probabilities** of the language (A.M. Treisman, DPhil thesis, University of Oxford, 1961; also [21]). When Hubel and Wiesel described cells specific to increasingly elaborate geometric features of the stimulus in the visual cortex, their hierarchical model inspired Konorski to extrapolate to higher levels and to postulate **gnostic units** specific for categories or indeed for individual objects, people, and so forth [22]. Sokolov and Nezlina later suggested that neurons, newly formed from stem cells, queued to be the next gnostic unit, becoming permanently dedicated to the set of sensory inputs that arrived in the ventral stream concurrently with a novelty signal from the hippocampus [23].

The plausibility of gnostic units or **grandmother cells** has been much debated elsewhere [5,24–28]. No one has convincingly found one such unit empirically – and it would be exceedingly difficult to do so. Neurons are indeed found in the face patches of the macaque temporal lobe that respond only to familiar faces as a class [29], and others in the human medial temporal lobe that respond to visual or auditory representations of individual people [30]. But these cells are found too readily to be gnostic units. If an electrophysiologist finds a cell responding to one member of the limited set of stimuli shown in a recording session, then there must be hundreds of

Psychotic illness: an illness in which the patient loses contact with reality.

Pyramidal cell: a common class of neurons that are characterized by the conical shape of the cell body.

Small-worldness: a network has the quality of small-worldness if most nodes can be reached from every other node by a small number of hops or steps.

Symbolic artificial intelligence (AI): the form of AI that models human cognitive processes as the manipulation of symbols.

Transitional probability: the probability that a given word follows another specific word.

Transmission control protocols (TCPs): standardized procedures that govern the exchange of signals between sender and receiver in a communications system.

White matter: the tracts of axons that carry signals between different parts of the brain. The myelin sheaths of the axons confer the whitish appearance.

thousands of other cells that would have responded if tested [31,32]; conversely, there are likely to be many other stimuli to which the cell would have responded if those stimuli had been shown.

The standard alternative to coding by gnostic units is 'ensemble coding' or 'population coding', where properties of the world, objects, or concepts are represented by the pattern of activity in a local population of neurons [28,33–35] (Box 1). Individual cells may contribute to several different representations. In some models of ensemble coding, the contributing neurons each represent an identifiable feature of the represented object or concept.

In the present paper, we emphasize a little-discussed limitation that faces both gnostic-unit coding and ensemble coding – if the representation is embodied in the activity of a specific cell, or a local subset of cells, then it can only be transmitted to other cortical sites by dedicated connections. First, however, we summarize psychophysical results that originally led us to question whether the entire brain is a neural net.

Box 1. Ensemble coding

Confusingly, the term 'ensemble coding' currently has two distinct meanings in neuroscience: it can refer to the encoding of an ensemble (a group) of similar stimuli (e.g., [70]), or it can refer to encoding of stimuli by a small population of neurons (e.g., [71]). Here, we are concerned with the latter meaning.

In the population of neurons in Figure I, one stimulus (A) is encoded by one subset of cells, here indicated in green and lime, and a second stimulus (B) is encoded by a second overlapping subset, indicated in green and blue. A major advantage of this 'distributed' coding is that a relatively small population of cells can in principle encode a large number of stimuli.

How then can the information be carried forward, for use by motor systems or manipulation in cognitive operations? In the case of some stimuli (e.g., a coherent pattern of motion), the vector encoded by an ensemble of perceptual neurons might be directly transformed to a vector in the neural population that controls the six extrinsic eye muscles.

In other cases, however, the ensemble might represent a discrete object, person, word, or concept, and the individual cells might correspond to the defining features of that entity. Further cognitive operations require that the pattern in the ensemble must be recognized – and identified as a unique entity. It would be possible to project the entire array of active cells forward to other cortical areas, but this would be costly in dedicated lines and the problem of recognition would only be postponed. Another solution would be to postulate 'reader neurons' (Figure IA,B) which draw weighted inputs from overlapping subsets of the ensemble. However, these 'reader neurons' would be no more or less than gnostic units, with the disadvantages discussed in the main text.

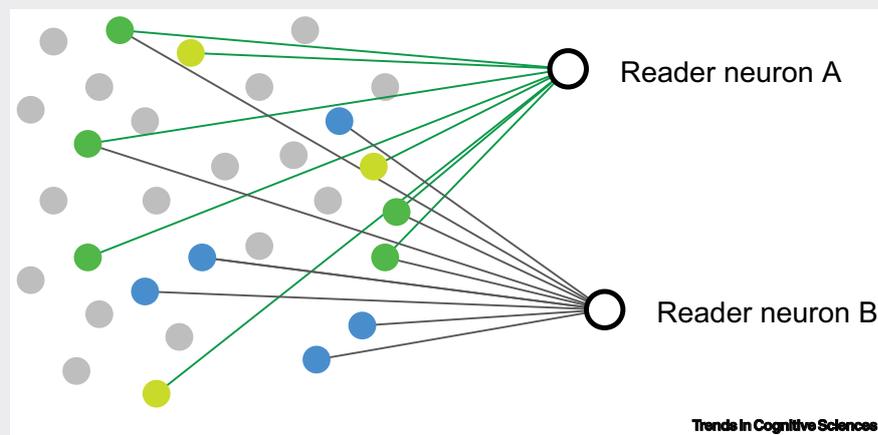


Figure I. Distributed coding requires reader neurons.

