The importance of being hierarchical
Nikola T Markov1,2,3 and Henry Kennedy1,2

Our understanding of cortical electrophysiology and anatomy at the single-cell level has led to the present day insight into the function of connections linking cortical areas. This made it possible to elaborate the cortical hierarchy in the early 1990s and was a prerequisite for the development of present day generative models of perception. These computational hierarchical models make strong predictions concerning the roles of feedforward (FF) and feedback (FB) pathways, including their segregation and topographical precision in both directions. This shows that instead of a single stream in the upper and lower compartments of the cortex there is in fact a bi-directional counter-stream in each compartment of the cortex. A significant advance in this field will require more detailed anatomy hand in hand with a network analysis of the directed and weighted cortical matrix.

Addresses
1 Stem Cell and Brain Research Institute, INSERM U846, 18 Avenue Doyen Lepine, 69500 Bron, France
2 Université de Lyon, Université Lyon I, 69003 Lyon, France

Corresponding author: Kennedy, Henry (Henry.Kennedy@inserm.fr)
3 Present address: Yale University, Department of Neurobiology, New Haven, CT 06520, USA.

Introduction
How structure relates to function has been at the forefront of research on the cortex since Hubel and Wiesel so successfully combined electrophysiology and anatomy in order to understand the physiology of the various morphological cell types and how this physiology differs across cortical layers and areas [1]. Relating these observations to the organization of connections between areas was an important step in the search for structure function regularities in the cortex. One such regularity was the repeated observation that in visual cortex, rostral directed connections stem from the supragranular layers and target the principal thalamic recipient layer, layer 4, while caudal directed connections stem from the infragranular layers and terminate outside of layer 4 [2,3]. The seminal paper of Rockland and Pandya tied these observations to the earlier observation that the rostral directed connections of area V1 are necessary for visual responsiveness [4,5], and were therefore to be thought of as feedforward (FF) driving projections. Because the infragranular layers house the feedback (FB) projections to the subcortical structures including the thalamus, the caudal directed connections stemming from infragranular layers were suggested to be analogous to cortico-thalamic FB projections and predicted to have a modulatory function. As we shall see these concepts of FF-driving and FB-modulatory have had a pervasive impact on theories of cortical function and are still very much with us to day.

The information processing concepts of Rockland and Pandya inspired Van Essen to undertake a meta study of some 305 pathways linking 32 areas. They showed it was possible to rank cortical areas thereby revealing the hierarchical organization of the macaque cortex [6,7]. The hierarchy was subsequently refined in order to better accommodate the what/ventral and where/dorsal streams of Ungerleider and Mishkin [8–10]. Most importantly, because the concept of cortical hierarchy is embedded in the single-cell anatomy and electrophysiology of the cortex it has led to present day generative models of brain function [11,12**,13]. Here we shall briefly summarize these theories, review present day shortcomings of the hierarchical concept and show how overcoming these limitations could lead to an important refinement of our understanding of the cortex. Before proceeding on the main course of this review we need to make a brief detour to consider an area of excitement in the theory of cortical function made possible by the anatomy studies mentioned above.

The lure of the small-world
The early hierarchical wiring diagrams and the data compilations that were required to construct them were powerful in focusing the understanding of the cortex as a network based on undirected graphs [14]. Graph theoretic approach had been very successful in elucidating the binary properties of complex networks, and in particular the so-called Small-World network architecture, the latter being thought to combine the short path length of random networks and high clustering of the regular lattice [15]. These properties of Small-World networks would optimize integration and segregation, complexity measures that could be particularly relevant for cortical function [14,16–18]. Modeling of the published anatomical data suggested that the cortical interareal network conformed to a Small-World network [19–25]. However these studies were carried out on collated data sets that had been partially extracted from the Felleman and Van Essen...
study, which acknowledged that not all connections had been tested at the time and therefore predicted that the density of the cortical network (i.e. the proportion of connections that exist with respect to the proportion that could exist) should be about 45%. For the purpose of constraining the connectivity in to a binary matrix these studies were assuming that the non-tested connections did not exist, the premise behind this approach is that the density was significantly lower than that anticipated by Felleman and Van Essen [7]. In a Small-World interareal architecture interactions between any two areas would be by means of unexpectedly short path lengths. However, a recent study has shown that in fact the density of the cortical matrix is 66% [26]. The small-world property, which chiefly implies that messages can be passed between any two nodes of a graph on very short paths is an independent property only in sparse networks [15]. For example in the social network, which is a very sparse network involving billions of people but only hundreds of edges connecting a single person (a density of $10^{-5.5}$%, or 0.00001%), it is unexpected to find paths of only six steps which could connect any two individuals. Hence, the famous six degrees of separation is a property emerging specifically from social behavior and the small-world property of the social network is a direct reflection of such behavior. However, in dense networks, such as the cortical matrix with a 66% link density, short paths between any two pairs of areas are expected, due uniquely to the abundance of links and therefore not an independent property of the cortical architecture.

