Canonical cortical circuits: current evidence and theoretical implications

Fioravante Capone1,2, Matteo Paolucci1,2, Federica Assenza1,2, Nicoletta Brunelli1,2, Lorenzo Ricci1,2, Lucia Florio1,2, Vincenzo Di Lazzaro1,2

1Unit of Neurology, Neurophysiology, Neurobiology, Department of Medicine, Università Campus Bio-Medico di Roma, Rome, Italy; 2Fondazione Alberto Sordi – Research Institute for Aging, Rome, Italy

Abstract: Neurophysiological and neuroanatomical studies have found that the same basic structural and functional organization of neuronal circuits exists throughout the cortex. This kind of cortical organization, termed canonical circuit, has been functionally demonstrated primarily by studies involving visual striate cortex, and then, the concept has been extended to different cortical areas. In brief, the canonical circuit is composed of superficial pyramidal neurons of layers II/III receiving different inputs and deep pyramidal neurons of layer V that are responsible for cortex output. Superficial and deep pyramidal neurons are reciprocally connected, and inhibitory interneurons participate in modulating the activity of the circuit. The main intuition of this model is that the entire cortical network could be modeled as the repetition of relatively simple modules composed of relatively few types of excitatory and inhibitory, highly interconnected neurons. We will review the origin and the application of the canonical cortical circuit model in the six sections of this paper. The first section (The origins of the concept of canonical circuit: the cat visual cortex) reviews the experiments performed in the cat visual cortex, from the origin of the concept of canonical circuit to the most recent developments in the modelization of cortex. The second (The canonical circuit in neocortex) and third (Toward a canonical circuit in agranular cortex) sections try to extend the concept of canonical circuit to other cortical areas, providing some significant examples of circuit functioning in different cytoarchitectonic contexts. The fourth section (Extending the concept of canonical circuit to economic decisions circuits) reviews the experiments conducted in humans by using transcranial magnetic stimulation to demonstrate the validity of the canonical cortical circuit model. The fifth section (Extending the concept of canonical circuit to economic decisions circuits) explores the hypothesis that also complex human behaviors such as economic decision-making could also be explained in terms of canonical cortical circuit. The final section (Conclusion) provides a critical point of view, evidencing the limits of the available data and tracking directions for future research.

Keywords: canonical cortical circuit, cortex, cortical layers, cortical physiology

Introduction

Neurophysiological studies and neuroanatomical investigations enabled to find that the same basic morphological and functional organization of excitatory and inhibitory neuronal circuits exists throughout the neocortex.1 This kind of cortical organization, renamed “canonical circuit”, has been functionally demonstrated primarily by studies involving visual striate cortex,2 because this area is the most developed granular cortex in terms of differentiation of cortical layers. In brief, a canonical cortical circuit is composed of superficial pyramidal neurons of layers II/III receiving different inputs
that are analyzed and processed. Deep pyramidal neurons of layer V are responsible for cortex output. Superficial and deep pyramidal neurons are reciprocally connected, and inhibitory interneurons participate in modulating the activity of the circuit. The general principle of functioning of the canonical cortical circuit is represented by strong recurrent, intralaminar connections and rather weak connections between layers. This principle represents the basis of such circuit and makes it powerful and flexible.

Here, we review the discovery and the application of the notion of canonical cortical circuit in functionally and structurally different areas of the human cortex. This review is composed of six sections.

The first section reviews the experiments performed in the cat visual cortex, from the origin of the concept of canonical circuit to the most recent developments in the modeling of cortex.

The second section tries to extend the concept of canonical circuit to other areas of the neocortex, providing some significant examples of circuit functioning in different contexts.

The third section evaluates the relationship between the cytoarchitectonic structure and the functionality of cortical circuitry, showing that the canonical circuit model can also be applied, with minor changes, to less cytoarchitectonically differentiated areas such as the primary olfactory cortex (PC).

The fourth section analyses the applicability of the canonical circuit model to the intact human brain by means of transcranial magnetic stimulation (TMS). The available data show that the effects of TMS on motor cortex can be explained by incorporating the available knowledge about the corticospinal volleys evoked by TMS into the anatomical and computational characteristics of the canonical circuit model of the neocortex.

The fifth section explores the hypothesis that complex human behaviors such as economic decision-making could also be explained in terms of canonical cortical circuit.

The final section provides a critical point of view, evidencing the limits and shortcomings of the available data and tracking directions for future research.

