Cortical Hierarchy, Dual Counterstream Architecture and The Importance of Top-Down Generative Networks

Julien Vezoli[a§], Loïc Magrou[b§], Rainer Goebel[c], Xiao-Jing Wang[d], Kenneth Knoblauch[b], Martin Vinck[a#], Henry Kennedy[b,e#]

[a] Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, 60528 Frankfurt, Germany

[b] Univ Lyon, Université Claude Bernard Lyon 1, Inserm, Stem Cell and Brain Research Institute U1208, 69500 Bron, France.

[c] Faculty of Psychology and Neuroscience, Department of Cognitive Neuroscience, Maastricht University, P.O. Box 616, 6200 MD, Maastricht, The Netherlands

[d] Center for Neural Science, New York University (NYU), New York, NY 10003, USA.

[e] Institute of Neuroscience, State Key Laboratory of Neuroscience, Chinese Academy of Sciences (CAS) Key Laboratory of Primate Neurobiology, CAS, Shanghai 200031, China

§ equal contributing

corresponding authors

Abstract

Hierarchy is a major organizational principle of the cortex and underscores modern computational theories of cortical function. The local microcircuit amplifies long-distance inter-areal input, which show distance-dependent changes in their laminar profiles. Statistical modeling of these changes in laminar profiles demonstrates that inputs from multiple hierarchical levels to their target areas show remarkable consistency, allowing the construction of a cortical hierarchy based on a principle of hierarchical distance. The statistical modeling that is applied to structure can also be applied to laminar differences in the oscillatory coherence between areas thereby determining a functional hierarchy of the cortex. Close examination of the anatomy of inter-areal connectivity reveals a dual counterstream architecture with well-defined distance-dependent feedback and feedforward pathways in both the supra- and infragranular layers, suggesting a multiplicity of feedback pathways with well-defined functional properties. These findings are consistent with feedback connections providing a generative network involved in a wide range of cognitive functions. A dynamical model

constrained by connectivity data shed insights into the experimentally observed signatures of frequency-dependent Granger causality for feedforward versus feedback signaling. Concerted experiments capitalizing on recent technical advances and combining tract-tracing, high-resolution fMRI, optogenetics and mathematical modeling hold the promise of a much improved understanding of lamina-constrained mechanisms of neural computation and cognition. However, because inter-areal interactions involve cortical layers that have been the target of important evolutionary changes in the primate lineage, these investigations will need to include human and non-human primates comparisons.

Keywords

Non-human primate, human, brain, electrophysiology, anatomy, modeling, connectivity, predictive coding, perception, consciousness

Plan

1. Introduction
2. Hierarchy – signatures of inputs to the local circuits.
3. Models of hierarchy
4. Hierarchy – input consistency
5. Dual stream architecture.
6. Functional characteristics of FF and FB pathways.
7. The predictive brain and the importance of top-down generative networks.
8. Conclusion.

1. Introduction

Hierarchy as an organizational feature of the brain has been a recurrent theme since the evolutionary theory of neurological disorders of John Hughlings Jackson (1835-1911), following his observations of positive and negative symptoms in his patients (York and Steinberg, 2011). The neurobiology of cortical hierarchy was explored by the pioneering work of David Hubel and Torsten Weisel when they characterized the receptive field properties of simple, complex and hypercomplex neurons across areas of the visual cortex (Hubel and Wiesel, 1962). Following the work of Rockland and Pandya (1979) a myriad of connectivity studies in the cortex found additional evidence of hierarchical organization, allowing Felleman and Van Essen to propose the first hierarchical model of the cortex (Felleman and Van Essen, 1991), thereby providing a framework for modern concepts of feedforward (FF) and feedback (FB) processes. The notion of hierarchy has become considerably generalized and for example can be considered to be the major concept linking biological and artificial intelligence (Hawkins and Blakeslee, 2004). Convolutional deep neural networks have a clear hierarchical organization, with convergent, FF connections passing information from lower to higher layers, and divergent FB connections shaping plasticity in the connections from lower layers (LeCun et al., 2015). But what exactly is the significance of hierarchy in the brain? Hierarchy has been extensively studied in terms of ranking of cortical areas with respect to a number of criteria including, gradients of structural and functional features, as a progression of scales or as a topological sequence of projections (Hilgetag and Goulas, 2020). Here we take a diametrically opposing view. Rather than simply considering hierarchy as a ranking of cortical areas, we address what it means in terms of monosynaptic inter-areal connectivity. In line with the tenet that the explanation of how the brain works demands an account of what neurons do, and that functional interactions of cortical areas is assured by neuronal activity relayed between areas by axons, we confine our discussion of hierarchy to the description of the neuronal properties of inter-areal relations. We propose that the structural and functional markers of hierarchy define the integration of long-range inputs into each local circuit. Future investigation of these markers are expected to provide insight to the cellular mechanisms underlying hierarchical processing. A critical aspect of these regularities concerns the spatial distances governing interactions between cortical neurons, which we and others have shown obey metabolic constraints in terms of wire minimization underlining the spatial embedding of the cortex (Markov et al., 2013). We discuss the functional implications of the Dual Stream Architecture (Markov et al., 2014b) in terms of FF and FB processing and where future investigations are expected to provide insight into the cellular mechanisms underlying hierarchical processing.

Finally, we cast FF and FB relations in terms of predictive processing theory and evoke the importance of top down generative networks.

2. Hierarchy – signatures of inputs to local circuits.

In 1989 Douglas, Martin and Whitteridge published a landmark study that proposed a canonical microcircuit for the neocortex (Douglas et al., 1989) (**Figure 1 and 10b**). A common misconception of the canonical microcircuit is that it constitutes solely a description of the inter-laminar wiring patterns of the cortex. In fact, it is much more a theory that sets out to explain the electrophysiological properties of the cortex in terms of local connectivity, in particular with respect to within-laminar connections. In an effort to nail down the transformation of the thalamic input, *in vivo* intracellular recordings were made in area V1 in cat cortex. This showed that minute inputs from the LGN are amplified by recurrent excitation in layer 4 neurons (Latawiec et al., 2000). Subsequent quantification of the synaptic components of the local circuit showed that high levels of within-layer recurrent excitation is a characteristic feature of the local circuit (**Figure 1**) (Binzegger et al., 2009). These experiments showed that the role of inhibition was not to carve out the selectivity of the neuron response but rather to exert a control over the amplitude of the response and therefore to maximize the inhibitory potentials in the optimal receptive field response (Douglas et al., 1995; Douglas et al., 1989). Subsequent work showed that there is a weak orientation response in the input to the cortex, meaning that the primary role of the recurrent excitation is the amplification of a signal and not its creation (Ferster et al., 1996).

For many years research on cortex was predominantly in carnivores and non-human primates, leading to the notion of the cortical column as a fundamental component of functional organization (Mountcastle, 1995). In these studies, electrophysiological recordings from electrode penetrations made perpendicular to the cortex found a conserved function in the depth of the cortex in passing from pia to white matter (Hubel and Wiesel, 1962; Mountcastle, 1957). In the visual system there were expectations that the columnar organization of the cortex would be both functional and structural, since physiologically demonstrated ocular-dominance columns appeared to co-localize with cortical territories labeled by transynaptic labeling following tracer injections in the eye (Hubel and Wiesel, 1977). However, close examination revealed important discrepancies in such a correspondence (da Costa and Martin, 2010)), suggesting that the link between cortical structure and function is to be found at a much finer scale. Thanks to work in the mouse visual cortex using the molecular tools that are available in this species, it has been confirmed that cortical responses to thalamic input are indeed the

consequence of an amplification (Harris and Mrsic-Flogel, 2013) (Lien and Scanziani, 2013) via the local recurrent excitation (Cossell et al., 2015; Douglas et al., 1995; Ko et al., 2011). These studies built on earlier findings of highly nonrandom features of synaptic connectivity in local cortical circuits, proposing that there is a skeleton of strong connections in a sea of weak connections (Song et al., 2005). Later it was shown that the rare strong connections in the lognormal distribution of synaptic weights are between neurons with similar receptive fields, meaning that neurons in the visual cortex listen most closely to a subset of their synaptic inputs (Cossell et al., 2015). These findings explain earlier observations showing that ongoing activity of a neuron (so called spontaneous activity) reflects the functional architecture (i.e. the backbone of strong connections) in which it is embedded (Tsodyks et al., 1999). The emerging picture is that layers 4, 3 and 2 neurons are organized into subnetworks so that the selectivity of the amplification is ensured by constraints at the scale of dendritic spines (Lee et al., 2016).

The principal wiring property of the canonical circuit is the recurrent excitation that is observed in all of the cortical layers including layer 4 (Binzegger et al., 2004). The relevance of the canonical microcircuit theory for understanding inter-areal processing became apparent when cortical connectivity was quantified. In fact, 80-90% of the connections of the cortex are in the local circuit spanning 1-2mm in the visual cortex (Markov et al., 2011). Except for the adjacent cortical area, the structural weight of the average input from a distant source area to a target area is several orders of magnitude *less* than the thalamic input (Markov et al., 2014a). These observations lead to the notion that amplification by local recurrent excitation is a general phenomenon, that allows selection and recombination of relatively small afferent signals (Douglas and Martin, 2007a, b). For instance, top-down signaling of selective attention multiplicatively modulates sets of sensory neurons (McAdams and Maunsell, 1999; Treue and Maunsell, 1996). In this manner, selective amplification by local circuit dynamics leads to all-or-none task switching (Ardid and Wang, 2013).

Early anatomists, working principally in non-human primates, distinguished between rostral directed connections that originate chiefly in the supragranular layers and terminate in layer 4 (Cragg, 1969; Kennedy and Bullier, 1985; Lund et al., 1975; Martinez-Millan and Hollander, 1975; Rockland and Pandya, 1979; Spatz et al., 1970; Van Essen and Zeki, 1978; Wong-Riley, 1978) and caudal directed connections that mostly originate from infragranular layers and terminate outside of layer 4 (Kaas and Lin, 1977; Kennedy and Bullier, 1985; Kuypers et al., 1965; Tigges et al., 1973; Wong-Riley, 1978). In a landmark study, Rockland and Pandya (Rockland and Pandya, 1979) were the first to formulate inter-areal connectivity in terms of hierarchy and suggested that the laminar organization of cortical afferents and their

terminals indicates the sequence of information processing in cortex. These authors proposed that connections originating from supragranular layers and terminating in layer 4 by analogy with the main thalamic input to cortex constitute the FF pathway channeling sensory information to cortical areas carrying out higher-order analyses. By contrast connections arising from the infragranular layers, by analogy with descending projections to subcortical structures, correspond to FB connections and were postulated to enable higher order areas to be able to modulate the activity of lower level areas (Rockland and Pandya, 1979).

3. Models of Hierarchy.

The classification of pathways between areas as FF and FB helped motivate the investigation of the role of the cortical network in terms of FF pathways shaping receptive fields in their target areas (Hubel, 1995) and FB pathways relaying contextual information (Gilbert and Li, 2013; Zipser et al., 1996). How the cortical network related to the mosaic of cortical areas was given substance by Felleman and Van Essen's demonstration that the layout of cortical areas corresponded to a distributed hierarchy (Felleman and Van Essen, 1991). In their seminal study these authors established a definition of FF and FB connections largely employing the criteria of Rockland and Pandya (1979), and although principally based on laminar patterns of anterograde labeling they were able to stipulate criteria so as to include retrograde labeling therefore enabling them to define pathways with respect to findings reported in a large number of publications (**Figure 2A**). Pairwise comparisons of the connections linking areas using these criteria revealed a high regularity; connections that were classified as FF were largely reciprocated by FB connections, allowing the authors to establish a distributed hierarchy across multiple streams in the macaque visual cortex shown in **Figure 2B**. Because of the many parallel pathways and given that hierarchical levels were defined arbitrarily, it meant that the precise ordering of cortical areas was ill-defined. Computational modeling showed that there were over 150,000 equally plausible solutions to the Felleman and Van Essen Model (Hilgetag et al., 1996) (**Figure 2C**).

A solution to the indeterminacy of the Felleman and Van Essen model could be overcome by an objective localization of hierarchical level. A suggestion that this might be the case was the observation that injections of retrograde tracers in a target area in one of the early visual areas generated a progressive decrease in the numbers of labeled FB neurons in supragranular layers in source areas with increasing physical rostral distance (**Figures 3A**) (Kennedy and Bullier, 1985). Quantitative measures of interareal connectivity showed that patterns of retrograde labeling were highly consistent across different brains, provided that

194 labeled neurons are summed across the full extent of a projection zone, defined as the region in
195 a particular source area which contains projections to an injected target area (**Figure 3B,C**).
196 Subsequently injections in cortical areas at higher hierarchical levels generated a progressive
197 increase in the numbers of labeled FB neurons in supragranular layers with increasing physical
198 caudal distance. In this manner FF and FB pathways exhibited opposing gradients of projection
199 neurons (Barone et al., 2000; Kennedy and Bullier, 1985; Markov et al., 2014b). These
200 observations led to the definition of an index of this gradient based on the proportion of
201 Supragranular Labelled Neurons or SLN (Barone et al., 2000; Vezoli et al., 2004). Because
202 these changes are highly consistent across brains, the smooth gradients of inputs from neurons
203 in different layers and areas to a target area lead to the derivation of a hierarchical distance rule
204 (**Figure 3B**).

205 The transition from a binary model of hierarchy to one based on hierarchical distance
206 had important consequences. One way of thinking about these connectivity gradients (**Figure**
207 **3A**) is that they represent gradual changes in the composition of inputs to the local microcircuit
208 of a cortical area that is dependent on physical distance. Elsewhere we have shown that input
209 strength to cortical areas declines exponential with distance (Ercsey-Ravasz et al., 2013). In
210 terms of hierarchy, the strong nearby connections are near-lateral and with increasing distance
211 connections are progressively more FB or FF in nature.

212 If these changing inputs to the local microcircuit represent the structural signature of
213 hierarchy it is legitimate to wonder if they have a functional correlate? If this is the case, then
214 Pascal Fries and his team reasoned that one can derive a functional hierarchy (Bastos et al.,
215 2015b). The hierarchical distance rule is based on the fact that supragranular layers primarily
216 send FF projections and infragrananular layers FB projections. In the visual system, superficial
217 and deep layers are characterized by relatively strong gamma and alpha/beta oscillations,
218 respectively (Buffalo et al., 2011). Furthermore, whereas in early visual areas, gamma
219 oscillations are relatively strong (Gray et al., 1989), beta oscillations tend to be strong in higher
220 areas like those in parietal cortex (Brovelli et al., 2004; Scherberger et al., 2005). These
221 observations lead to the prediction that in the visuo-parietal system interareal synchronization
222 in the gamma frequency band mediates FF and interareal synchronization in the alpha- and
223 beta-frequency band mediate FB influences. Granger-causality was used to quantify functional
224 FF and FB interactions, thereby allowing the strength of these interactions to be quantified for
225 individual frequency bands. Neuronal populations show spontaneous fluctuations over time that
226 are driven by brain rhythms in different frequency bands, such as theta (3-8Hz), alpha (8-12Hz),
227 beta (12-30Hz) and gamma (30-80Hz). Note, neocortical rhythms do not comprise band-limited

sinusoidal oscillation but typically describes noisy oscillations with energy in a wide frequency range (Burns et al., 2011; Spyropoulos et al., 2020). Frequency-specific directed influences of rhythmic synchronization are correlated with hierarchical distance, FF pathways are characterized by synchronization in the theta (4 Hz) and gamma-band (60-80 Hz) and FB in the beta-band (14-18 Hz) (Bastos et al., 2015b). These observations mean that the structural signatures to the microcircuit are indeed paralleled by functional signatures (Bastos et al., 2015b). However, whereas the structural hierarchy is fixed theses authors were able to show that activity patterns underlying functional hierarchy exhibit task dependent dynamics.

Using empirical connectivity data as a structural substrate, Mejias et al. (Mejias et al., 2016) built a dynamical model of multi-regional macaque monkey cortex endowed with a laminar structure. The model captures stochastic fast oscillations in the gamma frequency range in the superficial layers, and lower-frequency (alpha) oscillations in the deep layers. Importantly, in contrast to regular oscillators, such network rhythms are quite irregular and noisy, compatible with the notion of sparsely synchronous brain rhythms (Wang, 2010), which provides a unifying framework for explaining both population oscillations and Poisson-like highly variable spike trains of single neurons. The model accounts for the distinct signatures of frequency-dependent Granger causality that reflect FF versus FB signaling, and reproduces the experimentally deduced neural population dynamics that are consistent with the anatomically defined hierarchy. Therefore, this model provides a computational platform for theoretical investigations of the interplay between feedforward and feedback processes.

Both structural and functional hierarchies show that the regularities stemming from laminar distributions of connections and the signals they confer are remarkably consistent. In the following section, we address the extent of this consistency in order to formalize how hierarchy inputs to an area are shaped by distance.

4. Hierarchy – Input Consistency.

The notion of hierarchy implies order or rank, so that a prerequisite to determining if there are hierarchical relations between cortical areas requires determining if order relations can be defined between them. For example, the Felleman and Van Essen hierarchy was based on the binary classification of FB/FF relations between areas defined by laminar origin and termination of projections (Felleman and Van Essen, 1991). A FF projection from area A to B implied that B was ordered after A. Similarly, a FB projection from B to A would also be consistent with the above FF relation in assigning B after A. While in a hierarchy we would expect the two criteria to agree that might not in fact be the case. On a simple level,

disagreement could be taken to define equality of ordinal levels in the sense that equality is defined as those numbers, A and B, that simultaneously satisfy $A \geq B$ and $A \leq B$. Alternatively, distinct hierarchical orders might arise: one on the basis of FF laminar projections, the other on the basis of FB. These observations become important when the data supporting multiple laminar FB/FF pathways are analyzed.

The criteria for determining hierarchical rank described above are based on the properties of projections that define relations between areas. Alternatively, one can consider properties that are intrinsic to an area, such as cortical thickness, neuron density, soma size, spine count, spine density and dendritic tree size. These properties have been shown to display a smooth variation across the cortex that allows ranking of areas in accordance with the gradation principle of Sanides (Barbas, 2015; Sanides, 1972). Because these properties vary across the cortex, a hierarchical ranking can be established simply by ordering the areas according to the property. This distinction leads us to consider that criteria for building cortical hierarchies can be divided into two broad classes that we shall refer to as *node-based* and *link-based* (**Figure 4A**).

Here it is useful to draw an analogy with social networks. A hierarchy in a social network implies that the actors show specific kinds of interactions with each other (*link-based*). Hierarchy implies that those close to each other in a hierarchy show stronger interactions with each other than actors that are distant in the hierarchy. More information can be gauged from the nature of the interactions: We expect that someone high in the hierarchy (a boss) will show a more directive communication pattern to workers lower in the hierarchy. The workers, in turn, will show a different ascending communication pattern, e.g. asking more questions. Thus, a hierarchy can be constructed by studying the way in which people interact with each other, and knowing a hierarchy could in principal allow us to predict these interactions. By analogy, the SLN can be seen as a measure that directly quantifies the nature of the interactions between brain areas, based on the laminar projection patterns. Interestingly and as we described above, these laminar projection patterns also relate to certain functional interaction patterns (FF gamma and FB alpha/beta influences). In addition, social hierarchy might also be gauged from properties of the people in the hierarchy themselves. For instance, one expects the boss of the company to have the largest office, whereas the workers at the bottom to share an office and have smaller desks (*node-based*). In some sense, one could argue however, that the *node-based* view is based only on indirect markers and is ultimately grounded in the interaction *link-based* view.