The need for quantification

A quantitative approach to cortical connectivity allows for a better understanding of the fundamental parameters of the interareal network. Using retrograde tracers shows that 80% of the neurons projecting to a point on the cortical surface originate from within only one or two millimeters and comprise the local circuitry of the cortex [27]. The remaining 20% of neurons are distributed among the 25–80 areas that send projections to the injected cortical area, the majority being found in the few adjacent cortical areas and only a few percent are located in the large number of distant areas [26,27]. These considerations indicate that many of the connections in the Felleman and Van Essen (FVE) cortical hierarchy can be predicted from their weight characteristics to be unable to activate in isolation their target areas; the vast majority of long-distance cortical connections be they FF or FB must be acting in a concerted fashion (Figure 1b).

The FVE model ranks cortical areas on one of 11 levels using a pair-wise comparison of FF and FB connections between 32 visual areas [7]. Largely because of missing connections between levels the FVE model is indeterminate and there are 150,000 equally plausible solutions [28–30]. One solution to the indeterminacy problem is to use an index of hierarchical distance (SLN), which can be extracted from quantitative analysis of the laminar distribution of FF and FB projecting neurons [31,32]. This shows that short-distance FF connections stem from supragranular and infragranular layers, and the proportion of supragranular layers progressively increases with distance to become an SLN of 100% over very long distances (Figure 1) In the FB projections it is the proportion of infragranular layers that increases over long distances to give an SLN of 0%. Interestingly comparison of the SLN-based hierarchy with the FVE model gives essentially very similar results, with one exception concerning the frontal eye field (FEF) in the prefrontal cortex. In the

![Figure 1](image-url)

Functional wiring diagram of the macaque cortical hierarchy. (a) Injection in the middle area of a hierarchy reveals a hierarchical distance rule, long-distance feedforward (FF) originating uniquely from supragranular layers and progressively more infragranular layers are recruited on approaching the injection site, likewise long-distance feedback (FB) projections originate uniquely from the infragranular layers and more supragranular layers are recruited nearer the injection site. (b) The fraction of labeled supragranular layer neurons (SLN) provides an index of hierarchical distance, used here to construct a determinate model of cortical hierarchy [32]. Note the general similarity with the FVE model obtained with the exception of the frontal eye field (FEF) which has a FF relation to extrastrate visual areas of the dorsal and ventral streams. Box size indicates the 2-D dimensions of areas.
The SLN model the FEF has a FF and not a FB relationship to area V4 as it is the case in the FVE model [31,32,33]. The finding that area V4 and FEF exchange FF connections, instead of the classical FF reciprocated by a FB projection has a number of implications. Firstly, it argues against the hypothesis of Crick and Koch on the inter-diction of strong loops in the cortex [34], but more importantly, as we shall see in the next section, it questions the nature of the activity carried by FF and FB pathways.

The dogma of driving and modulating
The evidence for a driving role of FF and a modulatory action of FB pathways has been due to a large part obtained following attempts to inactivate specific pathways and recording the consequence on their target areas. Hence inactivation of area V1 was reported to silence numerous extrastriate areas including area V2 [35] reviewed in [36]. These results led to the conclusion that the V1 input to V2 is a driving input because it is necessary for V2 activation. Because visual response is known to persist in MT after inactivation of area V1, inactivation of V1 also addresses the role of FB projections; the absence of visual response in area V2 suggests that FB projections from area MT to area V2 were not sufficient for eliciting a visual response and that they are therefore modulatory. However, recently this experimental paradigm has been revisited with somewhat different results. After permanent long-term inactivation of V1 visual responses could be readily elicited form area V2 [37,38]. Although these later results are overall not surprising given the moderately strong input to area V2 from the lateral geniculate nucleus [27], they nevertheless pose a serious problem for earlier claims from V1 inactivation experiments on the driving influence to area V2 being uniquely due to area V1 inputs.

