

How deep is the brain? The shallow brain hypothesis

Mototaka Suzuki¹✉, Cyriel M. A. Pennartz¹ & Jaan Aru²✉

Abstract

Deep learning and predictive coding architectures commonly assume that inference in neural networks is hierarchical. However, largely neglected in deep learning and predictive coding architectures is the neurobiological evidence that all hierarchical cortical areas, higher or lower, project to and receive signals directly from subcortical areas. Given these neuroanatomical facts, today's dominance of cortico-centric, hierarchical architectures in deep learning and predictive coding networks is highly questionable; such architectures are likely to be missing essential computational principles the brain uses. In this Perspective, we present the shallow brain hypothesis: hierarchical cortical processing is integrated with a massively parallel process to which subcortical areas substantially contribute. This shallow architecture exploits the computational capacity of cortical microcircuits and thalamo-cortical loops that are not included in typical hierarchical deep learning and predictive coding networks. We argue that the shallow brain architecture provides several critical benefits over deep hierarchical structures and a more complete depiction of how mammalian brains achieve fast and flexible computational capabilities.

Sections

Introduction

Rationale for assuming hierarchical architectures

Neuroanatomical evidence under-represented by most hierarchical architectures

The shallow brain hypothesis

Implications of the shallow brain hypothesis

Summary and future directions

¹Department of Cognitive and Systems Neuroscience, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, The Netherlands. ²Institute of Computer Science, University of Tartu, Tartu, Estonia.

✉e-mail: mototaka@gmail.com; jaan.aru@gmail.com

Introduction

While the sustaining strength of the notion of hierarchical processing may be that it is rather simple, its fatal flaw is that it is overly simplistic¹.

Deep learning has not only revolutionized the field of machine learning and artificial intelligence, but also had a substantial effect on our daily life^{2,3}. Deep learning architectures commonly assume hierarchical structures; raw inputs (such as pixel values) are fed into the lowest hierarchical layer, consisting of an array of artificial neurons that provide (feed forward) their outputs to the next, one-step higher layer, and so on². Using tens of – sometimes more than a hundred⁴ – layers allows deep learning networks to encode progressively more abstract, complex features or information. The resulting hierarchical architectures have outperformed humans in tasks such as visual object categorization^{2,5}. Deep learning architectures have many variants and some have recurrent connections^{2,6} in addition to feedforward connections, yet they predominantly share the structure of a deep hierarchy. The layered, hierarchical structure that today's deep learning architectures commonly adopt was inspired by earlier neurobiological findings in the visual cortices that higher cortical areas progressively encode more abstract visual features such as motion, contours and faces^{7–9} (Fig. 1a). Vice versa, deep learning architectures are inspiring neuroscientists to interpret brain computations in terms of hierarchical models^{10–15}.

Predictive coding is a theory of brain function developed in theoretical neuroscience according to which the brain is constantly generating and updating internal models of the environment^{16–20}. The models are used to generate predictions of sensory input that are compared with actual sensory input. Similar to deep learning architectures, computational models of predictive coding^{17,21–23} commonly assume a hierarchical structure. In this Perspective, we refer to predictive coding in a broad sense that includes hierarchical inference (for example, Bayesian inference) and other error-coding hierarchical models. Although some models incorporate the contributions of subcortical areas^{24–26}, they also assume a hierarchy of cortical processing stages.

In this Perspective, we highlight that whereas deep learning and predictive coding assume an underlying hierarchy, neurobiological evidence suggests that hierarchical structures might not be as central to neural processing as commonly thought. Although we do not contest that hierarchical inference can be useful and is used in the cortex, we emphasize that there is another processing mechanism that is intertwined with the cortical hierarchy and is supported by neuroanatomy, which is what we call 'the shallow brain hypothesis'. Next, drawing on the collection of neuroanatomical and electrophysiological evidence, we describe the shallow brain architecture and the key computational benefits of the proposed processing regime. Last, we discuss the implications of the shallow brain hypothesis and propose future research directions with specific testable predictions.

Rationale for assuming hierarchical architectures

Hierarchical architectures have become so dominant for several reasons. First, neurobiological findings on the connectivity pattern between cortical areas illustrated a hierarchical structural organization of the cortex^{8,27} where cortical areas are ordered from lower to higher areas. Second, an influential theoretical proposal has been that these cortical connectivity patterns allow the brain to perform hierarchical Bayesian inference^{16,18}. Third, computer science spawned the idea that shallow architectures have computational limitations²⁸. We will

briefly consider these reasons below, demonstrating that the evidence they provide is not as strong as previously conceptualized.

The original notion of hierarchical processing in the cortex⁸ was based on anatomical and neurophysiological findings in the primary visual cortex and higher visual cortical areas^{8,29–31}. Of particular importance were the tracing studies by Rockland and Pandya³² and the wiring diagram of Felleman and Van Essen²⁷, which suggested a hierarchical structure between cortical areas based on the anatomical connectivity pattern (see also refs. 33,34) (Fig. 1b). Electrophysiological evidence supported this view^{35–37}, such as the presence of larger receptive fields and longer onset latencies in higher compared with intermediate or lower cortical areas. However, hierarchically distant cortical areas can be directly and reciprocally connected via cortico-cortical connections (see refs. 35,36). Moreover, a new analytic method for determining the hierarchy of mouse visual cortical areas revealed that many connectivity patterns that had previously been seen as hierarchical were non-hierarchical lateral connections^{35,38}, consistent with an older study of rat visual cortices³⁹. More importantly, as we detail in the next section, direct empirical evidence that cortical areas, higher or lower, project to and receive inputs from subcortical regions supports an alternative view to hierarchical processing, namely shallow processing.

The second reason for favouring hierarchical structures as models of neural processing has been theoretical. Influential papers argued that the hierarchical structure of the cortex allows the brain to perform hierarchical Bayesian inference^{16,18} – a way to learn from data by combining prior knowledge with newly arriving inputs. This idea of hierarchical Bayesian inference has been the dominant framework in neuroscience over the last decade^{17,40–43}. Over the years, empirical support has emerged for local computation of prediction errors in the sensory cortex⁴⁴ (for instance in the visual cortex (see refs. 45,46), but see also refs. 47,48 for different interpretations). However, only a few studies have used electrophysiological methods to explicitly test the key premises of hierarchical Bayesian inference theory⁴⁹, such as the propagation of predictions or prediction errors across several levels of the cortical hierarchy (for example, level n , $n + 1$, $n + 2$)^{50–52}. Furthermore, predictive coding-like algorithms can be implemented locally without the hierarchical propagation of errors and predictions^{21,53,54} (and also by lateral non-hierarchical connections⁴⁴). Last, and most relevant to the shallow brain hypothesis, it has been noted that hierarchical predictive coding can be mediated or facilitated by thalamic computations^{26,42}.

More functionally motivated support for hierarchical learning comes from deep learning. The predominant view in deep learning has been that deep architectures perform better than shallow ones^{55,56} and that shallow architectures cannot solve certain problems (such as classification problems where the data are not linearly separable)^{2,28}. However, the past two decades have shown a rise in the capabilities of shallow neural networks such as restricted Boltzmann machines^{57,58}. Instead of relying on a deep hierarchy, restricted Boltzmann machines consist of only two layers – one visible and one hidden. Recent work has demonstrated that restricted Boltzmann machines can perform similarly to deep architectures on some standard benchmarks⁵⁹, such as the MNIST database – a large database of handwritten digits commonly used for training and testing in the field of machine learning. In addition, current deep learning architectures have shown the ability to model and predict cortical neural responses, which is usually taken as evidence for hierarchical processing in the cortex^{10–13}. However, recent evidence shows that previous studies might have overestimated the hierarchical alignment⁶⁰.

Taken together, the findings in neurobiology and machine learning during the past two decades suggest that the reasons for preferring hierarchical architectures in models of cortical processing are not as convincing as they might have once seemed. This motivated us to look beyond hierarchical cortical networks for understanding brain function. Importantly, we do not question the principal relevance of predictive coding or deep learning algorithms but, rather, we simply question whether brain computations for cognition and behaviour are primarily processed by a hierarchical architecture of cortical areas.

Neuroanatomical evidence under-represented by most hierarchical architectures

The key reason to doubt the general plausibility of hierarchical architectures comes from the neurobiological evidence that all cortical areas, higher and lower, are directly connected to subcortical structures such as the thalamus^{61–64}, striatum^{65,66}, superior colliculus (SC)^{67–69}, claustrum^{70–72} and brainstem (including the pedunculopontine nucleus (PPN) and mesencephalic locomotor region, brainstem reticular formation and red nucleus)^{73–75} (Fig. 1c,d). Not only the primary sensory areas – the lowest in the hierarchy – but also even the higher associative cortical areas (such as the prefrontal cortex) project to these subcortical regions. Higher associative cortical areas are believed to be the origin of top-down signals that are sent downward to sensory and motor cortices to modulate those lower primary areas; however, higher cortical areas themselves also project to subcortical areas such as the thalamus, hypothalamus and brainstem^{73,74}. Inputs to these higher areas come not just from lower cortical areas but also from subcortical areas; for example, parallel inputs from the thalamic pulvinar to higher visual areas serve as the scaffold on which the cortical hierarchy develops⁷⁶.

Primary sensory cortices have been classically thought to constitute the first stage of sensory cortical processing, and their output has been thought to pass the computed outcome to higher cortical areas for further processing. By contrast, recent studies have found that subcortical projections directly from primary sensory areas are crucial for advanced sensorimotor control^{77,78}. For example, Tang and Higley found that cortical layer 5 pyramidal (L5p) neurons in the primary visual cortex (V1) that project to the brainstem pontine nuclei were necessary to learn a visually cued eye blink conditioning task⁷⁷. Similarly, Takahashi et al. found that subcortically projecting L5p neurons in the primary somatosensory area were crucial for whisker-based tactile detection⁷⁸. These studies suggest that not only higher cortical areas but also primary sensory areas – believed to be the lowest stage of the cortical hierarchy and to merely process sensory information and then pass the outcome to the next level – send their computational results subcortically and contribute substantially to sensorimotor control (Fig. 1d) (but see refs. 79,80 for the important functions of hierarchically higher cortical areas). Importantly, in both studies, the behaviourally relevant cortical neurons were subcortically projecting L5p neurons (Fig. 1c). In general, pyramidal neurons in cortical layers 5 and 6 project to the thalamic nuclei; those in layers 2/3, 5 and 6 project to the striatum; and only those in layer 5 project to the brainstem and spinal cord^{81,82}. Intermingled with cortico-cortically projecting neurons across layers 2/3, 5 and 6 (refs. 83–85), such subcortically projecting pyramidal neurons are distributed across layers 2/3, 5 and 6 of higher or lower cortical areas at every hierarchical level (see ‘The shallow brain hypothesis’ regarding how cortical L5p neurons integrate information from other cell types in the cortex and subcortical areas).

Higher and lower cortical regions not only send projections subcortically but also receive direct inputs from subcortical regions.

The thalamus broadly projects to cortical areas; for instance, the higher-order thalamic nuclei such as the pulvinar, ventromedial and intralaminar nuclei as well as the zona incerta widely project to higher and lower cortical areas^{86–90}. Higher-order thalamic nuclei have classically been thought to merely modulate cortical activity, but recent findings suggest that these thalamic nuclei strongly drive the activity of cortical areas^{91,92}; for example, the sustained neuronal activity in the prefrontal cortex depends on the mediodorsal thalamic nucleus⁹¹ and thalamic orphan receptors⁹². Higher-order thalamic nuclei also notably affect conscious processing by activating infragranular layer cortical neurons⁹³ or by coupling or decoupling thalamo-cortical loops and cortico-cortical loops^{94–96}. Accordingly, inactivation of these thalamic nuclei is associated with disorders of consciousness^{93,97,98}.