There are critically important differences for constructing hierarchies between node and link-based information. By definition, node-based relations depend only on a value intrinsic to the node, not the relations between nodes so they give no information on the symmetry or otherwise of inter-areal relations. By contrast, ranks based on links are expected to show reciprocity, so that if there is FF pattern from area A to area B, a FB pattern is expected from area B to area A. Node based criteria are defined between *any* two areas independently of whether or not a connection is actually present. Link-based criteria can provide information on asymmetric relations, provided they are directional and are strictly defined between areas *only* when there is a direct connection. Nevertheless, hierarchical ordering between unconnected areas can be achieved through indirect connections. Generally, link-based hierarchy describes the connections that are carrying information between areas and therefore the manner in which the connections and activity from source areas are integrated into the local circuit of the target area.

In order to define a hierarchical distance scale, i.e., that is not just ordinal, a distance function, d has to be defined. This function d should transform whatever anatomical or physiological measure one is using into a consistent measure of hierarchical distance across cortical areas. For example, **Figure 4B** shows a hypothetical distance scale on which 4 areas, A, B, C, D, are arranged. Suppose that hierarchical distances are estimated based on measures derived from tracer injections in areas A and B. The injection in area A provides information about hierarchical distances to areas B, C and D and the injection in B to areas A, C and D. A consistent measure of hierarchical distance, d, would generate the same estimate of distance, d_{AB}, between areas A and B or, formally, we would expect that $d_{CA} - d_{CB} = d_{DA} - d_{DB}$. This is easily derived from the two equations in **Figure 4B** that show for two areas, C and D, the expected relation between the hierarchical distances of a common area to C or D to two areas sites, A and B. For common projections X, plotting d_{XA} against d_{XB} should fall along a line of unit slope with intercept d_{AB}. The question is how to define the distance function d.

In contrast to the binary measure of hierarchy in the Felleman and Van Essen model, SLN is a continuous measure on the interval (0, 1), thus providing a measure of the degree of feedforwardness/feedbackness. A binary description treats a projection as FF (FB) if its SLN is greater (less) than 0.5. Using simply the SLN differences as a hierarchical distance measure, the Barone et al., 2000 study was able to reproduce nearly all of the features of the Felleman and Van Essen model based on the SLN values from injections in just two areas, V1 and V4. A notable exception is that the Barone et al. 2000 study placed the frontal eye field (FEF), this prefrontal area, *below* the early extrastriate visual area V4. The SLN value from FEF to V4 was

above 0.7 which placed V4 at higher hierarchical levels. Subsequent physiological studies confirmed an FEF role in early visual processing (Moore and Armstrong, 2003; Schall, 2015), thus justifying its relatively low hierarchical level. The unusual FF pattern for such a caudally directed projection was further confirmed in other studies (Pouget et al., 2009) and pertains to a specific link-category on which we expand later.

While differences in SLN establish a determinate hierarchical distance measure between areas, the measure is not necessarily consistent in the manner described in **Figure 4B**. As the measure is defined on the interval (0, 1), SLN differences for two areas projecting to a third area could be quite different from those to another more distant area. An ideal measure would project the interval (0, 1) to a scale where differences remain linear. This is commonly accomplished in statistical models, such as generalized linear models (GLM), by means of a logit or probit transformation (**Figure 4C**) that map the unit interval onto the real line. As the figure demonstrates, with the proper scaling both of these transformations yield rather similar mappings.

Figure 5 shows a set of scatter plots for SLN values of common projections for all area pairs between each of 11 visual areas injected with retrograde tracer (Markov et al., 2014b). The SLN values are plotted on probit transformed axes. For many of the area pairs, the point distributions are reasonably well described by a line of unit slope (dashed blue in each graph), as predicted by a consistent measure of distance, i.e., $g(SLN_j) = g(SLN_i) + c$, where c is a constant. Given the similarity of the transforms, it is not surprising that the logit transformation yields virtually the same linear patterns between area pairs. Thus, this indicates that the ratio of supra- and infra-granular projections follows a gradient across the cortical areas and constitutes a global signature of hierarchical distance among these areas.

Is this laminar pattern of connectivity specific to the visual system or is it a more general principle of organization of FF and FB pathways in the brain? In support of the latter hypothesis, **Figure 6** shows paired scatter plots of SLN values for a set of 8 somatosensory and motor areas in the same format as **Figure 5**. As in the visual system, the transformed SLN values, here by the similar logit function, provide evidence of a consistent distance measure in the hierarchical organization among these areas.

To quantify the consistency displayed in these pairs plots, we proposed a model to estimate hierarchical distances based on SLN values, but as we argue below, the model is quite general in its application. In short, we suppose that we can assign hierarchical levels, h_i and h_j, to all area pairs *i* and *j*, based on a measure of differences between properties linking the areas. For example, in the case of SLN, we suppose

$$g\left(\mathrm{SLN}_i^p\right) - g\left(\mathrm{SLN}_j^p\right) = h_i - h_j,$$

where g applies a logit or probit transformation to SLN values, from an injection into area p that receives projections from areas i and j. This suggests a formalism similar to a GLM with a binomial family. The SLN is taken as a binomial variable (neurons are found in the upper or lower cortical layers) and the sum of neurons in both compartments is used as a weight.

The key feature of the model that relates the estimates of hierarchy to the biological measure (i.e. the transformed SLN values) is the incidence matrix, $\boldsymbol{X}$, of the cortical graph. The incidence matrix of the graph is defined to have a column for each node and a row for each link. In each row, all values are 0 except for the two nodes of the link, taking on values of -1 and 1, respectively for source and target, if the link is directed. The product of the incidence matrix and the vector of hierarchical values, h, maps the differences in hierarchical value between two areas with the differences between the transformed SLN, such that:

$$g\left(\mathrm{SLN}_i^p\right) = \boldsymbol{Xh}$$

where the left side of the equation is the difference between transformed SLN values of the source area i and the injection site p. The vector h contains the hierarchical coefficients to estimate and its length is equal to the number of columns of the model matrix. The model as defined is not identifiable because the sum of every row equals 0, but by dropping one column, for example, area V1 in the visual hierarchy, the hierarchical coefficients can be estimated with the area corresponding to the dropped row fixed at a hierarchical level of 0.

This resembles a logistic or probit regression problem. However, these models yield standard errors for the estimated coefficients that are unrealistically small. Alternatively, we have used a beta-binomial model; this arises naturally as a binomial model in which the response, which in the current context is the expected SLN value, has variability greater than that specified by the binomial distribution and its distribution is described by a beta distribution (Lesnoff and Lancelot, 2012). For present purposes, the model can be reparameterized to include a dispersion parameter that better models the overdispersion typically observed in neural counts (see (Markov et al., 2014a) for further details). Once the statistical model is specified, the coefficients are estimated by maximum likelihood. Note that because numbers of neurons are used in the model and not just the SLN proportions, this method generates a weighted hierarchy.

The formalization is quite general. For example, if instead of SLN, a binary variable is used simply specifying whether a connection is of a FF or FB type, then the model corresponds to the original problem that Felleman and Van Essen solved. We have found that fitting the

model in this fashion leads to coefficients with much larger standard errors (**Figure 7A**), thus, providing an alternate demonstration of the indeterminacy or more limited information on hierarchy contained in purely binary relations. Thus, the use of a quantitative measure of hierarchy leads to a more exact solution (**Figure 7B**).

To summarize, a qualitative assessment of a hierarchical gradient is initially evaluated visually by choosing a (possibly transformed) measure of the hierarchical difference between area pairs and using pairs plots to assess the pairwise consistency of the distance measure. If the evidence looks promising, the hierarchical values are obtained by fitting a model that maps the hierarchical estimates to the biological measure of the gradient via the network incidence matrix. If a suitable probability distribution can be assigned to the biological measure, the solution can be obtained by maximum likelihood, but other methods, for example introducing Bayesian priors, might be applied in appropriate circumstances.

The visual hierarchy estimated from our model is shown in **Figure 8A** and resembles qualitatively the Felleman and Van Essen model (Felleman and Van Essen, 1991). In contrast, the levels are continuously distributed. Here we have split FEF into area 8L and 8M corresponding to regions involved in small and large saccades, respectively. Area 8L occupies a low position in the hierarchy while 8M is several levels higher. The goodness of fit of the model is indicated by plotting the empirical SLN values against those predicted by the model (**Figure 8B**) and shows that the model accounts for a high proportion of the variance in the data. The functional implications of this model have been explored in several contexts (Bastos et al., 2015b; Chaudhuri et al., 2015; Magrou et al., 2018) and observations indicate that it is applicable to anterograde tracing data from the mouse (D'Souza et al., 2020).

The use of a transformation of SLN to estimate hierarchical distance imposes strong constraints on inter-areal SLN relations. We demonstrate this in **Figure 8C** by using the logit transform, which performs quantitatively nearly identically to the probit transform (**Figure 4C**), but allows expression of the hierarchical relations in terms of ratios of projection strengths. The model defines hierarchical distance, h_{ij}, between two areas, *i* and *j*, as the difference between the logit of SLN values for their projections to a target area, p, shown in the top equation. For the logit, this distance is just the natural log of the ratio of ratios of supra- to infragranular projection strengths from areas *i* and *j* to p (orange arrows in **Figure 8C**, left). If the hierarchical distance measure is consistent, we expect that ratio to be the same for projections to any other area, q, (blue arrows in **Figure 8C**, left) as shown by the equation below the diagram. A simple rearrangement of this identity demonstrates that the ratio of projections from area *i* to areas p and q (orange arrows in **Figure 8C**, right) should be the same for any other area *j*, projecting to

areas p and q. Thus, the hierarchical model we propose implies strong invariances in the ratio of FF to FB projection strengths from common inputs and outputs across areas. We further hypothesize that these invariances impose constraints on the exchange and stability of information flow across the cortical hierarchy.

One might suppose that when simultaneous retrograde tracer injections are made in reciprocally connected areas that the pair of areas would display a reciprocal FF-FB relation. That is to say, the origin of the majority of projections from one area would arise from upper layers and the principal origin of the reciprocating projections from lower layers. This arrangement would naturally lead to the hierarchical regularities that we observe. However, this regularity is not imposed by our model, nor is it always found to occur. In effect, this is what explains the surprising observation (noted above in the hierarchy derived in **Figure 8A**) of a prefrontal area like 8L at the same hierarchical level as early visual areas V3 and V4. As expected, the projections from several lower order visual areas, e.g., V4 and TEO, to area 8L originate in upper layers signifying FF, projections. However, 8L projects back to these areas through FF projections also originating in upper layers (Barone et al., 2000; Markov et al., 2014b). We designate such FF-FF connections as strong loops (Crick and Koch, 1998). They correspond to the situation described earlier in an order relation that when $A \geq B$ and $B \geq A$, then $A = B$.

In the Felleman and Van Essen model, FF (FB) projections connect upstream (downstream) areas in a bottom-up (top-down) fashion placing hippocampus, anterior temporal and prefrontal areas at the top and occipital areas at the bottom of the visual hierarchy. As such, this model described two main counterstreams involved in hierarchical processing within the visual system: a rostro-caudal gradient of FB projections reciprocated by a caudo-rostral gradient of FF projections. Interestingly, in the data base of visual areas on which the Felleman and Van Essen model was based, only one pathway was reported – between FEF and CITd (dorsal part of inferior temporal cortex) that corresponds to a strong loop. This led Crick and Koch (Crick and Koch, 1998) to speculate that such configurations were forbidden in cortical organization. However, we have identified a significant number of strong loops in our data base. **Figure 8D** shows how the $\log_{10}$ FLN varies as a function of distance estimated through the white matter from the source area to the target injection site, referred to as white matter (WM) distance), replotted from Ercsey-Ravasz et al. (Ercsey-Ravasz et al., 2013). The beige points correspond to those connections that participate in strong-loops i.e. area pairs for which SLN is greater than 0.5 in both directions. For reference, the saturation of these points indicates their SLN values with less saturated points indicating a higher SLN, as shown by the inset color

bar. As expected, most of the SLN values near 1 cluster to the right as more long distance connections. There is a fair amount of scatter among the points but they do display a distance rule (blue regression line) just as the rest of the connections do (black regression line). Interestingly, the strength of the strong loop projections is on average greater than the rest of the projections. This suggests that they are likely to play a unique role in cortical computation. What that role is currently remains a mystery. However, there is experimental evidence in favor of an attentional role concerning the strong-loop between FEF and V4. FEF projections can drive attentional response in V4 (Moore and Armstrong, 2003), through selective gating of V4 receptive-fields (Armstrong et al., 2006). Further evidence points to the involvement of fast rhythmic synchronization during FEF attentional modulation of V4 visual responses (Gregoriou et al., 2012; Gregoriou et al., 2009), strongly suggesting a supragranular origin (Bastos et al., 2015a; Markov et al., 2014b).

The results reveal a high-degree of consistency of the structural regularities underlying cortical hierarchy in the sensory cortex. But how generalizable are these findings across the more anterior regions, particularly in frontal and prefrontal cortex (Badre and D'Esposito, 2009; Choi et al., 2018)? One of the few studies that has addressed the structural hierarchy with tract tracing of the prefrontal cortex (Goulas et al., 2014) found little evidence of the rich rostral to caudal hierarchical organization that has been reported in human imaging studies (Badre and D'Esposito, 2007; Koechlin et al., 2003). The controversial issue of frontal cortex and hierarchical control shows promise of resolution via a revision of the concept of a unitary model ensuring a unidimensional gradient. Recent reports favor distinct networks that are proposed to interact thereby ensuring a global hierarchical structure (Schumacher et al., 2019). Nevertheless, the mismatch between the multistage cascade architecture mediating a temporal organization of cognitive control and inter-areal connectivity contrasts with the situation in the visual cortex where there is a smooth increase in receptive field size ascending the Felleman and Van Essen cortical hierarchy (Roelfsema and de Lange, 2016). The mismatch reported in the prefrontal cortex is between the concept of a smooth and gradual rostral-caudal organization found in the imaging studies and the connectivity as found in the collated and binary data base. What about the relation of SLN across prefrontal areas? In **Figure 9A**, the SLN pairs plots for the prefrontal cortex show an encouraging agreement with that described in visual cortex. The hierarchical scale values estimated from the statistical model described above (**Figure 9B**) seem to support a rostral-caudal organization with F1 at the bottom and areas 10 and 24c (the most anterior limbic region) at the top. Note, analysis based on more complete coverage of the frontal cortex might give significantly improved results.

5. Dual stream Architecture and its functional significance.

In the preceding section we showed that the contributions of supra- and infragranular layers in the projections across hierarchical levels were highly consistent. The consistency of our measure of hierarchical distance, in conjunction with the known differences in oscillatory properties of laminar compartments of the cortex, suggests that FF and FB interactions are not only anatomically distinct, but (i) use specific frequencies for communication and (ii) play specialized roles in cortical computation. Here we address how these functional properties relate to the detailed anatomical properties of FF and FB pathways. In the light of recent findings showing FF and FB constitute distinct pathways in both the upper and lower layers constituting a Dual Stream Architecture (Markov et al., 2014b), leads to the hypothesis that FB signals in upper and lower layers have distinct roles in information processing.

There are a number of reasons for expecting that supra- and infragranular layers might house different FF and FB pathways. During corticogenesis the supragranular compartment is generated by a primate-specific germinal zone (Smart et al., 2002), exhibiting uniquely complex lineages (Betizeau et al., 2013; Dehay et al., 2015; Lukaszewicz et al., 2005), findings that have consequently been amply confirmed in human corticogenesis (Geschwind and Rakic, 2013). These specialized developmental origins of the supragranular layers are linked to the observed expansion of these layers in primates culminating in human (Cahalane et al., 2014; Sousa et al., 2017), and a diversification of cell-types, which we speculate underlies the observed coding properties of these layers in the adult primate cortex (Tang et al., 2018; Vinje and Gallant, 2000; Wang and Kennedy, 2016; Willmore et al., 2011). A number of studies have shown that supragranular layers exhibit sparse coding in which large numbers of cells are characterized by low levels of activity and a sensory stimulus activates only few neurons (Barth and Poulet, 2012; Crochet et al., 2011; Haider et al., 2013; Harris and Mrsic-Flogel, 2013; Petersen and Crochet, 2013; Tang et al., 2018). In a sparse code information is encoded at any instant by the spiking of a small number of neurons, as opposed to a dense code where overall activity is high and information is encoded by variation in firing rate as observed in the infragranular layers (Sakata and Harris, 2009). A sparse code reduces redundancy and is arguably more efficient. Studies indicating sparse coding in supragranular layers find evidence of higher levels of recruitment of inhibitory drive in these layers via fast spiking PV+ neurons (Hu et al., 2014), which supports the presence of distinct frequency channels for FB and FF communication (Bastos et al., 2018; Bastos et al., 2015b; Michalareas et al., 2016). In addition, sparse coding,

supragranular neurons in area V1 showed more complex and specific selectivity than expected for primary visual cortex (Bonnefond et al., 2017).

A more detailed description of the laminar organization of inter-areal connectivity suggests that variation of SLN with distance has complex origins concerning inter-areal connectivity in sublayers of the cortex. Exhaustive examination of inter-areal connectivity shows, that whereas canonical FB streams have been traditionally allocated to infragranular layers, a robust supragranular FB stream is in addition found in layer 2 in all areas examined in the extrastriate visual cortex of the macaque (**Figure 10A**) (Markov et al., 2014b). In addition to the layer 2 FB, we found some evidence of a layer 5 FF stream. Hence, in both upper *and* lower compartments there is a counter stream leading to the term dual counterstream architecture. Interestingly, the two FB streams in the supra and infragranular layers will impact differently the canonical microcircuit (**Figure 10B**) (Douglas et al., 1989). The strict segregation of FF and FB streams was hypothesized by Ullman in his counterstream model, which he proposed allows a bi-directional search for optimal matches between descending and ascending pathways (Ullman, 1995, 2000).

Closer examination of the individual streams showed that each obeys a unique distance rule. In all streams labeled cell counts decline with WM distance, however the rate of decline is characteristic for a given stream. In this way, the classical FB stream in layer 6 of the infragranular layers has the most gradual decline so that these FB connections span the biggest distance in the cortex. This contrasts with the layer 2 FB, which shows a much shorter space constant. Hence it is the combination of the space constants of the layers 2 and 6 FB streams that leads to the observed SLN values going up stream from near to far-distant areas (See **Figure 11**). The classical FF stream in L3 is also long-distance stream, but significantly less than the FB layer 6 stream, thus leading to the observation of the greater reach of FB pathways compared to FF pathways (Markov et al., 2014b).