The problem of comparison at a distance

Suppose that two stimuli are presented at different positions in our visual field but at the same eccentricity – the same distance from the fixation point – and suppose we are asked to discriminate some attribute in which the stimuli differ: for example, which is moving faster, or which is the more saturated red. Suppose too that the stimuli are so brief that we do not have time to make an eye movement from one to the other (Figure 2A).

In the case of some visual attributes, such as lightness, stereoscopic depth, and direction of motion, the precision of our discrimination deteriorates rapidly as the stimuli are separated in the visual field [36–38]. In these cases we may imagine that our responses depend on the signals arising from local differencing neurons – **comparator units**. For example, our discrimination of the lightness of adjacent patches may depend on signals that derive from retinal ganglion cells whose center-surround visual fields straddle the border between the two patches ([3] chapter 3): the cell extracts the difference, or the ratio, of the light intensities in adjacent parts of the visual field, and it is this contrast signal that is delivered centrally.

However, psychophysical experiments have shown that many visual attributes can be discriminated with the same precision when the two stimuli are juxtaposed or are separated by

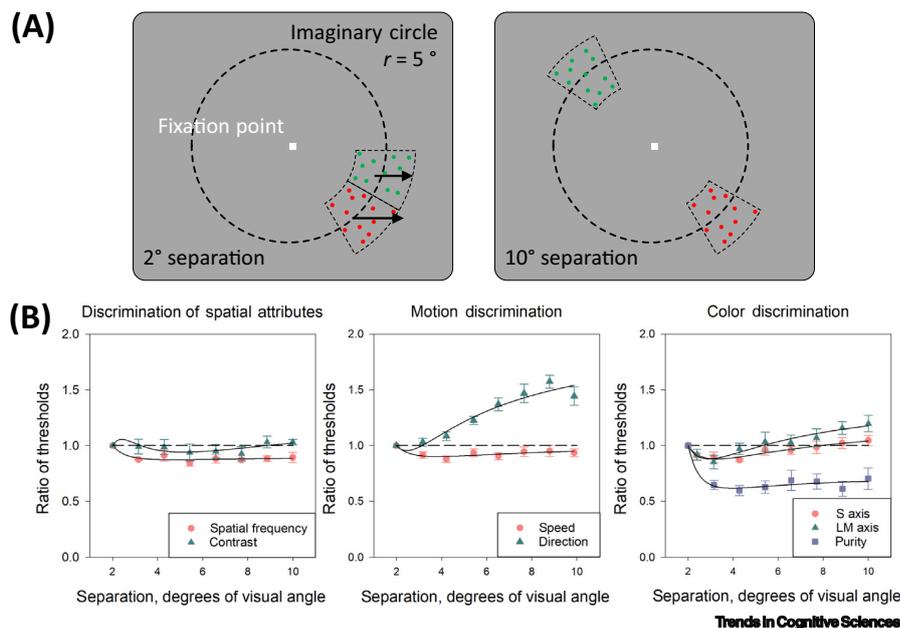
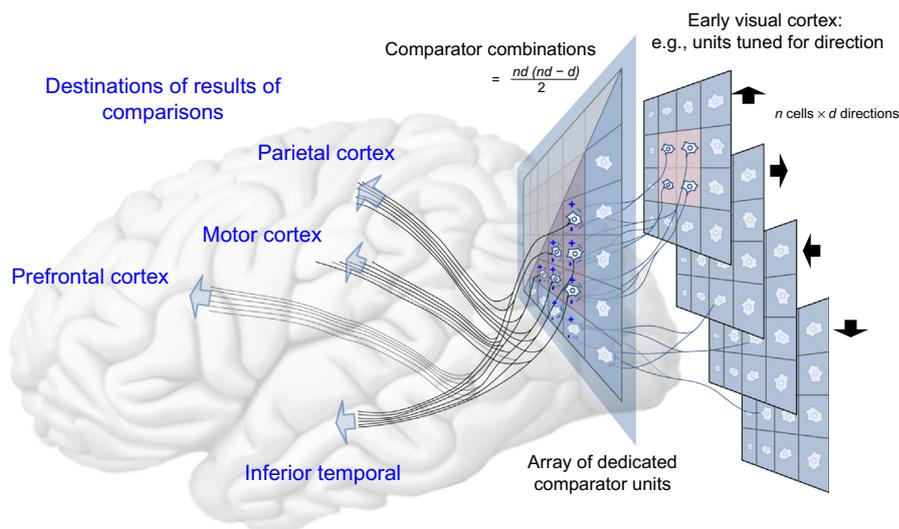


Figure 2. Comparison at a distance. (A) The stimulus arrangements for experiments on discrimination as a function of separation in the visual field. In the example illustrated, the participant is asked to judge which of the patches of random dots is moving at the higher speed. The eccentricity of the discriminanda is held constant, but their spatial separation can vary from 2 degrees of visual angle center-to-center (when the patches are juxtaposed) to 10 degrees (when they lie on opposite sides of an imaginary circle). (B) Normalized thresholds for several different visual attributes, plotted against spatial separation in the visual field. In many cases there is little variation of thresholds with separation (although often there is a 'gap effect', in that a small separation between the discriminanda improves performance). The marked exception is discrimination of the direction of motion, which deteriorates rapidly with increasing separation. In the latter case we have hypothesized that discrimination depends on detectors that are hard-wired to extract local differences in direction [36], whereas we doubt that there are arrays of dedicated 'comparator neurons' (Figure 3) to account for the flatter functions. Sources of data: spatial frequency and contrast [55], speed and direction [36], color axes [72], and colorimetric purity [54]. Abbreviations: LM, long/medium wavelength; S, short wavelength.

