ELSEVIER

Brain Research Bulletin 60 (2003) 387–393

BRAIN
RESEARCH
BULLETIN

www.elsevier.com/locate/brainresbull

The evolution of mammalian cortex, from lamination to arealization

Anna Montagnini*, Alessandro Treves

SISSA—Programme in Neuroscience, via Beirut 4, 34014 Trieste, Italy

Received 11 September 2002; received in revised form 7 November 2002; accepted 7 November 2002

Abstract

We analyse some of the most important anatomical and functional features emerging at different stages of mammalian brain evolution in terms of a possible computational advantage. At the transition from reptiles to mammals, a major anatomical change occurs in the originally sensory dorsal cortex. The principal layer of pyramidal cells is split by the insertion of a new layer of granule cells, giving rise to the laminated isocortex. It has been hypothesized that this qualitative change in the evolution of mammalian brains is necessary to support fine topography in their sensory maps. The simulation of neural network models demonstrates that a nonlaminated patch of cortex must compromise between transmitting "where" information, explicitly mapped, topographically, on the cortical sheet, and retrieving "what" information, represented by the distributed firing pattern across neurons. The differentiation of a granular layer is shown in the model to yield a small quantitative advantage, allowing to transmit a slightly better combination of both information types.

Along the same theoretical lines, we are investigating the multiplication of successive sensory areas coding for ever more composite stimuli, such as those in the visual and auditory temporal cortices in primates. In particular we analyse the possible computational advantage for a specific neural population devoted to encode the complex structure of whole stimuli, rather than relying on the coactivation of separate populations encoding their basic elements.

Keywords: Information theory; Neural networks; Cortical function; Modeling

1. The evolution of mammalian cortex

Mammals originate from early reptiles, and their brains [8] can be seen as continuing, with largely *quantitative* variations [13] the evolution of the basic organization blueprint seen in other vertebrates. In this general continuity, we would like to associate with the reptilian-to-mammal transition two major *qualitative* changes in the forebrain, two changes that subsequently define the cortex of mammals as distinct from that of reptilians and birds. In both cases, granule cells are introduced at the input end of a cortical network.

In the hippocampus, as it reorganizes in its mammalian form, the main step is the detachment of the most medial portion, that loses both its continuity with the rest of the cortex at the hippocampal sulcus, and its projections to dorso-lateral cortex [31]. The rest of the medial cortex becomes Ammon's horn, and retains typically cortical pyramidal cells, while the detached cortex becomes the dentate gyrus, with its population of granule cells, that project now, as a sort of preprocessing stage, to the pyramidal cells of field CA3 [3]. The dorsal cortex, instead, reorganizes internally, to become the cerebral neocortex. Most of mammalian neocortex shows the isocortical organization with six layers of cells; the crucial step is often considered to be, however, the insertion of layer IV, a layer of granule cells, in between infragranular and supragranular layers of pyramidal cells. This is prominent in primary sensory cortices, which we take to be good models, in simpler species where they have not specialized further, for what primordial isocortex might have looked like. The main thalamic inputs to cortex terminate in layer IV [9].

After these anatomical changes leading to the mammalian archicortex and neocortex, a further crucial reorganization of the neocortex, by now laminated, occurs in several mammalian species, setting the stage for the evolution of

---

* Corresponding author. Tel.: +39-040-3787531;
fax: +39-040-3787528.
E-mail addresses: montagni@sissa.it (A. Montagnini), ale@sissa.it (A. Treves).

0361-9230/03/$ – see front matter © 2003 Elsevier Science Inc. All rights reserved.
doi:10.1016/S0361-9230(03)00057-1