Hence, the dual counterstream architecture shows that the relative rate of decline in individual streams determines the way that SLN is modulated by WM distance. In the previous section (Hierarchy –input consistency) we showed that the agreement between SLN values across hierarchical levels is relatively constant across the extrastriate macaque cortex, but less so for the prefrontal cortex. These differences between frontal and posterior regions could be driven by two sets of factors. Firstly, quite simply the space constant can change in individual layers so that the two regions sample supra- and infragranular layers over different WM distances. For example, if the difference in space constants of layer 2 and 6 FB streams are

reduced then so are the SLN differences and there will be a reduction of the hierarchical distance as such between a group of areas with a common target. A second factor could be the identity of cells in the two streams. Comparison of the gene expression across species has revealed that some genes in rodents that are exclusively expressed in deep layers are expressed in the supragranular layers of primates (Zeng et al., 2012). Such changes in the laminar phenotype could perhaps occur across regions meaning that the layer 2 FB pathway in the prefrontal cortex may not correspond to the same identities as the FB pathway in extrastriate cortex.

6. Functional characteristics of FF and FB pathways.

In the present review we propose that cortical hierarchy can be gauged from the *nature* of the interactions between different brain areas, in the same manner that hierarchies in social networks reflect the nature of interactions between people. Crucially, our measure of hierarchical distance shows that SLN values of area pairs are highly consistent across multiple hierarchical levels. This consistency in conjunction with the known differences in oscillatory properties of laminar compartments of the cortex suggests that FF and FB interactions are not only anatomically distinct, but (i) use specific frequencies for communication and (ii) play specialized roles in cortical computation. Here we address how these functional properties relate to the detailed anatomical properties of FF and FB pathways.

As described above functional interactions between brain areas are distance dependent (Bastos et al., 2015b; D'Souza et al., 2016; Michalareas et al., 2016; Richter et al., 2018; van Kerkoerle et al., 2014). Granger-causality was used to quantify functional FF and FB interactions, thereby allowing the strength of these interactions to be quantified for individual frequency bands. Neuronal populations show spontaneous fluctuations over time that are driven by brain rhythms in different frequency bands, such as theta (3-8Hz), alpha (8-12Hz), beta (12-30Hz) and gamma (30-80Hz). As already noted above, neocortical rhythms do not comprise band-limited sinusoidal oscillation but typically describe noisy oscillations with energy in a wide frequency range (Burns et al., 2011; Spyropoulos et al., 2020). FF Granger-causality is particularly strong in the gamma-frequency band, while FB Granger is strong in the alpha and beta-frequency band (Bastos et al., 2015b; Michalareas et al., 2016; Richter et al., 2018; van Kerkoerle et al., 2014).

The finding that Granger-causality in the FF (FB) direction is strong in gamma (alpha/beta) frequencies is partially dictated by the cellular targets of inter-areal pathways. FF inputs target both excitatory and GABAergic interneurons (**Figure 12**). Importantly, the FF

projections to GABAergic interneurons target almost uniquely parvalbumin – PV interneurons, which are associated with gamma-frequency rhythms and respond to excitatory inputs with high temporal fidelity (Buzsaki and Wang, 2012; Cardin et al., 2009; Jouhanneau et al., 2018; Wang, 2010), and exhibit supra-threshold resonance at gamma-frequencies (Hasenstaub et al., 2005; Pike et al., 2000). Moreover FF projections skip the infragranular layers that are canonically associated with alpha and beta rhythms (Bollimunta et al., 2008; Buffalo et al., 2011; van Kerkoerle et al., 2014). By contrast, FB projections target multiple classes of GABAergic interneurons. Of these, somatostatin – SSt interneurons are associated with the generation of slower brain rhythms (Moore et al., 2010), and can directly influence activity of neurons in infragranular compartments, which are associated with alpha/beta rhythms.

Frequency-specific information transmission can also be predicted by the properties of the *sending* and *receiving* areas, so that the global distribution of brain rhythms across areas and layers contribute to the nature of functional interactions in the FF and FB directions. Gamma-rhythms are particularly strong in early visual areas *and* superficial layers (Bastos et al., 2015a; Buffalo et al., 2011; van Kerkoerle et al., 2014; Vinck and Bosman, 2016), whereas beta rhythms are most prominent in fronto-parietal areas *and* infragranular compartments (Bastos et al., 2015a; Buffalo et al., 2011). Consequently, one expects Granger-causal FF influences from early visual areas to parietal areas to be dominated by gamma frequencies, whereas FB influences to visual areas to be dominated by beta frequencies, in agreement with the fact that major long-range FF projections originate nearly exclusively from superficial layers, and FB from infragranular layers. Further, we note that gamma frequencies are generally associated with cortical activation, e.g. a visual stimulus or optogenetic activation, whereas alpha and beta frequencies are not clearly associated with a FF stimulus drive, consistent with the suggestion that FF provide the major driving influence (Covic and Sherman, 2011; De Pasquale and Sherman, 2011).

Above we hypothesized that these frequency specific channels constitute functional analogues of the SLN projections that we have quantified, leading to the expectation that they exhibit similar hierarchical properties. Bastos et al. (2015) defined the multiple Directed-influence Asymmetry Index (mDAI) based on the difference of gamma and theta vs beta frequency Granger-causality, obtained from ECog recordings in macaque. mDAI enabled these authors to reconstruct hierarchies that closely resembled the structural hierarchy derived from SLN values. The feasibility of accounting for their data with the model that we presented above is demonstrated by the strong tendency of mDAI values corresponding to common projections to lie along lines of unit slope (**Figure 13**). The mDAI index is already on a scale that appears

to show the consistency necessary to apply the model, so the hierarchical values could be estimated from the incidence matrix in a global fashion, as described above, rather than by the averaging method used by Bastos et al. (2015).

Future work will need to specifically address the nature of layer-specific functional interactions for individual FF and FB projections. The dual counterstream architecture postulates that a prominent short-range FB projection originates from L2 in the higher area. Consequently, we predict that FB from L2 is particularly strong at gamma-frequencies. Likewise, L5 is postulated to have a short-range FF projection, and to exhibit strong oscillations at alpha/beta frequencies. Hence, based on the dual counterstream architecture we predict a greater diversity of functional interactions than suggested by previous work, the elucidation of which will require multi-layer high-density recording across multiple cortical areas.

7. The predictive brain and the importance of topdown generative networks.

A large body of physiological studies has shown that FF pathways ascending the hierarchy generate increasingly complex representations of the world in higher areas, leading to the large range of receptive field properties observed at different levels of the hierarchy. Thus, at its core, convergent FF projections carry information from the outside world, and allow this information to be gradually transformed to low-dimensional representations that can be exploited for behavior. In this respect, it is worthwhile noting that the recent success of deep neural network architectures in solving complex tasks similarly demonstrates the power of FF transformations in computation (LeCun et al., 2015; Richards et al., 2019) e.g. by forming increasingly complex representations along the feedforward hierarchy in convolutional networks (Yamins and DiCarlo, 2016).

In contrast to FF-pathways, the neurobiology of the twice as numerous FB pathways (Markov et al., 2014a) remains elusive, forming a major impediment to understanding the brain. A clearly defined role of FB connections is proposed for attentional modulation, but FB pathways are likely critical in a host of complex operations including: the comparison of internally generated predictions of sensory input with actual inputs; imagining sensory-like representations from concepts of e.g. visual objects; carrying out mental simulations and finally gating synaptic plasticity. An early conceptualization of hierarchical processing in the cortex conceived of FF pathways as driving target areas, whereas FB pathways would merely exert a modulatory influence (Klink et al., 2017); some researchers, however, proposed a more nuanced view (Bastos et al., 2012). Indeed, the simple dichotomy of the roles of FF and FB pathways is difficult to reconcile with the multiple physiological effects that are imputed to FB

control. For example, in the case of perceptual completion (e.g. in illusory figures) or visual imagination, FB is usually conceived to *enhance* neural activity; by contrast in the case of filtering out self-generated sensory inputs, FB activity is expected to *suppress* neural activity. These forms of enhancement and suppression represent essentially distinct computational operations: a central concept is that of *gain modulation*, where an existing sensory representation is preserved, but in essence multiplied or divided, as in attentional mechanisms (McAdams and Maunsell, 1999). However, in the case of imagery, one expects FB to *"write-in"* a pattern of neural activity, i.e. operate a *driving* process, or alternatively selectively modulate the activity of specific spontaneously active cell assemblies. In the case of *cancelling out* self-generated sensory inputs through FB (as in self-tickling), FB activity is thought to be *subtractive* (Bastos et al., 2012). Finally, FB activity has been conceived to mediate *error* signals, playing a key role in shaping perceptual learning and synaptic plasticity. The notion of FB as a "swiss-army-knife" contrasts with FF processing which is essentially homogeneous and driving, as captured by the architecture of deep neural networks. These considerations underline the diversity of FB processes, which could be mediated by distinct neural circuits. In particular, we hypothesize that laminar and distance-determined diversity of FB pathways will exhibit anatomo-functional properties that characterize the cortical circuits underlying the diverse global influences engaged in different cognitive functions. Given the diversity of FB pathways, and the many functions in which FB is implicated, it is a daunting task to develop a unified theory of the function of cortical FB. Yet, our understanding of the brain depends crucially on the development of precise theories of cortical FB.

The core feature of FB that distinguishes them from FF is that their projections are more divergent; i.e. they project from few to many neurons. Interestingly, divergent projections are a core feature of the most popular kind of artificial neural network architectures, which are also rapidly becoming a mainstream model of sensory processing in the brain (LeCun et al., 2015; Richards et al., 2019). In FF (deep) neural networks, divergent error signals lead to an adjustment of synaptic weights of FF projections ("backprop"). In other words, in FF (deep) neural networks, the exclusive role of FB is to improve the data transformations implemented by the FF projections. For a biological organism, error signals could be provided for instance by multi-sensory signals or reward signals. However, it is an open question as to how FB would be able to adjust synaptic weights of FF projections (Whittington and Bogacz, 2019). A candidate pathway is the multiple FB projections converging onto layer 1 (Cauller, 1995). Here FB projections impact the distal dendrites of pyramidal neurons, activating non-linear NMDA-R-dependent dendritic integration mechanisms and voltage-gated calcium channels. A recent

study provides evidence that top-down FB to layer 1 might indeed be involved in perceptual learning in a primary sensory area: With reward reinforcement, mice rapidly learn a behavioral task in which their response is contingent on electric-current stimulation of S1. However, when layer 1 projections from perirhinal cortex to S1 are inhibited, mice fail to learn the same task, suggesting that top-down FB is instructive for learning (Doron et al., 2019). Another candidate pathway to modulate plasticity is the FB projection to VIP+ and SOM+ interneurons, given that SOM+ neurons can gate plasticity on the dendrites of pyramidal neurons (Batista-Brito et al., 2018).

Until very recently the dominant theory of brain function was that sensory information progressing up the cortical hierarchy undergoes successive abstractions generating increasingly complex receptive fields. In this feedforward processing model of cortex function, the characterization of the receptive fields at different levels of the hierarchy has been a neuroscience success story over the past 60 years (Hubel and Wiesel, 1962). The bottom up sensory processing driving receptive field elaboration leads to the notion of the importance of single neurons coding for perceptually significant features. According to this feature detection view the cortical hierarchy will house neurons ranging from edge detectors to the proverbial grandmother neurons (Martin, 1994). However, in recent years there has been a paradigm shift in cognitive neuroscience that takes account of the inherent uncertainty of the nature of the sensory input to the brain. In this view which goes back to Helmholtz and was later championed by Richard Gregory, making sense of input from the sensorium requires knowledge of the world, which allows the brain to develop hypotheses of the world that are tested against sensory evidence (Friston, 2010; Gregory, 1997). These hypotheses can be formalized as Bayesian inferences on the causes of our sensation and how these inferences induce perception and behavior (Dayan et al., 1995; Lee and Mumford, 2003). The implementation of inference in the cortex has more recently been recast in terms of the circuits underlying the hierarchical processing that formulate generative statistical models as predictive processing (Clark, 2013; de Lange et al., 2018; Friston, 2010; Keller and Mrsic-Flogel, 2018; Rao and Ballard, 1999).

According to the predictive processing model the brain possesses a model of the world that it seeks to optimize using sensory inputs. Predictive processing postulates that the prior information that resides at the different levels of the cortical hierarchy generate descending predictions that cascade down the cortical hierarchy, allowing interaction with bottom up information ascending the hierarchy. At each level of the hierarchy the interaction between the top-down predictions and the ascending input from the sensorium is such that only the residual, unexplained sensory information (or prediction error) proceeds to the next level. The overall

outcome of the hierarchical process is prediction error minimization. Predictive processing constitutes an inversion of the classical feedforward model by proposing that descending feedback pathways provide representations of the external world, which modifies the ascending signal that now indicates aspects of the world that are unexpected. Given its radical nature impacting every aspect of cortical function, predictive processing has attracted considerable attention from experimentalists which have gone some way to providing empirical support (Bastos et al., 2012; Clark, 2013; de Lange et al., 2018; Keller and Mrsic-Flogel, 2018; Walsh et al., 2020).

Much of the complexity of predictive coding concerns the circuitry underlying the interaction of FF and FB streams and this area of research is still hotly debated (Keller and Mrsic-Flogel, 2018). However, little research has focused on testing how FB connections in the brain could serve as the substrate of a top-down generative network (Bastos et al., 2012; Hinton, 2007; Kosslyn, 1994; Mumford, 1992; Senden et al., 2019). In FF visual processing, high-dimensional sensory representations such as an image are sequentially transformed into low-dimensional representations such as object categories, represented at higher hierarchical levels. Conceivably FB pathways *invert* this process by generating high-dimensional representations starting from low-dimensional variables (Hinton, 2007). We refer to such pathways as top-down generative networks. Whereas the FF projections in convolutional networks create a convergence of information from many to few, in generative networks information is relayed by divergent projections from few to many. For instance, the perception of a red apple depends on a transformation of the image of a red apple (represented by many neurons) into high-level representations in the temporal lobe (represented by few neurons). However, if we imagine a red apple, processing should start from high-level representations (few neurons) and generate a low level neural representation of the image of a red apple (represented by many neurons).

Top-down generative networks may play distinct functional roles in several cognitive processes such as predictive processing, mental simulation, imagery or selective attention. An attractive aspect of the generative top-down network hypothesis is that all of these functions are subserved by a relatively small number of anatomical FB pathways implementing a function-specific generative network. This network would then interact with distinct cellular components in individual target areas, thereby differentially impacting ongoing ascending FF activity in distinct pathways. As an example, we can take the processes of *imagination, expectancy and attention.* In the case of expectancy (e.g. walking to the kitchen and expecting to see a red apple in the fruit basket, but not in the sink), generative networks may cause a

767 reduction of neural response to an expected stimulus, whereas that to surprising stimuli would
768 not be suppressed by top-down predictions. In the case of attention (e.g. searching for a red
769 apple in the kitchen), generative networks may lead to an amplification of sensory activity when
770 we find the stimulus that we were seeking. In the case of *imagination* (e.g. thinking of a red
771 apple), generative networks would drive activity in lower areas based on the activation of
772 neurons in higher areas in the absence of any sensory stimulation. Since high-level brain areas
773 contain abstract representations of objects, how do the feedback pathways of the brain achieve
774 the remarkable feat of generating concrete sensory representations in the mind's eye during
775 mental imagery? An apple is an abstract concept, yet we can generate concrete instantiations of
776 apples by imagining for example, a Pink Lady or a green Granny Smith apple on a real or an
777 imagined table in front of us. Remarkably, we are also not limited to imagining objects as we
778 usually see them –it is equally possible for us to imagine a red apple that is as big as a football!
779 Remarkably, we can generate sensory experiences in environments of which we have no
780 experience, such as standing on Mars staring out through our space helmet at the red colored
781 landscape in front of us. These examples illustrate a cornerstone of our hypothesis: that
782 generative networks are competent to transform abstract concepts into concrete sensory-like
783 representations in our mind's eye.

784 Strong experimental support for top-down generative networks comes from findings in
785 human imaging experiments showing that imagined objects lead to corresponding spatial
786 temporal activation of area V1 (Emmerling et al., 2016; Naselaris et al., 2015; Senden et al.,
787 2019; Thirion et al., 2006). The activation of the top-down pathway is coherent with predictive
788 processing theory where FB pathways shape the FF pathway but not the inverse. These findings
789 of top-down generative networks creating a sensory-like representation in early visual areas
790 could be a particularity of the human brain, thereby providing the substrate for enhanced visual
791 imagery supporting simulation and imagination. However, this appears not to be the case; a
792 recent breakthrough publication from the Roelfsema lab revealed spiking activity in area V1 of
793 a macaque with respect to an object held in working memory (van Kerkoerle et al., 2017),
794 suggesting that the fMRI activation patterns reported in human experiments could correspond
795 to spiking activity generated by visual imagery of imagined objects (Emmerling et al., 2016;
796 Hinton, 2007; Naselaris et al., 2015; Senden et al., 2019; Thirion et al., 2006).

797 The dual counterstream architecture (see Section 5) suggests several features that are
798 congruent with the hypothesis of top-down generative networks: *(1)* the convergence of cortical
799 FF connections in the sensory stream is mirrored by the divergence of FB connections; *(2)*
800 source populations of FF and FB pathways are completely separate, which has been recognized

as a key requirement enabling distinct functionalities of generative top-down networks (Friston, 2018; Markov and Kennedy, 2013; Markov et al., 2014b; Shipp, 2016) *(3)* on average throughout the cortex there are twice as many FB pathways than FF projections to a given area. Studies in which human subjects are cued to generate a sensory percept also provide evidence for top-down generative networks in the cortex (Emmerling et al., 2016; Naselaris et al., 2015; Senden et al., 2019; Slotnick et al., 2005; Thirion et al., 2006).

Examining the dual counterstream suggests a possible division of labor between layer 2 and 6 FB projections: Layer 2 FB and layer 3 FF projections exhibit common features that distinguish them from layer 6 FB. Layer 2 FB and layer 3 FF are short distance, topographically organized and exhibit low rates of bifurcation; contrasting with layer 6 FB which are long-distance, diffuse and have high rates of bifurcation. Thus, the layer 2 FB system may mirror the layer 3 FF system and implement a generative top-down network in which high-dimensional sensory representations can be generated through sequential stages, starting from higher brain areas. The layer 6 FB pathway, on the other hand, may determine the way in which the layer 2 generative network interacts with the local microcircuit, sending contextual signals that reflect behavioral goals and reward signals. Based on this reasoning, we hypothesize that layer 6 FB has more modulatory effects that it exerts, for example, via targeting the apical layer 1 dendrites as well as GABAergic interneurons such as SOM+ and VIP+ interneurons that modulate the activity of local pyramidal neurons (Batista-Brito et al., 2018). Testing this hypothesis will require parallel anatomical and physiological investigations. Optogenetic techniques in non-human primates could allow the injection of precise spatiotemporal activation patterns into specific laminar compartments of higher brain areas, combined with physiological measurements of activity in lower brain areas.

Box 1

Experimental Exploration of the Dual Counterstream Architecture.

