REVIEWS

What we can do and what we cannot do with fMRI

Nikos K. Logothetis[1]

Functional magnetic resonance imaging (fMRI) is currently the mainstay of neuroimaging in cognitive neuroscience. Advances in scanner technology, image acquisition protocols, experimental design, and analysis methods promise to push forward fMRI from mere cartography to the true study of brain organization. However, fundamental questions concerning the interpretation of fMRI data abound, as the conclusions drawn often ignore the actual limitations of the methodology. Here I give an overview of the current state of fMRI, and draw on neuroimaging and physiological data to present the current understanding of the haemodynamic signals and the constraints they impose on neuroimaging data interpretation.

Magnetic resonance imaging (MRI) is the most important imaging advance since the introduction of X-rays by Conrad Röntgen in 1895. Since its introduction in the clinic in the 1980s, it has assumed a role of unparalleled importance in diagnostic medicine and more recently in basic research. In medicine, MRI is primarily used to produce structural images of organs, including the central nervous system, but it can also provide information on the physico-chemical state of tissues, their vascularization, and perfusion. Although all of these capacities have long been widely appreciated, it was the emergence of functional MRI (fMRI)—a technique for measuring haemodynamic changes after enhanced neural activity—in the early 1990s that had a real impact on basic cognitive neuroscience research. A recent database (ISI/Web of Science) query using the keywords 'fMRI' or 'functional MRI' or 'functional magnetic resonance imaging' returned over 19,000 peer-reviewed articles. Given that the first fMRI study without exogenous contrast agents was published in 1991, this corresponds to approximately 1,100 papers per year, or over 3 papers per day. This average obscures the actual rate of publications, as in 1992 there were four publications in total, increasing to about eight per day by 2007. About 43% of papers explore functional localization and/or cognitive anatomy associated with some cognitive task or stimulus—constructing statistical parametric maps from changes in haemodynamic responses from every point in the brain. Another 22% are region of interest studies examining the physiological properties of different brain structures, analogous to single-unit recordings; 8% are on neuropsychology; 5% on the properties of the fMRI signal; and the rest is on a variety of other topics including plasticity, drug action, experimental designs and analysis methods.

In humans, fMRI is used routinely not just to study sensory processing or control of action, but also to draw provocative conclusions about the neural mechanisms of cognitive capacities, ranging from recognition and memory to pondering ethical dilemmas. Its popular fascination is reflected in countless articles in the press speculating on potential applications, and seeming to indicate that with fMRI we can read minds better than direct tests of behaviour itself. Unsurprisingly, criticism has been just as vigorous, both among scientists and the public. In fact, fMRI is not and will never be a mind reader, as some of the proponents of decoding-based methods suggest, nor is it a worthless and non-informative 'neophrenology' that is condemned to fail, as has been occasionally argued.

Perhaps the extreme positions on both sides result from a poor understanding of the actual capacities and limitations of this technology, as well as, frequently, a confusion between fMRI shortcomings and potential flaws in modelling the organizational principles of the faculties under investigation. For example, a frequently made assumption is that the mind can be subdivided into modules or parts whose activity can then be studied with fMRI. If this assumption is false, then even if the brain's architecture is modular, we would never be able to map mind modules onto brain structures, because a unified mind has no components to speak of. Even if true, the challenge remains in coming up with the correct recursive decompositions— in each of which any given cognitive capacity, however abstract, is divided into increasingly smaller functional units that are localized to specific brain parts, which in turn can be detected and studied with fMRI. This is not a neuroimaging problem but a cognitive one. Hierarchical decompositions are clearly possible within different sensory modalities and motor systems. Their mapping, which reflects the brain's functional organization, is evidently possible and certainly meaningful beyond any reasonable doubt[1].

Here, I offer an assessment of fMRI methodology itself, leaving aside such epistemological and ontological issues. I take the modular organization of many brain systems as a well established fact, and discuss only how far fMRI can go in revealing the neuronal mechanisms of behaviour by mapping different system modules and their dynamic inter-relationships. In this context the term module captures the classical local neuronal circuits repeated iteratively within a structure (for example, the columns or swirling, slab-like tangential arrangements of the neocortex), as well as the entities within which modules might be grouped by sets of dominating external connections. The often used term functional segregation refers to such specialized and spatially separated modules. Segregated entities that are interconnected might further result in nested distributed systems, the activity of which, often termed functional integration, can only be visualized by large-scale neuroimaging.

The principal advantages of fMRI lie in its noninvasive nature, ever-increasing availability, relatively high spatiotemporal resolution, and its capacity to demonstrate the entire network of brain areas engaged when subjects undertake particular tasks. One disadvantage is that, like all haemodynamic-based modalities, it measures a surrogate signal whose spatial specificity and temporal response are subject to both physical and biological constraints. A more important

[1]Max Planck Institute for Biological Cybernetics, 72076 Tuebingen, Germany, and Imaging Science and Biomedical Engineering, University of Manchester, Manchester M13 9PL, UK.

shortcoming is that this surrogate signal reflects neuronal mass activity. Although this fact is acknowledged by the vast majority of investigators, its implications for drawing judicious conclusions from fMRI data are most frequently ignored. The aim of this review is first to describe briefly the fMRI technology used in cognitive neuroscience, and then discuss its neurobiological principles that very often limit data interpretation. I hope to point out that the ultimate limitations of fMRI are mainly due to the very fact that it reflects mass action, and much less to limitations imposed by the existing hardware or the acquisition methods. Functional MRI is an excellent tool for formulating intelligent, data-based hypotheses, but only in certain special cases can it be really useful for unambiguously selecting one of them, or for explaining the detailed neural mechanisms underlying the studied cognitive capacities. In the vast majority of cases, it is the combination of fMRI with other techniques and the parallel use of animal models that will be the most effective strategy for understanding brain function.