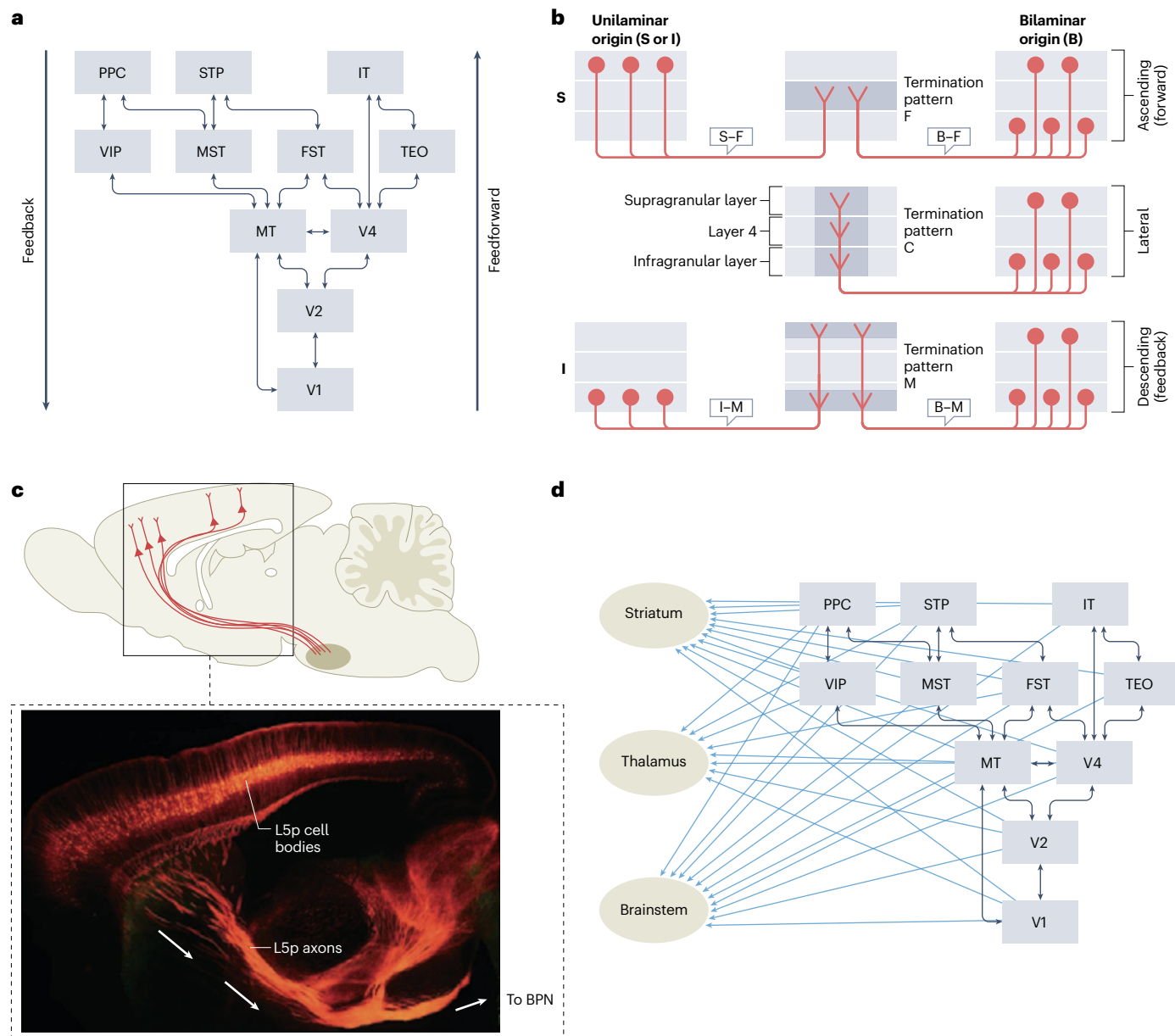
The claustrum – a subcortical structure that has been believed to be crucial for attention and consciousness⁹⁹ – also widely projects to higher and lower cortical areas to affect global cortical activity. Claustral neurons predominantly innervate cortical interneurons and inhibit global cortical activity^{71,72}. Recent studies also suggest that claustrum output has more diverse effects on cortical activity⁷⁰ and underscore its role in motor planning¹⁰⁰, fear memory encoding¹⁰¹, multisensory integration^{99,102}, salience processing^{103,104} and frontoposterior cortical connectivity¹⁰⁵. Other subcortical regions – such as the SC, the PPN and the basal forebrain – broadly affect cortical regions directly via neuromodulators^{106–109} and synaptically via direct connections¹¹⁰ or indirectly through the thalamus (SC → pulvinar → cortex^{111,112}). Neuromodulators such as acetylcholine, dopamine, noradrenaline and serotonin are also produced subcortically – acetylcholine in the nucleus basalis of Meynert in the basal forebrain; dopamine in the substantia nigra pars compacta, ventral tegmental area and retrorubral area (A8) in the midbrain and dorsal raphe nuclei in the brainstem; noradrenaline in the locus coeruleus in the brainstem; and serotonin in the dorsal and median raphe nuclei in the brainstem – and sent to the cortex. They strongly influence brain state, arousal and consciousness^{113–115}, properties of cortical neurons¹¹⁶, motivation and attentional state¹¹⁷, perception¹¹⁸ and memory^{119,120}.

Accumulating evidence also indicates that the hierarchical anatomical connectivity pattern of the cortex cannot explain some basic functional properties of visual cortical neurons^{31,121–123}. For instance, several brain areas at different hierarchical levels respond almost simultaneously to briefly presented visual stimuli^{124,125}. Moreover, direct, reciprocal cortico-cortical connections exist between hierarchically distant cortical areas (for example, V1 to MT and V4 to IT)^{35,36} (Fig. 1a), and connections previously seen as hierarchical in mouse cortex were non-hierarchical lateral connections according to the new analytical method^{35,38}, which has made the anatomical connectivity pattern a flatter hierarchy.

In sum, the discussed under-represented anatomical and electrophysiological findings in the brain suggest that the dominance of cortical hierarchical architectures as in deep learning and predictive coding is highly questionable; we believe such architectures are missing essential computational principles that the brain uses. The findings discussed in this section convinced us that these principles are much more than feature abstraction through the classical cortical hierarchy and led us to the shallow brain hypothesis introduced in the next section.

The shallow brain hypothesis

Although the cortical hierarchy is useful for representing abstract features in their context, we propose a processing regime, the shallow brain hypothesis, whereby the entire cortex is essentially one giant layer



providing contextualized input to subcortical motor and premotor centres in the brainstem and spinal cord (Fig. 1d). From the perspective of these subcortical centres, both higher and lower cortical areas would be part of an array of massively parallel computational units – cortical columns, cortical minicolumns or other types of module acting as a functional entity in the cortex with distinct inputs and outputs^{126–128} – all of which contribute to the selection of behaviour and thought (Fig. 2a, left).

Let us consider the primate oculomotor control network as an example. Among many cortical areas, V1, the posterior parietal cortex (PPC) and the frontal eye field (FEF) are known to be involved in oculomotor control. The classical, hierarchical view of these areas is that signals propagate in the order V1 → PPC → FEF²⁷. However, if the connections of these cortical areas with subcortical structures are considered, it becomes clear that these cortical areas together with the

subcortical areas are well described as a shallow architecture (Fig. 2a, right): V1, PPC and FEF are all reciprocally connected with thalamic regions^{129–134}, project to the SC^{135–138} and make distinct contributions to oculomotor control^{139–141}.

For the shallow brain architecture, the most important cells are subcortically projecting cortical L5p neurons^{77,78,94} (Fig. 1c). In many cortical areas they receive direct inputs from thalamic nuclei (first and/or higher order) and project back directly to the thalamus and subcortical motor and premotor centres, thus forming loops between cortical and subcortical areas (Fig. 2b). Brainstem and mid-brain nuclei such as the SC, PPN and red nucleus also receive the outputs of the cortex and project back – directly or polysynaptically – to the cortex as well as the thalamus^{90,111,142–144}. This massively parallel disynaptic thalamo-cortico-subcortical circuit has received little attention, but we hypothesize that it has a major role in the shallow

Fig. 1 | Deep and shallow architectures. **a**, Common view of the cortical hierarchy that is the basis for many deep learning and predictive coding hierarchical (deep) architectures. Cortical areas associated with visual information processing are shown as an example. The lowest area, the primary visual cortex (V1), processes a class of features and feeds the processed information forward to higher areas such as the inferior temporal cortex (IT), where more abstract information is processed and sent to subcortical areas. Note that cortical areas are connected bidirectionally – via feedforward and feedback connections – and laterally (between the middle temporal cortex (MT) and V4). **b**, Criteria for classifying connections between cortical areas (for example, V1, V2 and V4) as feedforward (top), lateral (middle) and feedback (bottom). Importantly, these criteria also reveal the direct, reciprocal connections between hierarchically distant cortical areas. A cortical column is depicted as the computational unit of any cortical area at any point within the hierarchy, with the six layers collapsed into three layers for simplicity. Termination patterns are depicted in the central column, preferentially in layer 4 (F pattern; feedforward), across all layers (C pattern; lateral), and in upper (supragranular (S)) and lower (infragranular (I)) layers avoiding layer 4 (M pattern; feedback). Laminar origin from a single layer (left column) is either S or I, and thus feedforward or feedback, respectively. Bilaminar (B) origins from two layers (right column) either terminate

in the middle layers (F pattern; feedforward), terminate in all layers (C pattern; lateral) or terminate predominantly in upper S and I layers (M pattern; feedback). The existence of these distinct termination patterns strongly supports the existence of a deep cortical hierarchy. **c**, Retrograde labelling in a sagittal section of the mouse brain showing that across all hierarchical cortical areas, cortical layer 5 pyramidal (L5p) neurons (I) send axonal fibres that project to the basal pontine nucleus (BPN) in the brainstem. Despite the existence of a cortical hierarchy, from the perspective of the BPN, there is no distinction between higher and lower cortical areas, as if the cortex is flat, shallow and just one giant layer. **d**, Cortical areas that are higher and lower in the hierarchy project directly to subcortical regions, including the brainstem, thalamus and striatum. From the perspective of these subcortical areas, the cortical areas are ‘flat’; thus, the architecture is shallow. For the sake of graphical clarity, lines between cortical areas indicate reciprocal connections; other subcortical areas as well as their connections with cortical and subcortical areas have been omitted. FST, fundus superior temporal cortex; MST, medial superior temporal cortex; PPC, posterior parietal cortex; STP, superior temporal polysensory; TEO, temporal occipital cortex; VIP, ventral intraparietal cortex. Part **a** adapted with permission from ref. 207, Elsevier. Part **b** adapted, with permission from ref. 27, OUP. Part **c** adapted with permission from ref. 75, Elsevier.

brain architecture, inspired by previous work on the anatomy of the thalamo-cortical system^{64,145}.