he species *homo*. Beside the global increase in size of the ieocortex, at least three major reorganizational trends lead o early hominid brains [17]: the first is the reduction of he primary visual striate cortex along with the enlargement f extrastriate visual areas, concerned with higher level ensory processing; second, the frontal lobe reorganizes, ncluding Broca's area, involved in the development of anguage; finally, emispheric specialization increases. All hese phenomena are thought to be crucial for the peculiar levelopment of human intelligent behavior, and are likely o have evolved under the pressure of social needs. We re particularly interested, here, in the first of the three hanges. The development of nonprimary visual areas in umans (accentuating a trend already prevalent among primates as well as among several other advanced mammalian pecies) has to be analyzed in the more general framework f the phenomenon of *arealization*. This corresponds to the mergence of new functionally specialized sensory areas, r *cortical fields*, for one or more modalities, as opposed to he mere enlargement of existing areas. In more evoluted nimals cortical fields multiply [21,22] that are sensitive to lifferent characteristics of sensory stimuli. This seems to appen both "horizontally," with respect to sets of features t equivalent levels of complexity, and "vertically," for hierarchies of increasing complexity (or level of abstraction). mong the newly developed sensory areas in hominids, the entral branch of visual cortex appears as a series of connected areas [12] coding for increasingly complex features. The final purely visual stage (before the areas where information from different sensory modalities is integrated) is in he inferotemporal (IT) cortex, where there is evidence for ighly selective neurons which are activated only by well lefined composite stimuli, such as images of human faces.

What is the evolutionary advantage, for mammals, rought about by these changes? We consider in this paper information-theoretical advantages, that is, quantitative mprovements in carrying out the generic function of the ortex, processing information.

Hippocampus

In the case of the hippocampus, attempts to account for ts remarkable internal organization have been based, since he seminal paper by David Marr [24], on the computational analysis of the role of the hippocampus in memory. The hippocampus is important for spatial memory also in irds. A reasonable hypothesis is that the "invention" of he dentate gyrus enhances its capability, in mammals, to erve as a memory store. Alessandro Treves and Edmund Rolls, building on the approach outlined by David Marr, ave proposed 10 years ago [30] that the new input to CA3 yramidal cells from the mossy fibers (the axons of the lentate granule cells) serves to create memory representations in CA3 richer in information content than they could ave been otherwise. The crucial prediction of this proposal was that the inactivation of the mossy fiber synapses should impair the formation of new hippocampal dependent memories, but *not* the retrieval of previously stored ones. This prediction has recently been verified [23] in mice. Thus a quantitative, information-theoretical advantage may have favored a qualitative change, such as the insertion of the dentate gyrus in the hippocampal circuitry. This idea, still to be tested further, is not explained in detail here. It raises the issue, however, of whether also the insertion of layer IV in the isocortex might be accounted for in quantitative, information-theoretical terms.

3. Neocortex

3. *Layers and maps*

It has long been hypothesized that isocortical lamination appeared together with fine topography in cortical sensory maps [2], pointing at a close relationship between the two phenomena. In early mammals, the main part of the cortex was devoted to the olfactory system, which is not topographic, and whose piriform cortex has never acquired isocortical lamination [16]. The rest of the cortex was largely allocated to the somatosensory, visual and auditory system, perhaps with just one topographic area, or map, each [10]. Each sensory map received its inputs directly from a corresponding portion of the thalamus, in which a distinction has been drawn [20] between its matrix and core nuclei. The matrix, the originally prevalent system, projects diffusely to the upper cortical layers; while the core nuclei, which specialize and become dominant in more advanced species [11], project with topographic precision to layer IV, although their axons contact, there, also the dendrites of pyramidal cells whose somata lie in the upper and deep layers.

3.2. *Higher level sensory areas*

As opposed to the relatively fixed, topographical connectivity between thalamus and sensory maps in primary cortex, the network of cortico–cortical connections has been greatly expanded [1,7], in mammals, by the evolution of multiple, hierarchically organized cortical areas in each sensory system [21,22]. Neural populations in higher level sensory areas (mainly in the visual cortex, which is by large the most developed in primates, but also in the auditory and somatosensory cortices) respond in general to complex combinations of features. Topographical organization is progressively lost, receptive fields broaden, and it becomes harder and harder to map neural activity with respect to intuitive external variables. Moreover, connections to and from subcortical regions become more diffuse; at the same time, perhaps, attentional and emotional modulatory effects become more important. At this stage there is no experimental evidence of a major anatomical abrupt change, driven by a hypothetical computational advantage, comparable with lamination. Nevertheless, it is tempting to speculate, here, too, what informational advantage might have favored a possibly continuous structural and functional trend towards arealization, along the visual and the other sensory pathways. For instance, neural assemblies selectively coding for whole stimuli with ecological meaning (say, faces) rather than for their components (such as mouth- or nose-shapes) could prove to be quantitatively advantageous in processing or in storing information, under certain boundary conditions.