The consequences of inactivation of higher order areas on the visual response of early up stream areas are if anything even more difficult to interpret than the lower order areas. Inactivation of area MT was found to lead to an arrest of visual response in a number of area V2 neurons [39]. One interpretation of these results is that the FB projection of MT to area V2 had a driving function, however an alternative interpretation that was favored by the authors was that the cessation of response in V2 was due to the interruption of a FB pathway that modulated the response of the V2 neurons to the driving influence of FF projections from area V1. However, an earlier study suggests that there are grounds to support the idea of a FB driving input from area V2 to V1 [40]. Following inactivation of the lateral geniculate nucleus visual response in area V1 is restricted to superficial layers, however this visual response disappears after destruction of area V2 leading the authors to conclude that the FB projection from area V2 to V1 has a driving function. Moreover, recent in vitro investigation of the synaptic properties of interareal connections is compatible with FB pathways having a driving function [41,42].

The conduction speeds of FB axons are considerably faster than the modulatory cortico-thalamic projections and equal the conduction velocities of FF axons [43]. This suggests that the FB connections do not merely constitute a slow and diffuse pathway for modulation of early order areas. Further, the extensive body of evidence that there is an activation of early visual areas during visual imagery can be only understood if one assumes a driving function of FB [44,45]. These activations support reverse hierarchy theory, where vision at a glance is thought to depend on a fast FF mechanism involving higher order areas and engendering an initial conscious percept that allows high level categorization, whereas vision with scrutiny involves FB pathways and early visual areas [46].

Failing to disentangle the driver/modulator conundrum of cortico-cortical pathways one could be tempted to search for the control of the cortical hierarchy in the thalamus [47]. The pulvinar possesses a topographical representation of the cortical sheet, that allows for overlap of the regions contacting and being contacted by neighboring areas. There is a general observation that the pulvinar exchanges inputs with neighboring cortical areas that are directly inter-connected via cortico-cortical projections and recent findings show that these thalamic loops serve to drive alpha synchronization between nearby interconnected areas [48,49].

The promise of generative models
The classical view is that perception is principally achieved by the analysis at successive levels of the visual system of activity derived from the retina. The neuro-physiological correlates of this process are the progressively more complex receptive field properties that are observed in hierarchically organized areas [50]. It corresponds to a FF process because the FF connections are thought to be directly responsible for the complexification of receptive fields [51]. The FF theory of cortical function assumes that FB projections play a role principally in selective attention by means of biased competition [52–54].

Visual perception as an extraction of information from the retinal input has been questioned since at least the time of Bishop Berkeley in the 18th century who noted that the unconstrained nature of the projection of the visual world on the retina leads to ambiguity that can only be resolved by inference derived from prior knowledge. These ideas were further developed by Helmholtz’s unconscious inference and more recently by Richard Gregory’s perception as hypothesis.
An alternative view is that sensory input and prior knowledge of the world are both required in order to construct internal representations (Figure 2). David Mumford tackled this issue from a hierarchical perspective appealing to traditions in computer vision, pattern theory and psychology. In this approach a major role is attributed to FB projections between hierarchical levels where they serve to disambiguate and ‘explain away’ the earlier representation [11]. In a Bayesian framework of hierarchical cortical computation, rather than viewing FB as biasing competition as in the FF model, FB is conceived to play the altogether more ambitious role of biasing inference [55]. Each area is an expert for computing specific aspects of the sensory data, and its inference is constrained by its FF and FB inputs. The FF to an area drives the generation of the hypothesis and is constrained by FB input derived from priors, which in generative models can be assimilated to stored high-level representations of the world. The issue of how the brain constructs the priors is resolved in empirical Bayes, a statistical process where priors are estimated from data by means of a hierarchical generative model [56]. Implementing hierarchical Bayesian inference using particle filtering and belief propagation means that each area develops multiple hypothesis and is not influenced uniquely or mainly by its direct neighbors, which as we see is important with regards the detailed structure of the cortical hierarchy. This allows numerous areas to work together through multiple interactive and concurrent cycles to allow hypothesis selection. In this viewpoint area V1 is not so much the head of the river, but rather a high-resolution buffer where the cascading FB priors are echoed in the complexity of the long-latency responses of its neurons [55]. The Bayesian framework includes the notion of predictive coding where in a hierarchical system it is the unexpected features that are signaled to the next stage and FB mechanisms are not limited to learning the FF input but integrate the generative FB projections into the perceptual inference process [13]. Rao and Ballard used this approach to model extra-classical receptive field effects. In their model the FF signals into a cortical area constitute prediction errors and the FB signals are the predictions, so that the conditional expectation units in an area integrate the excitatory ascending prediction error and the inhibitory descending prediction; when the prediction is fulfilled silence reigns. Recently Karl Friston has developed an extension of the Bayesian hierarchical computation to provide a global theory of brain function based on principals of homeostasis, where interactions with the environment serve to minimize free-energy, an information theory measure related to surprise. This not only sees perception but also memory, attention, value, reinforcement and salience as active exchanges with the environment and as such will subscribe to the minimization of free-energy principle [12**,57]. Applied to the cortex, when the descending prediction is fulfilled and free-energy is minimum, there is no surprise; silence again.