Thus, the basic unit of the shallow brain architecture is a single thalamo-cortico-subcortical loop with L5p neurons as the major driving force of the loops (Fig. 2b). Several further lines of evidence support this position. First, L5p neurons are the most active excitatory neurons in the cortex¹⁴⁶ and have been related to sensorimotor control^{147,148}, perception^{77,78} and consciousness^{94–96,149}. Second, subcortically projecting L5p neurons receive inputs from distant cortical areas and nearly all cell types across layers 1–6 within a cortical column (or a module in species where a cortical columnar structure is less clear)^{126–128} (Fig. 2c). Third, L5p cells are thought to learn and represent both simple and more abstract features of the external and internal environment^{94,150}. According to the shallow brain hypothesis, this is the main role of the cortex: to learn representations of different types of feature and object, which are then forwarded to subcortical structures. L5p cells can compute these features in a context-dependent manner, as they have contextual and top-down related processing targeting their apical dendrites^{94,150}. Fourth, it has been demonstrated that predictions and prediction errors can be computed at the level of single pyramidal cells¹⁵¹. Last, each of these thalamo-cortico-subcortical loops can be coupled and decoupled by projections from the higher-order thalamic nuclei to the cortical L5p cells^{94,96}, demonstrating how subcortical structures can actively control which cortical columns contribute to ongoing processing.

According to our hypothesis, every cortical column (or module), whether from a lower, intermediate or higher area, represents its respective features (simple edges, shapes and faces in visual cortical areas) and contributes its output to internal cognitive processing and the global selection of behaviour. Thus, this cortical representation of a hierarchy of features is still very useful, as it allows computations based on features at different levels of abstraction. However, this hierarchical representation of features needs to be supplemented by another computation: the combination of features across levels (Fig. 3). The distinguishing characteristic of our hypothesis is the shallowness of the neuronal network; cortical areas, higher or lower, directly send outputs to the subcortical motor and premotor centres and participate in parallel fast thalamo-cortico-subcortical loops (Figs. 1d

and 2a). The shallowness of the architecture is further supported by the higher-order thalamic nuclei that provide shorter (disynaptic) thalamo-cortico-subcortical connections between hierarchically distant cortical areas⁴² (Fig. 2b).

The shallow brain hypothesis further posits that any cortical area can directly contribute to the selection of action and internal cognitive operations. Thus, subcortical systems such as the striatum, thalamus or brainstem process massive parallel inputs from the whole cortex as from one big layer – a shallow brain (Fig. 2a). Next, we discuss three key benefits of the proposed processing regime.

Local learning

Learning in hierarchical cortical networks is complex, as any successful outcome needs to be learned across the cortical hierarchy¹⁵². In deep learning this is commonly done using a learning algorithm called backpropagation, which adjusts the weights of the artificial neural network according to the correctness of the outcome so that a correct output is more easily achieved upon the next trial. Backpropagation has several aspects, such as requiring the adjustment of weights throughout the hierarchy, which make it difficult (albeit not impossible) to be implemented in the brain^{153,154}. Synaptic strength (weight) adjustment would be easier to implement in the shallow brain architecture, where each cortical module can provide direct monosynaptic inputs to subcortical areas. For example, from the viewpoint of the striatum, the whole cortex is one very large layer providing excitatory input (Fig. 1c), so, in principle, only corticostriatal synapses need to be strengthened or weakened for reinforcement learning^{155,156}. This does not solve the question of how representation learning happens in the cortex but, rather, underscores the benefit of having the shallow brain processing regime to complement any learning based on cortical plasticity. It is likely that the cortex is needed for learning novel complex stimulus–response mappings^{157,158}, but once a mapping has been established between the output of a cortical column and the appropriate response, the shallow brain regime can take over.

Speed

There is no need for serial processing of inputs from the input layer all the way to the final output layer. Such serial processing would suffer

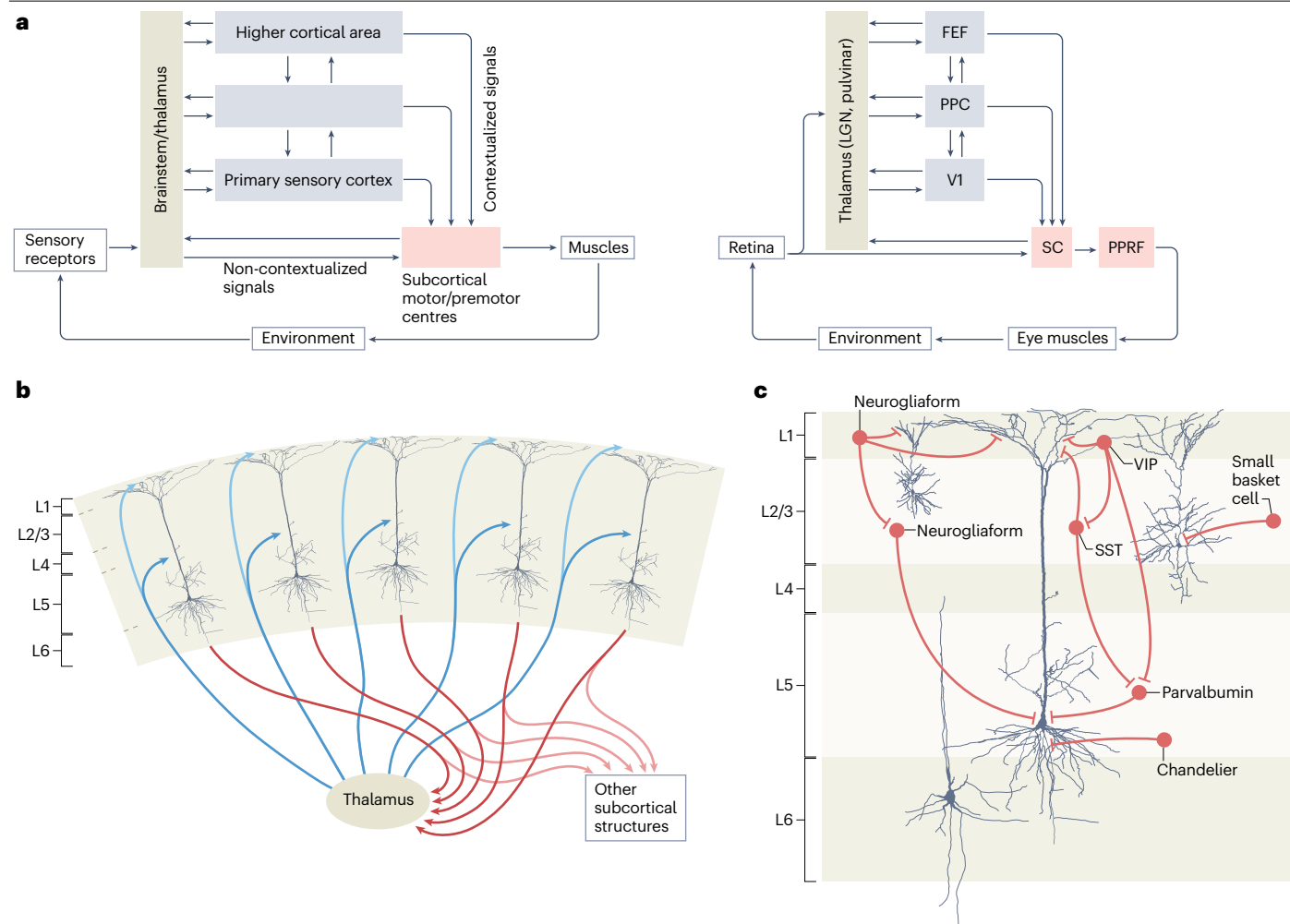


Fig. 2 | The shallow brain hypothesis. a, A distinguishing feature of the brain is the shallowness of its architecture; higher and lower cortical areas receive sensory signals from the environment through the thalamus and/or brainstem via a small number of synapses, and each cortical area can send its outputs directly to the sensorimotor and premotor centres (such as the brainstem reticular formation, red nucleus and spinal cord). For graphical clarity, other connections are omitted (left panel). The oculomotor control network provides an example of a shallow architecture. Hierarchically, the frontal eye field (FEF) is the highest whereas the primary visual cortex (V1) is the lowest. The classical, hierarchical view of these areas is that signals propagate in the order V1 → posterior parietal cortex (PPC) → FEF. However, all of these cortical areas are directly and reciprocally connected with subcortical (thalamic) regions and project to the subcortical superior colliculus (SC), which is well described in a shallow architecture (right panel). **b**, The fast, disynaptic thalamo-cortico-subcortical circuit that we hypothesize has an important role in the shallow architecture. Cortical layer 5 pyramidal (L5p) neurons that directly receive

(first-order and/or higher-order) thalamic inputs and project to subcortical areas are highlighted at the centre of such loops. Duplicated L5p neurons indicate exemplars in different cortical columns or other types of computational unit. For graphical clarity, other connections and cell types are omitted. **c**, Simplified cortical cell types in a columnar arrangement that contains various excitatory pyramidal (black) and inhibitory interneurons (coloured) across six layers. In the human brain, each cortical column contains in the order of 10,000 cells of various cell types. These various cell types form a highly sophisticated network; thus, a mini-hierarchy of processing is argued to exist within a cortical column. Moreover, dendrites of pyramidal neurons do not merely receive inputs from other cells but have substantial computational capacities, and there exist local recurrent connections within each cortical column. All these anatomical and physiological properties together endow each cortical column with an extraordinary computational capacity. LGN, lateral geniculate nucleus; PPRF, paramedian pontine reticular formation; SST, somatostatin; VIP, vasoactive intestinal polypeptide.

from the slowness of polysynaptic transmission and the long time constant of biological neurons¹⁵⁹. In this sense, a cortical hierarchy can perform complex, context-sensitive computations, but may pay the price in terms of sluggishness. The shallow brain hypothesis suggests that interactions between cortical and subcortical areas are shallower (and thus faster), regardless of how distant two areas are physically or

within the cortical hierarchy; interactions between cortical areas should be either direct or run via a minimum number of synapses (for example, via subcortical regions such as higher-order thalamic nuclei)^{42,160}. Moreover, to influence action and cognition, a cortical column simply has to affect the subcortical motor and premotor centres, which are one synapse away. For example, the primate FEF – a hierarchically high

cortical area in the frontal lobe – directly receives thalamic inputs and directly sends its outputs to the brainstem motor and premotor centres (Fig. 2a, right). Thus, the processing pathways – even those involving hierarchically high cortical areas – can be very shallow and fast.

Compositionality and flexible combination of features

The shallow brain architecture allows multiple representations across the cortical hierarchy to contribute to the selection of overt action and cognitive operations (Fig. 3). In some tasks and situations, lower levels of cortical processing may compute the details (for example, visual stimulus orientation) as necessary to make a behavioural decision. By contrast, a hierarchical computation would require the details from lower levels of cortical processing to be propagated throughout the cortical hierarchy. However, according to the shallow brain hypothesis, lower levels of cortical processing can contribute directly to the selection of thought and action by projecting to subcortical centres, where the output of lower and higher cortical areas can be combined (Fig. 3). More generally, the output of any cortical column (or module) has direct access to subcortical centres, thus allowing for compositionality, which is the ability to combine simpler features for complex representations^{161,162}. According to our hypothesis,

the outputs of cortical columns constitute the primitives, which are flexibly combined at the level of subcortical areas (Fig. 3 and Box 1). A concrete prediction of the shallow brain architecture is that when, for any given task, low-level (orientation) and high-level (face) information needs to be combined, this is done by combining the outputs of the respective cortical columns at the level of subcortical areas (Fig. 3).