3. *A functional hypothesis for any cortical patch*

The crucial aspect of fine topography in sensory cortices is the precise correspondence between the location of a cortical neuron and the location, on the array of sensory receptors, where a stimulus can best activate that neuron. Simple visual and somatosensory cortices thus comprise two-dimensional maps of the retina and of the body surface, while auditory cortices map sound frequency in one-dimension. Some of the parameters characterizing a stimulus, those corresponding to the position of the receptors it activates, are therefore represented continuously on the cortical sheet. We define them as providing *positional* information. Other parameters, which contribute to identify the stimulus, are not explicitly mapped on the cortex. For example, the exact nature of a tactile stimulus at a fixed spot on the skin, whether it is punctuate or transient or vibrating, and to what extent, are reflected in the exact pattern of activated receptors, and of activated neurons in the cortex, but not directly in the position on the cortical sheet. We define these parameters as providing *identity* information. Some additional parameters are also mapped on the cortical sheet of advanced species, like ocular dominance or orientation in the primary visual cortex of primates. This leads to the formation of so-called columns, or wrapped dimensions, and to the differentiation of layer IV in multiple sublayers. These should be regarded as specializations, which likely came much after the basic cortical lamination scheme had been laid out. Likewise, a modular organization in terms of columns may be present also in higher level cortical areas [14], though a precise functional mapping between stimulus characteristics and module activity has not been quite demonstrated, yet. If it will be, it should be regarded as another later specialization.

The sensory cortices of early mammals therefore received from the thalamus, and had to analyse, information about sensory stimuli of two basic kinds: positional or *where* information, I_p, and identity or *what* information, I_i. Note that cortical processing may at best maintain I_p as it is transmitted by thalamic neurons, whereas I_i can be actively improved. This is made possible by the storage of previous sensory events in terms of distributed efficacy modifications in synaptic systems, in particular on the recurrent collaterals connecting pyramidal cells in sensory cortex. Neural network models of autoassociative memories [18,24] have demonstrated how simple "Hebbian" rules modeling associative synaptic plasticity can induce weight changes that lead to the formation of dynamical attractors [4]. Once an attractor has been formed, a partial cue corresponding, for example, to a noisy or occluded version of a stimulus can take the recurrent network within its basin of attraction, and hence lead to a pattern of activation of cortical neurons, which represents the stored identity of the original stimulus. The storage of new memories and their cue-elicited retrieval is seen here as the generic function of the primitive cortex, blending perception with memory [32]. Specialized to the olfactory sense, this function does not seem to require new cortical machinery to be carried out efficiently. A novel circuitry may instead be advantageous when the generic function is specialized to topographic sensory systems, which have to relay both where and what information, I_p and I_i. This idea is analyzed in Ref. [28] and summarized in Section 4.

Even accepting the argument, sketched in Section 4, that lamination defines a new, computationally advantageous, cortical module, what is the meaning of multiplying such modules along the sensory pathways? What happens, in other words, with information processing in successive stages of sensory cortex? The question can be reformulated in other terms. In the general picture of the cortex performing encoding, storage and retrieval of identity information by means of serial and parallel neural networks, what are the convenient variables to be represented by neural activity in advanced cortical networks? More specifically: what is the sense of allocating neuronal units to code for more composite levels of stimulus identity? We shall try to sketch an approach to this issue in the last part of this paper.

4. Solving the what/where conflict through lamination

The simultaneous processing of what and where information produces an unavoidable conflict between preserving an accurate coding of stimulus position and analyzing stimulus identity. The argument is a quantitative one, and has been described in Ref. [28] using a suitable neural network model. It is summarized in this section.