Dual counterstream features	**Functional and structural correlates**
(1) Source populations of FF and FB pathways are completely separate, which has been recognized as a key requirement enabling distinct functionalities of generative top-down networks (Friston, 2018). A core feature of the dual counterstream architecture is that despite the ubiquity of bifurcation (Kennedy and Bullier, 1985), in no layer do individual neurons in FF and FB pathways possess	Molecular characterization of FF and FB neurons is very much on the agenda. This would lead to the development of markers of these two projection types and hold the promise of the development of genetic tools for independent manipulation of different FB pathways.

both up and downstream collaterals (Markov et al., 2014b). Similar findings hold for mouse (Berezovskii et al., 2011). This indicates that FF and FB cells indeed constitute distinct populations.	
(2) Large variability in the laminar configuration of FB pathways. FF connections are highly stereotypical, while FB projections show a large variability in their laminar origins and targets. Retrograde tracer in a midlevel target area labels up- and downstream areas with different proportions of supragranular neurons according to their hierarchical distance (**Figure 10C**). These laminar distributions of FF and FB neurons constitute their signatures that allow areas to be ranked in the Felleman and Van Essen hierarchy. Importantly, the different signatures of hierarchical distance reflect the changing laminar composition of the long-distance inputs to an area, meaning that "hierarchy" refers to the distance-dependent laminar-constrained integration of FF and FB inputs (Barone et al., 2000; Markov et al., 2014b) and signals (Bastos et al., 2015b; Michalareas et al., 2016) into the canonical microcircuit.	The multiplicity of the FB pathways in terms of laminar origins and targets over different distances suggest that compared to the singularity of FF pathways, individual FB pathways come in different flavors and fulfill different functions. These different FB pathways need to be investigated in human and non-human primates (NHP) in high-field, laminar resolution fMRI and in parallel using invasive electrophysiology in NHP. These functional investigations are of particular relevance to the supragranular counterstream given the unique features of the Primate supragranular layers which are generated by a primate-specific germinal zone (OSVZ) (Smart et al., 2002), that exhibits complex cell lineages (Lukaszewicz et al., 2005) that we have hypothesized generate unique cell features in primates (Dehay et al., 2015; Harris and Shepherd, 2015). These considerations support the notion that the two FB pathways (one in the supra- and one in infragranular layers) will have distinct functional roles in generative networks.
(3) FB and FF are not serially organized. Computational modeling studies e.g. (Bastos et al., 2012) invariably assume a serial hierarchical organization (**Figure 10C**) whereas in fact most areas receive and project to most other areas (Markov et al., 2014b).	The non-serial nature of inter-areal connectivity raises difficulty for ongoing attempts at large-scale computational modeling that needs to be given further attention.
(5) Lamination The dual counterstream architecture explicitly links pathways to layers, which is in-line with the increasing importance attached to cortical lamination and connectivity (Senzai et al., 2019). (l): effects on topography of FB. In addition to the low divergence-convergence values of supragranular layers (L2 FB and L3FF), and the high divergence-convergence values of infragranular layers (L5 FF and L6FB), high-rates of bifurcation by L6/L5 ensures innervation of multiple areas contrasting with and low levels of bifurcation by L2/L3 neurons (Kennedy and Bullier, 1985).	There are two interrelated issues here. Present understating of inter-areal connectivity in NHP, crucially lacks insight into the laminar restricted connectivity of these pathways. This requires using viral tracers in order to obtain precise information on the connectivity of individual pathways over different distances.

(6) Differences of FB and FF targets. The L3 FF projections primarily target upstream interneurons in L4. Presently we do not know the cellular targets of L2 FB projections although there is some suggestion that they could principally target L2 FB neurons in downstream areas, leading to long inter-areal FB chains. In parallel with differences in FF and FB cellular targets, do the cell targets change (and how) as a function of connection distance, possibly bringing a solution to addressing the problem raised in point (3)?	Laminar restricted connectivity to different cell types needs to be investigated in rodents and primates using next generation of viral tracers.
(7) FB show distinct development strategies. Differences in functionality are expected to require different developmental programs, which is indeed the case for FF and FB pathways. FF pathways develop precociously and exhibit directed-growth; FB pathway formation is delayed well into the postnatal period and exhibits diffuse growth followed by pruning (Barone et al., 1995; Barone et al., 1996; Kennedy et al., 1989).	These observations invite the investigation of inter-areal processing in NHP at different developmental periods and suggest that high-field, laminar resolution fMRI in infants and adults could reveal important developmental processes.

9. Conclusion and speculations.

We have shown that cortical hierarchy can be defined by connectivity gradients and the functional correlates of these gradients. In other words, inputs to a cortical area sample the cortical sheet in a principled manner defined by the differential space constants of the distance rules governing the individual cortical layers. This approach to understanding hierarchy is anchored in the recognition that it is the activity and connectivity linking neurons across the cortex that will ultimately reveal the process underlying the dynamics of cortical function. Link based investigation is complemented by characterization of the nodes. Helen Barbas has championed the structural model, which shows that laminar differentiation allows hierarchical ranking of cortical areas that correlates well with connectivity patterns (Barbas, 2015). The structural model is of particular interest because it allows ranking of cortical areas via gradual variations of cytoarchitecture and myeloarchitecture (Sanides, 1972), and has opened the interpretation of large-scale models of the cortex to investigation with non-invasive imaging techniques that can be applied to the human brain (Burt et al., 2018; Margulies et al., 2016; Paquola et al., 2019).

The central argument of this review is that cortical hierarchy can be usefully thought of as the gradual changes in the cortical input requirements of the local cortical circuit that in terms of synaptic mass constitutes the powerhouse of the cortex. Understanding the cellular

847 mechanisms underlying hierarchical processing require investigations of hierarchy in terms of
848 the laminar restricted connectivity and physiology that we have advocated in this review and
849 described in Box 1. It is nearly 15 years since Jean Bullier posed the question "What is fed
850 back?" (Bullier, 2006). The multiplicity of FB pathways and the complexity of their proposed
851 functions were deep issues that he felt needed to be addressed. In the last 14 years there has
852 been a spectacular development of three classes of techniques that now allow us to address Jean
853 Bullier's question. Firstly, optogenetics holographic stimulation makes it possible to address
854 causation (Carrillo-Reid et al., 2019; Marshel et al., 2019), thereby addressing how activation
855 of a given FB pathway influences a particular cognitive task. Secondly, viral tracing allows
856 cell-type and laminar-constrained connectivity (El-Shamayleh et al., 2016; Nassi et al., 2015;
857 Siu et al., 2020) making it possible to resolve the dual counterstream architecture. Thirdly, high-
858 resolution fMRI allows laminar and columnar resolution imaging in the human brain e.g.
859 (Kemper et al., 2018); furthermore, recent advances in MR methods and data analysis enable
860 investigation of directional laminar connectivity in hierarchical brain networks (Huber et al.,
861 2020). The key feature of future investigation will be to examine FB modulation of cognitive
862 tasks in animal models that can, in parallel, be applied in humans where perceptual
863 consequences can be reported (Kok et al., 2016; Schneider et al., 2019). These combined
864 approaches will address the complexity of the interaction of descending generative networks
865 with the local microcircuit. Interpreting the data coming out of these experiment in terms of
866 cortical dynamics will require computational modeling of the interactions of these top-down
867 effects with the canonical microcircuit using approaches developed by Maass (Haeusler and
868 Maass, 2007).

869 The interactions of descending with ascending pathways required by predictive
870 processing theory will be largely in supragranular cortical layers, which have been shown to be
871 a major target for human evolutionary adaptation (Heide et al., 2020; Won et al., 2019). The
872 evolutionary expansion of the supragranular layers is accompanied by an increase in
873 heterogeneity of glutamergic cell-types in terms of morphology, electrophysiology and gene
874 expression going from rodent to human (Berg et al., 2020) in turn supporting a complexification
875 of the circuits in these layers (Hodge et al., 2019). The amplification and diversification of
876 supragranular intratelencephalic-projecting neurons in primates suggest that the investigation
877 of the biology of the generative networks advocated here may well exploit reductionist
878 approaches in the rodent model, but will need nonetheless to be studied with a particular
879 emphasis on human and non-human primates.

Finally, the observation that top-down networks generate sensory-like activity in low levels of the cortical hierarchy in the *absence* of sensory input raises interesting issues of representation in the brain. According to predictive coding theory, sensory perception requires predicting the present. These top-down generative networks can be important for predicting or simulating the future (Grezes and Decety, 2001), in interplay with the hippocampus (Buckner, 2010; Kay et al., 2020), as well as revisiting the past allowing for instance the minds eye to scrutinize detail in visual images (Kosslyn, 1994). In this way the top-down generative networks could have much to do with what we call thinking. In addition these networks would support an integral aspect of human consciousness, namely the seamless and continuous navigation of the mental narrative of self across past, present and future (Jaynes, 1976; Nørretranders, 1991). Importantly, the observation that these networks are active in the NHP (van Kerkoerle et al., 2017), indicates that we can now plan experimental investigation of their integrative neurobiology at the cellular level.

Acknowledgments.

We would like to thank Kevan Martin and Rodney Douglas for numerous and often passionate discussions on some of the central issues discussed here and for their thoughtful comments on the text, as well as Stewart Ship for his deep insight. Funding gratefully acknowledged by XJW (ONR Grant N00014-17-1-2041, US National Institutes of Health NIH grant 062349, Simons Collaboration on the Global Brain program grant 543057SPI) ; KK (DUAL_STREAM ANR-19-CE37-0025); MV (ERC Starting Grant SPATEMP, BMF BINDA-031L0167); HK (LABEX CORTEX ANR-11-LABX-0042; Université de Lyon ANR-11-IDEX-0007), A2P2MC ANR-17-NEUC-0004, CORTICITY ANR-17-HBPR-0003, SCUSI, Région Auvergne-Rhône-Alpes 1700933701, FRC APE13 “Rotary-Espoir en Tête”.

Author contribution. Data acquisition JV, LM, HK; Analysis of data JV, LM, KK, HK; Statistical modeling of hierarchy KK; All authors conjointly wrote the first draft and revised and edited the completed document; Proposed the study MV, HK

References

Ardid, S., Wang, X.J., 2013. A tweaking principle for executive control: neuronal circuit mechanism for rule-based task switching and conflict resolution. J Neurosci 33, 19504-19517.

914 Armstrong, K.M., Fitzgerald, J.K., Moore, T., 2006. Changes in visual receptive fields with
915 microstimulation of frontal cortex. Neuron 50, 791-798.
916 Badre, D., D'Esposito, M., 2007. Functional magnetic resonance imaging evidence for a
917 hierarchical organization of the prefrontal cortex. J Cogn Neurosci 19, 2082-2099.
918 Badre, D., D'Esposito, M., 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? Nat
919 Rev Neurosci 10, 659-669.
920 Barbas, H., 2015. General cortical and special prefrontal connections: principles from structure
921 to function. Annu Rev Neurosci 38, 269-289.
922 Barone, P., Batardiere, A., Knoblauch, K., Kennedy, H., 2000. Laminar distribution of neurons
923 in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical
924 rank and indicates the operation of a distance rule. J Neurosci 20, 3263-3281.
925 Barone, P., Dehay, C., Berland, M., Bullier, J., Kennedy, H., 1995. Developmental remodeling
926 of primate visual cortical pathways. Cereb Cortex 5, 22-38.
927 Barone, P., Dehay, C., Berland, M., Kennedy, H., 1996. Role of directed growth and target
928 selection in the formation of cortical pathways: prenatal development of the projection of
929 area V2 to area V4 in the monkey. J Comp Neurol 374, 1-20.
930 Barth, A.L., Poulet, J.F., 2012. Experimental evidence for sparse firing in the neocortex. Trends
931 Neurosci 35, 345-355.
932 Bastos, A.M., Litvak, V., Moran, R., Bosman, C.A., Fries, P., Friston, K.J., 2015a. A DCM
933 study of spectral asymmetries in feedforward and feedback connections between visual
934 areas V1 and V4 in the monkey. Neuroimage 108, 460-475.
935 Bastos, A.M., Loonis, R., Kornblith, S., Lundqvist, M., Miller, E.K., 2018. Laminar recordings
936 in frontal cortex suggest distinct layers for maintenance and control of working memory.
937 Proc Natl Acad Sci U S A 115, 1117-1122.
938 Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., Friston, K.J., 2012.
939 Canonical microcircuits for predictive coding. Neuron 76, 695-711.
940 Bastos, A.M., Vezoli, J., Bosman, C.A., Schoffelen, J.M., Oostenveld, R., Dowdall, J.R., De
941 Weerd, P., Kennedy, H., Fries, P., 2015b. Visual Areas Exert Feedforward and Feedback
942 Influences through Distinct Frequency Channels. Neuron 85, 390-401.
943 Batista-Brito, R., Zagha, E., Ratliff, J.M., Vinck, M., 2018. Modulation of cortical circuits by
944 top-down processing and arousal state in health and disease. Curr Opin Neurobiol 52,
945 172-181.
946 Berezovskii, V.K., Nassi, J.J., Born, R.T., 2011. Segregation of feedforward and feedback
947 projections in mouse visual cortex. J Comp Neurol 519, 3672-3683.

948 Berg, J., Sorensen, S.A., Ting, J.T., Miller, J.A., Chartrand, T., Buchin, A., Bakken, T.E.,
949 Budzillo, A., Dee, N., Ding, S.-L., Gouwens, N.W., Hodge, R.D., Kalmbach, B., Lee, C.,
950 Lee, B.R., Alfiler, L., Baker, K., Barkan, E., Beller, A., Berry, K., Bertagnolli, D.,
951 Bickley, K., Bomben, J., Braun, T., Brouner, K., Casper, T., Chong, P., Crichton, K.,
952 Dalley, R., de Frates, R., Desta, T., Dingman Lee, S., D'Orazi, F., Dotson, N., Egdorf,
953 T., Enstrom, R., Farrell, C., Feng, D., Fong, O., Furdan, S., Galakhova, A.A., Gamlin, C.,
954 Gary, A., Glandon, A., Goldy, J., Gorham, M., Goriounova, N.A., Gratiy, S., Graybuck,
955 L., Gu, H., Hadley, K., Hansen, N., Heistek, T.S., Henry, A.M., Heyer, D.B., Hill, D.,
956 Hill, C., Hupp, M., Jarsky, T., Kebede, S., Keene, L., Kim, L., Kim, M.-H., Kroll, M.,
957 Latimer, C., Levi, B.P., Link, K.E., Mallory, M., Mann, R., Marshall, D., Maxwell, M.,
958 McGraw, M., McMillen, D., Melief, E., Mertens, E.J., Mezei, L., Mihut, N., Mok, S.,
959 Molnar, G., Mukora, A., Ng, L., Ngo, K., Nicovich, P.R., Nyhus, J., Olah, G., Oldre, A.,
960 Omstead, V., Ozsvar, A., Park, D., Peng, H., Pham, T., Pom, C.A., Potekhina, L.,
961 Rajanbabu, R., Ransford, S., Reid, D., Rimorin, C., Ruiz, A., Sandman, D., Sulc, J.,
962 Sunkin, S.M., Szafer, A., Szemenyei, V., Thomsen, E.R., Tieu, M., Torkelson, A., Trinh,
963 J., Tung, H., Wakeman, W., Ward, K., Wilbers, R., Williams, G., Yao, Z., Yoon, J.-G.,
964 Anastassiou, C., Arkhipov, A., Barzo, P., Bernard, A., Cobbs, C., de Witt Hamer, P.C.,
965 Ellenbogen, R.G., Esposito, L., Ferreira, M., Gwinn, R.P., Hawrylycz, M.J., Hof, P.R.,
966 Idema, S., Jones, A.R., Keene, C.D., Ko, A.L., Murphy, G.J., Ng, L., Ojemann, J.G.,
967 Patel, A.P., Phillips, J.W., Silbergeld, D.L., Smith, K., Tasic, B., Yuste, R., Segev, I., de
968 Kock, C.P.J., Mansvelder, H.D., Tamas, G., Zeng, H., Koch, C., Lein, E.S., 2020. Human
969 cortical expansion involves diversification and specialization of supragranular
970 intratelencephalic-projecting neurons. bioRxiv, 2020.2003.2031.018820.
971 Betizeau, M., Cortay, V., Patti, D., Pfister, S., Gautier, E., Bellemin-Menard, A., Afanassieff,
972 M., Huissoud, C., Douglas, R.J., Kennedy, H., Dehay, C., 2013. Precursor diversity and
973 complexity of lineage relationships in the outer subventricular zone (OSVZ) of the
974 primate. Neuron 80, 442-457.
975 Binzegger, T., Douglas, R.J., Martin, K.A., 2004. A quantitative map of the circuit of cat
976 primary visual cortex. J Neurosci 24, 8441-8453.
977 Binzegger, T., Douglas, R.J., Martin, K.A., 2009. Topology and dynamics of the canonical
978 circuit of cat V1. Neural Netw 22, 1071-1078.
979 Bollimunta, A., Chen, Y., Schroeder, C.E., Ding, M., 2008. Neuronal mechanisms of cortical
980 alpha oscillations in awake-behaving macaques. J Neurosci 28, 9976-9988.

981 Bonnefond, M., Kastner, S., Jensen, O., 2017. Communication between Brain Areas Based on
982 Nested Oscillations. eNeuro 4.
983 Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., Bressler, S.L., 2004. Beta
984 oscillations in a large-scale sensorimotor cortical network: directional influences revealed
985 by Granger causality. Proc Natl Acad Sci U S A 101, 9849-9854.
986 Buckner, R.L., 2010. The role of the hippocampus in prediction and imagination. Annu Rev
987 Psychol 61, 27-48, C21-28.
988 Buffalo, E.A., Fries, P., Landman, R., Buschman, T.J., Desimone, R., 2011. Laminar
989 differences in gamma and alpha coherence in the ventral stream. Proc Natl Acad Sci U S
990 A 108, 11262-11267.
991 Bullier, J., 2006. What is Fed Back? In: van Hemmen, J.L., Sejnowski, T.J. (Eds.), 23 Problems
992 in Systems Neuroscience. Oxford University Press USA, pp. 103-132.
993 Burns, S.P., Xing, D., Shapley, R.M., 2011. Is gamma-band activity in the local field potential
994 of V1 cortex a "clock" or filtered noise? J Neurosci 31, 9658-9664.
995 Burt, J.B., Demirtaş, M., Eckner, W.J., Navejar, N.M., Ji, J.L., Martin, W.J., Bernacchia, A.,
996 Anticevic, A., Murray, J.D., 2018. Hierarchy of transcriptomic specialization across
997 human cortex captured by structural neuroimaging topography. Nat Neurosci 21, 1251-
998 1259.
999 Buzsaki, G., Wang, X.J., 2012. Mechanisms of gamma oscillations. Annu Rev Neurosci 35,
1000 203-225.
1001 Cahalane, D.J., Charvet, C.J., Finlay, B.L., 2014. Modeling local and cross-species neuron
1002 number variations in the cerebral cortex as arising from a common mechanism. Proc Natl
1003 Acad Sci U S A 111, 17642-17647.
1004 Cardin, J.A., Carlen, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.H.,
1005 Moore, C.I., 2009. Driving fast-spiking cells induces gamma rhythm and controls sensory
1006 responses. Nature 459, 663-667.
1007 Carrillo-Reid, L., Han, S., Yang, W., Akrouh, A., Yuste, R., 2019. Controlling Visually Guided
1008 Behavior by Holographic Recalling of Cortical Ensembles. Cell 178, 447-457 e445.
1009 Cauller, L., 1995. Layer I of primary sensory neocortex: where top-down converges upon
1010 bottom-up. Behav Brain Res 71, 163-170.
1011 Chaudhuri, R., Knoblauch, K., Gariel, M.A., Kennedy, H., Wang, X.J., 2015. A Large-Scale
1012 Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. Neuron
1013 88, 419-431.