The aim of this paper is to provide a deeper knowledge of the canonical cortical circuits to facilitate the translation of this model from in vitro cellular studies to the interpretation of complex human behaviors.

The origins of the concept of canonical circuit: the cat visual cortex

The idea that the basic architecture of the cortex could be understood in terms of “canonical” circuits built of relatively few types of excitatory and inhibitory neurons arises from the neurophysiological study of the cat visual cortex.³⁴

One of the most important works about this topic was performed by Douglas and Martin⁵ more than 20 years ago. In this pioneering study, the authors developed a microcircuit model in order to simulate the data obtained from in vivo intracellular recordings of visual cortex (Brodmann area 17). Their intuition was that the cortical network of highly interconnected neurons could be modeled as the repetition of relatively simple modules.

Experimental data were obtained by recording the intracellular activity evoked by electrical stimulation of the lateral geniculate afferents. Responses consisted of a short-lasting depolarizing potential (excitatory postsynaptic potential or EPSP), followed by a long-lasting hyperpolarizing response (inhibitory postsynaptic potential or IPSP) (Figure 1). Further analysis revealed two different patterns of response correlated with the cortical layer in which they were found. In particular, neurons located in superficial layers (layers II and III) show a longer EPSP and a slower evolution of the IPSP compared to neurons located in deeper layers (layers V + VI). The authors, using the data obtained from the literature about the connections between the different cortical layers, attempted to build a model to simulate data originating from the intracellular recordings.

![Figure 1 EPSP and IPSP postsynaptic potentials. Notes: Electrical stimulation of the lateral geniculate afferents induces an electrical intracellular activity consisting of a short-lasting depolarizing response (EPSP), followed by a long-lasting hyperpolarizing potential (IPSP). Abbreviations: EPSP, excitatory postsynaptic potential; IPSP, inhibitory postsynaptic potential.](https://www.dovepress.com/)

The model circuit that more closely resembles the actual responses was composed of three interacting populations of neurons: one group of inhibitory cells and two groups of excitatory cells, representing superficial (layers II + III) and deep (layers V + VI) pyramidal neurons. Simplifying the functioning of such a microcircuit, a short thalamic input reaches both cortical excitatory and inhibitory neurons; superficial excitatory neurons in layers II/III have a long-lasting depolarization, while deeper excitatory neurons receive inhibition quickly mediated by GABA$_A$ receptors; a GABA$_A$ inhibition is responsible of the latest phase of the response, reflected by a sustained hyperpolarization. The activation of the superficial layers is then responsible for the amplification of the thalamic input. The mutual columnar connections between the superficial layer, deeper layer, and inhibitory interneurons are used to adjust the balance between excitation and inhibition and finely select the output of the circuit.

The experimental data obtained in cat visual cortex validated by the aforementioned computer-based model has substantially contributed to the development of a common “canonical” circuit for the entire neocortex.

However, the interest in cat visual cortex has not ended yet. Recently, in fact, Binzegger et al. used the primary visual cortex as a model attempting to assign a source to every synapse in a single cortical area. They developed a quantitative description of the circuits in cat primary visual cortex through a complete three-dimensional reconstruction of axons and dendrites of single neurons and thalamic afferents labeled with horseradish peroxidase during intracellular recordings in vivo. Their data confirmed the concept of the canonical circuit as a functional module based on interlayer connections and also added some complexity to this simplification. Indeed, they showed that a significant number of synapses (~34%) in cat visual cortex are mediated by axons running horizontally within a cortical layer. Thus, according to this view, the functional architecture of visual cortex relies on two main pathways connecting inhibitory and excitatory neurons: a vertical one (canonical circuit) based on the connections between neurons located in different layers and a horizontal one (self-innervation) based on the connections between neurons located in the same layer.

In conclusion, the experiments involving cat’s visual cortex have helped trace a significant path in the discovery of cortical circuitry, raising the possibility to define a functional neurophysiological model that could allow us to draw a definitive map of the almost-infinite universe of cortical networks.

The canonical circuit in neocortex

The promising results obtained in the study of visual cortex have created great interest regarding the possibility to extend the concept of canonical circuit to other areas of the neocortex.