A patch of cortex (as shown in Fig. 1) is modeled as a wafer of three arrays in which, after an initial uniform network, used as a nonlaminated control, different properties and connectivities are introduced among the arrays, in order to model laminated cortical tissue. The three arrays will then model supragranular, granular and infragranular layers of the isocortex. Feed-forward connections model the afferent inputs from the thalamus, and have a certain cortical spread S_{ff}; recurrent connections, with a spread S_{rc}, link each unit to other units in the same and in the other two layers. Both sets of connections are modifiable through learning, on the basis of a simple Hebbian associative rule. A local pattern of activation is applied to the thalamic units, fed-forward to the cortical patch and processed through the recurrent connections, and then the activity of some of the units in the patch is read out. To separate out "what" and

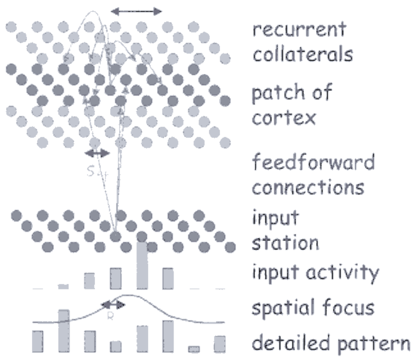

Fig. 1. The model patch that was used in the simulations. At the bottom, a typical input activity pattern is sketched, as a distributed vector of activation across a population of cells. It is obtained as the convolution of the original global pattern of activity (carrying information about the identity of the stimulus) with a spatial focus of activation of width R. Projections from the thalamic input station have a spread S_{ff}, whereas S_{rc} is the spread of collaterals in the three layered cortical patch. In most cases we used 20×20 units per layer, $R = 2$, $S_{\mathrm{rc}} = 8$, $S_{\mathrm{ff}} = 2$–8 and 12 patterns.

"where" information, the input activation, during training, is generated as the product of one of a set of predetermined global patterns by a local focus of activation, defined as a Gaussian function. During a subsequent testing phase, only a partial cue of a pattern is presented in the thalamus, at a certain position, and the cortical network has to retrieve both the position and the identity of the corresponding pattern. Failure is possible of course, and a cue belonging to pattern i might be decoded as stemming from pattern i_{d}. I_{i} is extracted from the frequency table $P(i, i_{\mathrm{d}})$ reporting how many times input pattern i has been decoded as i_{d}

$$I_{\mathrm{i}} = \sum_{i,i_{\mathrm{d}}} P(i, i_{\mathrm{d}}) \log_2 \frac{P(i, i_{\mathrm{d}})}{P(i)P(i_{\mathrm{d}})} \qquad (1)$$

and a similar formula is used for I_{p}. Among the several parameters of the network, the crucial one is S_{ff}. In the nonlaminated, uniform model (in which all layers in the wafer are equivalent), it is intuitive that if the feed-forward connections are focused, S_{ff} small, "where" information can be substantially preserved, but the cortical patch is activated over a limited, almost point-like extent, and it may fail to use efficiently its recurrent collaterals to retrieve "what" information. If, on the other hand, S_{ff} is large, the recurrent collaterals can better use their attractor dynamics, leading to higher I_{i} values, but the spread of activity from thalamus to cortex means degrading I_{p}. This conflict between I_{p} and I_{i} (depicted in Fig. 2) results in a $I_{\mathrm{p}} - I_{\mathrm{i}}$ boundary. Is it possible to go beyond this boundary?

4.1. Differentiation of a granular layer

The insertion of a granule layer (approaching the model of a laminated cortex) is made through three main steps in our simulation

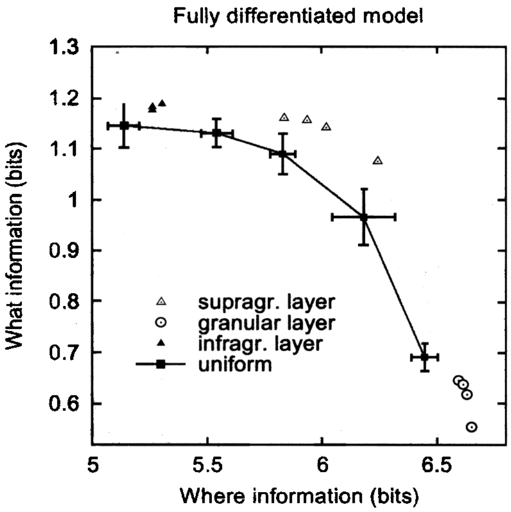

Fig. 2. The curve gives the boundary applying to the uniform, non laminated control model, that is, the I_{i} and I_{p} values obtained with different values of S_{ff}, the spread of feed-forward connections. Standard errors obtained with multiple simulations are also reported. The points are the values obtained with the fully laminated model, for different choices of certain parameters (see Ref. [28]), by decoding the activity in each of the three layers. Three of the data points for the infragranular layer (black triangles) are nearly superimposed.