Adaptive changes in the brain that minimize free-energy in order to recapitulate the environmental causes of sensory input also take place over evolutionary time [58]. This leads to the expectation that the hierarchical organization of the cortex (i.e. the number of areas, their shape, sizes and interconnectivity) would mirror the hierarchical causal structure of the environment [59].

**Generative models and cortical counter-streams**

Theoreticians of the generative models have the merit of giving very detailed consideration of cortical anatomy. Ascending (FF) pathways carrying predictions are thought to originate from the supragranular layers and target layer 4 of higher order areas, while descending pathways reporting prediction errors (or free-energy) originate in the infragranular layers and avoid layer 4 in there upstream targets. An essential constraint that the generative models will have to fulfill is that the FB prediction and the FF prediction error have to remain highly segregated, and combining these signals before their interaction with the evaluation units of the individual areas would play havoc...
Counter-stream organization of FF and FB pathways. (a) Current theory postulates that there are two categories of connections linking cortical areas, FF connections linking lower to higher areas (blue in A), and FB linking higher to lower (red). However, the anatomy shows that there is a FF and FB in both the supragranular and infragranular layers. Given the recent demonstration of different oscillatory coherence in these two compartments the upper and lower counter-streams could have different roles. (b, c) Common physiological properties indicated by cell shape. Physiological property is sublayer specific (b), if layer 3B is a driving type, then there is a small contingent of driving neurons located in layer 3A of level 2 projecting to level 1. Physiological property is not layer specific (c); here the contingent of FB in layer 3B is predicted to have driving characteristics.

The importance of being hierarchical

with the generation of percepts. Relegating these two streams to the upper and lower pathways would seem to be one way of assuring their segregation. However, such a configuration does not take full account of the known laminar organization of FF and FB pathways (Figure 3). The cumulative work of numerous anatomists, notably Rockland and the Tiges, suggests that there is a clearly defined FB pathway in layer 3A, radially separated from the FF pathway in layer 3B. Likewise in the infragranular layer there is a FF pathway centered on layer 5, partially distinct from the classical FB pathway centered in layer 6. Bifurcation of axons is ubiquitous in the cortex, and it is rare for adjacent populations of cortical neurons not to send collaterals to each other’s targets [60]. The question then is, given the close proximity of these pathways are they in fact segregated, as the generative hierarchical computational models would require? This has been investigated using a double-tracing paradigm with simultaneous tracing of FF and FB projecting neurons and shows that indeed the two populations are segregated and do not have common targets [61**]. This high degree of segregation is accompanied by the fact that the FF and FB supragranular pathways both share similar levels of high topographical precision that therefore distinguishes them from the diffuse FF and FB pathways of the infragranular layers. The two compartments housing the two counter-streams exhibit distinct oscillation features (gamma coherence in the supragranular layers and alpha coherence in the infragranular layers) [62**]. These findings could be important as the basis of improved interareal communication through coherence [63].