10 degrees of visual angle and fall in opposite hemifields (Figure 2B). Such attributes include the speed of motion, the hue or the purity of colors, contrast, and spatial frequency (the fineness of texture) [36,39,40]. Whether a layman or neuroscientist, the owner of the visual system takes this remarkable ability very much for granted. The question is seldom asked what brain processes underlie such comparisons, but it is an interesting question. If pressed, many psychophysicists would postulate higher-order comparator units – cells that draw signals of opposite sign from lower-order cells that correspond to the two positions in the visual field that are being compared (Figure 3). There would be multiple arrays of these comparator units – one array for each of the visual attributes (hue, purity, speed, and so on). A decision, and appropriate action, could then be based on the output of the relevant comparator unit.

The principal problem here is the combinatorial explosion in the number of comparator units that are required (Figure 3). In the case of speed discrimination, assuming that there is one comparator unit for every possible pair of the ~5300 **hypercolumns** in the primary visual cortex [41], and assuming around 4×10^4 neurons in 1 mm^3 of human visual cortex [42], we have calculated that a cortical volume of ~5600 mm^3 would be needed for the array of comparator units underlying this



Trends in Cognitive Sciences

Figure 3. An implausible model of sensory comparison. In the case of many visual attributes, the precision of discrimination changes very little as the discriminanda are separated in space. Could these discriminations be performed by dedicated 'comparator neurons'? We take speed discrimination as an example [36,39]. At the right are shown arrays of directionally selective neurons at an early stage of the visual cortex (black arrows). We show the minimal number of arrays that are needed (four) to encode all possible directions. Suppose that, at a subsequent level, there is an array of hard-wired comparator neurons that draw inputs of opposite sign (excitatory and inhibitory) from each possible pair of lower-level detectors. A combinatorial explosion arises here in the number of comparators required. An advocate for long-range comparator neurons might point out that receptive fields become larger in more anterior visual areas (e.g., [73–75]), and therefore, only a small population of comparator cells would be needed. Each such cell might collect the decisions of local comparators at intermediate stages of a neural net. However, if a cell is to subservise the psychophysical task, it is not enough for it to integrate inputs for a given sensory attribute over a large area. It would instead need to signal the ratio of the speeds registered in two specific, local, and arbitrarily chosen regions of its receptive field and would need to preserve the sign of the difference while being indifferent to the directions of motion of the two discriminanda [39]. This figure illustrates a second problem that arises if representations are embodied in dedicated neurons: if the information is to be carried forward to other cortical areas that might require it, then many dedicated long-range axons are required. The alternative is to encode the information in symbolic form, thus allowing it to be carried on a shared 'cerebral bus' (Figure 4).

one arbitrary psychophysical task [36]; in addition, further arrays, of similar magnitude, would be needed for other visual attributes.

There are, of course, several uncertainties associated with such estimates [39], but there is a deeper issue. If comparisons are achieved by individual comparator units, and if the resulting decisions are represented only by the responses of these units, then each comparator unit requires a further array of dedicated connections to deliver the information to the other cortical modules that need to know the outcome of the comparison (Figure 3). It would clearly be more efficient to transmit the raw data from the primary sensory neurons (e.g., quantitative values for spatial frequency, speed, saturation, or hue) in symbolic form to the frontal lobes together with addresses that represent the positions in the visual field of the **discriminanda**.

Gnostic units and ensembles require dedicated output connections that carry information elsewhere

The problem introduced in the preceding section holds much more generally if objects, people, voices, words, concepts, and even decisions are represented centrally by gnostic units or grandmother cells. The information represented by such cells will be required in many different modules elsewhere in the brain. The babushka's presence may prompt a hug or a spoken greeting; it may trigger a recollection of the blinis she used to cook when you were a child; it may trigger a variety of emotional responses. If the brain has no other form of representation, then the information can be delivered elsewhere only by an extravagant array of dedicated connections. A gnostic unit is thus a clumsy token for higher-level cognitive operations. In itself, the token cannot move – the representation is frozen in position. The information is imprisoned in its cell.

A secondary problem is one that we have called the problem of junk mail [39]. If the only form of representation is the activity of a gnostic unit, then an active unit will broadcast costly [43,44] action potentials over many paths to many areas, only a subset of which currently require the information.

Could both of these problems be solved by a switching system, as in a classical telephone network? This is unlikely. Whereas in a telephone network the information is transmitted symbolically by a stream of phonemes, and the routing is determined by a sequence of digits, the gnostic unit offers to the network only a train of action potentials on its dedicated axon. To achieve the switching, there would need to be another array of dedicated connections in the upstream direction. If one were to suppose that the gnostic units have symbolic addresses (equivalent to the telephone number), one would have already admitted that there is a second form of representation in the brain.

Ensemble coding does not offer a helpful alternative solution. The first problem is – what mechanism identifies a given ensemble? In a special subset of cases, where there is a simple mapping from sensory input to motor output via an analog representation, there may be no requirement for explicit identification of a stimulus. Thus, a coherent pattern of **optic flow** might be represented in the brain by a population of neurons that are individually tuned to different directions. The vector encoded within this population might be mapped directly onto the activity of the six external eye muscles so as to allow accurate ocular tracking.

In many other cases, however, a discrete category – or indeed a unique object – must be available as input to a variety of further cognitive operations. In these cases, explicit recognition is required. If different objects are represented by different patterns in the same ensemble, then some further process is necessary that can recognize each pattern to allow motor responses or cognitive operations (Box 1) (the problem is analogous to that faced by the Gestalt doctrine of **isomorphism**

[45]). It is possible, of course, to postulate 'collector units', each of which identifies a given pattern in the ensemble – for example, Buzsáki's 'reader neurons' [34] or the 'dot product decoders' of Rolls [33]. However, these would be gnostic units. Any alternative method of identification seems to imply a symbolic representation.