1. The thalamic afferents to the granular layer are focused, while those to the two pyramidal layers are diffuse.
2. The recurrent collateral system of the granular units is severely restricted. That is, those arriving at layer IV units are fewer in number with respect to the other layers.
3. Model layer IV units follow a nonadaptive dynamics, and are suppressed during training.

The three modifications, combined, produce a slight quantitative advantage in the joint I_{p} and I_{i} values that can be read off pyramidal cell activity (similar to the one seen in Fig. 2, which however refers to the fully laminated model introduced next). The advantage is small, but the model cortical patch used is also tiny, and the expectation is that the difference between uniform and granulated patches would scale up, as the size of the patch reaches realistic values.

A prediction immediately testable in vivo is that recordings of activity in layer IV should, relative to the pyramidal layers, yield more positional information, and less identity information.

4.2. Differentiating infragranular from supragranular connections

Why does isocortex have pyramidal layers both above and below the granule layer? In the granulated model considered above, the supragranular and infragranular layers are still identical in all their properties. In the real cortex, however, the supragranular and infragranular layers differ in several ways. One difference which likely goes back hundreds of millions of years is in their efferent projections. The supragranular layers project mainly onward, to the next stage of processing. In advanced mammalian species, this means they project to the next cortical areas in the sensory or motor stream [5]. The infragranular layers project mainly backward [6], or subcortically. Among their chief target structures are the very thalamic nuclei from which projections arise to layer IV. It is clear that having different preferential targets would in principle favour different mixes of what and where information. In particular, cortical units that project back to the thalamus would not need to repeat to the thalamus "where" a stimulus is, since this information is already coded, and more accurately, in the activity of thalamic units. They would rather report identity information. Units that project to further stages of cortical processing, on the other hand, should balance the "what" added value with the preservation of positional information—the mix that we have so far considered optimal for pyramidal units in general.

In addition to the difference in extrinsic projections, the *intrinsic* connectivity of supragranular and infragranular layers also differs. In order to capture these differences, albeit at a very crude level, in a fully laminated version of the model patch, all direct projections from layer IV to layer V were removed, and replaced with an equal number of projections from layer III to layer V. All other parameters remained as in the granulated model. The effect of this differentiation can be appreciated by decoding the activity in the three layers, separately, as shown in Fig. 2. From layer IV one can extract, as before, a large I_p but limited I_i; from layer III one obtains a balanced mix (similar to what was obtained with the granulated model—not reported here). From layer V, on the other hand, one can extract predominantly "what" information, I_i, at the price of a rather reduced I_p content. These results also translate into obvious predictions for neurophysiological recording experiments.

5. Faces as a salient example of composite stimuli

Electrophysiological recordings in IT cortex in monkeys show that most cells are selectively activated by composite stimuli of the type shown in Fig. 3. They may also be organized in some columnar structure. Among the critical stimuli represented in this area, a very important class is that of primate and human faces. The so-called face cells are actually found to be widespread in IT and in other areas of cortex, in monkeys. Moreover, face cells are present in a larger fraction than any other "specific-shape–cell" (say, a hand–cell or, more unlikely, a chair–cell). The peculiarity of faces with respect to other natural visual stimuli can be intuitively understood in terms of their relevance in social interaction. A number of behavioral and neuropsychological studies point at the salience of face processing among other sensory-cognitive processes. In what exactly face processing differs from other complex object processing, though, is still an open question. Moreover, there may be two opposite explanations about the origins of this functional distinction: the first is based on an evolution driven genetic predisposition [19] of a specific face area. A second, more recent point of view [15] is that a face processing specialization is achieved only through repeated daily experience; any other object would be processed in the same way if we had the same expertise with it. A lot of experimental work (both in physiology and behaviour) is still needed to gain insight in these issues. What we propose here is a computational model to investigate some generic aspects of processing composite stimuli along multiple modules, as faces might be processed along IT cortex, disregarding all the crucial social, emotional and semantic implications of face processing.