The counter-stream organization shown in Figure 3 was anticipated by Ullman and in detail it fulfills a number of features that are required by existing generative models including topographical precision in FF and FB streams and complete segregation of these streams [64]. The two FF streams are found in close proximity to the principal input layer to the cortex, layer 4. The outermost FB layer is in close proximity to layer 2, which we have found does not project out of the cortex [61**]. Layer 2 has been described as the external granular layer [65], suggesting that each of the two streams of the supragranular layers is closely associated with its particular granular layer, the FF layer 3B with layer 4 and the FB in layer 3A with layer 2. Interestingly two recent electrophysiology studies exploring layer 2 report neuron response that would be highly appropriate for a FB granular layer; compared to lower layers they had larger receptive fields, were less restricted in their selectivity and exhibited a high incidence of combined feature selectivity [66,67].

Building hierarchy

Given that hierarchy is such a strong principle of areal organization and function, what are the principles that govern its emergence during the development of the
individual? Work in the fetal nonhuman primate suggests that the first-formed interareal pathways are FF projections, where supragranular axons exhibit rapid directed growth and target selection forming highly segregated pathways [68]. During the early phase of rapid axonal growth there is an important increase in the density of the projections of supragranular neurons, followed by a protracted period remodeling of FB connections via axon elimination that extends well into the postnatal period [69]. Importantly however, the hierarchical layout of the cortex is overall similar in the immature and adult cortex despite these important differences in the tempo of FF and FB pathway formation [70].

The extensive prolongation in the maturation of the FB pathways observed in the nonhuman primate appears to be a universal feature and has been observed in the development of rodent and human cortex [71,72,73]. The late maturation of the axonal trajectories of FB projections is accompanied by late maturation of inhibitory responses which is specific to the FB pathways [73]. These developmental processes have important consequences for normal development. A number of studies have implicated FB connections in figure-ground segmentation [74] possibly via push–pull effects [75]. During development the psychophysical figure-ground response matures late in childhood in agreement with the late development of the FB pathways [76].

**Perspectives**

There is a present need to take on board the complexity of the cortical hierarchy. Most theoretical studies of hierarchy consider the existence of just two types of interareal connections: FB and FF. Here the definition of the FB pathway is that it goes from a higher to a lower area and the reverse for FF (Figure 3a). However, this binary classification ignores the different laminar origins of these pathways, which are, as we have seen, highly distinct. This is all the more curious given that these differences in laminar origin are what enable us to define the hierarchical ranking in the first place. Laminar considerations show that there are at least two sets of FB and FF pathways (i.e. a supragranular FF and FB and an infragranular FF and FB) with as we have shown very different topological features. We can go further, each of these pathways has weak reverse pathways, so for instance a small contingent of neurons in layer 3A (a FF sublayer) actually projects back to lower order areas, so that we have four and not two sets of FF and FB pathways (Figure 3b,c). Taking into consideration these different instances of FF and FB pathways will only make sense if we have some understanding of their individual anatomy and physiology. At the moment we do not, but the very fact that SLN gives a measure of hierarchical distance (Figure 1) indicates that the relative proportions of these pathways in some way define a unique contribution from each of the pathways.

Future progress in our understanding of the structural underpinnings of hierarchy points to two directions. Firstly there is an urgent need for more refined tract tracing data, involving better understanding of large-scale models of the cortex with single-cell resolution, and in particular investigation at the laminar level. This will necessarily involve molecular characterization of neurons and a concerted effort to span different scales in the enquiry. Secondly, there is a need to use these data to develop a graphical approach for the analysis of weighted, directed matrices. This will make it possible to establish a network level analysis of cortical hierarchies and to better constrain the generative models they support. While the modeling of this data will provide novel and exciting insights with a strong appeal to a large community, the detailed anatomy that will in fact constitute the data will be a highly specialized but nevertheless necessary endeavor.

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


An important development of hierarchical theory of brain function.

The importance of being hierarchical
Markov and Kennedy


8 Macrocircuits

Shows that there are topographically organized feedforward and feedback pathways in the supragranular layers that confirm to a counter-stream architecture. These counter-streams will allow fine detailed information to travel both up and down the cortical hierarchy.
Reports differences of the oscillatory coherence in the supragranular and infragranular layers which will result in important consequences on the counter-streams housed in these compartments.
Shows that in rodent contrary to primate, feedforward and feedback neurons are distributed in a salt and pepper fashion, but like in monkey few cortico-cortical cells have axon collaterals that ascend and descend the cortical hierarchy.