A brief overview of fMRI

The beautiful graphics MRI and fMRI produce, and the excitement about what they imply, often mask the immense complexity of the physical, biophysical and engineering procedures generating them. The actual details of MRI can only be correctly described via quantum mechanics, but a glimpse of the method's foundation can be also afforded with the tools of classical physics using a few simple equations. (See refs 2 and 3 for a comprehensive account of the theoretical and practical aspects of MRI, and ref. 4 for its functional variants.) Here I offer a brief overview that permits an understandable definition of the terms and parameters commonly used in magnetic resonance imaging (see 'MRI and fMRI principles' in the Supplementary Information for a description of the principles and terms of anatomical and functional MRI). Functional activation of the brain can be detected with MRI via direct measurements of tissue perfusion, blood-volume changes, or changes in the concentration of oxygen. The blood-oxygen-level-dependent (BOLD) contrast mechanism[5,6] is currently the mainstay of human neuroimaging.

Critical factors determining the utility of fMRI for drawing conclusions in brain research are signal specificity and spatial and temporal resolution. Signal specificity ensures that the generated maps reflect actual neural changes, whereas spatial and temporal resolution determine our ability to discern the elementary units of the activated networks and the time course of various neural events, respectively. The interpretability of BOLD fMRI data also depends critically on the experimental design used.

Spatiotemporal properties of BOLD fMRI. The spatiotemporal properties of fMRI are covered in some detail in the Supplementary Information. Briefly, spatial specificity increases with increasing magnetic field strength and for a given magnetic field can be optimized by using pulse sequences that are less sensitive to signals from within and around large vessels (see Fig. 1 and 'Spatial and temporal specificity' in the Supplementary Information). Spatiotemporal resolution is likely to increase with the optimization of pulse sequences, the improvement of resonators, the application of high magnetic fields, and the invention of intelligent strategies such as parallel imaging, for example, sensitivity encoding (SENSE) method (see 'Spatial resolution' section in the Supplementary Information).

Human fMRI can profit a great deal from the use of high-field scanners and by the optimization of the pulse sequences used. Surprisingly, only a minority of the studies in the cognitive sciences seem to exploit the technical innovations reported from laboratories working on magnetic resonance methodologies. Most of the top-cited cognitive neuroscience studies (approximately 70%) were carried out at 1.5 T scanners, 20% were carried out at 3 T scanners, and very few at 2 T or 4 T field strengths. About 87% of all studies used the conventional gradient-echo echoplanar imaging (GE-EPI), whereas the rest used different variants of the spin-echo echoplanar imaging (SE-EPI) sequence. This combination of low magnetic field and traditional GE-EPI is prone to many localization errors. However, as of the beginning of the twenty-first century the percentage of middle-field (3 T) studies has increased, to reach about 56% in 2007. High magnetic fields are likely to dominate magnetic resonance research
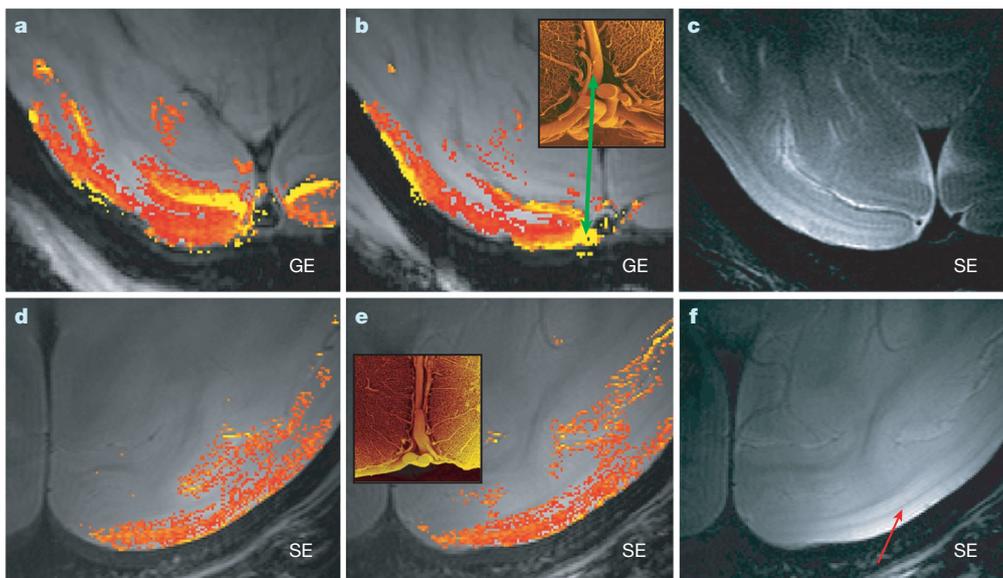
Figure 1 | Specificity of GE-EPI and SE-EPI. Examples of high-resolution GE-EPI and SE-EPI (courtesy J. Goense, MPI for Biological Cybernetics). **a, b,** Two slices of GE-EPI demonstrating the high functional signal-to-noise ratio (SNR) of the images, but also the strong contribution of macrovessels. The yellow areas (indicated with the green arrows) are pia vessels, an example of which is shown in the inset scanning electron microscopy image (total width of inset, 2 mm). For the functional images red indicates low and yellow indicates high. In-plane resolution 333 × 333 μm²; slice thickness 2 mm. **c,** Anatomical scan, SE-EPI, 250 × 188 μm², 2 mm slice, with time to echo (TE) and repetition time (TR) 70 and 3,000 ms respectively. **d, e,** Two slices of SE-EPI showing the reduction of vascular contribution at the pial side of the cortex. In-plane resolution 250 × 175 μm², slice thickness 2 mm. **f,** The anatomical scan is the SE-EPI used for obtaining the functional scans (TE/TR = 48/2,000 ms) but at different greyscale and contrast. The resolution of the anatomical scan permits the clear visualization of the Gennari line (red arrow), the characteristic striation of the primary visual