Recent evidence suggests that very few changes are needed to adapt such a model to other cortical regions such as the frontal eye fields (FEF). This area, located in the frontal cortex, controls the voluntary eye movements. However, the local cortical circuit of the FEF is not well known. Recently, Heinzle et al. proposed a model of the local circuit of FEF that follows the layered architecture of the neocortex and obeys the principles of canonical circuit. This model has been applied to explain the saccadic eye movements. Such a function requires several computational steps: 1) the selection of a target; 2) the allocation of attentional resources to the location of the intended target; and 3) the generation of a motor output that drives the eye movements. According to the authors, cortical layers within the FEF are arranged in a functional order and perform different tasks. The neurons located in layer IV receive the visual input. They form a visual saliency map for rapid scanning. The neurons located in layer II/III transform the phasic signal of layer IV into an attentional signal for the position of the selected target and store it until the time of the saccade. Moreover, they are in connection with other cognitive areas, different from FEF, involved in the recognition of the visual features of the target. When a target is “recognized”, the neurons of layer II/III receive a signal-back from these areas and activate the motor neurons in layer V. Therefore, layer II/III could be interpreted as generating both an attentional signal and a motor plan, while layer V represents the motor output of the circuit. In addition, layer V sends afferents to the layer VI, which, in turn, projects back to layer IV and influences the visual selection. The validity of this model was tested by performing different visual tasks such as saccades, antisaccades, and scanning. The model was able to reliably reproduce experimental data.

Auditory cortex (AI) represents another excellent example to better understand the functional architecture of the canonical microcircuit. Indeed, the information processing in AI is progressive and becomes more complex and synergistic, as the auditory information moves from thalamic input to cortical output layers. The processing strategy in AI fits well with the model of canonical circuit, because different types of cells interact with each other within a columnar distribution in which each layer performs a specific processing function. A recent study provided new insights about these mechanisms.
The authors recorded simultaneously from all cortical layers the activity induced by acoustic stimulus and estimated several parameters to describe the functionality of the AI. They found that supragranular (layers II/IIIa), granular (layers IIIb/IV), and infragranular (V/VI) layers may be considered as separate nuclei, with significantly different processing in each, although still maintaining interdependence. The neurons process acoustic stimuli following an ordinal functional progression from granular to supragranular and then to infragranular layers, with a progressive increase in the complexity and cooperativity of the computational properties. Interestingly, this study demonstrates that different layers within the canonical circuit differ more in processing strategies than in processing content. The demonstration that processing strategies are layer-dependent but largely independent from the stimulus content further confirms the canonical circuit as a model valid for the entire cortex.

Pyramidal neurons and different kinds of interneurons are considered the building blocks of the canonical circuit, and many attempts were made to identify the different cell classes and their functions in the circuit. In particular, this issue has been addressed in the study of the prefrontal cortex (PFC). Cell classes in this area may be classified according to the shape of the action potential as the main feature to distinguish putative (narrow spiking [NS]) interneurons and (broad spiking [BS]) pyramidal cells. Therefore, these cells may be distinguished according to four highly different sets of firing properties, which could have fundamental implications for mediating the balance between excitation and inhibition that control information flow and ultimately shape the function of the prefrontal microcircuit.

Ardid et al have identified seven cell classes in the prefrontal region of macaques engaged in an attention task, based on electrophysiological measures such as spike waveform, firing rate, and firing variability. First of all, they classified four BS and three NS cell classes dissociated by how sparse, bursty, or regular they fire.

BS cell classes were considered to be pyramidal cells, and NS cell classes were considered to be inhibitory cells. Then, they tried to assign to each cell class a specific function within the canonical circuit. For instance, two BS classes (B1 and B4) show sparse, bursty firing, and a phase synchronization at θ (4–7 Hz) and β (16–31 Hz) bands of the local field potential. Because of these properties, B1 and B4 cells can flexibly respond to network activation at varying frequencies and may be involved in integrating long-range inputs, in locally processing information, and in generating the outputs of the microcircuit. Two NS (N3 and N4) and two BS (B2 and B3) cell classes show regular firing and higher rate with marginal synchronization phase. These properties make these cells set the balance between excitation and inhibition in the microcircuit. Two NS classes (N1 and N2) have irregular firing and selective synchronization to either θ or β bands. The main contribution of such kind of studies is to start to establish a link, within the canonical circuit, between the in vitro cell-type characterization and in vivo brain functions.