1014 Choi, E.Y., Drayna, G.K., Badre, D., 2018. Evidence for a Functional Hierarchy of Association
1015 Networks. J Cogn Neurosci 30, 722-736.
1016 Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive
1017 science. Behav Brain Sci 36, 181-204.
1018 Cossell, L., Iacaruso, M.F., Muir, D.R., Houlton, R., Sader, E.N., Ko, H., Hofer, S.B., Mrsic-
1019 Flogel, T.D., 2015. Functional organization of excitatory synaptic strength in primary
1020 visual cortex. Nature.
1021 Covic, E.N., Sherman, S.M., 2011. Synaptic properties of connections between the primary and
1022 secondary auditory cortices in mice. Cereb Cortex 21, 2425-2441.
1023 Cragg, B.G., 1969. The topography of the afferent projections in the circumstriate visual cortex
1024 of the monkey studied by the Nauta method. Vision Res 9, 733-747.
1025 Crick, F., Koch, C., 1998. Constraints on cortical and thalamic projections: the no-strong-loops
1026 hypothesis. Nature 391, 245-250.
1027 Crochet, S., Poulet, J.F., Kremer, Y., Petersen, C.C., 2011. Synaptic mechanisms underlying
1028 sparse coding of active touch. Neuron 69, 1160-1175.
1029 D'Souza, R.D., Meier, A.M., Bista, P., Wang, Q., Burkhalter, A., 2016. Recruitment of
1030 inhibition and excitation across mouse visual cortex depends on the hierarchy of
1031 interconnecting areas. Elife 5.
1032 D'Souza, R.D., Wang, Q., Ji, W., Meier, A.M., Kennedy, H., Knoblauch, K., Burkhalter, A.,
1033 2020. Canonical and noncanonical features of the mouse visual cortical hierarchy.
1034 bioRxiv, 2020.2003.2030.016303.
1035 da Costa, N.M., Martin, K.A., 2010. Whose Cortical Column Would that Be? Front Neuroanat
1036 4, 16.
1037 Dayan, P., Hinton, G.E., Neal, R.M., Zemel, R.S., 1995. The Helmholtz machine. Neural
1038 Comput 7, 889-904.
1039 de Lange, F.P., Heilbron, M., Kok, P., 2018. How Do Expectations Shape Perception? Trends
1040 Cogn Sci 22, 764-779.
1041 De Pasquale, R., Sherman, S.M., 2011. Synaptic properties of corticocortical connections
1042 between the primary and secondary visual cortical areas in the mouse. J Neurosci 31,
1043 16494-16506.
1044 Dehay, C., Kennedy, H., Kosik, K.S., 2015. The Outer Subventricular Zone and Primate-
1045 Specific Cortical Complexification. Neuron 85, 683-694.

1046 Doron, G., Shin, J.N., Takahashi, N., Bocklisch, C., Skenderi, S., Drüke, M., de Mont, L.,
1047 Toumazo, M., von Heimendahl, M., Brecht, M., Naud, R., Larkum, M.E., 2019.
1048 Perirhinal input to neocortical layer 1 controls learning. bioRxiv, 713883.
1049 Douglas, R.J., Koch, C., Mahowald, M., Martin, K.A., Suarez, H.H., 1995. Recurrent excitation
1050 in neocortical circuits. Science 269, 981-985.
1051 Douglas, R.J., Martin, K.A., 1991. A functional microcircuit for cat visual cortex. J Physiol
1052 440, 735-769.
1053 Douglas, R.J., Martin, K.A., 2007a. Mapping the matrix: the ways of neocortex. Neuron 56,
1054 226-238.
1055 Douglas, R.J., Martin, K.A., 2007b. Recurrent neuronal circuits in the neocortex. Curr Biol 17,
1056 R496-500.
1057 Douglas, R.J., Martin, K.A.C., Whitteridge, D., 1989. A canonical microcircuit for neocortex.
1058 Neural Comput 1, 480-488.
1059 El-Shamayleh, Y., Ni, A.M., Horwitz, G.D., 2016. Strategies for targeting primate neural
1060 circuits with viral vectors. J Neurophysiol 116, 122-134.
1061 Emmerling, T.C., Zimmermann, J., Sorger, B., Frost, M.A., Goebel, R., 2016. Decoding the
1062 direction of imagined visual motion using 7T ultra-high field fMRI. Neuroimage 125, 61-
1063 73.
1064 Ercsey-Ravasz, M., Markov, N.T., Lamy, C., Van Essen, D.C., Knoblauch, K., Toroczkai, Z.,
1065 Kennedy, H., 2013. A predictive network model of cerebral cortical connectivity based
1066 on a distance rule. Neuron 80, 184-197.
1067 Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate
1068 cerebral cortex. Cereb Cortex 1, 1-47.
1069 Ferster, D., Chung, S., Wheat, H., 1996. Orintation delectivity of thalamic input to simple cells
1070 of cat visual cortex. Nature 380, 249-252.
1071 Friston, K., 2010. The free-energy principle: a unified brain theory? Nat Rev Neurosci 11, 127-
1072 138.
1073 Friston, K., 2018. Does predictive coding have a future? Nat Neurosci 21, 1019-1021.
1074 Geschwind, D.H., Rakic, P., 2013. Cortical evolution: judge the brain by its cover. Neuron 80,
1075 633-647.
1076 Gilbert, C.D., Li, W., 2013. Top-down influences on visual processing. Nat Rev Neurosci 14,
1077 350-363.
1078 Goulas, A., Uylings, H.B., Stiers, P., 2014. Mapping the hierarchical layout of the structural
1079 network of the macaque prefrontal cortex. Cereb Cortex 24, 1178-1194.

1080 Gray, C.M., Konig, P., Engel, A.K., Singer, W., 1989. Oscillatory responses in cat visual cortex
1081 exhibit inter-columnar synchronization which reflects global stimulus properties. Nature
1082 338, 334-337.
1083 Gregoriou, G.G., Gotts, S.J., Desimone, R., 2012. Cell-type-specific synchronization of neural
1084 activity in FEF with V4 during attention. Neuron 73, 581-594.
1085 Gregoriou, G.G., Gotts, S.J., Zhou, H., Desimone, R., 2009. High-frequency, long-range
1086 coupling between prefrontal and visual cortex during attention. Science 324, 1207-1210.
1087 Gregory, R.L., 1997. Knowledge in perception and illusion. Philos Trans R Soc Lond B Biol
1088 Sci 352, 1121-1127.
1089 Grezes, J., Decety, J., 2001. Functional anatomy of execution, mental simulation, observation,
1090 and verb generation of actions: a meta-analysis. Hum Brain Mapp 12, 1-19.
1091 Haeusler, S., Maass, W., 2007. A statistical analysis of information-processing properties of
1092 lamina-specific cortical microcircuit models. Cereb Cortex 17, 149-162.
1093 Haider, B., Hausser, M., Carandini, M., 2013. Inhibition dominates sensory responses in the
1094 awake cortex. Nature 493, 97-100.
1095 Harris, K.D., Mrsic-Flogel, T.D., 2013. Cortical connectivity and sensory coding. Nature 503,
1096 51-58.
1097 Harris, K.D., Shepherd, G.M., 2015. The neocortical circuit: themes and variations. Nat
1098 Neurosci 18, 170-181.
1099 Hasenstaub, A., Shu, Y., Haider, B., Kraushaar, U., Duque, A., McCormick, D.A., 2005.
1100 Inhibitory postsynaptic potentials carry synchronized frequency information in active
1101 cortical networks. Neuron 47, 423-435.
1102 Hawkins, J., Blakeslee, S., 2004. On intelligence: How A New Understanding of the brain will
1103 lad to the creation of truely intelligent machines. Owl books.
1104 Heide, M., Haffner, C., Murayama, A., Kurotaki, Y., Shinohara, H., Okano, H., Sasaki, E.,
1105 Huttner, W.B., 2020. Human-specific ARHGAP11B increases size and folding of primate
1106 neocortex in the fetal marmoset. Science, eabb2401.
1107 Hilgetag, C.C., Goulas, A., 2020. 'Hierarchy' in the organization of brain networks. Philos Trans
1108 R Soc Lond B Biol Sci 375, 20190319.
1109 Hilgetag, C.C., O'Neill, M.A., Young, M.P., 1996. Indeterminate organization of the visual
1110 system. Science 271, 776-777.
1111 Hinton, G.E., 2007. Learning multiple layers of representation. Trends Cogn Sci 11, 428-434.
1112 Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close,
1113 J.L., Long, B., Johansen, N., Penn, O., Yao, Z., Eggermont, J., Hollt, T., Levi, B.P.,

1114 Shehata, S.I., Aevermann, B., Beller, A., Bertagnolli, D., Brouner, K., Casper, T., Cobbs,
1115 C., Dalley, R., Dee, N., Ding, S.L., Ellenbogen, R.G., Fong, O., Garren, E., Goldy, J.,
1116 Gwinn, R.P., Hirschstein, D., Keene, C.D., Keshk, M., Ko, A.L., Lathia, K., Mahfouz,
1117 A., Maltzer, Z., McGraw, M., Nguyen, T.N., Nyhus, J., Ojemann, J.G., Oldre, A., Parry,
1118 S., Reynolds, S., Rimorin, C., Shapovalova, N.V., Somasundaram, S., Szafer, A.,
1119 Thomsen, E.R., Tieu, M., Quon, G., Scheuermann, R.H., Yuste, R., Sunkin, S.M.,
1120 Lelieveldt, B., Feng, D., Ng, L., Bernard, A., Hawrylycz, M., Phillips, J.W., Tasic, B.,
1121 Zeng, H., Jones, A.R., Koch, C., Lein, E.S., 2019. Conserved cell types with divergent
1122 features in human versus mouse cortex. Nature 573, 61-68.
1123 Hu, H., Gan, J., Jonas, P., 2014. Interneurons. Fast-spiking, parvalbumin(+) GABAergic
1124 interneurons: from cellular design to microcircuit function. Science 345, 1255263.
1125 Hubel, D., 1995. Eye, brain and vision. Freeman and Co, New York.
1126 Hubel, D.H., Wiesel, T.N., 1962. Receptive fields binocular interaction and functional
1127 architecture in the cat visual cortex. J Physiol 160, 106-154.
1128 Hubel, D.H., Wiesel, T.N., 1977. Ferrier lecture. Functional architecture of macaque monkey
1129 visual cortex. Proc R Soc Lond B Biol Sci 198, 1-59.
1130 Huber, L., Finn, E.S., Chai, Y., Goebel, R., Stirnberg, R., Stocker, T., Marrett, S., Uludag, K.,
1131 Kim, S.G., Han, S., Bandettini, P.A., Poser, B.A., 2020. Layer-dependent functional
1132 connectivity methods. Prog Neurobiol, 101835.
1133 Jaynes, J., 1976. The origin of consciousness in the breakdown of the bicameral mind.
1134 Houghton Mifflin Company, Boston.
1135 Jouhanneau, J.S., Kremkow, J., Poulet, J.F.A., 2018. Single synaptic inputs drive high-precision
1136 action potentials in parvalbumin expressing GABA-ergic cortical neurons in vivo. Nat
1137 Commun 9, 1540.
1138 Kaas, J.H., Lin, C.S., 1977. Cortical projections of area 18 in owl monkeys. Vision Res 17,
1139 739-741.
1140 Kay, K., Chung, J.E., Sosa, M., Schor, J.S., Karlsson, M.P., Larkin, M.C., Liu, D.F., Frank,
1141 L.M., 2020. Constant Sub-second Cycling between Representations of Possible Futures
1142 in the Hippocampus. Cell 180, 552-567 e525.
1143 Keller, G.B., Mrsic-Flogel, T.D., 2018. Predictive Processing: A Canonical Cortical
1144 Computation. Neuron 100, 424-435.
1145 Kemper, V.G., De Martino, F., Emmerling, T.C., Yacoub, E., Goebel, R., 2018. High resolution
1146 data analysis strategies for mesoscale human functional MRI at 7 and 9.4T. Neuroimage
1147 164, 48-58.

1148 Kennedy, H., Bullier, J., 1985. A double-labeling investigation of the afferent connectivity to
1149 cortical areas V1 and V2 of the macaque monkey. J Neurosci 5, 2815-2830.
1150 Kennedy, H., Bullier, J., Dehay, C., 1989. Transient projections from the superior temporal
1151 sulcus to area 17 in the newborn macaque monkey. Proc Natl Acad Sci USA 86, 8093-
1152 8097.
1153 Klink, P.C., Dagnino, B., Gariel-Mathis, M.A., Roelfsema, P.R., 2017. Distinct Feedforward
1154 and Feedback Effects of Microstimulation in Visual Cortex Reveal Neural Mechanisms
1155 of Texture Segregation. Neuron 95, 209-220 e203.
1156 Ko, H., Hofer, S.B., Pichler, B., Buchanan, K.A., Sjostrom, P.J., Mrsic-Flogel, T.D., 2011.
1157 Functional specificity of local synaptic connections in neocortical networks. Nature 473,
1158 87-91.
1159 Koechlin, E., Ody, C., Kouneiher, F., 2003. The architecture of cognitive control in the human
1160 prefrontal cortex. Science 302, 1181-1185.
1161 Kok, P., Bains, L.J., van Mourik, T., Norris, D.G., de Lange, F.P., 2016. Selective Activation
1162 of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. Curr
1163 Biol 26, 371-376.
1164 Kosslyn, S.M., 1994. Image and Brain. MIT Press.
1165 Kuypers, H.G.J.M., Szwarcbart, M.K., Mishkin, M., Rosvold, H.E., 1965. Occipitotemporal
1166 corticortical connections in the rhesus monkey. Exp Neurol 11, 245-262.
1167 Latawiec, D., Martin, K.A., Meskenaite, V., 2000. Termination of the geniculocortical
1168 projection in the striate cortex of macaque monkey: a quantitative immunoelectron
1169 microscopic study. J Comp Neurol 419, 306-319.
1170 LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436-444.
1171 Lee, T.S., Mumford, D., 2003. Hierarchical Bayesian inference in the visual cortex. J Opt Soc
1172 Am A Opt Image Sci Vis 20, 1434-1448.
1173 Lee, W.C., Bonin, V., Reed, M., Graham, B.J., Hood, G., Glattfelder, K., Reid, R.C., 2016.
1174 Anatomy and function of an excitatory network in the visual cortex. Nature 532, 370-
1175 374.
1176 Lesnoff, M., Lancelot, R., 2012. aod: Analysis of Overdispersed Data. R package version 1.3
1177 http://cran.r-project.org/package=aod.
1178 Lien, A.D., Scanziani, M., 2013. Tuned thalamic excitation is amplified by visual cortical
1179 circuits. Nat Neurosci 16, 1315-1323.

1180 Lukaszewicz, A., Savatier, P., Cortay, V., Giroud, P., Huissoud, C., Berland, M., Kennedy, H.,
1181 Dehay, C., 2005. G1 phase regulation, area-specific cell cycle control, and
1182 cytoarchitectonics in the primate cortex. Neuron 47, 353-364.
1183 Lund, J.S., Lund, R.D., Hendrickson, A.E., Bunt, A.H., Fuchs, A.F., 1975. The origin of
1184 efferent pathways from the primary visual cortex of the macaque monkey as shown by
1185 retrograde transport of horseradish peroxydase. J Comp Neurol 164, 287-304.
1186 Magrou, L., Barone, P., Markov, N.T., Killackey, H.P., Giroud, P., Berland, M., Knoblauch,
1187 K., Dehay, C., Kennedy, H., 2018. How Areal Specification Shapes the Local and
1188 Interareal Circuits in a Macaque Model of Congenital Blindness. Cereb Cortex 28, 3017-
1189 3034.
1190 Margulies, D.S., Ghosh, S.S., Goulas, A., Falkiewicz, M., Huntenburg, J.M., Langs, G., Bezgin,
1191 G., Eickhoff, S.B., Castellanos, F.X., Petrides, M., Jefferies, E., Smallwood, J., 2016.
1192 Situating the default-mode network along a principal gradient of macroscale cortical
1193 organization. Proc Natl Acad Sci U S A 113, 12574-12579.
1194 Markov, N.T., Ercsey-Ravasz, M., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., Kennedy,
1195 H., 2013. Cortical high-density counter-stream architectures. Science 342, 1238406.
1196 Markov, N.T., Ercsey-Ravasz, M.M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J.,
1197 Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud,
1198 C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z.,
1199 Knoblauch, K., Van Essen, D.C., Kennedy, H., 2014a. A weighted and directed interareal
1200 connectivity matrix for macaque cerebral cortex. Cereb Cortex 24, 17-36.
1201 Markov, N.T., Kennedy, H., 2013. The importance of being hierarchical. Curr Opin Neurobiol
1202 23, 187-194.
1203 Markov, N.T., Misery, P., Falchier, A., Lamy, C., Vezoli, J., Quilodran, R., Gariel, M.A.,
1204 Giroud, P., Ercsey-Ravasz, M., Pilaz, L.J., Huissoud, C., Barone, P., Dehay, C.,
1205 Toroczkai, Z., Van Essen, D.C., Kennedy, H., Knoblauch, K., 2011. Weight Consistency
1206 Specifies Regularities of Macaque Cortical Networks. Cereb Cortex 21, 1254-1272.
1207 Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C.,
1208 Misery, P., Giroud, P., Barone, P., Dehay, C., Ullman, S., Knoblauch, K., Kennedy, H.,
1209 2014b. The Anatomy of Hierarchy: Feedforward and feedback pathways in macaque
1210 visual cortex. J Comp Neurol 522, 225-259.
1211 Marshel, J.H., Kim, Y.S., Machado, T.A., Quirin, S., Benson, B., Kadmon, J., Raja, C.,
1212 Chibukhchyan, A., Ramakrishnan, C., Inoue, M., Shane, J.C., McKnight, D.J.,

1213 Yoshizawa, S., Kato, H.E., Ganguli, S., Deisseroth, K., 2019. Cortical layer-specific
1214 critical dynamics triggering perception. Science 365.
1215 Martin, K.A.C., 1994. A Brief-History of the Feature Detector. Cereb Cortex 4, 1-7.
1216 Martinez-Millan, L., Hollander, H., 1975. Cortico-cortical projections from striate cortex of the
1217 squirrel monkey (saimiri sciureus). A radioautographic study. Brain Res 83, 405-417.
1218 McAdams, C.J., Maunsell, J.H., 1999. Effects of attention on orientation-tuning functions of
1219 single neurons in macaque cortical area V4. J Neurosci 19, 431-441.
1220 Mejias, J.F., Murray, J.D., Kennedy, H., Wang, X.J., 2016. Feedforward and feedback
1221 frequency-dependent interactions in a large-scale laminar network of the primate cortex.
1222 Sci Adv 2, e1601335.
1223 Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.M., Kennedy, H., Fries, P., 2016. Alpha-
1224 Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among
1225 Human Visual Cortical Areas. Neuron 89, 384-397.
1226 Moore, C.I., Carlen, M., Knoblich, U., Cardin, J.A., 2010. Neocortical interneurons: from
1227 diversity, strength. Cell 142, 189-193.
1228 Moore, T., Armstrong, K.M., 2003. Selective gating of visual signals by microstimulation of
1229 frontal cortex. Nature 421, 370-373.
1230 Mountcastle, V., 1995. The evolution of ideas concerning the function of the neocortex. Cerbral
1231 cortex 5, 289-295.
1232 Mountcastle, V.B., 1957. Modality and topographic properties of single neurons of cat's
1233 somatic sensory cortex. J Neurophysiol 20, 408-434.
1234 Mumford, D., 1992. On the computational architecture of the neocortex. II. The role of cortico-
1235 cortical loops. Biol Cybern 66, 241-251.
1236 Naselaris, T., Olman, C.A., Stansbury, D.E., Ugurbil, K., Gallant, J.L., 2015. A voxel-wise
1237 encoding model for early visual areas decodes mental images of remembered scenes.
1238 Neuroimage 105, 215-228.
1239 Nassi, J.J., Cepko, C.L., Born, R.T., Beier, K.T., 2015. Neuroanatomy goes viral! Front
1240 Neuroanat 9, 80.
1241 Nørretranders, T., 1991. The user illusion: Cutting consciousness down to size. Viking, New
1242 York, NY, US.
1243 Paquola, C., Vos De Wael, R., Wagstyl, K., Bethlehem, R.A.I., Hong, S.J., Seidlitz, J.,
1244 Bullmore, E.T., Evans, A.C., Misic, B., Margulies, D.S., Smallwood, J., Bernhardt, B.C.,
1245 2019. Microstructural and functional gradients are increasingly dissociated in transmodal
1246 cortices. PLoS Biol 17, e3000284.