facilities in the future, and this should definitely improve the quality of data obtained in human magnetic resonance studies. At the same time, high magnetic field scanners are likely to require even tighter interaction between magnetic resonance physicists and application scientists, as the much larger inhomogeneity of both B0 (main static field) and B1 (the field generated by the excitation pulses) at high field will demand a great deal of expertise and experimental skill to achieve the desired image quality.

All in all, MRI may soon provide us with images of a fraction of a millimetre (for example, $300 \times 300\ \mu m^2$ with a couple of millimetres slice thickness or $500 \times 500 \times 500\ \mu m^3$ isotropic), which amount to voxel volumes of about two–three orders of magnitude smaller than those currently used in human imaging (see 'Developments and perspectives' in the Supplementary Information). With an increasing number of acquisition channels such resolution may ultimately be attained in whole-head imaging protocols, yielding unparalleled maps of distributed brain activity in great regional detail and with reasonable—a couple of seconds—temporal resolution. Would that be enough for using fMRI to understand brain function?

The answer obviously depends on the scientific question and the spatial scale at which this question could be addressed—"it makes no sense to read a newspaper with a microscope", as neuroanatomist Valentino Braitenberg once pointed out. To understand the functioning of the microcircuits in cortical columns or of the cell assemblies in the striosomes of basal ganglia, one must know a great deal about synapses, neurons and their interconnections. To understand the functioning of a distributed large-scale system, such as that underlying our memory or linguistic capacities, one must first know the architectural units that organize neural populations of similar properties, and the interconnections of such units. With 10^{10} neurons and 10^{14} connections in the cortex alone, attempting to study dynamic interactions between subsystems at the level of single neurons would probably make little sense, even if it were technically feasible. It is probably much more important to understand better the differential activity of functional subunits—whether subcortical nuclei, or cortical columns, blobs and laminae—and the instances of their joint or conditional activation. If so, whole-head imaging with a spatial resolution, say, of $0.7 \times 0.7\ mm^2$ in slices of 1-mm thickness, and a sampling time of a couple of seconds, might prove optimal for the vast majority of questions in basic and clinical research. More so, because of the great sensitivity of the fMRI signal to neuromodulation (see below and Supplementary Information). Neuromodulatory effects, such as those effected by arousal, attention, memory, and so on, are slow and have reduced spatiotemporal resolution and specificity[7,8].

Designs and analyses. Many studies initially used block designs, reminiscent of earlier positron emission tomography (PET) paradigms. These designs use time-integrated averaging procedures, and usually analyse the data by means of subtraction methods. The central idea is to compare a task state designed to place specific demands on the brain with an investigator-defined control state. Under these conditions, both enhancements and reductions of the fMRI signal are observed. In the early cognitive fMRI studies the prevailing block design was cognitive subtraction, with an emphasis on serial subtraction designs[9]. Such designs rely strictly on pure insertion, which asserts that a single cognitive process can be inserted into a task without affecting the remainder, an assumption that all too often is not tenable (see 'On pure insertion' in the Supplementary Information). Even if an experimental design could satisfy this assumption at the cognitive level, the assumption would be condemned to fail at the level of its neuronal instantiation[10] owing to the highly nonlinear nature of most brain processes. To overcome this kind of problem and ensure better interpretation of the neuroimaging data it is necessary to perform a detailed task analysis to determine subtraction components and their interactions. Yet most neuroimaging studies provide no formal task analysis that would ensure that the particular cognitive process of interest is indeed being isolated by the subtraction[11]. Traditional block designs have excellent functional contrast-to-noise ratio (that is, signal difference between test and control epochs, normalized to the mean signal of all epochs), but they are usually long (from 20 to 60 s), and may be confounded by the general state of arousal of the subject. High-speed fMRI methods, capable of whole-brain imaging with a temporal resolution of a few seconds, enabled the employment of so-called event-related designs[12]. The time course of the response in such experiments is closer to the underlying neural activity.

The block designs discussed so far may reveal differential patterns of activation only in those cases in which different stimulus attributes or different cognitive processes have distinct, non-overlapping spatial organizations. Overlapping networks of neurons subserving different functions are likely to go unnoticed owing to the spatial averaging that characterizes the blocked subtraction paradigms. Functional MRI adaptation designs were conceived as tools that might, at least to some extent, tackle the problem of spatially overlapping neural networks[13]. In this experimental design, a stimulus is presented repeatedly with the expectation that it will eventually induce response adaptation in neurons selective for its various properties. In general, repetition of an identical stimulus does indeed produce a reduction in the fMRI signal. After adaptation, the subject is presented with a stimulus that is varied along one dimension (for example, the direction of a moving pattern or the view of a human face) and the possibility of a response rebound is examined. If the underlying neural representation is insensitive to the changes in the stimulus then the fMRI signal will be reduced, similar to the reduction produced by the repetition of identical stimuli. Alternatively, if the neurons are sensitive to the transformation, the signal will show a clear rebound to its original, pre-adaptation level.