**Toward a canonical circuit in agranular cortex**

The canonical cortical microcircuit described in above-mentioned works has been developed in neocortex, or granular cortex, which represents the most differentiated cortical region in which all six cortical layers are represented. Agranular cortical regions have a less differentiated architecture. Cytoarchitectonic differentiation gradually changes across the cortex: the spectrum of differentiation ranges from visual striate cortex, the most clearly eulaminated area even in the rodent brain, to agranular areas that lack the inner granular layer (layer IV) and have few identifiable layers as well as very low neuron density. The variation in local cortical structure needs to be taken into account when considering canonical circuit because it is unlikely for the patterns of intra- and interlaminar connections to be uniform in spite of strong cytoarchitectural variations. The available literature on intrinsic interlaminar circuitry in the agranular frontal cortex has been conducted mainly on the rodent brain because of the relative abundance of experimental data available for this animal model, and enables the comparison to a recent rendering of the intrinsic circuitry in the striate cortex. Primary somatosensory cortex, although still clearly eulaminated, is already much less dense and comprises fewer distinguishable sublayers, while primary motor cortex is even less cytoarchitectonically differentiated, but it is still granular. Kätzel et al reported relatively uniform patterns of intralaminar inhibition across these three cortical regions, while interlaminar inhibitory-to-excitatory connectivity differed substantially. In striate cortex, a considerable interlaminar inhibition was observed between all layers. In primary somatosensory cortex, a similar pattern of interlaminar inhibition was reported, but without inhibition between layers II/III and layers V/VI. In primary motor cortex, in contrast, no substantial inhibition between different layers was evident. Thus, across the three sampled regions, interlaminar inhibitory-to-excitatory connectivity was progressively weaker in less cytoarchitectonically
Differentiated areas. It is attractive to investigate how some areas of the cortex, despite their relatively simple architecture, accomplish the sophisticated task of sensory processing and integration. In mammals, the PC or piriform is a trilaminar paleocortex, which is thought to perform a high-level synthetic role, constructing unified “odor images” from the odor components identified by the olfactory bulb. Suzuki and Bekkers studied in the PC the major classes of neurons and the microcircuits in which they are embedded, by using in vitro patch-clamp recordings from slices of PC of mice. They showed that semilunar cells (SL) and superficial pyramidal cells (SP) – the two morphologically distinctive types of neurons contained in layer II of the PC – differ in their contribution to cortical microcircuit: SL cells receive stronger afferent excitation from the olfactory bulb, whereas SP cells receive stronger associational excitation from other principal neurons within the cortex. These results suggest that SL and SP cells are specialized to provide two levels of synaptic processing (afferent and associational, respectively). Godlove et al investigated whether an agranular frontal cortex (supplementary eye field, SEF) exhibits features of laminar circuitry similar to those observed in granular cortex. They showed that the laminar pattern of current source density derived from the local field potential corresponds in many ways to the pattern of the canonical circuit model developed from data collected in granular primary visual cortex. More recently, Beul and Hilgetag hypothesized a speculative model of intrinsic circuitry in cortex by assimilating data from the available literature on inter- and intralaminar connectivity in the agranular frontal cortex of the rodent brain: intralaminar connectivity in agranular cortex is similar to that in granular cortex, but interlaminar connectivity significantly differs because of a weakening of inhibition. This study represents an important starting point for discussion and for further experimental efforts to elucidate circuitry outside of visual striate cortex.

**Canonical circuits in human motor cortex: the contribution of TMS**

In primary motor cortex (M1) – an area of granular cortex with a well-differentiated architecture (although less of visual cortex) – the organization of cortical circuits could be studied in vivo by using TMS experiments. TMS can activate the human brain through the intact scalp. Usually, the effects induced by TMS on M1 are measured as motor-evoked potentials (MEPs) by recording electromyography of the muscles of the limbs. An alternative way of assessing the effects of TMS is to directly record the synaptic activity evoked by stimulation of M1 from epidural electrodes implanted chronically in the spinal cord for relief from pain. From 1998 onward, Di Lazzaro and Ziemann have performed an extensive series of studies using this approach.

A magnetic stimulation of M1 produces a volley composed of a first wave (D-wave) that is determined by the direct activation of pyramidal neurons, and subsequent waves (indirect or I-waves) that are an expression of indirect, trans-synaptic activation of pyramidal cells.

Interestingly, the direction and the intensity of magnetic stimulation influence the evocation of these waves. For instance, when the direction of the stimulating current is reversed from posterior–anterior to an anterior–posterior direction, the descending volleys have different latencies and duration.