Of course, the pattern of activity in the ensemble could be delivered to other brain modules by arrays of dedicated connections, but this is extravagant in the bulk of axons and, most importantly, it only postpones the problem of recognition.

White matter

We now return to our central question – is the brain as a whole no more than a network of neural nets, or do the central tracts of white matter constitute a communications network that carries representations in a symbolic code over a common substrate?

No less than 42% of the bulk of each hemisphere consists of white matter [46], and this might suggest rich possibilities for dedicated long-range connections. However, the larger part of the white matter comprises U-fibers that link adjacent regions of neocortex. Below these local connections lie the long-range tracts, among which the most important are the superior and inferior occipitofrontal fascicles, the superior and inferior longitudinal fascicles, and the uncinate fascicle [46,47]. For our present purposes, Schüz and Braitenberg's estimate of the proportion of cortical neurons that contribute axons to the long-range tracts is highly pertinent. Excluding callosal neurons, efferent neurons, and all non-pyramidal cells, they estimate that the total number of neurons making **ipsilateral** connections within one hemisphere is 6×10^9 . However, they estimate that there are only $\sim 10^8$ axons in all the major long-range tracts combined. Thus, of all the cells that make cortico-cortical connections, most are local in their projections, and only $\sim 2\%$ have access to the long-range tracts within one hemisphere (and the proportion having access to any individual tract is likely to be still smaller) [46]. The proportion of non-efferent cells contributing axons to the corpus callosum is similarly $\sim 2\%$.

The estimates obtained by Schüz and Braitenberg were based on classical histology, but they draw confirmation from a recent analysis of diffusion MRI (dMRI) data. Rosen and Halgren [48] analyzed tractography data for 1065 individuals in the Human Connectome Project. They calibrated their dMRI measure by reference to the known density of axons in the corpus callosum and the cross-sectional area of the corpus callosum of each individual (obtained by structural MRI). They then used this conversion factor to estimate the number of axons in the long-distance fasciculi. For each of the 360 'parcels' [49] of cortex, they calculated the number of fibers connecting to every other parcel. Long-range connections (callosal plus intra-hemispheric) were sparse, about 3.7% in total – a value close to Schüz and Braitenberg's estimate of 4%.

The limited capacity of the long-distance tracts is difficult to reconcile with models that suppose the brain is a meta-net [9] or with accounts of memory in which cell assemblies depend on many long-range excitatory connections [50]. It would be possible to suppose that only 'gnostic units' in the output layers of local neural nets have access to the long-range tracts, but, at least in the case of vision, we seem to have access to low-level, local, features when we wish to attend to them.

Rosen and Halgren [48] make one striking point. Traffic between Broca's and Wernicke's areas is thought to integrate the expressive and receptive aspects of speech, but only $\sim 25\,000$ axons link the core cortical regions (area 44 and PSL, the posterior perisylvian language area), whereas an average 60-year-old speaker of American English has some minimal knowledge of 60 000 lemmas (base words), a value that excludes proper names and the inflexions of the lemmas [51] (and

does not take into account the fact that much of the human population is bilingual). It is therefore difficult to suppose that 60 000 gnostic units – 'dictionary units' – each enjoy a relevant axon in the corresponding tracts. It is true that the extents of Broca's and Wernicke's areas are debated [52] and that estimates of vocabulary size vary considerably according to age, intellectual level, the corpus sampled, and the criterion for recognition [53]. Nevertheless, the example shows how limited the potential for wholesale transmission of arrays of gnostic units or ensembles is.

The cerebral bus and its unknown code

We previously introduced the term 'cerebral bus' to refer to a shared communications substrate in the brain [54,55]: the term **bus** is used here in the computer-science sense of a common channel carrying signals between different devices. This modern use of 'bus' probably derives, by a wonderfully tortuous route, from the dative plural of the Latin *omnis*, via the 18th century French *voiture omnibus* and the late 19th century use of 'omnibus bar' or 'bus' to denote a source of current that supplies several devices.

A communications network within the brain will not resemble in detail the man-made Internet, where fiber optics allows high rates of serial transmission. Axonal transmission is slow, and action potentials are limited in their temporal precision and in the minimal interval between them. Thus, coding within white matter is likely to rely on many parallel lines. If this is so, however, the code could be seen as no more than a form of ensemble coding. Can a firm line be drawn between the ensembles envisaged in, for example, current models of face perception and the fully symbolic 'language of thought' envisaged in symbolic AI?

There is clearly a continuum here, rather than a dichotomy, but an analogy might be made with alphabetic scripts. At one extreme, one might draw on a clay tablet two schematic cows to represent a tax of two cows, whereas at the other extreme, much of current human knowledge can be represented using 26 Roman letters and ten Arabic numerals – and ultimately, of course, by patterns of ones and zeros. A fundamental question therefore concerns the generality of the cerebral code. If the modules of the cortex are indeed linked by a communications network, we can ask whether the same code is used by different tracts – or whether each tract has a private form of representation matched to the information that it carries. It is plausible, for example, that the inferior longitudinal fasciculus [47] might be optimized to transmit visual data, while the arcuate fasciculus [56] might adopt a code that favors phonological and syntactic data. Indeed, at one extreme of specificity, each long-range fascicle may itself be no more than a bundle of local subsets of fibers, each transmitting local ensembles from one module to another without transformation.