Fig. 3. Examples of the composite visual stimuli which typically elicit selective activation of IT neurons. Derived from Tamura and Tanaka (2001) (Cerebral Cortex, vol. 11, p. 384).

5.1. Allocating new modules to represent whole stimuli

We implemented two simple modular neural networks and tested them on two different tasks: a recall task, requiring detailed memory of the elements (nose, mouth, eyes, ...) comprising the complete stimulus (the face); and a recognition task, directly associated with identifying the face. Rather than an experimental validation based on elecrophysiological recordings (which is quite difficult to obtain at this stage) we propose to compare the quantitative performance of the artificial neural network with that of a simulated optimal subject performing the same kind of test [26]. This *ideal subject* is shown with a sequence of examples of noisy versions of the complex stimuli (faces) and it is able to keep them in memory, and to use this information in an optimal way to solve both the recall and the recognition task. The performance comparison between *ideal subject* and the neural networks is not described here in detail.

In our network simulations, faces are represented as ordered vectors of values, each value identifying a component class. The fixed spatial relationship among face components has been shown to play a crucial role in face processing, and we include it in the input vector structure. Both networks are one layered networks, with each module dedicated to code information about one single element of the complete stimulus. This should ideally reproduce the modular organization in IT cortex, with columns representing basic shapes, whose geometrical complexity could approximate that of elements like a nose, a mouth, the eyes, etc. The neural populations are fully connected by means of both intramodular and intermodular collaterals.

We trained and tested the network in a face processing test in which we present partial and noisy cues of the stimulus vectors. The units belonging to the modules, which code elements presented in the partial cue, receive a direct external input, which dominates their activity. The units processing the elements to be guessed, on the other side, update their activity summing up the inputs coming from the active collaterals. After pattern completion (the recall task), the network can learn, in the sense that the weights of the connections are modified according to the correct complete pattern of activity. The two architectures differ in their connectivity. In the first one (the *independent modules* network) intermodular connections are active only to set the initial state and are subsequently set to 0, so that every module evolves independently to an attractor state. This network model reproduces a visual object coding area with no face cells and, in practice, no unitary representation for composite stimuli. In the second architecture, connections between modules remain active as long as the intramodular ones, and the network dynamics evolves rather as a whole. Once again, no face cells are considered, but now processing the composite object dominates with respect to the single components, in the population activity. We tested both networks with the recall and recognition task. The correlated modules network performance turns out to be better both for recall (as Fig. 4 illustrates) and, across most simulations, also for recognition (not shown).

We are implementing now a third network architecture, in which the number of active collaterals is differently balanced, and there are some units (face cells, finally!) devoted to coding the whole stimulus. Note that these face units, perhaps like real face cells [27], comprise a distinct neural population supporting *distributed* representations of different faces. By no means these should be interpreted as "grandmother cells," that is, an ensemble of neurons each of which fires only to the sight of the face of a specific person.

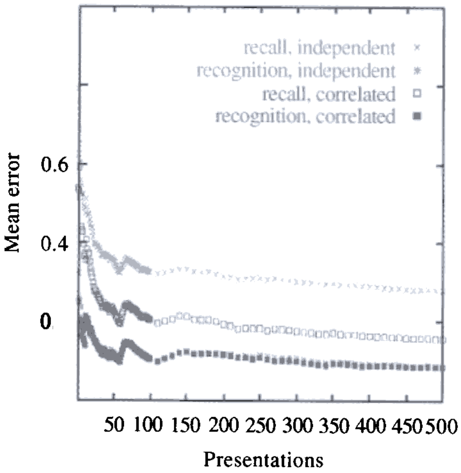

Fig. 4. Average normalized error for the independent and the correlated network in the recall and in the recognition task. The dashed horizontal line represents an intrinsic lower bound for this error.