Functional MRI adaptation designs have been widely used in cognitive neuroscience, but they also have shortcomings, as any area receiving input from another region may reveal adaptation effects that actually occurred in that other region, even if the receiving area itself has no neuronal specificity for the adapted property[13]. Moreover, the conclusions of experiments relying on adaptation designs strongly rely on existing electrophysiological evidence, which itself may hold true for one area and not for another[72].

Finally, clever analysis is required to exploit clever design. Most studies so far have used voxel-based conventional analyses of MRI time series from one or more subjects[14]. The approach is predicated on an extension of the general linear model that allows for correlations between error terms owing to physiological noise or correlations that ensue after temporal smoothing. The method is reliable and, when well implemented, offers the best analysis strategy for most studies. Another approach is to take into account the full spatial pattern of brain activity, measured simultaneously at many locations[15]. Such multivariate analyses or pattern-classification-based techniques (decoding techniques) can often detect small differences between two task or stimulus conditions—differences that are not picked up by conventional univariate methods. However, this is not equivalent to saying that they unequivocally reveal the neural mechanisms underlying the activation patterns. The presence, for instance, of voxels selective to two different stimulus attributes could be potentially detected by modern classifiers, yet the existence of two types of patterns does not necessarily imply the existence of two different types of neural populations[72].

What do activation maps represent?

Does the activation of an area mean that it is truly involved in the task at hand? This question implies that we understand what neural activity in a given area would unequivocally show its participation in the studied behaviour. But do we? It is usually alleged that cognitive capacities reflect the 'local processing of inputs' or the 'output' of a region, instantiated in the patterns of action potentials, with their characteristic frequency and timing. In principle, brain structures can be conceptualized as information processing entities, with an input, a local-processing capacity, and an output. Yet, although such a scheme may describe the function of subcortical nuclei, its implementation

in different areas of cortex is anything but straightforward. In fact, we now know that the traditional cortical input–elaboration–output scheme, commonly presented as an instantiation of the tripartite perception–cognition–action model, is probably a misleading over-simplification[16]. Research shows that the subcortical input to cortex is weak; the feedback is massive, the local connectivity reveals strong excitatory and inhibitory recurrence, and the output reflects changes in the balance between excitation and inhibition, rather than simple feedforward integration of subcortical inputs[17]. In the context of this review, the properties of these excitation–inhibition networks (EIN) deserve special attention, and are briefly discussed below.

Feedforward and feedback cortical processing. Brain connectivity is mostly bidirectional. To the extent that different brain regions can be thought of as hierarchically organized processing steps, connections are often described as feedforward and feedback, forward and backward, ascending and descending, or bottom-up and top-down[18]. Although all terms agree on processing direction, endowing backward connections with a role of engineering-type or functional 'feedback' might occasionally be misleading, as under a theoretical generative model perspective on brain function, it is the backward connections that generate predictions and the forward connections that convey the traditional feedback, in terms of mismatch or prediction error signals[19].

In the sensory systems, patterns of long-range cortical connectivity to some extent define feedforward and feedback pathways[20]. The main thalamic input mainly goes to middle layers, whereas second-order thalamic afferents and the nonspecific diffuse afferents from basal forebrain and brain-stem are, respectively, distributed diffusely regionally or over many cortical areas, making synapses mainly in superficial and/or deep layers. Cortical output has thalamic and other subcortical projections originating in layers VI and V, respectively, and corticocortical projections mostly from supragranular layers. The primary thalamic input innervates both excitatory and inhibitory neurons, and communication between all cell types includes horizontal and vertical connections within and between cortical layers. Such connections are divergent and convergent, so that the final response of each neuron is determined by all feedforward, feedback and modulatory synapses[17].

Very few of the pyramid synapses are thalamocortical (less than 10–20% in the input layers of cortex, and less than 5% across its entire depth; in the primary visual cortex the numbers are even lower, with the thalamocortical synapses on stellate cells being about 5%[21]), with the rest originating from other cortical pyramidal cells. Pyramidal axon collateral branches ascend back to and synapse in superficial layers, whereas others distribute excitation in the horizontal plane, forming a strongly recurrent excitatory network[17].

The strong amplification of the input signal caused by this kind of positive feedback loop is set under tight control by an inhibitory network interposed among pyramidal cells and consisting of a variety of GABAergic interneurons[22,23]. These can receive both excitatory and inhibitory synapses on to their somata, and have only local connections. About 85% of them in turn innervate the local pyramidal cells. Different GABAergic cells target different subdomains of neurons[22,24]. Some (for example, basket cells) target somata and proximal dendrites, and are excellent candidates for the role of gain adjustment of the integrated synaptic response; others (for example, chandelier cells) target directly the axons of nearby pyramidal neurons, and appear to have a context-dependent role[25]—they can facilitate spiking during low activity periods, or act like gatekeepers that shunt most complex somatodendritic integrative processes during high activity periods (for example, see up- and down states below). Such nonlinearities might generate substantial dissociations between subthreshold population activity and its concomitant metabolic demand and the spiking of pyramidal cells.