1247 Petersen, C.C., Crochet, S., 2013. Synaptic computation and sensory processing in neocortical
1248 layer 2/3. Neuron 78, 28-48.
1249 Pike, F.G., Goddard, R.S., Suckling, J.M., Ganter, P., Kasthuri, N., Paulsen, O., 2000. Distinct
1250 frequency preferences of different types of rat hippocampal neurones in response to
1251 oscillatory input currents. J Physiol 529 Pt 1, 205-213.
1252 Pouget, P., Stepniewska, I., Crowder, E.A., Leslie, M.W., Emeric, E.E., Nelson, M.J., Schall,
1253 J.D., 2009. Visual and motor connectivity and the distribution of calcium-binding
1254 proteins in macaque frontal eye field: implications for saccade target selection. Front
1255 Neuroanat 3, 2.
1256 Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation
1257 of some extra-classical receptive-field effects. Nat Neurosci 2, 79-87.
1258 Richards, B.A., Lillicrap, T.P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., Clopath,
1259 C., Costa, R.P., de Berker, A., Ganguli, S., Gillon, C.J., Hafner, D., Kepecs, A.,
1260 Kriegeskorte, N., Latham, P., Lindsay, G.W., Miller, K.D., Naud, R., Pack, C.C., Poirazi,
1261 P., Roelfsema, P., Sacramento, J., Saxe, A., Scellier, B., Schapiro, A.C., Senn, W.,
1262 Wayne, G., Yamins, D., Zenke, F., Zylberberg, J., Therien, D., Kording, K.P., 2019. A
1263 deep learning framework for neuroscience. Nat Neurosci 22, 1761-1770.
1264 Richter, C.G., Coppola, R., Bressler, S.L., 2018. Top-down beta oscillatory signaling conveys
1265 behavioral context in early visual cortex. Sci Rep 8, 6991.
1266 Rockland, K.S., Pandya, D.N., 1979. Laminar origins and terminations of cortical connections
1267 of the occipital lobe in the rhesus monkey. Brain Res 179, 3-20.
1268 Roelfsema, P.R., de Lange, F.P., 2016. Early Visual Cortex as a Multiscale Cognitive
1269 Blackboard. Annu Rev Vis Sci 2, 131-151.
1270 Sakata, S., Harris, K.D., 2009. Laminar structure of spontaneous and sensory-evoked
1271 population activity in auditory cortex. Neuron 64, 404-418.
1272 Sanides, F., 1972. Representation in the cerebral cortex and its areal lamination patterns. In:
1273 Bourne, G.H. (Ed.), The structure and function of the nervous tissue. Academic Press,
1274 New York and London, pp. 329-453.
1275 Schall, J.D., 2015. Visuomotor Functions in the Frontal Lobe. Annu Rev Vis Sci 1, 469-498.
1276 Scherberger, H., Jarvis, M.R., Andersen, R.A., 2005. Cortical local field potential encodes
1277 movement intentions in the posterior parietal cortex. Neuron 46, 347-354.
1278 Schneider, M., Kemper, V.G., Emmerling, T.C., De Martino, F., Goebel, R., 2019. Columnar
1279 clusters in the human motion complex reflect consciously perceived motion axis. Proc
1280 Natl Acad Sci U S A 116, 5096-5101.

Schumacher, F.K., Schumacher, L.V., Schelter, B.O., Kaller, C.P., 2019. Functionally dissociating ventro-dorsal components within the rostro-caudal hierarchical organization of the human prefrontal cortex. Neuroimage 185, 398-407.

Senden, M., Emmerling, T.C., van Hoof, R., Frost, M.A., Goebel, R., 2019. Reconstructing imagined letters from early visual cortex reveals tight topographic correspondence between visual mental imagery and perception. Brain Struct Funct 224, 1167-1183.

Senzai, Y., Fernandez-Ruiz, A., Buzsaki, G., 2019. Layer-Specific Physiological Features and Interlaminar Interactions in the Primary Visual Cortex of the Mouse. Neuron 101, 500-513 e505.

Shipp, S., 2016. Neural Elements for Predictive Coding. Front Psychol 7, 1792.

Siu, C., Balsor, J., Federer, F., Angelucci, A., 2020. A direct interareal feedback-to-feedforward circuit in primate visual cortex. bioRxiv, 2020.2007.2007.192450.

Slotnick, S.D., Thompson, W.L., Kosslyn, S.M., 2005. Visual mental imagery induces retinotopically organized activation of early visual areas. Cereb Cortex 15, 1570-1583.

Smart, I.H., Dehay, C., Giroud, P., Berland, M., Kennedy, H., 2002. Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. Cereb Cortex 12, 37-53.

Song, S., Sjostrom, P.J., Reigl, M., Nelson, S., Chklovskii, D.B., 2005. Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol 3, e68.

Sousa, A.M.M., Meyer, K.A., Santpere, G., Gulden, F.O., Sestan, N., 2017. Evolution of the Human Nervous System Function, Structure, and Development. Cell 170, 226-247.

Spatz, W.B., Tigges, J., Tigges, M., 1970. Subcortical projections, cortical associations, and some intrinsic interlaminar connections of the striate cortex in the squirrel monkey (saimiri). J Comp Neurol 140, 155-174.

Spyropoulos, G., Dowdall, J.R., Schölvinck, M.L., Bosman, C.A., Lima, B., Peter, A., Onorato, I., Klon-Lipok, J., Roese, R., Neuenschwander, S., Singer, W., Vinck, M., Fries, P., 2020. Spontaneous variability in gamma dynamics described by a linear harmonic oscillator driven by noise. bioRxiv, 793729.

Tang, S., Zhang, Y., Li, Z., Li, M., Liu, F., Jiang, H., Lee, T.S., 2018. Large-scale two-photon imaging revealed super-sparse population codes in the V1 superficial layer of awake monkeys. Elife 7.

Thirion, B., Duchesnay, E., Hubbard, E., Dubois, J., Poline, J.B., Lebihan, D., Dehaene, S., 2006. Inverse retinotopy: inferring the visual content of images from brain activation patterns. Neuroimage 33, 1104-1116.

Tigges, J., Spatz, W.B., Tigges, M., 1973. Reciprocal point-to-point connections between parastriate and striate cortex in the squirrel monkey (saimiri). J Comp Neurol 148, 481-490.

Treue, S., Maunsell, J.H., 1996. Attentional modulation of visual motion processing in cortical areas MT and MST. Nature 382, 539-541.

Tsodyks, M., Kenet, T., Grinvald, A., Arieli, A., 1999. Linking spontaneous activity of single cortical neurons and the underlying functional architecture. Science 286, 1943-1946.

Ullman, S., 1995. Sequence seeking and counter streams: a computational model for bidirectional information flow in the visual cortex. Cereb Cortex 5, 1-11.

Ullman, S., 2000. Sequence seeking and counter streams: A model for information flow in the visual cortex. High-Level Vision. Bradford / MIT Press.

Van Essen, D.C., Zeki, S., 1978. The topographic organization of rhesus monkey prestriate cortex. J Physiol 277, 193-226.

van Kerkoerle, T., Self, M.W., Dagnino, B., Gariel-Mathis, M.A., Poort, J., van der Togt, C., Roelfsema, P.R., 2014. Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. Proc Natl Acad Sci U S A 111, 14332-14341.

van Kerkoerle, T., Self, M.W., Roelfsema, P.R., 2017. Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. Nat Commun 8, 13804.

Vezoli, J., Falchier, A., Jouve, B., Knoblauch, K., Young, M., Kennedy, H., 2004. Quantitative analysis of connectivity in the visual cortex: extracting function from structure. Neuroscientist 10, 476-482.

Vinck, M., Bosman, C.A., 2016. More Gamma More Predictions: Gamma-Synchronization as a Key Mechanism for Efficient Integration of Classical Receptive Field Inputs with Surround Predictions. Front Syst Neurosci 10, 35.

Vinje, W.E., Gallant, J.L., 2000. Sparse coding and decorrelation in primary visual cortex during natural vision. Science 287, 1273-1276.

Walsh, K.S., McGovern, D.P., Clark, A., O'Connell, R.G., 2020. Evaluating the neurophysiological evidence for predictive processing as a model of perception. Ann N Y Acad Sci 1464, 242-268.

Wang, X.-J., Kennedy, H., 2016. Brain structure and dynamics across scales: in search of rules. Curr Opin Neurobiol 37, 92-98.

1348 Wang, X.J., 2010. Neurophysiological and computational principles of cortical rhythms in
1349 cognition. Physiol Rev 90, 1195-1268.
1350 Whittington, J.C.R., Bogacz, R., 2019. Theories of Error Back-Propagation in the Brain. Trends
1351 Cogn Sci 23, 235-250.
1352 Willmore, B.D., Mazer, J.A., Gallant, J.L., 2011. Sparse coding in striate and extrastriate visual
1353 cortex. J Neurophysiol 105, 2907-2919.
1354 Won, H., Huang, J., Opland, C.K., Hartl, C.L., Geschwind, D.H., 2019. Human evolved
1355 regulatory elements modulate genes involved in cortical expansion and
1356 neurodevelopmental disease susceptibility. Nat Commun 10, 2396.
1357 Wong-Riley, M., 1978. Reciprocal connections between striate and prestriate cortex in the
1358 squirrel monkey as demonstrated by combined peroxydase histochemistry and
1359 autoradiography. Brain Res 147, 159-164.
1360 Yamins, D.L., DiCarlo, J.J., 2016. Using goal-driven deep learning models to understand
1361 sensory cortex. Nat Neurosci 19, 356-365.
1362 York, G.K., 3rd, Steinberg, D.A., 2011. Hughlings Jackson's neurological ideas. Brain 134,
1363 3106-3113.
1364 Zeng, H., Shen, E.H., Hohmann, J.G., Oh, S.W., Bernard, A., Royall, J.J., Glattfelder, K.J.,
1365 Sunkin, S.M., Morris, J.A., Guillozet-Bongaarts, A.L., Smith, K.A., Ebbert, A.J.,
1366 Swanson, B., Kuan, L., Page, D.T., Overly, C.C., Lein, E.S., Hawrylycz, M.J., Hof, P.R.,
1367 Hyde, T.M., Kleinman, J.E., Jones, A.R., 2012. Large-scale cellular-resolution gene
1368 profiling in human neocortex reveals species-specific molecular signatures. Cell 149,
1369 483-496.
1370 Zipser, K., Lamme, V.A., Schiller, P.H., 1996. Contextual modulation in primary visual cortex.
1371 J Neurosci 16, 7376-7389.
1372
1373
1374

Figure & Legends

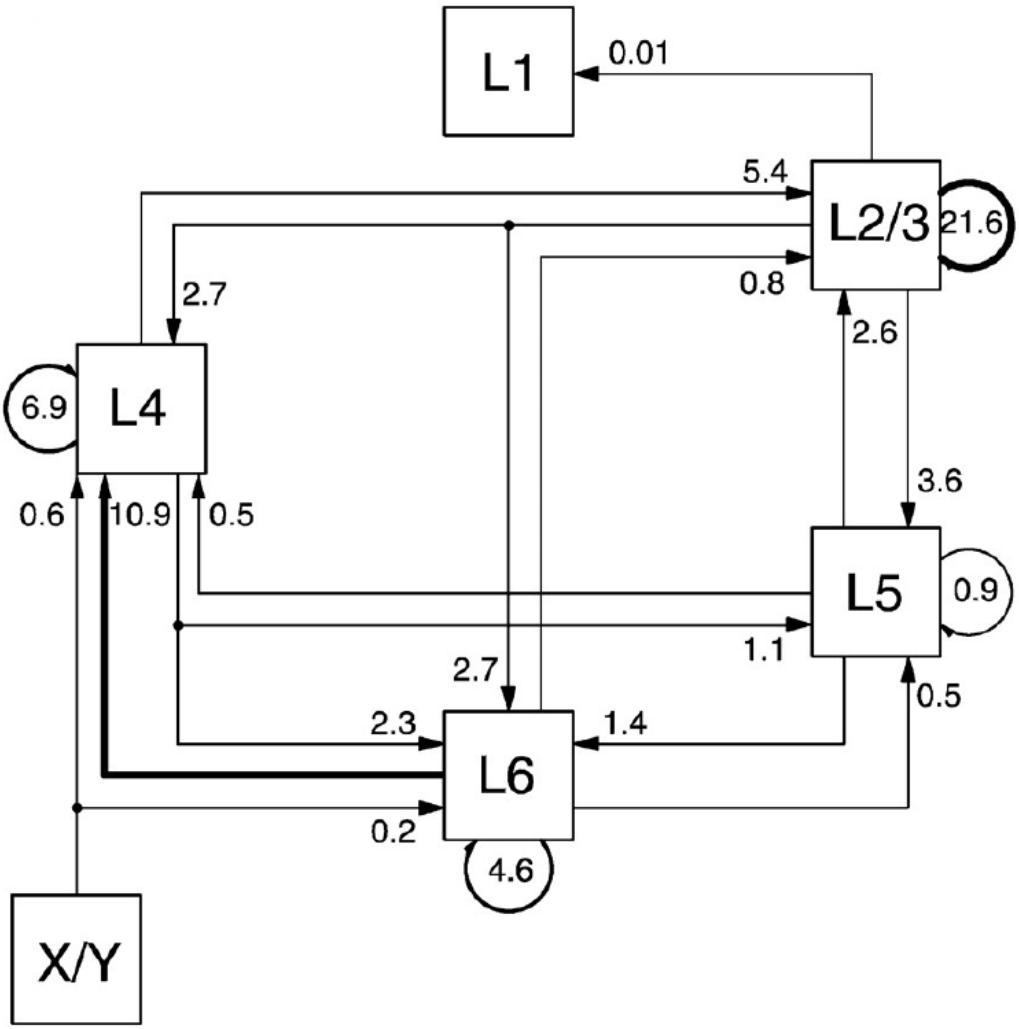


Figure 1. Quantitative map of excitatory synapses between excitatory neurons of the local microcircuit in visual cortex (area 17) of the cat. Numbers indicate proportions of excitatory synapses, note the dominance of within layer recurrent connectivity with 21.6 peak values in Layers 2/3. The FF loop starts in layer 4, the major thalamic recipient layer and then extends to layers 2/3, 5 and 6 with recurrent inputs back to layer 4. This FF loop corresponds to a little less than half of synapses involved in self-innervation of individual cortical layers. X/Y refers to the component cells of the lateral geniculate nucleus, the major thalamic relay. The original canonical microcircuit is shown in Figure 9B. L refers to layer. From (Binzegger et al., 2004) with permission.

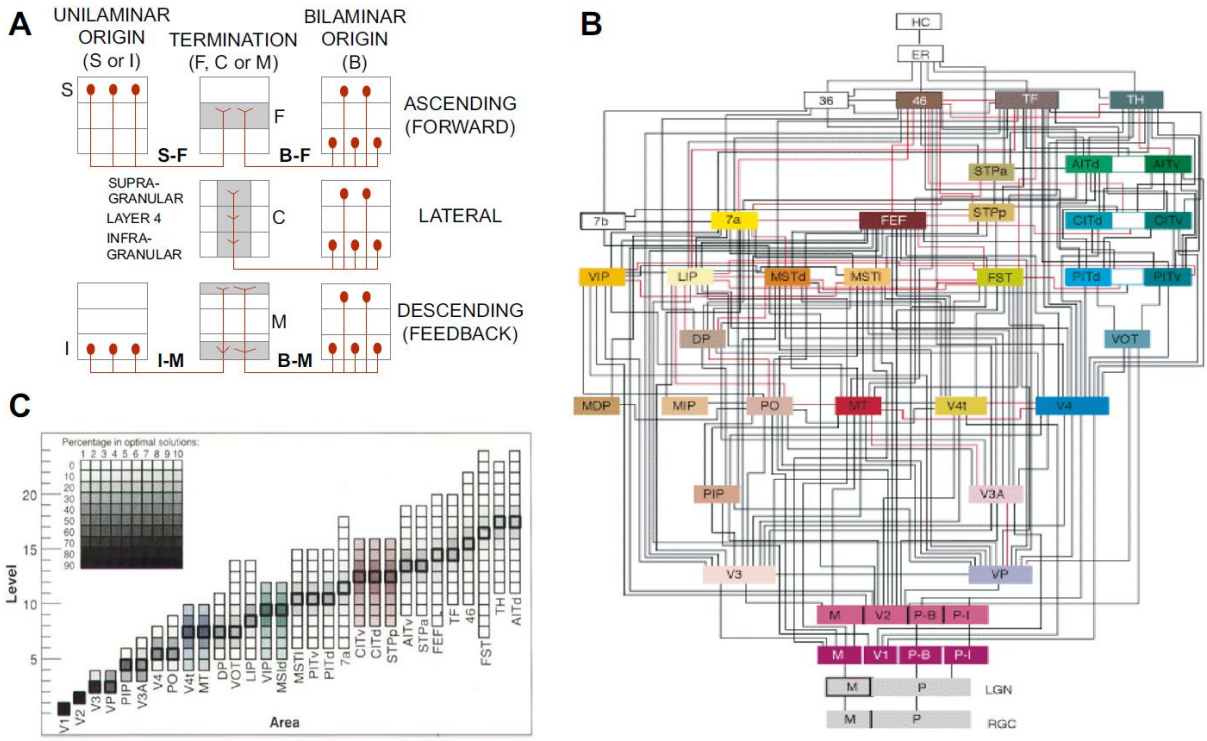


Figure 2 The Felleman and Van Essen binary model of cortical hierarchy. A) Criteria for classifying connections between areas as FF (top), lateral (middle) and FB (bottom) row. Termination patterns are depicted in the central column, preferentially in layer 4 (F pattern) FF, across all layers (C pattern) lateral, in upper and lower layers avoiding layer 4 (M pattern) FB. Laminar origin from a single layer (left column), is either supragranular (S) and therefore FF, or infragranular (I) and therefore FB. Bilaminar (B) origins (right column) either terminate in the middle layers (F pattern) and are therefore FF, terminate in all layers (lateral) or terminate predominantly in upper supra- and infragranular layers (M pattern) and therefore FB. **B)** The binary hierarchical model. **C**: Area frequency distributions for 150,000 optimal hierarchical orderings (Hilgetag et al., 1996).