Similarly, when low-level and high-level processed information would compete to invigorate conflicting decisions, this competition can be resolved in subcortical areas, not in the cortex. This may be more efficient, as low and high cortical areas are anatomically more remote from each other than their outputs at the level of the thalamus or other subcortical regions, where many cortical outputs are funnelled together (Figs. 1d and 3). Thus, cortical computation adds a layer of contextual, combinatorial complexity, which can be integrated into subcortical decision-making. Moreover, the communication lines between the cortex and subcortical centres may enable the brain to avoid conflict in case of discrepancy between cortical and subcortical computations when solving a task. Cortical contextual, combinatorial processing is especially needed when dealing with novel stimuli or situations^{157,163}. However, if the cortex has already developed the representations required for solving the task, the subcortical structures

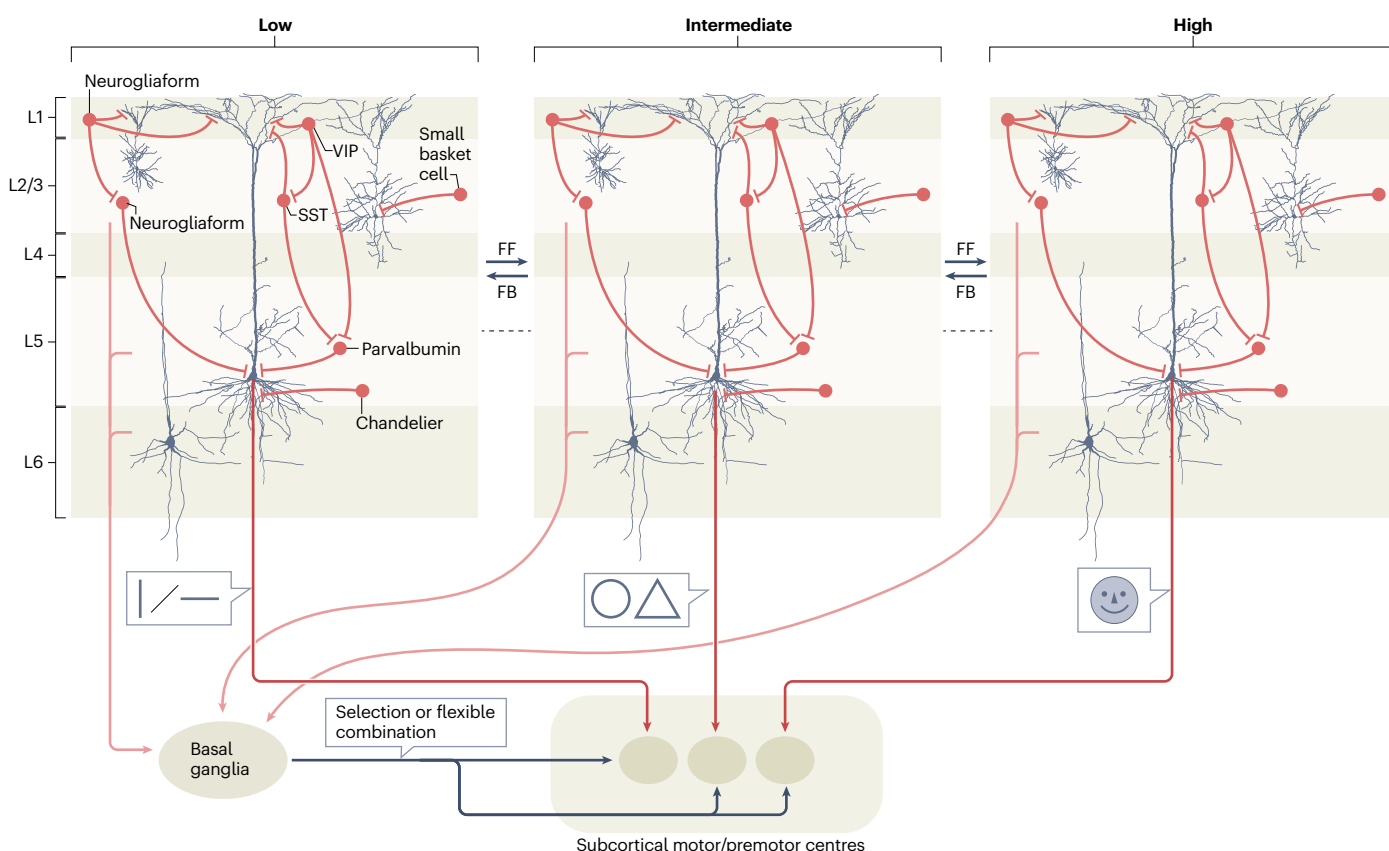


Fig. 3 | New computational potential in shallow architectures. The human neocortex consists of approximately 100,000 cortical columns. The shallow architecture we propose combines such massive outputs potentially from all – higher and lower – cortical columns, which is markedly different from the conventional hierarchical structure of deep learning and predictive coding networks, in which only the highest layer can provide the output. Here, this potential for combinations is illustrated by the convergence of low-level

(lines, edges), intermediate-level (simple shapes) and high-level (face) information. Inhibitory projections (black) from the basal ganglia (BGa) to subcortical motor and premotor centres (in the thalamus and brainstem) are proposed to serve as the selection mechanism for the convergent information from excitatory outputs from cortical columns in lower, intermediate and higher cortical areas (red). FB, feedback; FF, feedforward; SST, somatostatin; VIP, vasoactive intestinal polypeptide.

can directly use them in a shallow brain architecture. Hence, although there is progression of more abstract features when going up the cortical hierarchy, subcortical areas can process and combine information from all these levels flexibly (Box 1).

A question that arises from the shallow brain hypothesis is how the subcortical structures select and combine all this information. This is discussed in the next section.

Implications of the shallow brain hypothesis

A fast, robust computational capacity of a parallel computational architecture – the subsumption architecture¹⁶⁴ – was proposed early on in artificial intelligence and robotics¹⁶⁵. Instead of successively computing abstract features from sensory areas (as assumed in predictive coding and deep learning), a subsumption architecture decomposes the complete robot behaviour into behavioural competences (for example, to avoid an obstacle, to reach a specific goal), each of which is implemented by a control layer. Importantly, all layers of the subsumption architecture receive sensory information and produce behavioural outputs in parallel¹⁶⁴ – similar to what we propose for the shallow brain architecture (Figs. 1d and 2a). At this point, the computational principles of the shallow brain architecture are relatively unknown, but what is already clear is that it should be completely different from principles used in purely hierarchical computation (such as deep learning). Thus, implications of the shallow brain hypothesis are wide-reaching.

First, a longstanding mystery in neuroscience is how our train of thought, intelligent behaviour and cognitive functions are made possible despite a slow synaptic transmission and membrane time constant – followed by an even longer refractory period after spiking¹⁵⁹ (see ref. 166 for a proposal to overcome this slowness). If our intelligent behaviour and cognitive functions depend on hierarchical processing within the cortex, processing time increases as the number of areas (and thus

the number of synapses to pass through) increases (Fig. 1a). By contrast, if the brain's network is essentially shallow, as we hypothesize, the problem of neuronal sluggishness disappears. Here, we are not claiming that hierarchical processing would be irrelevant but that its sluggishness can be avoided via a set of fast pathways – cortico-subcortical projections and direct cortico-cortical or trans-thalamic connections between hierarchically distant areas (Fig. 2a).

One might ask how it is possible for such a shallow architecture to produce or contribute to our rich internal representations of the world, including our body. A deep hierarchical architecture permits progressive encoding of abstract features for categorization, which is why a deep hierarchical architecture is more intuitively appealing for many people and has been widely adopted in machine learning. However, deep learning hierarchical architectures commonly assume that each layer consists of simplistic point neurons that merely summate inputs, and this assumption substantially underestimates the computational capacity of the microcircuit in each cortical column. Each cortical microcircuit contains a large variety of neurons, excitatory and inhibitory, that form a highly sophisticated network across the six layers in a cortical column^{83,167} (Fig. 2c). The argument can be made that, even within each cortical microcircuit, a mini-hierarchy of processing exists (namely from layer 4 to layer 2/3 to layers 5 and 6 (ref. 168), but see refs. 169–171 for other pathways). Moreover, dendrites have been classically considered as structures that merely receive inputs from other cells, but abundant evidence indicates that they possess substantial computational capacities^{172,173} (see refs. 174–179 for models that incorporate dendritic computational capacities). Consequently, single cortical pyramidal neurons have been approximated by a deep neural network with five to eight layers¹⁸⁰. Thus, the shallow architecture we present here (Fig. 2a) highlights the substantial computational capacity of the microcircuit in each cortical column.

Box 1

Flexibility in cortical and subcortical computations

Computational flexibility can be considered at two timescales. The first is the slow timescale that applies to learning and adaptation. The second is the fast timescale of behavioural decision-making and selection of actions and cognitive operations (in the milliseconds to seconds range). Here we focus mostly on the fast timescale and define flexibility as the ability of a neural system to adjust its operations adaptively, on the basis of task demands, behavioural feedback and properties of the stimulus and context reacted upon.

Various candidate mechanisms may contribute to computational flexibility. A key question is how different combinations of cortical feature representations can be rapidly selected. We suggest that attentional brain mechanisms can drive this selection process, aided by prefrontal regions coding task rules^{208–211}. A frontal–parietal network has been identified that harbours key resources for selective attention^{212,213}. Influences of attention are expressed in the visual cortex by, for instance, reduced cross-correlations²¹⁴, increased gain modulation²¹⁵ and regional filling effects in figure–ground segmentation²¹⁶, but it is notable that this frontal–parietal network also exerts effects on at least some subcortical structures^{213,217}.

Vice versa, many subcortical structures also regulate attention, cognitive flexibility and/or arousal. For instance, the thalamic pulvinar and reticular nuclei regulate visual attention²¹⁸ and switching between cortical representations²¹⁹, whereas the intralaminar thalamic nuclei function in cortical and subcortical arousal^{220,221}. Moreover, the neuromodulatory cell groups in the locus coeruleus, raphe nuclei and pedunculopontine nucleus (PPN) receive cortical inputs and exert arousing and attentional effects throughout the cortex, basal ganglia (BGa) and other subcortical structures. Recent studies reveal modular targeting of neuromodulatory outputs to the cortex and related areas, suggesting selective modulation of particular cortical regions and representations^{222–224}. Finally, when considering how the BGa may help to select which cortical inputs gain access to subcortical processing, it is of note that they have been implicated in attention and cognitive flexibility as well^{225,226}. Thus, although there is no shortage of mechanisms underlying computational flexibility in cortico-subcortical architectures, further studies are required to examine which mechanisms are most prominent under different circumstances.

Glossary

Bayesian inference

A method of statistical analysis that is grounded in Bayes' theorem, which describes how the probability of a hypothesis (posterior probability) is updated as new data (evidence) become available, given prior knowledge about the hypothesis (prior probability).

Cortico-cortical loops

Neural circuits that connect different regions of the cerebral cortex to one another, allowing communication and integration of information across various cortical areas. These loops can be either short range, connecting adjacent or nearby cortical regions, or long range, linking distant regions of the cortex.

Deep hierarchy

A hierarchical structure consisting of many layers (roughly analogous to cortical areas) through which information from the external world is processed step by step.

Deep learning architectures

Structured configurations of hierarchical, interconnected layers of artificial neurons, or nodes, in a neural network. Common types of

deep learning architecture include feedforward convolutional neural networks and recurrent neural networks.

Hierarchical inference

The process of drawing conclusions from data wherein parameters are organized into different levels or layers. In hierarchical Bayesian inference, Bayesian statistics are employed within a layered framework, integrating prior knowledge at multiple levels to refine posterior distributions.

Higher-order thalamic nuclei

Thalamic nuclei can be categorized anatomically into first-order and higher-order nuclei. First-order nuclei receive driving afferents from ascending pathways, whereas the higher-order nuclei receive driving afferents from cortical layer 5 pyramidal (L5p) neurons. Notable examples of higher-order thalamic nuclei include the pulvinar and the medial dorsal nucleus.

Non-hierarchical lateral connections

Connections made between two cortical areas that are not distinguished hierarchically (for instance, primary

auditory and visual cortex). This connectivity pattern is illustrated in Fig. 1b.