By incorporating the available knowledge about the corticospinal volleys evoked by TMS into the anatomical and computational characteristics of the canonical circuit, the main characteristics of the D- and I-waves, including their regular and rhythmic nature, their stimulus intensity dependence, and their pharmacological modulation, can be elegantly explained (Figure 2). According to this perspective, the first I-wave (I1) represents the monosynaptic activation of the pyramidal neuron by neurons of layer II/III; the following I-waves (later I-waves) are produced by the reverberant circuit between pyramidal neurons located in layer V, excitatory neurons located in layer II/III, and GABAergic inhibitory interneurons. The D-wave originates from direct excitation of corticospinal axons in the subcortical white matter at some distance from the cell body of neurons located in layer V. I-waves are separated from each other by approximately 1.5 ms, that is the mean transmission delay between layer II/III neurons and layer V pyramidal neurons. TMS induces a strong depolarization of the excitatory cells of layer II/III, which, in turn, leads to the recruitment of fully synchronized clusters of pyramidal neurons of layer V and inhibitory neurons producing a high-frequency (~670 Hz) repetitive discharge of the corticospinal axons. The role of the inhibitory circuits is crucial in controlling the firing of the excitatory networks to produce a high-frequency discharge.

**Extending the concept of canonical circuit to economic decisions circuits**

One fascinating hypothesis is that complex human behaviors such as economic decision-making could also be explained in terms of canonical cortical circuit. Indeed, the network described in the studies of Padoa-Schioppa
et al. meets both the criteria required for a canonical microcircuit: 1) the existence of a stereotyped relationship between different kinds of neurons, each one charged with a specific task; and 2) the presence of an adequate anatomical background.

Regarding the first point, the authors draw a model that reproduces the relationship between different types of cells. According to their view, the network for economic decisions is composed of three types of cells: the “offer value” cells, which encode the value of individual presented goods; the “chosen juice” cells, which encode the identity of the chosen good and thus the binary decision outcome; and the “chosen value” cells, which encode the value of the chosen goods. In terms of canonical circuit, the “offer value” cells represent the input layer, while the “chosen juice” cells represent the output layer of this network; finally, the “chosen value” cells may work as inhibitory interneurons. It is interesting that the authors themselves have sought a parallel between their “economic decisions” network and Wang’s model for perceptual decisions, as a further demonstration that the basic mechanisms of cortical circuits could be applicable in different contexts. Applying Wang’s model, the authors suggest that the mechanisms of “economic decisions” network are also based on the recurrent excitation of the pyramidal neurons that act as “chosen juice” cells, and inhibition is mediated by the “chosen value” cells.

For that which concerns anatomical requirements, the “economic decisions” circuit has been located in the orbitofrontal cortex, the ventromedial portion of prefrontal cortex. This area exhibits several progressive stages of laminar development, with a more differentiated layer organization of cortex in a posterior–anterior direction, with greater and sufficient complexity toward the frontal pole. Thus, we can hypothesize that the cytoarchitecture of this area may provide an adequate anatomical background for the functioning of a canonical circuit.

Although the data provided by Padoa-Schioppa et al. are very interesting, further experimental studies are needed to definitely extend the concept of canonical circuit to cortical areas involved in complex human behaviors such as economic decision-making.
Conclusion
The information required to construct a detailed and specific configuration of the neocortex, which contains approximately $10^{12}$ connections, exceeds by far the ~10^6 bits of information available in the genome for specification of the entire organism. On these grounds alone, it appears that nature’s strategy for construction of the neocortex must depend on the dynamic assembly of rather specific but simple modules. The surprisingly uniform microscopic structure of the neocortex and its laminar developmental process support this view. A possible question that could arise is why different systems (like motor and decision making) should share a basic circuit with the same organization. The first thing to consider is the biological advantage derived from the use of the same pattern in several areas, both in anatomical and adaptability-plasticity terms; in addition to the anatomical advantages (which ensure that innumerable synaptic connections are maintained in order and in reduced spaces), having a basic canonical circuit also allows easy plastic changes according to environmental requirements. This point seems valid, considering the issue also from an evolutionary perspective: nonhuman primates share many of the decision-making processes typical of humans.

It is likely that a canonical circuit that can easily be applied to different functions in different areas of the same brain is also an easily transmittable, hence phylogenetically beneficial, model.

Disclosure
The authors report no conflicts of interest in this work.

References