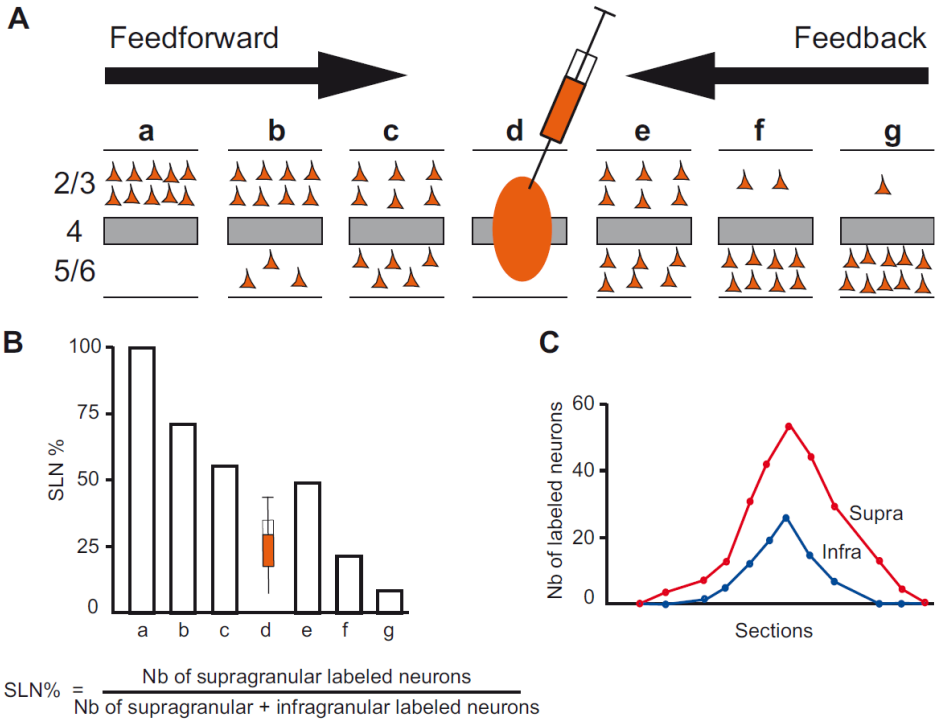


Figure 3. Quantitative parameters characterizing the hierarchy. A) The laminar distribution of parent neurons in each pathway, referred to as SLN (fraction of supragranular neurons) is determined by high frequency sampling and quantitative analysis of labeling. Supra- and infragranular layer neurons contribute to both FF and FB pathways, and their relative proportion is characteristic for each type of pathway. For a given injection there is a gradient of SLN of the labeled areas, between purely FF (SLN = 100%, all the parent neurons are in the supragranular layers) to purely FB (SLN = 0%, all the parent neurons in the infragranular layers) and a spectrum of intermediate proportions; **B**) All labeled areas can then be ordered by decreasing SLN values and this order is consistent with hierarchical order according to Felleman and Van Essen. SLN is thus used as an indicator of hierarchical distance between areas from the same injection; **C**) Reliable estimation of SLN crucially requires sampling labeling throughout the full extent of the projection zone in each area.

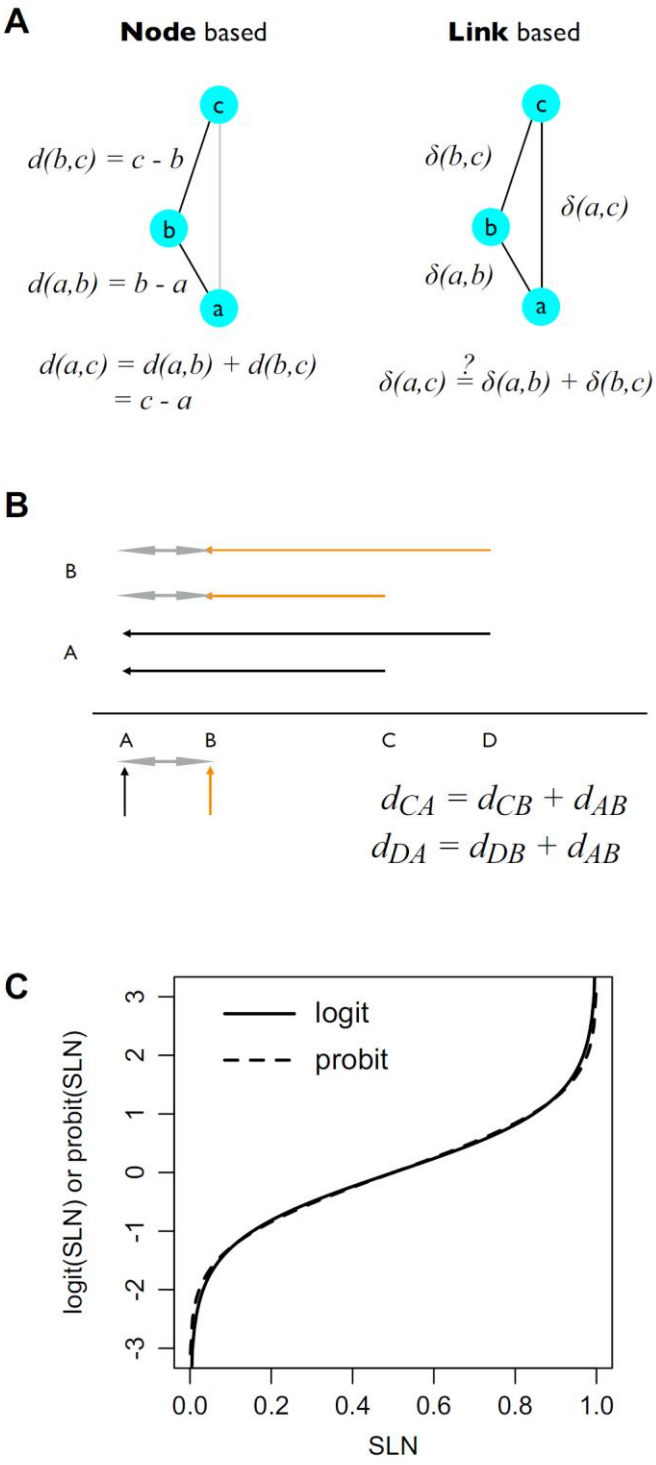


Figure 4. Properties of nodes and links. Nodes have fixed properties, a, b, c which in turn fixes their order and any distance measure, d, calculated from these properties. b. Link properties depend on the relations between node pairs, ab, bc, ac. The distance measures, δ, for ab, and bc do not necessarily fix that for ac. The above graphs are unidirectional, but in bi-directional graphs the distances between nodes need not by symmetric. **B) Hierarchical scales.** Suppose a hierarchical scale between areas A, B, C, D, with the ordering and distances as illustrated on the bottom line. We expect measures of distance to be consistent measured

between any pairs of areas. For example, injections in areas A and B lead to distances defined with respect to each of these areas, i.e., distances AB (double headed grey arrow), AC and AD (black arrows) for injection in area A, and BA (double headed grey arrow), BC and BD (orange arrows) for injection in area B. Consistency would imply, for example, that for a distance measure, d, the estimate of $d_{AB} = d_{BA}$ would be the same for both injections, i.e., $d_{CA} - d_{CB} = d_{DA} - d_{DB}$. **C) SLN Transformation**. Comparison of logit (solid) and probit (dashed) transformations of SLN values on the interval (0, 1). The logit SLN is defined as ln(SLN/(1 – SLN)). The probit is defined as the inverse of the Gaussian cumulative distribution function and is often notated by Φ^{-1}. The scale factor of the logit curve has been adjusted by a factor of 0.588 to match it to the probit curve.

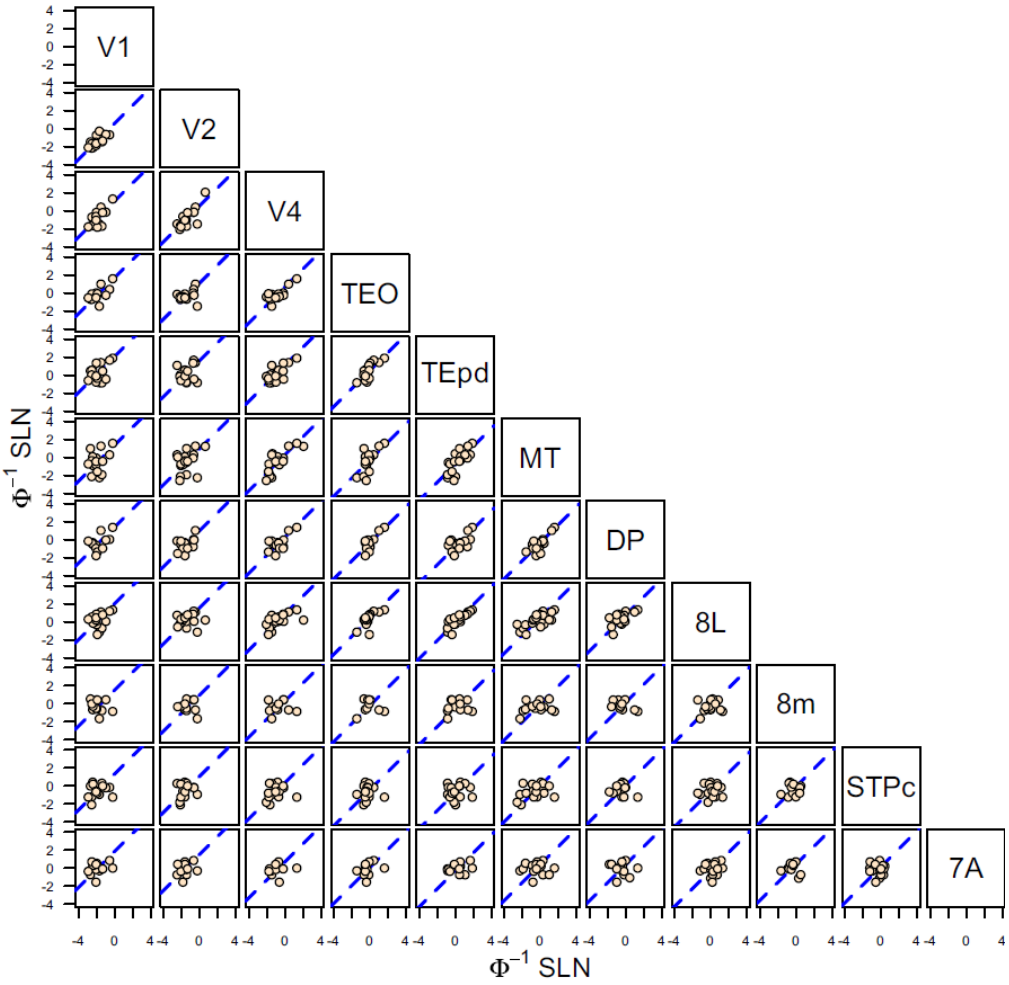


Figure 5. Probit transformation. Scatter plots of probit transformed SLN values of common source areas from pairs of 11 visual areas, obtained from retrograde tracer injections. The absicssa of each graph corresponds to the transformed SLN values of area *i*, indicated on the diagonal at the top of the column and the ordinate values are the transformed SLN values of

area *j* indicated on the diagonal at the right of the row. The dashed blue line in each plot is the best fit line of unit slope (replotted from (Markov et al., 2014b)).

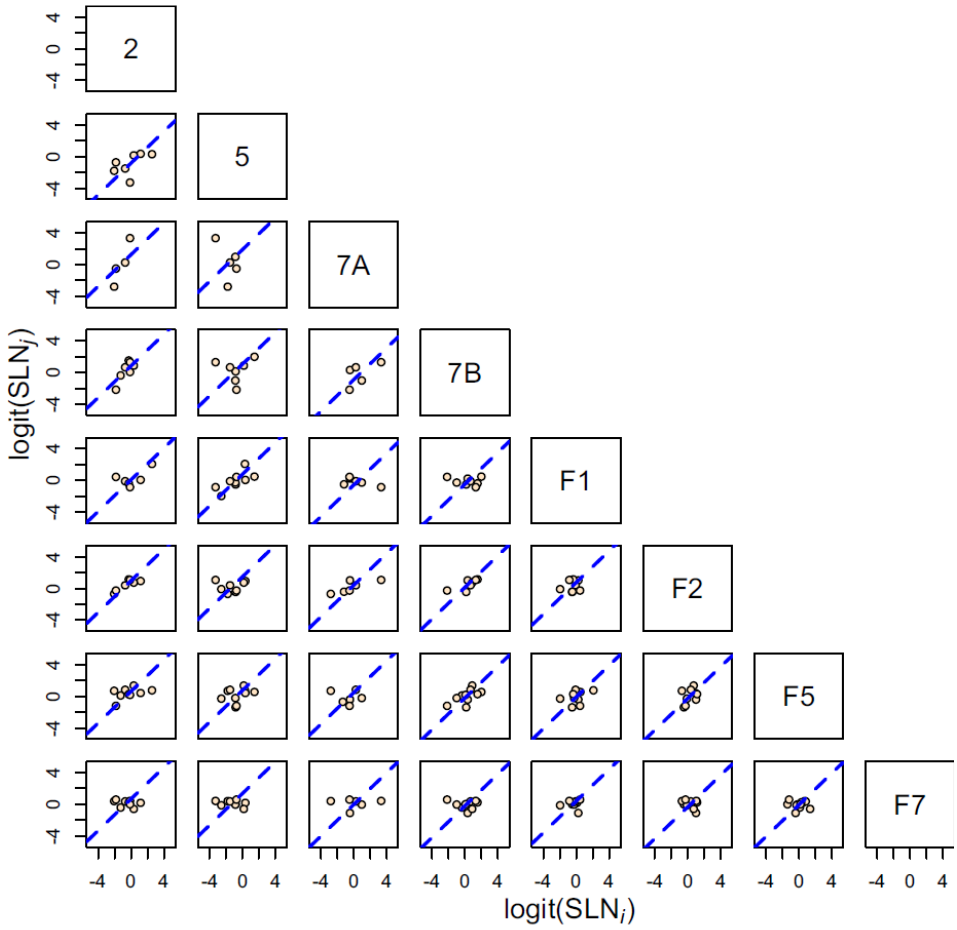


Figure 6. Logit transformation. Scatter plots of logit transformed SLN values of common source areas from pairs of 8 somatosensory and motor areas, obtained from retrograde tracer injections. The plots follow the same format as in Figure 6 except that the SLN values from each axis are transformed by the logit function. The dashed blue line in each plot is the best fit line of unit slope.

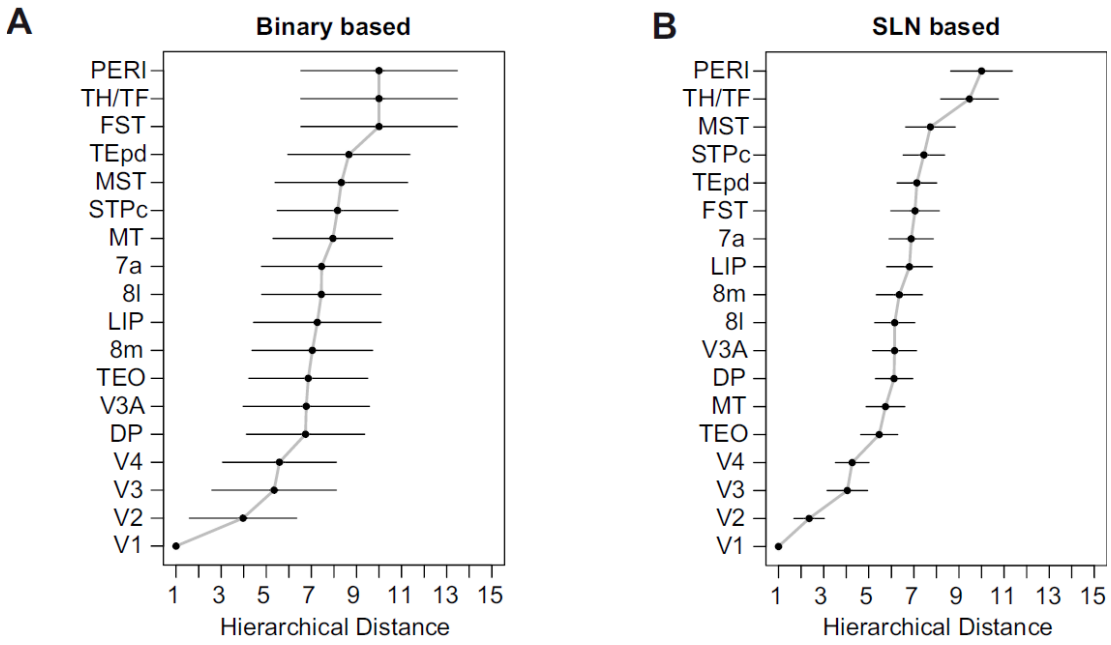


Figure 7. Precision of estimated hierarchy based on hierarchical index. A. Estimated hierarchy obtained using logit transformed SLN values as a measure of hierarchical distance. The counts of supra- and infragranular neurons are used as weights. The error bars are 95% confidence intervals estimated from the covariance matrix of the fitted model. **B**. Estimated hierarchy using a binary variable as an indicator of the hierarchical relation between area pairs. A logit link was also used in this case. The larger 95% confidence intervals demonstrate the loss of precision in estimating the hierarchical distance when using only binary information about connectivity.