By contrast, the ARPANET, the forerunner of the Internet designed by the American military during the Cold War, was deliberately constructed to allow transmission by alternative routes in case of damage to particular nodes or paths in the network [57]. The present-day Internet directs packets by alternative routes to minimize congestion. In the case of the brain, it would be advantageous to be able to send information by alternative routes, either to avoid temporary congestion or to compensate for damage [58,59]. This would require a code that does not depend on dedicated point-to-point connections. Once information is in symbolic form, the symbols can be manipulated and combined in novel ways. Importantly, they can be stored independently of the processor. In [Box 2](#), we draw attention to the relationship between the issues discussed in this paper and a current debate in the field of memory.

The regulation of the cerebral bus and the TCP

If the central tracts of white matter do constitute a communications network, this would raise secondary questions intrinsic to all man-made communications systems and their TCPs [6]. These

Box 2. The relationship between the cortical network and long-term memory

The nature of the **engram** – the physical basis of memory – remains one of the major unsolved problems of biology. The distinction we make in this review, between neural nets and communications networks, bears on a current debate in the field of memory (e.g., [76–80]). In a neural net, memory is intrinsically embodied in the structure of the net, in its connections and weights: memory is 'content-addressable' in that an input pattern reactivates the relevant array of activity in the network. Already in the 18th century, in the connectionist model of Charles Bonnet, the memory of an object or a word depended on the reactivation of the fiber – the gnostic unit – that corresponded to the actual perception of the word, and there were excitatory links between associated fibers [2,81]. In modern times, the predominant view has continued to be that memory depends on the growth and attenuation of synaptic connections [34,82–84]. A popular view is that an engram supporting a specific experience may be widely distributed throughout the brain in an 'engram cell ensemble' [50], the elements being connected by long-range dedicated connections. This view must be judged in the context (discussed above) of the severe limits to the numbers of long-range axons [46,48]. Particularly striking is the estimate by Rosen and Halgren that the dorsolateral prefrontal cortex is connected to the hippocampus by <10 axons/mm² [48].

In alternative models, the storage and the processor(s) are separate. One possibility is that declarative memory has a molecular basis, and is encoded, for example, in nucleic acid [79,80,85]. In an independent literature, computer scientists have shown the feasibility of storing large quantities of data in tiny volumes of DNA [14,86,87]. If such storage were used in the brain, the challenge would be to understand how information is read into a molecular form and how it is read out. That question is relevant to our present issue: if communication over the white matter tracts is by abstract, symbolic code, then one might expect that there will be a close matching between the neural code and the molecular code so as to allow efficient conversion from one to the other.

issues are strangely neglected in the case of the brain, but they might yield to investigation before the actual data code is deciphered. Some of these secondary questions are listed below.

- (i) Is information transmitted only on request? One way to conceptualize selective attention is as a request for particular data that are needed elsewhere. The serious costs of neural transmission [43,44] argue against the unwanted broadcasting of information.
- (ii) Are the addresses of source and destination encoded by the same fibers as the data, or are they transmitted in parallel by separate fibers?
- (iii) Does data transmission occur in fixed packets, or is it continuous [58,59]? If a given representation is spread across many parallel fibers, we might expect that some form of clock is used to coordinate each byte or word of the message.
- (iv) How is the speed of transmission matched to the receiver? What **handshaking protocols** are in place to prevent overflow of the receiving buffers? Do pathologies arise from errors in this handshaking?
- (v) What protocols are in place for error checking? Are there analogs of the **parity bits** that are used in computer networks?

The possible role of large axons

All man-made communications networks require TCP signals to control the transmission of data – signals that indicate, for example, 'acknowledge request', 'data ready', 'input buffer full'. It is difficult to believe that the brain lacks some analog of TCP – given the limitations to serial transmission within individual axons (see above), the control signals are likely to be carried by fibers that run in parallel to those that carry data.

White matter tracts invariably include a range of axons of different diameters [60–64], and DTI techniques have been developed to estimate the distribution of sizes *in vivo* [65–67]. All tracts contain a minority of large fibers, and transmission time is inversely related to the diameter of axons [63]. Large axons have the highest peak firing rates, but they are costly in bulk and energy. Perge and colleagues [60] have suggested that large axons are used when maximal rates of information transmission are required. In the present context, it is interesting to note that the range of diameters, and thus of transmission speeds, may offer a means to convert between spatial and temporal codes: if parallel axons of different calibers converge on the dendrites of a single

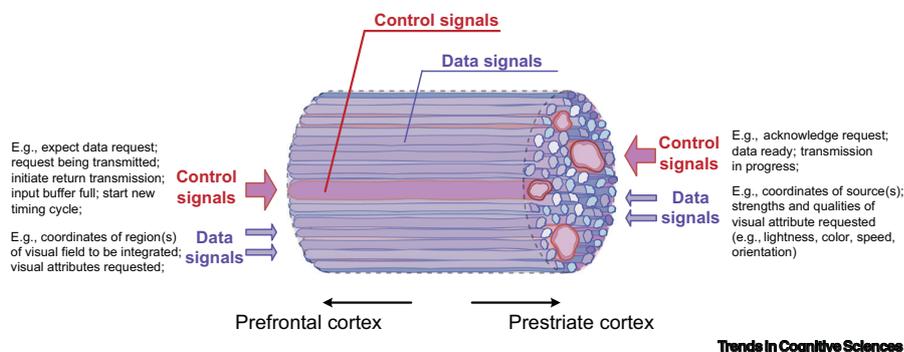


Figure 4. Anatomy of a cerebral bus. A cross-section of part of one of the long-range white matter tracts of the brain (e.g., the inferior occipitofrontal fasciculus), indicating the large range of axon diameters. In this version of the 'cerebral bus' hypothesis, we suggest that the largest (and most rapidly transmitting axons) carry the transmission control protocol (TCP), the handshaking and feedback signals that are required in any communications system, while the smaller fibers carry the data – and the addresses of source and destination. Importantly, the same physical substrate carries different data from moment to moment. The data code is one of the great unsolved issues of neuroscience, although we can guess that it depends on a concurrent pattern distributed across multiple parallel axons.

receiver neuron, a spatial pattern at the origin becomes a temporal pattern at the destination. A similar conversion could occur in reverse.