Recurrent connections

Connections in which the output of a neuron at a given layer is fed back as an input to either the same layer or a previous layer. This creates a loop in the network, allowing information, for instance, to persist and be reused across sequential steps.

Recurrent neural network

A class of neural networks in which connections between nodes form directed cycles, enabling the retention of information from previous inputs. This sequential memory feature makes recurrent neural networks suitable for tasks involving time-series or sequential data.

Reinforcement learning

A machine learning method in which an agent makes decisions and receives reinforcing feedback to train the network to improve its output (for example, reward for desired behaviours, punishment for behaviour resulting in undesirable output).

Shallow architectures

Architectures that do not consist of a deep hierarchy. Shallow architectures instead have a minimum number of layers.

Shallow processing

Computations carried out by a shallow architecture, namely in a few steps instead of tens or hundreds of layers of processing.

Thalamo-cortical loops

Bidirectional pathways between the thalamus and the cerebral cortex. Thalamo-cortical loops play a vital role in the regulation of consciousness, attention and sensory processing, and have been implicated in several neurological and psychiatric disorders.

Trans-thalamic connections

Connections made between two brain regions via the thalamus.

In the human brain, each cortical column contains in the order of 10,000 cells of various cell types, and there are approximately 100,000 cortical columns¹²⁸. Given the existence of local recurrent connections within each cortical column^{181,182}, each cortical column can be seen as a recurrent neural network. The shallow architecture can be represented as a massive array of parallel recurrent neural networks that not only project subcortically but also connect with each other via cortico-cortical and trans-thalamic connections. In the shallow brain architecture, this array of cortical columns processes information in parallel and provides their outputs through the L5p neurons to subcortical motor and premotor centres (Fig. 2). The cortical hierarchy is used to represent progressively more abstract features in their context^{7–9}. The shallow architecture we propose flexibly combines many outputs from cortical columns of every hierarchical level (Fig. 3), which is strikingly different from conventional hierarchical structures in deep learning and predictive coding networks, where only the highest layer provides the output. Another component that may further enhance the computational capacity of the shallow architecture is the non-hierarchical lateral connections between cortical areas^{35,183–185}. These connections, in theory, enable integration of multimodal information processing at multiple hierarchical levels. Thus, these non-hierarchical intracortical connections provide an additional degree of flexibility at the cortical and subcortical levels, which potentially facilitates multimodal sensorimotor control.

The shallow architecture of the brain we propose also raises novel and interesting questions. For instance, how is the solution to any given task computed if it is dispersed among so many parallel processors? Are there subcortical mechanisms for implementing internal competition to select an optimal action choice – for example, winner takes all (Fig. 3)? On the basis of anatomical data, the most likely candidate we propose for this mechanism is the set of inhibitory projections from the basal ganglia (BGa) to the diencephalon, mesencephalon and brainstem. Besides the extensively studied inhibitory projections from the BGa to thalamic nuclei such as the anterior, centromedian, parafascicular and mediodorsal nuclei^{186–188}, less exhaustively studied yet important for our hypothesis are BGa projections to the brainstem^{189–191}. BGa targets in the brainstem include the PPN and mesencephalic locomotor region, retrorubral area (A8 dopaminergic cell group), red nucleus and SC^{108,191–194}. For example, in the oculomotor control system (Fig. 2a, right), the substantia nigra pars reticulata sends GABAergic projections to the SC as well as the thalamus^{108,195}. Together with the inhibitory BGa projections to the thalamus, these inhibitory BGa projections to the brainstem are proposed to provide the mechanism for flexible selection of which cortically receptive cell groups in the motor and premotor centres are activated, and thus which cortico-subcortical projections remain shut down by inhibition (Fig. 3). For instance, BGa projections that inhibit all except one descending motor output from the cortex

result in the selection of a particular target cell group for effectuating a particular action. In addition, the BGa exert inhibitory control over widespread release of neuromodulators, such as acetylcholine from the PPN and dopamine from the retrorubral area.

However, not only can the subcortical areas perform selection of cortico-subcortical projections at the level of subcortical areas, but also subcortical areas can considerably influence the selection of cortico-subcortical projections at the level of cortical processing (see also Box 1). One key pathway for this influence at the level of cortical processing comes from projections from higher-order thalamic nuclei to the cortex, which can bias whether cortical columns are strongly or weakly activated. Furthermore, two different types of projection from the higher-order thalamic nuclei to different layers of the cortex exist (Fig. 2b). The thalamo-cortical projection into layer I not only activates distal apical dendrites of subcortically projecting pyramidal neurons but also activates interneurons that strongly influence the activity of the receiving cortical column^{196–198} (Fig. 2c). The thalamo-cortical projection into the middle layer has been shown to couple or decouple cortical processing streams^{94,96}. Thus, the higher-order thalamus is in a powerful position to select which cortical computations reach subcortical areas at the level of cortical processing.

Summary and future directions

In this Perspective, we have questioned the current dominance of hierarchical cortical structures used in deep learning and predictive coding. Instead, on the basis of neuroanatomical and electrophysiological evidence, we propose the shallow brain hypothesis, which provides a new perspective on massively parallel computations in the brain. We have highlighted how each higher or lower cortical area is reciprocally connected with subcortical regions and that we have not yet fully understood the extraordinary computational capacity of each cortical microcircuit. Our hypothesis posits that the brain has a shallow architecture – consisting of a massive array of parallel recurrent neural networks, each of which not only projects subcortically but also possesses a highly sophisticated microcircuit – which enables fast yet powerful computations that exploit the shallow cortico-subcortical loops and the computational capacity of cortical hierarchy that is ‘laterally’ formed between the parallel networks (Fig. 3). We hope that the shallow brain hypothesis inspires researchers to focus more on developing computational models that take into account this shallow architecture with a massive array of parallel thalamo-cortico-subcortical loops. To stimulate such developments, we make some testable predictions below.

We have emphasized the computational power of cortical microcircuits. However, the key to the shallow brain hypothesis is that the genuine power of cortical microcircuits can be understood only when their interactions with the subcortex are fully taken into account. Some exciting work has demonstrated how cortical computations are maintained and supported by the thalamus^{91,199,200}, but the shallow brain hypothesis predicts that this phenomenon is widespread – when cortico-subcortico-cortical loops are specifically manipulated, cortical activity will be substantially affected. Systematic studies using optogenetic and pharmacogenetic manipulations must be done to examine the contributions of different subcortical areas and loops. For example, although the contribution of cortical areas to learning is widely accepted, the contribution of the higher-order thalamus (for example, pulvinar, posteromedial and mediodorsal nuclei) to learning is far less recognized^{201,202}. Our hypothesis predicts a crucial role of such thalamo-cortical loops in learning; for example, optogenetic or

pharmacological disruption of these specific thalamo-cortical pathways may impair learning, even when cortico-cortical pathways are intact.

The shallow brain hypothesis posits that subcortically projecting L5p neurons are the key components that enable fast yet powerful computation by the shallow architecture because they directly receive subcortical inputs and send their outputs to subcortical regions (Fig. 2b). During learning, slow hierarchical computations involving various cell types in multiple cortical areas take place, but we predict that as learning progresses, plastic changes in synapses shape cortical circuits such that these subcortically projecting L5p neurons take over the hierarchical computation and perform it faster. This prediction is consistent with previous observations in imaging studies during rodent motor learning¹⁴⁷ but might be generalized to other cortical areas and other types of learning (such as representation learning).

We predict that computational models and artificial intelligence implementations of the shallow brain hypothesis will outperform conventional architectures in some applications – in particular, in tasks that require compositionality^{203,204} and a flexible combination of features that are computed at different levels of the hierarchy. Also, the architecture presented here can lay the backbone for novel artificial intelligence algorithms such as modular deep learning architectures^{205,206}. Moreover, our hypothesis predicts that when low-level information and high-level information need to be combined, then this is done not in the cortex but, rather, by combining the outputs of the respective cortical processing areas at the level of subcortical areas (Fig. 3). Similarly, when low-level information and high-level information compete, this competition is resolved in subcortical areas, not in the cortex. Last, the BGa are known to project to the thalamus and brainstem, but little is known as to whether these projections are independent of each other; our hypothesis predicts that these BGa–thalamus and BGa–brainstem connections are not independent of each other but, rather, are orchestrated to flexibly combine or regulate massively parallel cortical outputs to these subcortical regions.

A deeper understanding of the computational capacity of cortical microcircuits, each of which reciprocally interacts with subcortical regions, as well as subcortical mechanisms for flexible combination of cortical outputs and the functional role of direct cortico-cortical or trans-thalamic connections between hierarchically distant areas, may revise, prove or disprove the shallow brain hypothesis.

Published online: 27 October 2023

References

- Hegde, J. & Felleman, D. J. Reappraising the functional implications of the primate visual anatomical hierarchy. *Neuroscientist* **13**, 416–421 (2007).
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
- Stokel-Walker, C. & Van Noord, R. What ChatGPT and generative AI mean for science. *Nature* **614**, 214–216 (2023).
- He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 770–778 (2016).
- Russakovsky, O. et al. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
- Xu, K. et al. Show, attend and tell: neural image caption generation with visual attention. In *32nd Int. Conf. on Machine Learning* (eds F. Bach. & D. Blei) 2048–2057 (2015).
- Fukushima, K. Neocognitron—a self-organizing neural network model for a mechanism of pattern-recognition unaffected by shift in position. *Biol. Cybern.* **36**, 193–202 (1980).
- Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154 (1962).
- Chang, L. & Tsao, D. Y. The code for facial identity in the primate brain. *Cell* **169**, 1013–1028.e14 (2017).
- Yamins, D. L. et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl Acad. Sci. USA* **111**, 8619–8624 (2014).
- Guclu, U. & van Gerven, M. A. Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *J. Neurosci.* **35**, 10005–10014 (2015).

12. Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A. & Oliva, A. Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Sci. Rep.* **6**, 27755 (2016).
13. Schrimpf, M. et al. The neural architecture of language: integrative modeling converges on predictive processing. *Proc. Natl Acad. Sci. USA* **118**, e2015646118 (2021).
14. Kriegeskorte, N. Deep neural networks: a new framework for modeling biological vision and brain information processing. *Annu. Rev. Vis. Sci.* **1**, 417–446 (2015).
15. Yamins, D. L. & DiCarlo, J. J. Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* **19**, 356–365 (2016).
16. Friston, K. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360**, 815–836 (2005).
17. Rao, R. P. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
18. Lee, T. S. & Mumford, D. Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **20**, 1434–1448 (2003).
19. Srinivasan, M. V., Laughlin, S. B. & Dubs, A. Predictive coding: a fresh view of inhibition in the retina. *Proc. R. Soc. Lond. B Biol. Sci.* **216**, 427–459 (1982).
20. Dayan, P., Hinton, G. E., Neal, R. M. & Zemel, R. S. The Helmholtz machine. *Neural Comput.* **7**, 889–904 (1995).
21. Dora, S., Bohte, S. M. & Pennartz, C. M. A. Deep gated Hebbian predictive coding accounts for emergence of complex neural response properties along the visual cortical hierarchy. *Front. Comput. Neurosci.* **15**, 666131 (2021).
22. McDermott, J. H., Wroblewski, D. & Oxenham, A. J. Recovering sound sources from embedded repetition. *Proc. Natl Acad. Sci. USA* **108**, 1188–1193 (2011).
23. Mill, R. W., Bohm, T. M., Bendixen, A., Winkler, I. & Denham, S. L. Modelling the emergence and dynamics of perceptual organisation in auditory streaming. *PLoS Comput. Biol.* **9**, e1002925 (2013).
24. Kanai, R., Komura, Y., Shipp, S. & Friston, K. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20140169 (2015).
25. Schwartenbeck, P., FitzGerald, T. H., Mathys, C., Dolan, R. & Friston, K. The dopaminergic midbrain encodes the expected certainty about desired outcomes. *Cereb. Cortex* **25**, 3434–3445 (2015).
26. Rikhye, R. V., Wimmer, R. D. & Halassa, M. M. Toward an integrative theory of thalamic function. *Annu. Rev. Neurosci.* **41**, 163–183 (2018).
27. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
28. Minsky, M. & Papert, S. *Perceptrons; An Introduction to Computational Geometry* (MIT Press, 1969).
29. Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. Visual properties of neurons in inferotemporal cortex of the macaque. *J. Neurophysiol.* **35**, 96–111 (1972).
30. Tsao, D. Y., Schweers, N., Moeller, S. & Freiwald, W. A. Patches of face-selective cortex in the macaque frontal lobe. *Nat. Neurosci.* **11**, 877–879 (2008).
31. Hegde, J. & Van Essen, D. C. A comparative study of shape representation in macaque visual areas V2 and V4. *Cereb. Cortex* **17**, 1100–1116 (2007).
32. Rockland, K. S. & Pandya, D. N. Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Res.* **179**, 3–20 (1979).
33. Markov, N. T. et al. Cortical high-density counterstream architectures. *Science* **342**, 1238406 (2013).
34. Markov, N. T. & Kennedy, H. The importance of being hierarchical. *Curr. Opin. Neurobiol.* **23**, 187–194 (2013).
35. D'Souza, R. D. et al. Hierarchical and nonhierarchical features of the mouse visual cortical network. *Nat. Commun.* **13**, 503 (2022).
36. Siegle, J. H. et al. Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* **592**, 86–92 (2021).
37. Nakamura, H., Gattass, R., Desimone, R. & Ungerleider, L. G. The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. *J. Neurosci.* **13**, 3681–3691 (1993).
38. Burkhalter, A., D'Souza, R. D., Ji, W. & Meier, A. M. Integration of feedforward and feedback information streams in the modular architecture of mouse visual cortex. *Annu. Rev. Neurosci.* **46**, 259–280 (2023).
39. Coogan, T. A. & Burkhalter, A. Hierarchical organization of areas in rat visual cortex. *J. Neurosci.* **13**, 3749–3772 (1993).
40. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).
41. Bastos, A. M. et al. Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711 (2012).
42. Pennartz, C. M. A., Dora, S., Muckli, L. & Lorteije, J. A. M. Towards a unified view on pathways and functions of neural recurrent processing. *Trends Neurosci.* **42**, 589–603 (2019).
43. Findling, C. et al. Brain-wide representations of prior information in mouse decision-making. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.07.04.547684> (2023).
44. Keller, G. B. & Mrsic-Flogel, T. D. Predictive processing: a canonical cortical computation. *Neuron* **100**, 424–435 (2018).
45. Keller, G. B., Bonhoeffer, T. & Hubener, M. Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. *Neuron* **74**, 809–815 (2012).
46. Jordan, R. & Keller, G. B. Opposing influence of top-down and bottom-up input on excitatory layer 2/3 neurons in mouse primary visual cortex. *Neuron* **108**, 1194–1206.e5 (2020).
47. Padamsey, Z. & Rochefort, N. L. Defying expectations: how neurons compute prediction errors in visual cortex. *Neuron* **108**, 1016–1019 (2020).
48. Muzzu, T. & Saleem, A. B. Feature selectivity can explain mismatch signals in mouse visual cortex. *Cell Rep.* **37**, 109772 (2021).
49. Walsh, K. S., McGovern, D. P., Clark, A. & O'Connell, R. G. Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Ann. N. Y. Acad. Sci.* **1464**, 242–268 (2020).
50. Schwiedrzik, C. M. & Freiwald, W. A. High-level prediction signals in a low-level area of the macaque face-processing hierarchy. *Neuron* **96**, 89–97.e4 (2017).
51. Issa, E. B., Cadieu, C. F. & DiCarlo, J. J. Neural dynamics at successive stages of the ventral visual stream are consistent with hierarchical error signals. *eLife* **7**, e42870 (2018).
52. Chao, Z. C., Takaura, K., Wang, L., Fujii, N. & Dehaene, S. Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. *Neuron* **100**, 1252–1266.e3 (2018).
53. Spratling, M. W. A review of predictive coding algorithms. *Brain Cogn.* **112**, 92–97 (2017).
54. Spratling, M. W. Fitting predictive coding to the neurophysiological data. *Brain Res.* **1720**, 146313 (2019).
55. Bianchini, M. & Scarselli, F. On the complexity of neural network classifiers: a comparison between shallow and deep architectures. *IEEE Trans. Neural Netw. Learn. Syst.* **25**, 1553–1565 (2014).
56. Cohen, N., Sharir, O. & Shashua, A. On the expressive power of deep learning: a tensor analysis. In *29th Annual Conference on Learning Theory* (eds. Feldman, V., Rakhlin, A. & Shamir, O.) 698–728 (2016).
57. Hinton, G. E. Training products of experts by minimizing contrastive divergence. *Neural Comput.* **14**, 1771–1800 (2002).
58. Dabelow, L. & Ueda, M. Three learning stages and accuracy-efficiency tradeoff of restricted Boltzmann machines. *Nat. Commun.* **13**, 5474 (2022).
59. Liao, R., Kornblith, S., Ren, M., Fleet, D. J. & Hinton, G. Gaussian–Bernoulli RBMs without tears. Preprint at *arXiv* <https://doi.org/10.48550/ARXIV.2210.10318> (2022).
60. Hilgetag, C. C. & Goulas, A. 'Hierarchy' in the organization of brain networks. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190319 (2020).
61. Sherman, S. M. & Guillery, R. W. Functional organization of thalamocortical relays. *J. Neurophysiol.* **76**, 1367–1395 (1996).
62. Jones, E. G. The thalamic matrix and thalamocortical synchrony. *Trends Neurosci.* **24**, 595–601 (2001).
63. Sherman, S. M. & Guillery, R. W. *Exploring the Thalamus and Its Role in Cortical Function* 2nd edn (MIT Press, 2009).
64. Halassa, M. *Thalamus* 1st edn (Cambridge Univ. Press, 2023).
65. Kemp, J. M. & Powell, T. P. The cortico-striate projection in the monkey. *Brain* **93**, 525–546 (1970).
66. Oka, H. Organization of the cortico-caudate projections. A horseradish peroxidase study in the cat. *Exp. Brain Res.* **40**, 203–208 (1980).
67. Ito, S. & Feldheim, D. A. The mouse superior colliculus: an emerging model for studying circuit formation and function. *Front. Neural Circuits* **12**, 10 (2018).
68. Basso, M. A. & May, P. J. Circuits for action and cognition: a view from the superior colliculus. *Annu. Rev. Vis. Sci.* **3**, 197–226 (2017).
69. May, P. J. The mammalian superior colliculus: laminar structure and connections. *Prog. Brain Res.* **151**, 321–378 (2006).
70. McBride, E. G. et al. Influence of claustrum on cortex varies by area, layer, and cell type. *Neuron* **111**, 275–290.e5 (2022).
71. Nariyoshi, K. et al. The claustrum coordinates cortical slow-wave activity. *Nat. Neurosci.* **23**, 741–753 (2020).
72. Jackson, J., Karnani, M. M., Zemelman, B. V., Burdakov, D. & Lee, A. K. Inhibitory control of prefrontal cortex by the claustrum. *Neuron* **99**, 1029–1039.e4 (2018).
73. Legg, C. R., Mercier, B. & Glickstein, M. Corticopontine projection in the rat: the distribution of labelled cortical cells after large injections of horseradish peroxidase in the pontine nuclei. *J. Comp. Neurol.* **286**, 427–441 (1989).
74. Habas, C. & Cabanis, E. A. Cortical projections to the human red nucleus: a diffusion tensor tractography study with a 1.5-T MRI machine. *Neuroradiology* **48**, 755–762 (2006).
75. Tervo, D. G. et al. A designer AAV variant permits efficient retrograde access to projection neurons. *Neuron* **92**, 372–382 (2016).
76. Murakami, T., Matsui, T., Uemura, M. & Ohki, K. Modular strategy for development of the hierarchical visual network in mice. *Nature* **608**, 578–585 (2022).
77. Tang, L. & Higley, M. J. Layer 5 circuits in V1 differentially control visuomotor behavior. *Neuron* **105**, 346–354.e5 (2020).
78. Takahashi, N. et al. Active dendritic currents gate descending cortical outputs in perception. *Nat. Neurosci.* **23**, 1277–1285 (2020).
79. Fuster, J. M. *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe* (Raven, 1980).
80. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
81. Oswald, M. J., Tantirigama, M. L., Sonntag, I., Hughes, S. M. & Empson, R. M. Diversity of layer 5 projection neurons in the mouse motor cortex. *Front. Cell Neurosci.* **7**, 174 (2013).
82. Akintunde, A. & Buxton, D. F. Origins and collateralization of corticospinal, corticopontine, corticorubral and corticostriatal tracts: a multiple retrograde fluorescent tracing study. *Brain Res.* **586**, 208–218 (1992).
83. Harris, K. D. & Shepherd, G. M. The neocortical circuit: themes and variations. *Nat. Neurosci.* **18**, 170–181 (2015).
84. Musall, S. et al. Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. *Nat. Neurosci.* **26**, 495–505 (2023).