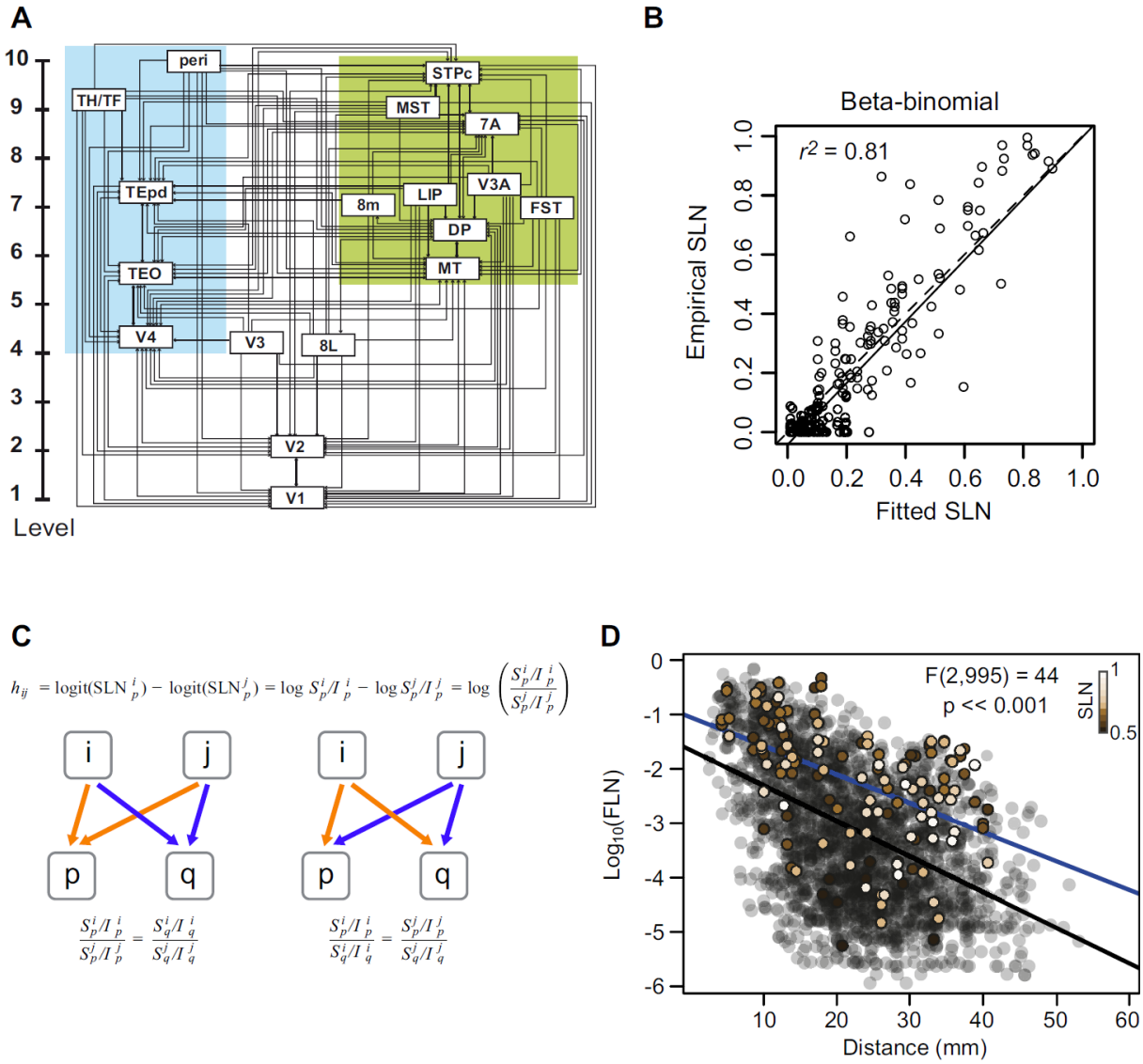


Figure 8. Hierarchical organization of visual areas (A) estimated from the beta-binomial model. The model only provides the vertical level of the areas with respect to the lowest level. For clarity of presentation, we have separated them laterally into ventral and dorsal stream areas. The estimated values are only unique up to adding a constant and multiplying by a coefficient. Here, we have the areas to span the range 1-10. **B**) The scatter plot shows the empirical SLN values plotted against those predicted by the model. The solid line is the unit slope line through the origin and the dashed line is the best fit linear regression. **C) Hierarchical distance.** The hierarchical distance, h_{ij}, between common projections from areas *i* and *j* to area p, defined as the difference of logits of their SLN values, is equivalent to the log of the ratio of their supra- to infra-granular projection strengths to area *p*. a. This definition implies that the ratio between the laminar ratios of areas *i* and *j* to area *p* (orange arrows) is the same as that for any other target area *q* receiving projections from the same source areas (blue arrows), as formalized in the equation below the diagram. This is because the hierarchical distance from *i*

to *j* should be the same for injections in both areas *p* and *q*. b. A rearrangement of the equation (below) implies, also, that the ratio between the laminar ratios of projections from a common source area, *i*, to areas *p* and *q*, will be the same for any other common source area, *j*, to the same target areas. **D) Cortical-cortical strong loops.** The strength-distance relation of 1615 projections from 91 to 29 cortical areas obtained from retrograde tracer injections. The transparent black points indicate all of the projections except those that participate in strong-loops in beige. The color gradient on these symbols corresponds to SLN strength as indicated by the inset color bar. The black line is the best fit linear regression to the transparent black points and the blue line is the best fit to the strong-loops. The F-statistic indicates the result of a nested likelihood test indicating the probability of a difference in strength between the two sets of points as large as that obtained under the null hypothesis that the true difference is zero, when physical distance via the WM is taken into account.

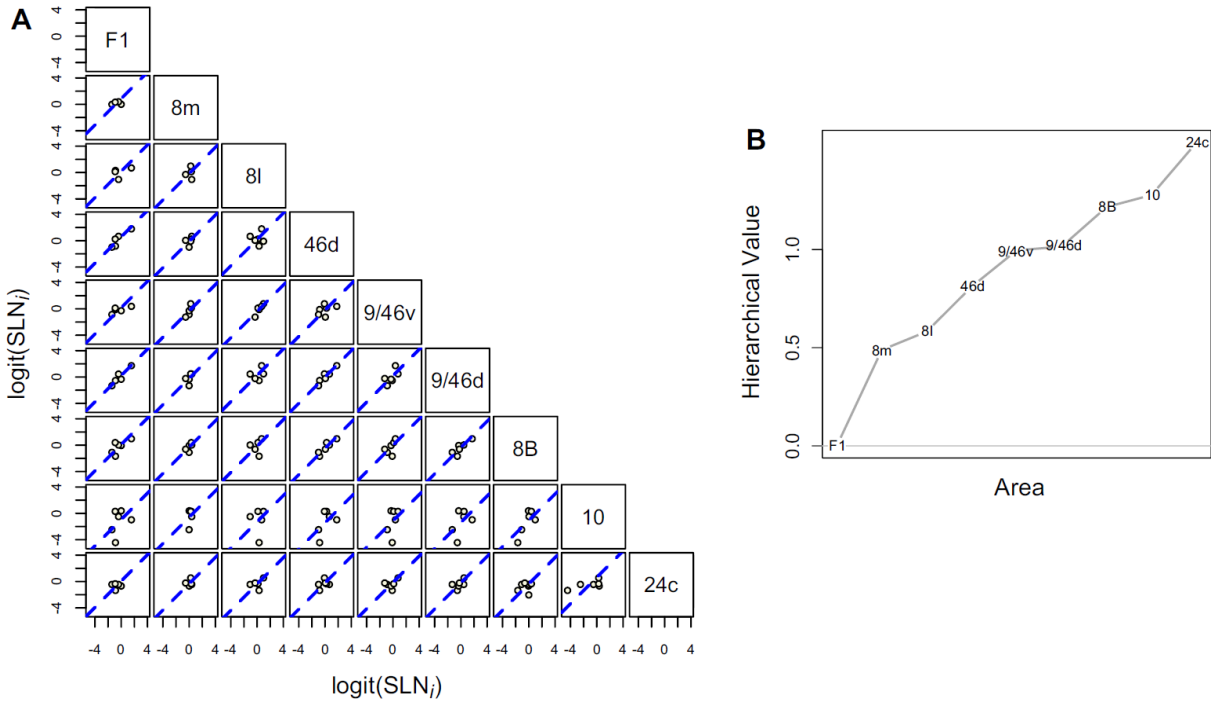


Figure 9. A) Scatter plots of logit transformed SLN values of common source areas from pairs of 9 frontal and pre-frontal areas, obtained from retrograde tracer injections. The plots follow the same format as previous in Figures 6 and 7. The dashed blue line in each plot is the best fit line of unit slope. **B)** Hierarchical scale values estimated for the 9 areas based on the proposed statistical model. Area F1 was assigned a value of 0 for model identifiability.

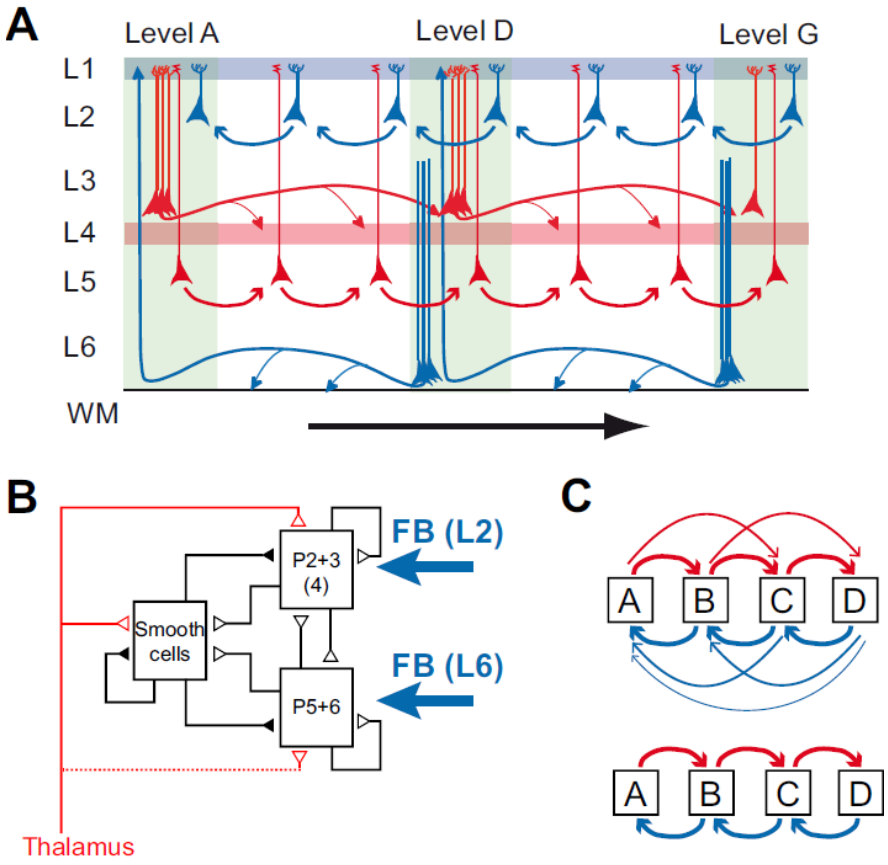


Figure 10. Dual counterstream architecture of inter-areal pathways. A) Parent neurons at L3 and L5 have FF projections (red) to higher order areas reciprocated by FB projections (blue) in L2 and the L6. Simultaneous tracer injections in high and low areas show that the upper layer counterstream has near 100% segregation, i.e. the FF (FB) neurons do not send axon collaterals to lower (higher) order areas. However the evidence that the FF and FB pathways form continuous streams, as depicted here is indirect; what crucially remains to be elucidated are the laminar details on the connectivity and the cellular targets. **B)** the canonical microcircuit showing the two FB pathways targeting L2 and L6. Modified from (Douglas and Martin, 1991); **C)** the incorrectly assumed serial processing (lower) between areas that is not observed in the cortex, where instead each areas project to all upper and lower stream areas (all to all). (panel A from (Markov et al., 2014b));

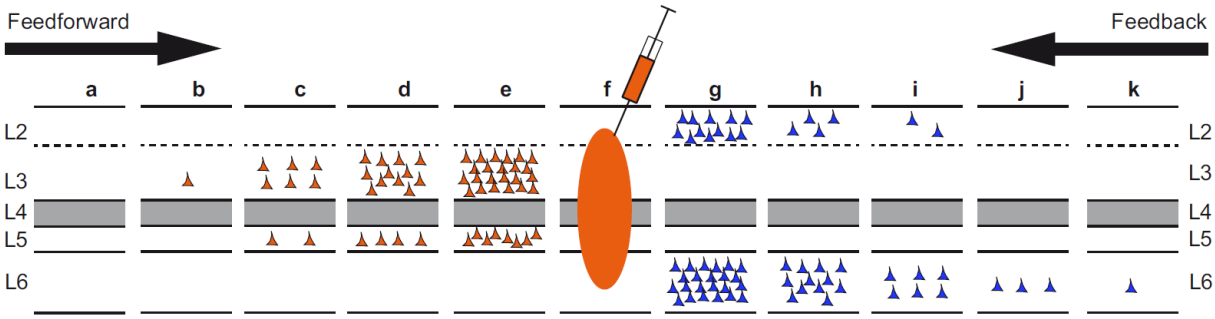
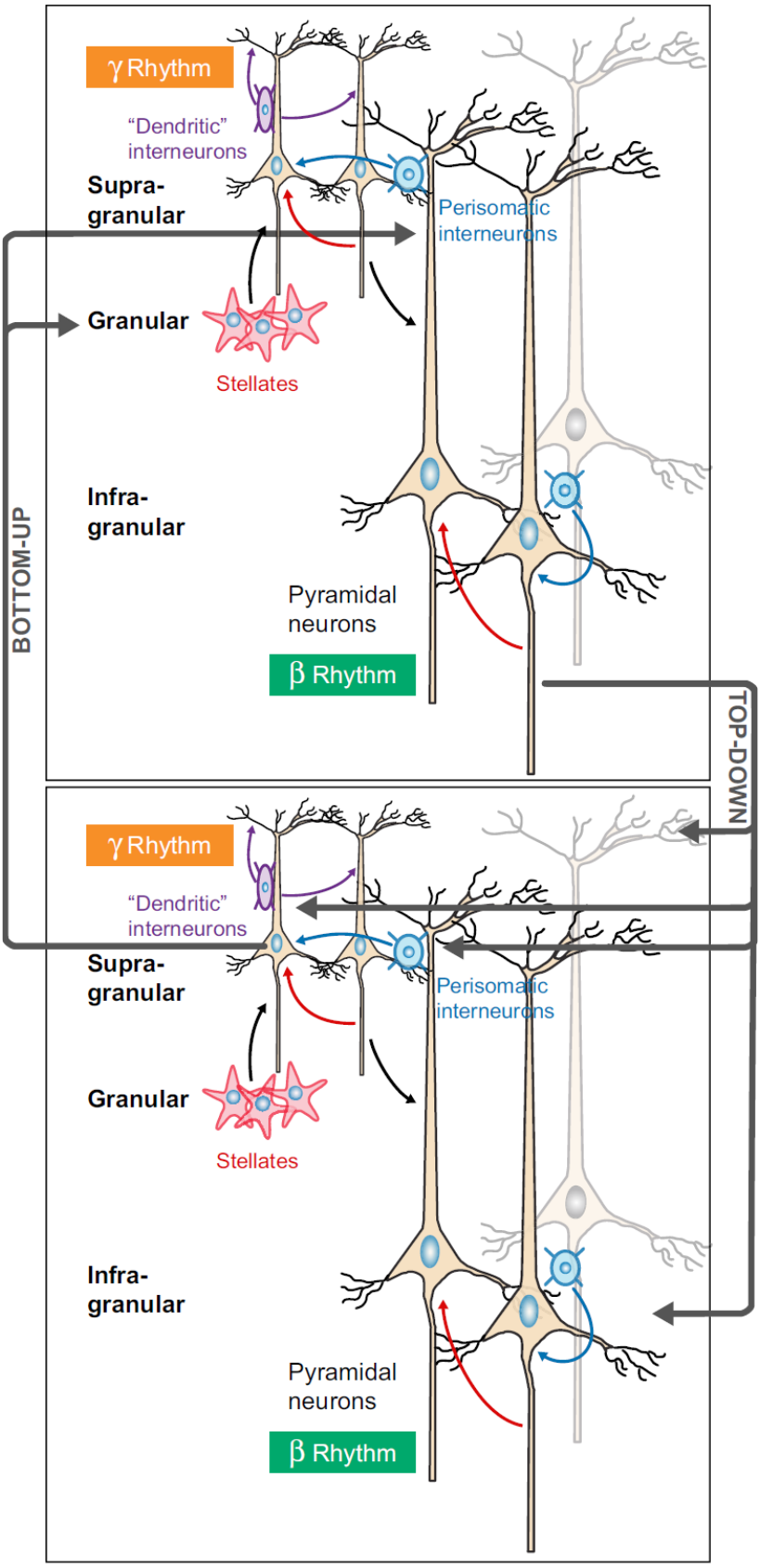


Figure 11. Distance effects of labeling in individual layers. This figure how FB projecting neurons are differentially distributed in L2 and 6 and FF in L3 and 5. The characteristic SLN gradient found in up- and down stream areas shown in figure 2 is due to different distance rules operating in individual layers. Hence the short-distance spread of labeled neurons in L2 coupled with the long-distance spread in L6 leads to the observed decrease in SLN with increasing FB hierarchical distances. Likewise the long-distance spread of labeled neurons in L3 coupled with the short-distance spread in L5 leads to the observed increase in SLN with increasing FF hierarchical distances.



1503

1504 **Figure 12. Schematic circuit for the interplay between bottom-up and top-down signaling**

1505 **characterized by differential frequency-band synchrony.** In a reciprocally connected loop

between a sensory-type area and a cognitive-type area, neural circuits in the superficial layers are endowed with strong intrinsic synaptic connections and generate stochastic oscillations in the gamma frequency range, whereas the deep layers have a propensity to display slower oscillations in the lower beta or alpha frequency range. Top-down projections originate in the deep layers and innervate pyramidal cells (brown), as well as dendrite-targeting (purple) and perisoma-targeting (blue) inhibitory interneurons. In this scenario, low beta/alpha oscillations are directly involved in top-down signaling, which interacts with locally generated gamma oscillations. Adopted with permission from Wang (Wang, 2010).

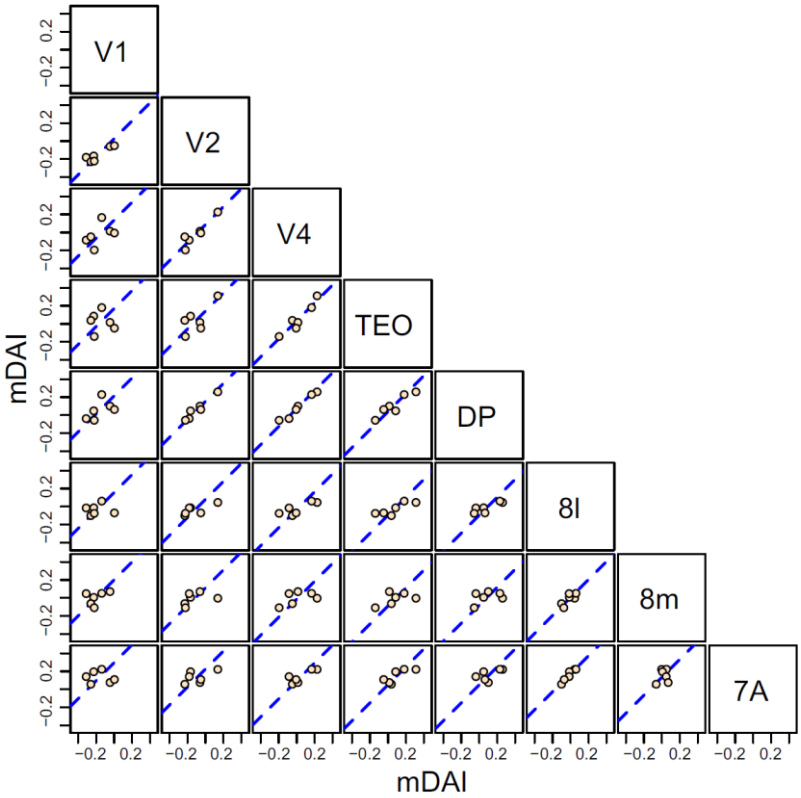


Figure 13. Scatter plots of a hierarchical measure of cortical distance (mDAI) derived by Bastos et al., 2015 of common source areas for pairs of 8 visual areas obtained from contrasting Granger Causality measures in gamma, theta and beta bands. The abscissa of each graph corresponds to the value calculated for the area at the top of the column and the ordinate to the area at the right of the row. Dashed blue line in each plot is the best fit line of unit slope.