For the very largest axons, we recently suggested that their role is to carry the control signals for the cerebral bus [40] (Figure 4). Because the control signals are likely to be stereotyped and to be repeated at intervals, they might prove easy to identify and yield to experimental investigation earlier than the code that carries the actual data.

The largest axons may derive from the giant **pyramidal cells** that have been described at the terminations of major tracts. Striking examples are the unusual pyramidal cells observed at the origins of callosal fibers at the border between areas 17 and 18 – cells that are characterized by apical dendrites almost devoid of spines and by many axo-somatic contacts [68,69]. It is at the origins and terminations of the tracts – the potential hubs of a communications network – that empirical insights might be gained into the questions we have raised. Histological examination might reveal specialized circuits that suggest the encoding of data, and electrophysiological recording from such regions in primates might identify stereotyped patterns of large action potentials that could be the control signals governing data transmission.

Concluding remarks

Our purpose in this review is to encourage discussion of what exactly is being transmitted by the major fasciculi of the brain (see [Outstanding questions](#)). Are the white matter tracts only extensions of local neural networks, or, at the other extreme, do they constitute a communications network in which the same physical substrate can carry a variety of different messages from moment to moment? We also ask whether the central network of the brain is governed by a TCP such as the protocols that are needed in almost all man-made networks. It is doubtful that conventional imaging methods could readily answer the questions that we raise, but we have suggested that insight might come from histological and electrophysiological study of the potential hubs at the origins and terminations of the fasciculi.

Acknowledgments

We dedicate this paper to the memory of the network scientist Richard Gibbens (1962–2018) to whom we are indebted for guidance. Preparation of the paper was supported by Biotechnology and Biological Sciences Research Council (BBSRC)

Outstanding questions

How is information encoded for transmission over the major white matter tracts of the brain?

Is the code common to different tracts, allowing rerouting in the case of congestion or local failure?

Are the data and the addresses of source and destination encoded separately?

Are data transmitted only on request so as to minimize energy consumption by the brain? Is selective attention essentially a request for specific information?

Are there brain analogs of the TCPs that are employed in man-made communication systems to regulate the flow of data and prevent the overflow of buffers? If so, are the control signals transmitted separately from the data, and are they carried by the largest axons? Do errors in the TCP underlie some pathologies, such as speech defects?

Which subset of axons correspond to the unusual giant pyramidal cells found at the terminations of some tracts and characterized by apical dendrites almost devoid of spines? It would be valuable to have more histological information on the immediate circuits of these neurons. The local circuitry may offer strong hints as to what information the giant pyramidal cell is collecting and how the information is encoded.

grant BB/S000623/1 to Cambridge University and by Russian Ministry of Science and Higher Education grant to the Pavlov Center 'Integrative Physiology for Medicine, High-Tech Healthcare and Stress Tolerance Technologies' (agreement 075-1502020-921). We thank the Editor and reviewers for many helpful suggestions.

Declaration of interests

No interests are declared.