Modules and their microcircuits. A large number of structural, immunochemical and physiological studies, in all cortical areas examined so far, suggested that the functional characteristics of a cortical module are instantiated in a simple basic EIN, referred to as a canonical microcircuit[17] (see also Fig. 2a). Activation of a microcircuit sets in motion a sequence of excitation and inhibition in every neuron of the module, rather than initiating a sequential activation of separate neurons at different hypothetical processing stages. Re-excitation is tightly controlled by local inhibition, and the time evolution of excitation–inhibition is far longer than the synaptic delays of the circuits involved. This means the magnitude and timing of any local mass activation arise as properties of the microcircuits.

Computational modelling suggested that EIN microcircuits, containing such a precisely balanced excitation and inhibition, can account for a large variety of observations of cortical activity, including amplification of sensory input, noise reduction, gain control[26], stochastic properties of discharge rates[27], modulation of excitability with attention[28], or even generation of persisting activity during the delay periods of working memory tasks[29].

The principle of excitation–inhibition balance implies that microcircuits are capable of large changes in activity while maintaining proportionality in their excitatory and inhibitory synaptic conductances. This hypothesis has been tested directly in experiments examining conductance changes during periods of high (up) and low (down) cortical activity. Alternating up states and down states can be readily observed in cerebral cortex during natural sleep or anaesthesia[30], but they can be also induced *in vitro* by manipulating the ionic concentrations in a preparation so that they match those found *in situ*. Research showed that the up state is characterized by persisting synaptically mediated depolarization of the cell membranes owing to strong barrages of synaptic potentials, and a concomitant increase in spiking rate, whereas the down state is marked by membrane hyperpolarization and reduction or cessation of firing[31,32]. Most importantly, the excitation–inhibition conductances indeed changed proportionally throughout the duration of the up state despite large changes in membrane conductance[31,32].

Microcircuits therefore have the following distinct features: (1) the final response of each neuron is determined by all feedforward, feedback and modulatory synapses; (2) transient excitatory responses may result from leading excitation, for example, due to small synaptic delays or differences in signal propagation speed, whereupon inhibition is rapidly engaged, followed by balanced activity[31,32]; (3) net excitation or inhibition might occur when the afferents drive the overall excitation–inhibition balance in opposite directions; and (4) responses to large sustained input changes may occur while maintaining a well balanced excitation–inhibition. In the latter case, experimentally induced hyperpolarization of pyramidal cells may abolish their spiking without affecting the barrages of postsynaptic potentials (see ref. 31 and references therein). It is reasonable to assume that any similar hyperpolarization under normal conditions would decrease spiking of stimulus-selective neurons without affecting presynaptic activity. In visual cortex, recurrent connections among spiny stellate cells in the input layers can provide a significant source of recurrent excitation[26]. If driven by proportional excitation–inhibition synaptic currents, the impact of their sustained activity might, once again, minimally change the spiking of the pyramidal cells. This last property of microcircuits suggests that changes with balanced excitation–inhibition are good candidates for mechanisms adjusting the overall excitability and the signal-to-noise ratio (SNR) of the cortical output. Thus microcircuits—depending on their mode of operation—can, in principle, act either as drivers, faithfully transmitting stimulus-related information, or as modulators, adjusting the overall sensitivity and context-specificity of the responses[28]. Figure 2b summarizes the different types of excitation-inhibition changes and their potential effect on the haemodynamic responses.

This interesting and important driver/modulator distinction was initially drawn in the thalamus[33], in which the afferents in the major sensory thalamic relays were assigned to one of two major classes on the basis of the morphological characteristics of the axon terminals, the synaptic relationships and the type of activated receptors, the
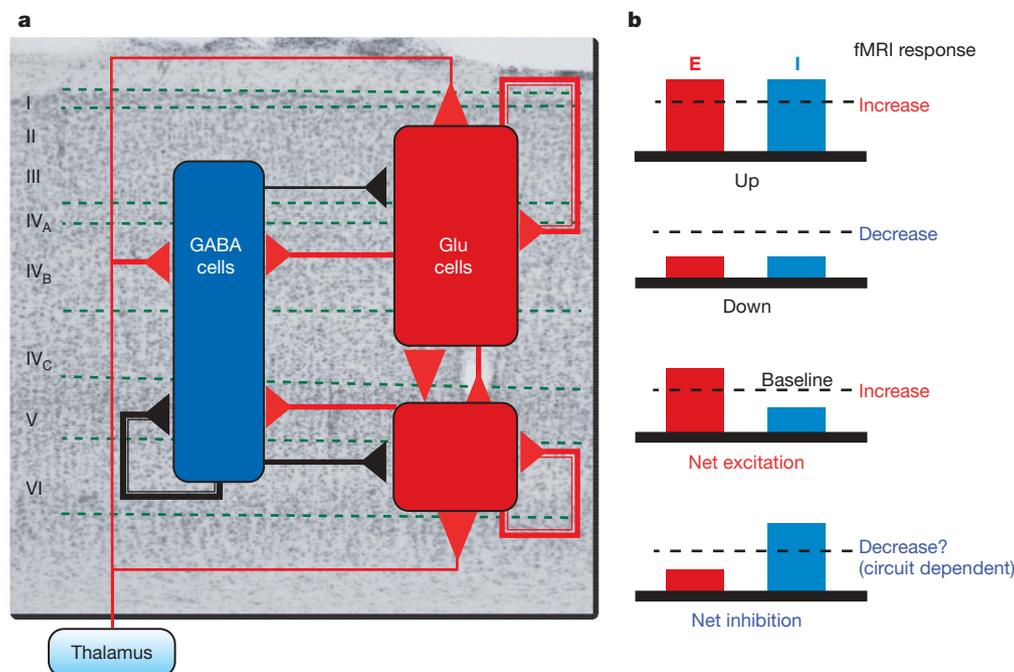
Figure 2 | Principles of excitation–inhibition circuits. a, Model of a canonical cerebral microcircuit (adapted from ref. 71). Three neuronal populations interact with each other: supragranular–granular and infragranular glutamatergic spiny neurons, and GABAergic cells. Excitatory synapses are shown in red and inhibitory synapses in black. All groups receive excitatory thalamic input. The line width indicates the strength of connection. The circuit is characterized by the presence of weak thalamic input and strong recurrence (see text for details). Glu, glutamatergic.