6. Discussion

A small quantitative advantage in relaying combined positional and identity information is proposed to have driven the differentiation of neocortical layers in mammals. The proposal is in line with speculations arising from traditional comparative neuroanatomy, and is supported by the simulation of a simple neural network model. The truly important elements, here, are the mutual constraint between relaying where information and retrieving what information, evident in the uniform model, and the quantitative comparison with a laminated model with no more units and/or connections. The what/where conflict is manifest also in rather different models, like those developed by Meffin and Treves [25], and essentially, it requires only the separate measurability of I_i and I_p to be demonstrated, whatever the remaining details of the model. The small advantage of the laminated patch is expected to scale up, as mentioned above, when the model, and in particular synaptic numbers, are scaled up to realistic values. It should be considered, however, that even a slight quantitative advantage may be selected for, once replicated over millions of sensory experiences per individuals, and over millions of generations in the course of mammalian evolution. One attractive point of this model is that some of the results can be experimentally checked in a relatively straightforward way: with an appropriate experimental design, I_p and I_i can be measured in vivo from populations of tens of units [29] recorded in well identified layers. Differences in the information content of the activity of populations of cells in different layers can be used to try and falsify the model.

On the contrary, an experimental confirmation of whatever quantitative result of the IT cortex model would be hard to obtain, now, from electrophysiology. Quantitative theoretical approaches to higher level functions, such as those ascribed to IT cortex, can probably find a more immediate experimental check through some psychophysical test, at this stage. Even though this second model needs still to be refined, it shares with the first, about laminated cortex, the same general perspective. Understanding specific mammalian mechanisms of information representation and retrieval requires, we propose, quantitative (information-theoretical) analyses at the level of populations of individual neurons. Far from being exhaustive, this kind of approach has the advantage of dealing with notions of sufficient abstraction and generality as to apply to each sensory cortex. This abstraction is crucial, we believe, to explain the appearance, in evolution, of what we refer to as the universal neocortical microchip; and to explain its ever more extensive use, through the process of arealization.

Acknowledgements

We are indebted to Hamish Meffin, Mathew Diamond and Bill Bialek, for discussions on different aspects of this work. Partial support from the Human Frontier Science Programme grant RG0110/1998-B. AT is currently on sabbatical at the ITP in Santa Barbara, whose hospitality is gratefully acknowledged.

References

[1] M. Abeles, Corticonics: Neural Circuits of the Cerebral Cortex, Cambridge University Press, Cambridge, 1991.

[2] J. Allman, Evolution of neocortex, in: E.G. Jones, A. Peters (Eds.), Comparative Structure and Evolution of Cerebral Cortex. Cerebral Cortex, vol. 8A, Plenum Publishing Co., New York, 1990, pp. 269–283.

[3] D.G. Amaral, N. Ishizuka, B. Claiborne, Neurons, numbers and the hippocampal network, Prog. Brain Res. 83 (1990) 1–11.

[4] D.J. Amit, The Hebbian paradigm reintegrated: local reverberations as internal representations, Behav. Brain Sci. 18 (1995) 617–657.

[5] H. Barbas, N. Rempel-Clower, Cortical structure predicts the pattern of cortico-cortical connections, Cereb. Cortex 7 (1997) 635–646.

[6] A. Batardiere, P. Barone, C. Dehay, H. Kennedy, Area-specific laminar distribution of cortical feedback neurons projecting to cat area 17: quantitative analysis in the adult and during ontogeny, J. Comp. Neurol. 396 (1998) 493–510.

[7] V. Braitenberg, A. Schütz, Anatomy of the Cortex, Springer-Verlag, Berlin, 1991.

[8] R.L. Carroll, Vertebrate Paleontology and Evolution, W H Freeman & Co., New York, 1988.

[9] I.T. Diamond, M. Conley, K. Itoh, D. Fitzpatrick, Laminar organization of geniculocortical projections in *Galago senegalensis* and *Aotus trivirgatus*, J. Comp. Neurol. 242 (1985) 610.

[10] I.T. Diamond, W.C. Hall, Evolution of neocortex, Science 164 (1969) 251–262.