References

- Savioz, R. (1948) *Mémoires Autobiographiques de Charles Bonnet de Genève*, Librairie Philosophique J. Vrin
- Bonnet, C. (1755) *Essai de Psychologie; ou Considérations sur les Opérations de l'Âme, sur l'Habitude et sur l'Éducation*. Publisher not specified
- Masland, R. (2020) *We Know It When We See It: What the Neurobiology of Vision Tells Us About How We Think*. Basic Books
- Pülvemüller, F. et al. (2021) Biological constraints on neural network models of cognitive function. *Nat. Rev. Neurosci.* 22, 488–502
- Bowers, J.S. (2017) Parallel distributed processing theory in the age of deep networks. *Trends Cogn. Sci.* 21, 950–961
- Tanenbaum, A.S. et al. (2021) *Computer Networks* (6th edn), Pearson
- Peterson, L.L. and Davie, B.S. (2021) *Computer Networks: A Systems Approach* (6th edn), Morgan Kaufmann
- Boden, M.A. (1995) AI's half-century. *AI Mag.* 16, 96–99
- Herbet, G. and Duffau, H. (2020) Revisiting the functional anatomy of the human brain: toward a meta-networking theory of cerebral functions. *Physiol. Rev.* 100, 1181–1228
- van den Heuvel, M.P. and Sporns, O. (2019) A cross-disorder connectome landscape of brain dysconnectivity. *Nat. Rev. Neurosci.* 20, 435–446
- Wig, G.S. (2017) Segregated systems of human brain networks. *Trends Cogn. Sci.* 21, 981–996
- Farahani, F.V. et al. (2019) Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review. *Front. Neurosci.* 13, 585
- Sporns, O. (2018) Graph theory methods: applications in brain networks. *Dialogues Clin. Neurosci.* 20, 111–121
- Organick, L. et al. (2018) Random access in large-scale DNA data storage. *Nat. Biotechnol.* 36, 242–248
- Varela, F. et al. (2001) The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239
- Sadaghiani, S. et al. (2022) Connectomics of human electrophysiology. *Neuroimage* 247, 118788
- Dubois, J. et al. (2018) A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 373
- Baker, J.T. et al. (2019) Functional connectomics of affective and psychotic pathology. *Proc. Natl. Acad. Sci. U. S. A.* 116, 9050–9059
- Dong, D. et al. (2018) Dysfunction of large-scale brain networks in schizophrenia: a meta-analysis of resting-state functional connectivity. *Schizophr. Bull.* 44, 168–181
- Beatty, R.E. et al. (2018) Robust prediction of individual creative ability from brain functional connectivity. *Proc. Natl. Acad. Sci. U. S. A.* 115, 1087–1092
- Treisman, A.M. (1960) Contextual cues in selective listening. *Q. J. Exp. Psychol.* 12, 242–248
- Konorski, J. (1967) Some new ideas concerning the physiological mechanisms of perception. *Acta Neurobiol. Exp.* 27, 147–161
- Sokolov, E.N. and Nezlina, N.I. (2003) Long-term memory, neurogenesis, and novelty signal. *Zh. Vyss. Nerv. Deyat. Im. I. P. Pavlova* 53, 451–463 (in Russian)
- Bowers, J.S. (2017) Grandmother cells and localist representations: a review of current thinking. *Lang. Cogn. Neurosci.* 32, 257–273
- Thorpe, S.J. (2012) Grandmother cells and distributed representation. In *Visual Population Codes* (Kriegeskorte, N. and Kreiman, G., eds), pp. 23–51, MIT Press
- Coltheart, M. (2017) Grandmother cells and the distinction between local and distributed representation. *Lang. Cogn. Neurosci.* 32, 350–358
- Thomas, E. and French, R. (2017) Grandmother cells: much ado about nothing. *Lang. Cogn. Neurosci.* 32, 342–349
- Rolls, E.T. (2017) Cortical coding. *Lang. Cogn. Neurosci.* 32, 316–329
- Landi, S.M. et al. (2021) A fast link between face perception and memory in the temporal pole. *Science* 373, 581–585
- Rey, H.G. et al. (2020) Single neuron coding of identity in the human hippocampal formation. *Curr. Biol.* 30, 1152–1159
- Waydo, S. et al. (2006) Sparse representation in the human medial temporal lobe. *J. Neurosci.* 26, 10232–10234
- Quiroga, R. (2019) Akakievitch revisited: comment on 'The unreasonable effectiveness of small neural ensembles in high-dimensional brain' by Alexander N. Gorban et al. *Phys. Life Rev.* 29, 111–114
- Rolls, E.T. (2016) *Cerebral Cortex: Principles of Operation*. Oxford University Press
- Buzsáki, G. (2010) Neural syntax: cell assemblies, synapses, and readers. *Neuron* 68, 362–685
- Valdez, A.B. et al. (2015) Distributed representation of visual objects by single neurons in the human brain. *J. Neurosci.* 35, 5180–5186
- Takahashi, C. et al. (2020) Separation in the visual field has divergent effects on discriminating the speed and the direction of motion. *Curr. Biol.* 30, R1250–R1251
- Sharpe, L.T. and Wyszecki, G. (1976) Proximity factor in color-difference evaluations. *J. Opt. Soc. Am.* 66, 40–49
- Fendick, M. and Westheimer, G. (1983) Effects of practice and the separation of test targets on foveal and peripheral stereocuity. *Vis. Res.* 23, 145–150
- Daniilova, M.V. et al. (2020) How does the human visual system compare the speeds of spatially separated objects? *PLoS One* 15, e0231959
- Mollon, J. and Daniilova, M. (2019) Cortical communication and the comparison of colors. *Curr. Opin. Behav. Sci.* 30, 203–209
- Sakitt, B. and Barlow, H.B. (1982) A model for the economical encoding of the visual image in cerebral cortex. *Biol. Cybern.* 34, 97–108
- Cragg, B.G. (1967) The density of synapses and neurones in the motor and visual areas of the cerebral cortex. *J. Anat.* 101, 639–654
- Sterling, P. and Laughlin, S.B. (2015) *Principles of Neural Design*. MIT Press
- Levy, W.B. and Calvert, V.G. (2021) Communication consumes 35 times more energy than computation in the human cortex, but both costs are needed to predict synapse number. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2008173117
- Gregory, R.L. (2005) The Medawar Lecture 2001. Knowledge for vision: vision for knowledge. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 360, 1231–1251
- Schüz, A. and Braitenberg, V. (2002) The human cortical white matter: quantitative aspects of cortico-cortical long-range connectivity. In *Cortical Areas: Unity and Diversity* (Schüz, A. and Miller, R., eds), pp. 377–385, Taylor and Francis
- Herbet, G. et al. (2018) Functional anatomy of the inferior longitudinal fasciculus: from historical reports to current hypotheses. *Front. Neuroanat.* 12, 77
- Rosen, B.Q. and Halgren, E. (2021) Human cortical areas are sparsely connected: combining histology with diffusion MRI to estimate the absolute number of axons. *bioRxiv* Published online June 14, 2021. <https://doi.org/10.1101/2021.06.07.447453>
- Glasser, M.F. et al. (2016) A multi-modal parcellation of human cerebral cortex. *Nature* 536, 171–178
- Josselyn, S.A. and Tonegawa, S. (2020) Memory engrams: recalling the past and imagining the future. *Science* 367, eaaw4325