b, Potential proportional and opposite-direction changes of cortical excitation (E) and inhibition (I). Responses to large sustained input changes may occur while maintaining a well balanced excitation–inhibition (up and down). The commonly assumed net excitation or inhibition might occur when the afferents drive the overall excitation–inhibition balance in opposite directions. The balanced proportional changes in excitation–inhibition activity, which occur as a result of neuromodulatory input, are likely to strongly drive the haemodynamic responses.

degree of input convergence, and the activity patterns of postsynaptic neurons. The same concept also broadly applies to the afferents of the cerebral cortex[34], wherein the thalamic or corticocortical axons terminating in layer IV can be envisaged as drivers, and other feedback afferents terminating in the superficial layers as modulators. It can also be applied to the cortical output, whereby the projections of layer VI back to the primary relays of the thalamus are modulatory, whereas the cortico-thalamo-cortical paths originating in layer V of cortex, reaching higher-order thalamic nuclei (for example, pulvinar), and then re-entering cortex via layer IV, are drivers[33].

The initial information reaching a cortical region is elaborated and evaluated in a context-dependent manner, under the influence of strong intra- and cross-regional cortical interactions. The cortical output reflects ascending input but also cortico-thalamo-cortical pathways, whereas its responsiveness and SNR reflect the activity of feedback, and likely input from the ascending diffuse systems of the brain-stem. The neuromodulation (see 'Neurotransmission and neuromodulation' in Supplementary Information) afforded by these systems, which is thought to underlie the altered states of cognitive capacities, such as motivation, attention, learning and memory, is likely to affect large masses of cells, and potentially induce larger changes in the fMRI signal than the sensory signals themselves.

Excitation–inhibition networks and fMRI. The organization discussed above evidently complicates both the precise definition of the conditions that would justify the assignment of a functional role to an 'active' area, and interpretation of the fMRI maps. Changes in excitation–inhibition balance—whether they lead to net excitation, inhibition, or simple sensitivity adjustment—inevitably and strongly affect the regional metabolic energy demands and the concomitant regulation of cerebral blood flow (CBF) (that is, they significantly alter the fMRI signal). A frequent explanation of the fMRI data simply assumes an increase in the spiking of many task- or stimulus-specific neurons. This might be correct in some cases, but increases of the BOLD signal may also occur as a result of balanced proportional

increases in the excitatory and inhibitory conductances, potential concomitant increases in spontaneous spiking, but still without a net excitatory activity in stimulus-related cortical output. In the same vein, an increase in recurrent inhibition with concomitant decreases in excitation may result in reduction of an area's net spiking output, but would the latter decrease the fMRI signal? The answer to this question seems to depend on the brain region that is inhibited, as well as on experimental conditions.

Direct haemodynamic measurements with autoradiography suggested that metabolism increases with increased inhibition[35]. An exquisite example is the inhibition-induced increase in metabolism in the cat lateral superior olive (LSO). This nucleus, which contains the representations of low-, middle- and high-tone frequencies, receives afferents from both ears: over a two-neuron pathway from the ipsilateral ear and over a three-neuron pathway from the contralateral ear. Furthermore, it has no presynaptic axo-axonic endings that might mediate presynaptic inhibition via excitatory terminals. Electrophysiology showed that the LSO afferents from the ipsilateral ear are excitatory whereas the afferents from the contralateral ear are inhibitory. This unusual combination of anatomical and physiological features suggests that if one ear is surgically deafened and the animal is exposed to a high-frequency pure tone, a band of tissue in the LSO on the side opposite to the remaining active ear is subjected to strictly inhibitory synaptic activity without complications by presynaptic inhibition, concurrent lateral excitation, disinhibition/excitation, or other kinds of possibly excitatory action. Under these conditions, maps obtained with [^{14}C]2-deoxyglucose (2DG) autoradiography[36] demonstrated clear increases in metabolism in the contralateral LSO[37], suggesting that the presynaptic activity in that area is sufficient to show strong energy consumption despite the ensuing spiking reduction. Similar increases in metabolism during the reduction of spike rates were observed during long-lasting micro-stimulation of the fornix, which induces sustained suppression of pyramidal cell firing in hippocampus[38].