[11] R.P. Erickson, W.C. Hall, J.A. Jane, M. Snyder, I.T. Diamond, Organization of the posterior dorsal thalamus of the hedgehog, J. Comp. Neurol. 131 (1967) 103–130.

[12] D.J. Felleman, D.C. Van Essen, Distributed hierarchical processing in the primate cerebral cortex, Cereb. Cortex 1 (1991) 1–47.

[13] B.L. Finlay, R.B. Darlington, Linked regularities in the development and evolution of mammalian brains, Science 268 (1995) 1578–1584.

[14] I. Fujita, K. Tanaka, M. Ito, K. Cheng, Columns for visual features of objects in monkey inferotemporal cortex, Nature 360 (1992) 343–346.

[15] I. Gauthier, M.J. Tarr, Becoming a "Greeble" expert: exploring the face recognition mechanism, Vision Res. 37 (12) (1997) 1673–1682.

[16] L.B. Haberly, Comparative aspects of olfactory cortex, in: E.G. Jones, A. Peters (Eds.), Comparative Structure and Evolution of Cerebral Cortex. Cerebral Cortex, vol. 8B, Plenum Publishing Co., New York, 1990, pp. 137–166.

[17] R.L. Holloway, Origins of the Human Brain, Clarendon Press, Oxford, 1996.

[18] J.J. Hopfield, Neural networks and physical systems with emergent collective computational capabilities, Proc. Natl. Acad. Sci. U.S.A. 79 (1982) 2554–2558.

[19] M.H. Johnson, J. Morton, Biology and Cognitive Development. The Case of Face Recognition, Blackwell, Oxford, 1991.

[20] E.G. Jones, Viewpoint: the core and matrix of thalamic organization, Neuroscience 85 (1998) 331–345.

[21] J.H. Kaas, Evolution of multiple areas and modules within neocortex, Perspect. Dev. Neurobiol. 1 (1993) 101–107.

[22] L. Krubitzer, The organization of neocortex in mammals: are species differences really so different? Trends Neurosci. 18 (1995) 408–417.

[23] J.-M. Lassalle, T. Bataille, H. Halley, Reversible inactivation of the hippocampal mossy fiber synapses in mice impairs spatial learning, but neither consolidation nor memory retrieval, in the Morris navigation task, Neurobiol. Learn. Mem. 73 (2000) 243–257.

[24] D. Marr, Simple memory: a theory for archicortex, Philos. Trans. R. Soc. Lond. B 262 (1971) 24–81.

[25] H. Meffin, A. Treves, Information about location versus content in a neocortical model of autoassociation memory, in: Proceedings of the Fifth IBRO World Congress of Neuroscience, 1993 (abstract 123).

[26] A. Montagnini, Statistical learning of human faces, in: Proceedings of the Fourth International Conference on Cognitive Modelling, July 26–28, 2001, George Mason University, Fairfax-VA, Lawrence Erlabum Associates Publishers, Mahawah, NJ, 2001, pp. 285–286.

[27] E.T. Rolls, M.J. Tovée, Sparseness of the neuronal representation of stimuli in the primate temporal cortex, J. Neurophys. 73 (1995) 713–726.

[28] A. Treves, Computational constraints that may have favoured the lamination of sensory cortex, J. Comput. Neurosci. 14 (2003) 271–282.

[29] A. Treves, Information coding in higher sensory and memory areas, in: F. Moss, S. Gielen (Eds.), Neuro-Informatics and Neural Modelling. Handbook of Biological Physics, vol. 4, Elsevier, Amsterdam, 2001, pp. 825–852.

[30] A. Treves, E.T. Rolls, Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network, Hippocampus 2 (1992) 189–199.

[31] P.S. Ulinski, The cerebral cortex of reptiles, in: E.G. Jones, A. Peters (Eds.), Comparative Structure and Evolution of Cerebral Cortex, Cerebral Cortex, vol. 8A, Plenum Publishing Co., New York, 1990, pp. 139–215.

[32] I.C. Whitfield, The object of the sensory cortex, Brain Behav. Evol. 16 (1979) 129–154.