85. Mohan, H. et al. Cortical glutamatergic projection neuron types contribute to distinct functional subnetworks. *Nat. Neurosci.* **26**, 481–494 (2023).
86. Kuramoto, E. et al. Ventral medial nucleus neurons send thalamocortical afferents more widely and more preferentially to layer 1 than neurons of the ventral anterior–ventral lateral nuclear complex in the rat. *Cereb. Cortex* **25**, 221–235 (2015).
87. Cruikshank, S. J. et al. Thalamic control of layer 1 circuits in prefrontal cortex. *J. Neurosci.* **32**, 17813–17823 (2012).
88. Schroeder, A. et al. Inhibitory top-down projections from zona incerta mediate neocortical memory. *Neuron* **111**, 727–738.e8 (2023).
89. Ahmadlou, M. et al. A cell type-specific cortico-subcortical brain circuit for investigatory and novelty-seeking behavior. *Science* **372**, eabe9681 (2021).
90. Brenner, J. M., Beltramo, R., Gerfen, C. R., Ruediger, S. & Scanziani, M. A genetically defined tecto-thalamic pathway drives a system of superior-colliculus-dependent visual cortices. *Neuron* **111**, 2247–2257.e7 (2023).
91. Guo, Z. V. et al. Maintenance of persistent activity in a frontal thalamocortical loop. *Nature* **545**, 181–186 (2017).
92. Hsiao, K. et al. A thalamic orphan receptor drives variability in short-term memory. *Cell* **183**, 522–536.e19 (2020).
93. Redinbaugh, M. J. et al. Thalamus modulates consciousness via layer-specific control of cortex. *Neuron* **106**, 66–75.e12 (2020).
94. Aru, J., Suzuki, M. & Larkum, M. E. Cellular mechanisms of conscious processing. *Trends Cogn. Sci.* **24**, 814–825 (2020).
95. Aru, J., Suzuki, M., Rutiku, R., Larkum, M. E. & Bachmann, T. Coupling the state and contents of consciousness. *Front. Syst. Neurosci.* **13**, 43 (2019).
96. Suzuki, M. & Larkum, M. E. General anesthesia decouples cortical pyramidal neurons. *Cell* **180**, 666–676.e13 (2020).
97. Schiff, N. D. et al. Behavioural improvements with thalamic stimulation after severe traumatic brain injury. *Nature* **448**, 600–603 (2007).
98. Bastos, A. M. et al. Neural effects of propofol-induced unconsciousness and its reversal using thalamic stimulation. *eLife* **10**, e60824 (2021).
99. Crick, F. C. & Koch, C. What is the function of the claustrum. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360**, 1271–1279 (2005).
100. Chevee, M., Finkel, E. A., Kim, S. J., O'Connor, D. H. & Brown, S. P. Neural activity in the mouse claustrum in a cross-modal sensory selection task. *Neuron* **110**, 486–501.e7 (2022).
101. Huang, W., Qin, J., Zhang, C., Qin, H. & Xie, P. Footshock-induced activation of the Huang–entorhinal cortical pathway in freely moving mice. *Physiol. Res.* **71**, 695–701 (2022).
102. Smythies, J. On the function of object cells in the claustrum—key components in information processing in the visual system? *Front. Cell Neurosci.* **9**, 443 (2015).
103. Tsumoto, T. & Suda, K. Effects of stimulation of the dorsocaudal claustrum on activities of striate cortex neurons in the cat. *Brain Res.* **240**, 345–349 (1982).
104. Remedios, R., Logothetis, N. K. & Kayser, C. A role of the claustrum in auditory scene analysis by reflecting sensory change. *Front. Syst. Neurosci.* **8**, 44 (2014).
105. Qadir, H. et al. The mouse claustrum synaptically connects cortical network motifs. *Cell Rep.* **41**, 111860 (2022).
106. Taylor, N. L. et al. Structural connections between the noradrenergic and cholinergic system shape the dynamics of functional brain networks. *Neuroimage* **260**, 119455 (2022).
107. Deutch, A. Y. & Roth, R. H. in *Fundamental Neuroscience* (eds M. J. Zigmond et al.) 193–234 (Academic, 1999).
108. Chevalier, G. & Deniau, J. M. Disinhibition as a basic process in the expression of striatal functions. *Trends Neurosci.* **13**, 277–280 (1990).
109. Voorn, P., Vanderschuren, L. J., Groenewegen, H. J., Robbins, T. W. & Pennartz, C. M. Putting a spin on the dorsal–ventral divide of the striatum. *Trends Neurosci.* **27**, 468–474 (2004).
110. Budinger, E., Heil, P., Hess, A. & Scheich, H. Multisensory processing via early cortical stages: connections of the primary auditory cortical field with other sensory systems. *Neuroscience* **143**, 1065–1083 (2006).
111. Benavidez, N. L. et al. Organization of the inputs and outputs of the mouse superior colliculus. *Nat. Commun.* **12**, 4004 (2021).
112. Beltramo, R. & Scanziani, M. A collicular visual cortex: neocortical space for an ancient midbrain visual structure. *Science* **363**, 64–69 (2019).
113. Constantinople, C. M. & Bruno, R. M. Effects and mechanisms of wakefulness on local cortical networks. *Neuron* **69**, 1061–1068 (2011).
114. Aru, J., Siclari, F., Phillips, W. A. & Storm, J. F. Apical drive—a cellular mechanism of dreaming? *Neurosci. Biobehav. Rev.* **119**, 440–455 (2020).
115. Wainstein, G., Muller, E. J., Taylor, N., Munn, B. & Shine, J. M. The role of the locus coeruleus in shaping adaptive cortical melodies. *Trends Cogn. Sci.* **26**, 527–538 (2022).
116. Polack, P. O., Friedman, J. & Golshani, P. Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nat. Neurosci.* **16**, 1331–1339 (2013).
117. Harris, K. D. & Thiele, A. Cortical state and attention. *Nat. Rev. Neurosci.* **12**, 509–523 (2011).
118. Parikh, V., Kozak, R., Martinez, V. & Sarter, M. Prefrontal acetylcholine release controls cue detection on multiple timescales. *Neuron* **56**, 141–154 (2007).
119. Puig, M. V. & Gullledge, A. T. Serotonin and prefrontal cortex function: neurons, networks, and circuits. *Mol. Neurobiol.* **44**, 449–464 (2011).
120. Buhot, M. C., Martin, S. & Segu, L. Role of serotonin in memory impairment. *Ann. Med.* **32**, 210–221 (2000).
121. Petroni, F., Panzeri, S., Hilgetag, C. C., Kotter, R. & Young, M. P. Simultaneity of responses in a hierarchical visual network. *Neuroreport* **12**, 2753–2759 (2001).
122. Zeki, S. The rough seas of cortical cartography. *Trends Neurosci.* **41**, 242–244 (2018).
123. Silvano, J. Why is “blindsight” blind? A new perspective on primary visual cortex, recurrent activity and visual awareness. *Conscious. Cogn.* **32**, 15–32 (2015).
124. Schmolesky, M. T. et al. Signal timing across the macaque visual system. *J. Neurophysiol.* **79**, 3272–3278 (1998).
125. Bullier, J. & Nowak, L. G. Parallel versus serial processing: new vistas on the distributed organization of the visual system. *Curr. Opin. Neurobiol.* **5**, 497–503 (1995).
126. Douglas, R. J. & Martin, K. A. Mapping the matrix: the ways of neocortex. *Neuron* **56**, 226–238 (2007).
127. Rockland, K. S. & Ichinohe, N. Some thoughts on cortical minicolumns. *Exp. Brain Res.* **158**, 265–277 (2004).
128. Molnár, Z. & Rockland, K. S. in *Neural Circuit and Cognitive Development* Ch. 5 (eds J. Rubenstein, P. Rakic, B. Chen & K. Y. Kwan) 103–126 (Academic, 2020).
129. Trojanowski, J. Q. & Jacobson, S. Medial pulvinar afferents to frontal eye fields in rhesus monkey demonstrated by horseradish peroxidase. *Brain Res.* **80**, 395–411 (1974).
130. Baizer, J. S., Desimone, R. & Ungerleider, L. G. Comparison of subcortical connections of inferior temporal and posterior parietal cortex in monkeys. *Vis. Neurosci.* **10**, 59–72 (1993).
131. Stanton, G. B., Goldberg, M. E. & Bruce, C. J. Frontal eye field efferents in the macaque monkey: I. Subcortical pathways and topography of striatal and thalamic terminal fields. *J. Comp. Neurol.* **271**, 473–492 (1988).
132. Lynch, J. C., Hoover, J. E. & Strick, P. L. Input to the primate frontal eye field from the substantia nigra, superior colliculus, and dentate nucleus demonstrated by transneuronal transport. *Exp. Brain Res.* **100**, 181–186 (1994).
133. Berman, R. A. & Wurtz, R. H. Exploring the pulvinar path to visual cortex. *Prog. Brain Res.* **171**, 467–473 (2008).
134. Huerta, M. F., Krubitzer, L. A. & Kaas, J. H. Frontal eye field as defined by intracortical microstimulation in squirrel monkeys, owl monkeys, and macaque monkeys: I. Subcortical connections. *J. Comp. Neurol.* **253**, 415–439 (1986).
135. Leichnetz, G. R., Smith, D. J. & Spencer, R. F. Cortical projections to the paramedian tegmental and basilar pons in the monkey. *J. Comp. Neurol.* **228**, 388–408 (1984).
136. Andersen, R. A., Asanuma, C., Essick, G. & Siegel, R. M. Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *J. Comp. Neurol.* **296**, 65–113 (1990).
137. Lynch, J. C., Graybiel, A. M. & Lobeck, L. J. The differential projection of two cytoarchitectonic subregions of the inferior parietal lobule of macaque upon the deep layers of the superior colliculus. *J. Comp. Neurol.* **235**, 241–254 (1985).
138. Schall, J. D., Morel, A., King, D. J. & Bullier, J. Topography of visual cortex connections with frontal eye field in macaque: convergence and segregation of processing streams. *J. Neurosci.* **15**, 4464–4487 (1995).
139. Vernet, M., Quentin, R., Chanes, L., Mitsumasa, A. & Valero-Cabre, A. Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations. *Front. Integr. Neurosci.* **8**, 66 (2014).
140. Liu, Y., Yttri, E. A. & Snyder, L. H. Intention and attention: different functional roles for LIPd and LIPv. *Nat. Neurosci.* **13**, 495–500 (2010).
141. Coe, B. C. & Munoz, D. P. Mechanisms of saccade suppression revealed in the anti-saccade task. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160192 (2017).
142. Milardi, D. et al. Red nucleus connectivity as revealed by constrained spherical deconvolution tractography. *Neurosci. Lett.* **626**, 68–73 (2016).
143. Na, J., Kakei, S. & Shinoda, Y. Cerebellar input to corticothalamic neurons in layers V and VI in the motor cortex. *Neurosci. Res.* **28**, 77–91 (1997).
144. Martinez-Gonzalez, C., Bolam, J. P. & Mena-Segovia, J. Topographical organization of the pedunculopontine nucleus. *Front. Neuroanat.* **5**, 22 (2011).
145. Sherman, S. M. & Guillery, R. W. Distinct functions for direct and transthalamic corticocortical connections. *J. Neurophysiol.* **106**, 1068–1077 (2011).
146. de Kock, C. P., Bruno, R. M., Spors, H. & Sakmann, B. Layer- and cell-type-specific suprathreshold stimulus representation in rat primary somatosensory cortex. *J. Physiol.* **581**, 139–154 (2007).
147. Masamizu, Y. et al. Two distinct layer-specific dynamics of cortical ensembles during learning of a motor task. *Nat. Neurosci.* **17**, 987–994 (2014).
148. Guo, K., Yamawaki, N., Svoboda, K. & Shepherd, G. M. G. Anterolateral motor cortex connects with a medial subdivision of ventromedial thalamus through cell type-specific circuits, forming an excitatory thalamo-cortico-thalamic loop via layer 1 apical tuft dendrites of layer 5b pyramidal tract type neurons. *J. Neurosci.* **38**, 8787–8797 (2018).
149. Bharioke, A. et al. General anesthesia globally synchronizes activity selectively in layer 5 cortical pyramidal neurons. *Neuron* **110**, 2024–2040.e10 (2022).
150. Larkum, M. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.* **36**, 141–151 (2013).
151. Brea, J., Gaal, A. T., Urbanczik, R. & Senn, W. Prospective coding by spiking neurons. *PLoS Comput. Biol.* **12**, e1005003 (2016).
152. Roelfsema, P. R. & Holtmaat, A. Control of synaptic plasticity in deep cortical networks. *Nat. Rev. Neurosci.* **19**, 166–180 (2018).
153. Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J. & Hinton, G. Backpropagation and the brain. *Nat. Rev. Neurosci.* **21**, 335–346 (2020).
154. Whittington, J. C. R. & Bogacz, R. Theories of error back-propagation in the brain. *Trends Cogn. Sci.* **23**, 235–250 (2019).