In contrast, human fMRI studies reported haemodynamic and metabolic downregulation accompanying neuronal inhibition in motor[39] and visual cortices[40], suggesting that the sustained negative BOLD response (NBR) is a marker of neuronal deactivation. Similarly, combined fMRI and electrophysiological experiments showed a clear correspondence of NBR and decreased population spiking in haemodynamically 'negative' areas in the monkey primary visual cortex[41]. Decreases in blood oxygenation and volume were also found to be co-localized with predominant neuronal inhibition and arteriolar vasoconstriction during somatosensory stimulation in rats[42]. Thus, without understanding the intrinsic correlation between direct or indirect inhibitory activity and concomitant changes in energy metabolism in a given situation, conclusions cannot be drawn. Unfortunately, the few published theoretical estimates of energy budget have not considered the metabolic costs of spikes in interneurons and of the inhibitory postsynaptic potentials (IPSPs) they produce[43]. Modelling of inhibition is unlikely to be straightforward. On the one hand, the density of cortical inhibitory neurons is 10–15 times lower than excitatory neurons[16], and for each one of them the electrochemical gradient, down which Cl^- moves postsynaptically at inhibitory synapses, is weaker than that of Na^+ at excitatory synapses, requiring less energy to pump Cl^- back. In fact, the transport cycles of the cation–chloride co-transporters, which have a key role in intracellular Cl^- regulation, are driven without the direct hydrolysis of ATP, by using the energy from the cation gradients generated by the Na,K-ATPase[44]. On the other hand, inhibitory interneurons are fast spiking[45,46]. For example, the firing of pyramidal cells in hippocampus is 1.4 Hz, whereas that of interneurons in the strata pyramidale and oriens is 15 Hz and 10 Hz, respectively. Similarly, cortical inhibitory interneurons may discharge 2–3 times faster than pyramidal cells[47]. In principle, inhibition may increase or decrease energy consumption depending on the contribution of the aforementioned factors (for a recent comprehensive review on inhibitory neurons and brain metabolism, see ref. 48). Last but not least, neurons directly affect microvessels. Pericytes, the flat, contractile connective-tissue cells, often attached to the abluminal surface of the capillary endothelial cells, might directly alter CBF in response to changes in neural activity[49]. Moreover, a body of evidence suggests that increased activity of single inhibitory interneurons results in precise vasomotor responses in neighbouring brain microvessels, and these contractile or dilatory responses were attributed to arteriole smooth muscle[50].

The diversity of the haemodynamic responses to neural inhibition obtained in different types of experiments is therefore hardly surprising: it is primarily due to the fact that regional inhibition itself might have a number of different causes, including early shunting of the weak cortical input, leading to a reduction of recurrent excitation rather than an increase in summed inhibition; increased synaptic inhibition; shunting of the cortical output through the axo-axonic connections of the chandelier cells; or any combination thereof. In the first case inhibition might result in a clear NBR; in the other two it might reflect the local metabolism increases induced by the unaffected input and its ongoing processing, resulting in fMRI activations. The fMRI responses might further blur the origin of inhibition owing to the direct effects of the latter on the arterioles and microvessels.

Evidently much research is needed to characterize the actual state of an area and its participation in behaviour, but quite independent of this fact, the nature of the EIN suggests that mass action and its surrogate haemodynamics are ambiguous signals, the interpretation of which must be constrained by the concurrent use of other methodologies.

Neurophysiological correlates of the BOLD signal

EIN and mesoscopic neural signals. The active regions of the membrane of a discharging neuron at any given time are considered to act as a current sink, whereas the inactive ones act as a current source for the active regions (see 'Neural signals' in Supplementary Information). The linear superposition of currents from all sinks and sources forms the extracellular field potential measured by microelectrodes. The extracellular field potential captures at least three different types of EIN activity: single-unit activity representing the action potentials of well isolated neurons next to the electrode tip, multiple unit activity reflecting the spiking of small neural populations in a sphere of 100–300 μm radius, and perisynaptic activity of a neural population within 0.5–3 mm of the electrode tip, which is reflected in the variation of the low-frequency components of the extracellular field potential. Multiple unit activity and local field potentials (LFPs) can be reliably segregated by frequency band separation. A high-pass filter cutoff in the range of 500–1,000 Hz is used in most recordings to obtain the multiple unit activity, and a low-pass filter cutoff of approximately 250 Hz to obtain LFP. A large number of experiments have presented data indicating that such a band separation does indeed underlie different neural events (see 'Neural signals' in Supplementary Information).

LFP signals and their different band-limited components (alpha, beta, gamma, and so on) are invaluable for understanding cortical processing, as they are the only signs of integrative EIN processes. In fact, LFPs do not, as initially thought, solely reflect population postsynaptic potentials, but also integrative soma–dendritic processes—including voltage-dependent membrane oscillations and afterpotentials following soma–dendritic spikes—that all together represent the local (perisynaptic) activity in a region (see 'Neural signals' in Supplementary Information). A shortcoming of the LFP is its ambiguity. A change in the power of LFP in a particular frequency band most likely occurs for any mode of operations of the EIN. As most of the excitatory input into an area is local, LFPs will also indirectly reflect some of the postsynaptic effects of pyramidal cell activity. In addition, LFPs have a certain neural-class bias, which in this case is determined by geometry and regional architecture. The arrangement of the pyramidal and Purkinje cells will give rise to large LFP modulations; in contrast, interneurons will contribute only weakly because of their star-shaped dendrites and their geometrical disorder. Finally, inhibitory synapses may occasionally act as 'shunts' for the excitatory currents through low-resistance channels, in which case large synaptic conductance changes may produce little effect in the membrane potential, and result in weak and hard-to-measure multiple unit activity and LFPs.