51. Brysbaert, M. *et al.* (2016) How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Front. Psychol.* 7, 1116
52. Tremblay, P. and Dick, A.S. (2016) Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain Lang.* 162, 60–71
53. Milton, J. and Treffers-Daller, J. (2013) Vocabulary size revisited: the link between vocabulary size and academic achievement. *Appl. Linguist. Rev.* 4, 151–172
54. Danilova, M.V. and Mollon, J.D. (2018) Cerebral iconics: how are visual stimuli represented centrally in the human brain? *J. Opt. Technol.* 85, 515–520
55. Danilova, M.V. and Mollon, J.D. (2003) Comparison at a distance. *Perception* 32, 395–414
56. Aboitiz, F. (2018) A brain for speech. Evolutionary continuity in primate and human auditory-vocal processing. *Front. Neurosci.* 12, 174
57. Tanenbaum, A.S. and Wetherall, D.J. (2014) *Computer Networks* (5th edn), Pearson
58. Graham, D.J. (2014) Routing in the brain. *Front. Comput. Neurosci.* 8, 44
59. Graham, D. and Rockmore, D. (2011) The packet switching brain. *J. Cogn. Neurosci.* 23, 267–276
60. Perge, J.A. *et al.* (2012) Why do axons differ in caliber? *J. Neurosci.* 32, 626–638
61. Wang, S.S.H. *et al.* (2008) Functional trade-offs in white matter axonal scaling. *J. Neurosci.* 28, 4047–4056
62. Liewald, D. *et al.* (2014) Distribution of axon diameters in cortical white matter: an electron-microscopic study on three human brains and a macaque. *Biol. Cybern.* 108, 541–557
63. Innocenti, G.M. (2017) Network causality, axonal computations, and Poffenberger. *Exp. Brain Res.* 235, 2349–2357
64. Caminiti, R. *et al.* (2009) Evolution amplified processing with temporally dispersed slow neuronal connectivity in primates. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19551–19556
65. Harkins, K.D. *et al.* (2021) A simple estimate of axon size with diffusion MRI. *Neuroimage* 227, 117619
66. Lee, H.H. *et al.* (2019) Along-axon diameter variation and axonal orientation dispersion revealed with 3D electron microscopy: implications for quantifying brain white matter microstructure with histology and diffusion MRI. *Brain Struct. Funct.* 224, 1469–1488
67. Assaf, Y. *et al.* (2008) AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. *Magn. Reson. Med.* 59, 1347–1354
68. Glickstein, M. and Whitteridge, D. (1976) Degeneration of layer III pyramidal cells in area 18 following destruction of callosal input. *Brain Res.* 104, 148–151
69. Braak, E. and Braak, H. (1985) On layer III pyramidal cells in the parastriate borderzone of man. *J. Hirnforsch.* 26, 117–125
70. Robson, M.K. *et al.* (2018) Ensemble coding of face identity is present but weaker in congenital prosopagnosia. *Neuropsychologia* 111, 377–386
71. Ohla, K. *et al.* (2019) Recognizing taste: coding patterns along the neural axis in mammals. *Chem. Senses* 44, 237–247
72. Danilova, M.V. and Mollon, J.D. (2006) The comparison of spatially separated colours. *Vis. Res.* 46, 823–836
73. Pigarev, I.N. *et al.* (2001) Neurons with large bilateral receptive fields in monkey prelunate gyrus. *Exp. Brain Res.* 136, 108–113
74. Allman, J. *et al.* (1985) Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* 14, 105–126
75. Shelepin, Y.E. and Gerchikova, V.F. (1982) Representation of two visual hemifields in lateral suprasylvian area of one hemisphere. *Fiziol. Zh. SSSR Im. I. M. Sechenova* 68, 763–767 (in Russian)
76. Langille, J.J. and Gallistel, C.R. (2020) Locating the engram: should we look for plastic synapses or information-storing molecules? *Neurobiol. Learn. Mem.* 169, 107164
77. Trettenbrein, P.C. (2016) The demise of the synapse as the locus of memory: a looming paradigm shift? *Front. Syst. Neurosci.* 10, 88
78. Queenan, B.N. *et al.* (2017) On the research of time past: the hunt for the substrate of memory. *Ann. N. Y. Acad. Sci.* 1396, 108–125
79. Gallistel, C.R. (2021) The physical basis of memory. *Cognition* 213, 104533
80. Gold, A.R. and Glanzman, D.L. (2021) The central importance of nuclear mechanisms in the storage of memory. *Biochem. Biophys. Res. Commun.* 564, 103–113
81. Bonnet, C. (1770) *La Palingénésie Philosophique, ou Idées sur l'État Passé et sur l'État Futur des Êtres Vivants*, Claude Philibert & Barthelemi Chirol
82. Langille, J.J. and Brown, R.E. (2018) The synaptic theory of memory: a historical survey and reconciliation of recent opposition. *Front. Syst. Neurosci.* 12, 52
83. Asok, A. *et al.* (2019) Molecular mechanisms of the memory trace. *Trends Neurosci.* 42, 14–22
84. Choi, J.H. *et al.* (2018) Interregional synaptic maps among engram cells underlie memory formation. *Science* 360, 430–435
85. Akhlaghpour, H. (2021) An RNA-based theory of natural universal computation. *J. Theor. Biol.* 110984
86. Erlich, Y. and Zielinski, D. (2017) DNA Fountain enables a robust and efficient storage architecture. *Science* 355, 950–954
87. Ceze, L. *et al.* (2019) Molecular digital data storage using DNA. *Nat. Rev. Genet.* 20, 456–466