155. Xiong, Q., Znamenskiy, P. & Zador, A. M. Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* **521**, 348–351 (2015).
156. Cox, J. & Witten, I. B. Striatal circuits for reward learning and decision-making. *Nat. Rev. Neurosci.* **20**, 482–494 (2019).
157. Park, J. M. et al. Deep and superficial layers of the primary somatosensory cortex are critical for whisker-based texture discrimination in mice. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.08.12.245381> (2022).
158. Hong, Y. K., Lacefield, C. O., Rodgers, C. C. & Bruno, R. M. Sensation, movement and learning in the absence of barrel cortex. *Nature* **561**, 542–546 (2018).
159. Von Neumann, J. *The Computer and the Brain* (Yale Univ. Press, 1958).
160. Mo, C. & Sherman, S. M. A sensorimotor pathway via higher-order thalamus. *J. Neurosci.* **39**, 692–704 (2019).
161. Lake, B. M., Ullman, T. D., Tenenbaum, J. B. & Gershman, S. J. Building machines that learn and think like people. *Behav. Brain Sci.* **40**, e253 (2017).
162. Baroni, M. Linguistic generalization and compositionality in modern artificial neural networks. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190307 (2020).
163. Ruediger, S. & Scanziani, M. Learning speed and detection sensitivity controlled by distinct cortico-fugal neurons in visual cortex. *eLife* **9**, e59247 (2020).
164. Brooks, R. A. A robust layered control-system for a mobile robot. *IEEE T Robot Automat.* **2**, 14–23 (1986).
165. Brooks, R. A. New approaches to robotics. *Science* **253**, 1227–1232 (1991).
166. Haider, P., Ellenberger, B., Kriener, L., Jordan, J., Senn, W. & Petrovici, M. A. Latent equilibrium: a unified learning theory for arbitrarily fast computation with arbitrarily slow neurons. *Adv. Neural Inf. Process. Syst.* **34**, 17839–17851 (2021).
167. Narayanan, R. T. et al. Beyond columnar organization: cell type- and target layer-specific principles of horizontal axon projection patterns in rat vibrissa cortex. *Cereb. Cortex* **25**, 4450–4468 (2015).
168. Chen, G., Scherr, F. & Maass, W. A data-based large-scale model for primary visual cortex enables brain-like robust and versatile visual processing. *Sci. Adv.* **8**, eabq7592 (2022).
169. Guest, J. M., Bast, A., Narayanan, R. T. & Oberlaender, M. Thalamus gates active dendritic computations in cortex during sensory processing. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.10.21.465325> (2021).
170. Constantinople, C. M. & Bruno, R. M. Deep cortical layers are activated directly by thalamus. *Science* **340**, 1591–1594 (2013).
171. Pluta, S. et al. A direct translaminar inhibitory circuit tunes cortical output. *Nat. Neurosci.* **18**, 1631–1640 (2015).
172. Stuart, G., Spruston, N. & Häusser, M. *Dendrites* 3rd edn (Oxford Univ. Press, 2016).
173. Major, G., Larkum, M. E. & Schiller, J. Active properties of neocortical pyramidal neuron dendrites. *Annu. Rev. Neurosci.* **36**, 1–24 (2013).
174. Mikulasch, F. A., Rudelt, L., Wibrál, M. & Priesemann, V. Where is the error? Hierarchical predictive coding through dendritic error computation. *Trends Neurosci.* **46**, 45–59 (2022).
175. Richards, B. A. & Lillicrap, T. P. Dendritic solutions to the credit assignment problem. *Curr. Opin. Neurobiol.* **54**, 28–36 (2019).
176. Guerguiev, J., Lillicrap, T. P. & Richards, B. A. Towards deep learning with segregated dendrites. *eLife* **6**, e22901 (2017).
177. Hawkins, J. & Ahmad, S. Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Front. Neural Circuits* **10**, 23 (2016).
178. Schiess, M., Urbanczik, R. & Senn, W. Somato-dendritic synaptic plasticity and error-backpropagation in active dendrites. *PLoS Comput. Biol.* **12**, e1004638 (2016).
179. Poirazi, P. & Papoutsi, A. Illuminating dendritic function with computational models. *Nat. Rev. Neurosci.* **21**, 303–321 (2020).
180. Beniauev, D., Segev, I. & London, M. Single cortical neurons as deep artificial neural networks. *Neuron* **109**, 2727–2739.e3 (2021).
181. Cossell, L. et al. Functional organization of excitatory synaptic strength in primary visual cortex. *Nature* **518**, 399–403 (2015).
182. Seeman, S. C. et al. Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex. *eLife* **7**, e37349 (2018).
183. Garner, A. R. & Keller, G. B. A cortical circuit for audio-visual predictions. *Nat. Neurosci.* **25**, 98–105 (2022).
184. Ghazanfar, A. A. & Schroeder, C. E. Is neocortex essentially multisensory? *Trends Cogn. Sci.* **10**, 278–285 (2006).
185. Fetsch, C. R., DeAngelis, G. C. & Angelaki, D. E. Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nat. Rev. Neurosci.* **14**, 429–442 (2013).
186. Graybiel, A. M. The basal ganglia. *Curr. Biol.* **10**, R509–R511 (2000).
187. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381 (1986).
188. Parent, A. et al. Organization of the basal ganglia: the importance of axonal collateralization. *Trends Neurosci.* **23**, S20–S27 (2000).
189. Takakusaki, K., Saitoh, K., Harada, H. & Kashiwayanagi, M. Role of basal ganglia-brainstem pathways in the control of motor behaviors. *Neurosci. Res.* **50**, 137–151 (2004).
190. Graybiel, A. M., Aosaki, T., Flaherty, A. W. & Kimura, M. The basal ganglia and adaptive motor control. *Science* **265**, 1826–1831 (1994).
191. Roseberry, T. K. et al. Cell-type-specific control of brainstem locomotor circuits by basal ganglia. *Cell* **164**, 526–537 (2016).
192. Parent, M., Levesque, M. & Parent, A. Two types of projection neurons in the internal pallidum of primates: single-axon tracing and three-dimensional reconstruction. *J. Comp. Neurol.* **439**, 162–175 (2001).
193. Parent, M. & Parent, A. The pallidofugal motor fiber system in primates. *Parkinsonism Relat. Disord.* **10**, 203–211 (2004).
194. Pennartz, C. M., Groenewegen, H. J. & Lopes da Silva, F. H. The nucleus accumbens as a complex of functionally distinct neuronal ensembles: an integration of behavioural, electrophysiological and anatomical data. *Prog. Neurobiol.* **42**, 719–761 (1994).
195. Di Chiara, G., Porceddu, M. L., Morelli, M., Mulas, M. L. & Gessa, G. L. Evidence for a GABAergic projection from the substantia nigra to the ventromedial thalamus and to the superior colliculus of the rat. *Brain Res.* **176**, 273–284 (1979).
196. Williams, L. E. & Holtmaat, A. Higher-order thalamocortical inputs gate synaptic long-term potentiation via disinhibition. *Neuron* **101**, 91–102.e4 (2019).
197. Gambino, F. et al. Sensory-evoked LTP driven by dendritic plateau potentials in vivo. *Nature* **515**, 116–119 (2014).
198. Anastasiades, P. G., Collins, D. P. & Carter, A. G. Mediodorsal and ventromedial thalamus engage distinct L1 circuits in the prefrontal cortex. *Neuron* **109**, 314–330.e4 (2021).
199. Schmitt, L. I. et al. Thalamic amplification of cortical connectivity sustains attentional control. *Nature* **545**, 219–223 (2017).
200. Inagaki, H. K. et al. A midbrain–thalamus–cortex circuit reorganizes cortical dynamics to initiate movement. *Cell* **185**, 1065–1081.e23 (2022).
201. Wang, M. B. & Halassa, M. M. Thalamocortical contribution to flexible learning in neural systems. *Netw. Neurosci.* **6**, 980–997 (2022).
202. La Terra, D. et al. The role of higher-order thalamus during learning and correct performance in goal-directed behavior. *eLife* **11**, e77177 (2022).
203. Ruis, L., Andreas, J., Baroni, M., Bouchacourt, D. & Lake, B. M. A benchmark for systematic generalization in grounded language understanding. In *Proc. 34th Int. Conf. Neural Information Processing Systems* (eds. Larochelle, H. et al.) 19861–19872 (Curran, 2020).
204. Lake, B. M. & Baroni, M. Generalization without systematicity: on the compositional skills of sequence-to-sequence recurrent networks. In *Int. Conf. Machine Learning* (eds. Dy, J. & Krause, A.) 2879–2888 (2018).
205. Pfeiffer, J., Ruder, S., Vulic, I. & Ponti, E. M. Modular deep learning. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2302.11529> (2023).
206. Goyal, A. et al. Recurrent independent mechanisms. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.1909.10893> (2020).
207. Albright, T. D., Jessell, T. M., Kandel, E. R. & Posner, M. I. Neural science: a century of progress and the mysteries that remain. *Neuron* **25**, S1–S55 (2000).
208. Wallis, J. D., Anderson, K. C. & Miller, E. K. Single neurons in prefrontal cortex encode abstract rules. *Nature* **411**, 953–956 (2001).
209. Verschure, P. F., Pennartz, C. M. & Pezzulo, G. The why, what, where, when and how of goal-directed choice: neuronal and computational principles. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130483 (2014).
210. Dias, R., Robbins, T. W. & Roberts, A. C. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* **380**, 69–72 (1996).
211. Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267–279 (2014).
212. Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I. & Posner, M. I. The activation of attentional networks. *Neuroimage* **26**, 471–479 (2005).
213. Womelsdorf, T. & Everling, S. Long-range attention networks: circuit motifs underlying endogenously controlled stimulus selection. *Trends Neurosci.* **38**, 682–700 (2015).
214. Cohen, M. R. & Maunsell, J. H. Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* **12**, 1594–1600 (2009).
215. Reynolds, J. H. & Desimone, R. Interacting roles of attention and visual salience in V4. *Neuron* **37**, 853–863 (2003).
216. Poort, J. et al. The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* **75**, 143–156 (2012).
217. Reep, R. L. & Corwin, J. V. Posterior parietal cortex as part of a neural network for directed attention in rats. *Neurobiol. Learn. Mem.* **91**, 104–113 (2009).
218. Saalmann, Y. B., Pinsk, M. A., Wang, L., Li, X. & Kastner, S. The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* **337**, 753–756 (2012).
219. Rikhye, R. V., Gilra, A. & Halassa, M. M. Thalamic regulation of switching between cortical representations enables cognitive flexibility. *Nat. Neurosci.* **21**, 1753–1763 (2018).
220. Van der Werf, Y. D., Witter, M. P. & Groenewegen, H. J. The intralaminar and midline nuclei of the thalamus. Anatomical and functional evidence for participation in processes of arousal and awareness. *Brain Res. Brain Res. Rev.* **39**, 107–140 (2002).
221. Groenewegen, H. J. & Berendse, H. W. The specificity of the ‘nonspecific’ midline and intralaminar thalamic nuclei. *Trends Neurosci.* **17**, 52–57 (1994).
222. Breton-Provencher, V., Drummond, G. T., Feng, J., Li, Y. & Sur, M. Spatiotemporal dynamics of noradrenaline during learned behaviour. *Nature* **606**, 732–738 (2022).
223. Ren, J. et al. Anatomically defined and functionally distinct dorsal raphe serotonin sub-systems. *Cell* **175**, 472–487.e20 (2018).
224. Lohani, S. et al. Spatiotemporally heterogeneous coordination of cholinergic and neocortical activity. *Nat. Neurosci.* **25**, 1706–1713 (2022).
225. Morris, L. S. et al. Fronto-striatal organization: defining functional and microstructural substrates of behavioural flexibility. *Cortex* **74**, 118–133 (2016).
226. Apicella, P., Legallet, E., Nieoullon, A. & Trouche, E. Neglect of contralateral visual stimuli in monkeys with unilateral striatal dopamine depletion. *Behav. Brain Res.* **46**, 187–195 (1991).

Acknowledgements

M.S. discloses support for the research of this work from the Brain Science Foundation and the Sumitomo Foundation (2200084). C.M.A.P. discloses support for the research of this work from the European Union's Horizon 2020 Framework Program for Research and Innovation (Human Brain Project SGA3, 945539). J.A. discloses support for the research of this work from the European Social Fund through the 'ICT programme' measure and the Estonian Research Council grant (PSG728).

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature Reviews Neuroscience* thanks Andreas Burkhalter and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2023