When individual LFP bands are examined separately, local spiking activity is occasionally found to affect certain frequency bands, whereas that of neuromodulation affects others[51–53]. It is evident that the most useful information will not be derived by one type of signal alone, but rather by the study of relative changes in one signal or the other. Electrophysiological studies examining the individual contributions of different LFP frequency bands, multiple unit activity, and spiking of individual neurons are probably our only realistic chance of gaining insights into the neural mechanisms of haemodynamic responses and their meaning in the context of different cognitive tasks.

Mesoscopic signals and the BOLD signal. The relationship of neocortical LFPs and spiking activity to the BOLD signal itself was examined directly in concurrent electrophysiology and fMRI experiments in the visual system of anaesthetized[54] and alert[55] monkeys. These studies found that the BOLD responses reflect input and intracortical processing rather than pyramidal cell output activity. Initially, both LFPs and spiking seemed to be correlated with the BOLD response, although quantitative analysis indicated that LFPs are better predictors of the BOLD response than multiple-unit or single-unit spiking. The decisive finding leading to the papers' conclusion, however, was not the degree of correlation between the neural and the fMRI responses or the differential contribution of any type of signal into the BOLD responses[55], but rather the striking, undiminished haemodynamic responses in cases where spiking was entirely absent despite a clear and strong stimulus-induced modulation of the field

potentials[54,55]. Similar dissociations between spikes and CBF had been demonstrated earlier and very recently in a number of studies using other techniques[56–58].

The findings are in close agreement with a number of older auto-radiography studies, also showing that regional glucose utilization is directly related to neuronal synaptic activity[35]. For example, the greatest 2-DG uptake occurs in the neuropil (that is, in areas rich in synapses, dendrites and axons, rather than in cell bodies). During orthodromic and antidromic electrical microstimulation, only orthodromic microstimulation, which involves presynaptic terminals, increases glucose consumption. Similarly, the highest density of cytochrome oxidase (an enzyme of the respiratory chain) is found in somato-dendritic regions that are adjacent to axon terminals. Finally, as mentioned earlier, presynaptic activity increases metabolism even if the output is inhibited (that is, the spiking activity is abolished).

Despite all this evidence, some discussion still concentrates on the importance of the firing rate of action potentials of projection neurons in the generation of the haemodynamic responses, perhaps stemming from the fact that important early studies of neural correlates of behaviour took the mean spiking rate to be the gold standard for quantifying neuronal activation. These discussions, however, often suffer from a certain amount of contention seeking where none is warranted. In many cases, spikes do indeed correlate with LFPs, and they will also correlate with the BOLD signal. In addition, unusually high correlations between multiple unit activity and BOLD signal (or LFP and multiple unit activity) may result from excessive signal-smoothing owing to sampling rates of several seconds rather than a fraction of a second, as well as inter-subject averaging when simultaneous physiology and fMRI measurements are not possible (see ref. 55 for discussion).

Predicting neural activity from the fMRI signals. Functional MRI signals are presumed to result from changes in the activity of the neuronal populations responsible for the functions in question (for example, stimulus- or task-selective neurons). This assumption is mainly based on decades of electrophysiology research with recordings from isolated single neurons in experimental animals, in which particular sensory stimuli that the animal perceives or tasks that it performs were found to increase the firing rate of certain cells but not of others. The psychologist or cognitive neuroscientist who finds cortical area X to be activated by the task at hand implicitly or explicitly assumes that—if an electrode were placed in the subject's brain—an increase in the spiking rate of those specialized neurons underlying the subject's behaviour would be observed. This might well be true in some cases, but not in all. When attempting to interpret the fMRI signal by modelling, or when comparing the results of human neuroimaging to those obtained in monkey physiology experiments, it is useful to take the following facts into consideration.

In humans, there are about 90,000–100,000 neurons under 1 mm^2 of cortical surface. This number is relatively constant for all structurally and functionally distinct areas, including the somatosensory, temporal, parietal, frontal and motor cortical areas[16,59]. An exception is the primary visual cortex of certain primates, including monkey and human, which has approximately twice as many neurons. The number of cortical neurons under unitary cortical surface is also similar across many species, including mouse, rat, cat, monkey and human. Its small variability is the result of a trade-off between cortical thickness and neural density. The former varies from area to area and from species to species (for example, from mouse to human the cortex becomes approximately three times thicker). Neural density varies inversely to cortical thickness. On average, density is 20,000 to 30,000 neurons per mm^3; it peaks in the primary visual cortex by a factor of 4, and it is minimal in the motor cortex[59,60]. Synaptic density ranges from 0.4 to 1×10^9 per mm^3. Depending on the thickness of the cortex (2–4 mm), the number of synapses beneath 1 mm^2 surface is around 10^9 (0.8–4×10^9). Although the number of synapses and the axonal length per neuron increases with increasing cortical

thickness[61], the overall length of neuronal processes remains relatively constant, with axonal length being approximately 4 km mm^{-3} and dendrite length 0.4 km mm^{-3}. Overall, synaptic density and the ratio of excitatory to inhibitory synapses also remain constant.

Given these neuro-statistical data, what are the actual contents of a neuroimaging voxel? An examination of the 300 top-cited cognitive fMRI studies suggests that the commonly used in-plane resolution is 9–16 mm^2, for slice thicknesses of 5–7 mm. The average voxel size before any pre-processing of the data is thus 55 µl (or 55 mm^3). Often the effective size is 2–3 times larger due to the spatial filtering that most investigators apply to improve the functional SNR. Less than 3% of this volume is occupied by vessels and the rest by neural elements (see Fig. 3) A typical unfiltered fMRI voxel of 55 µl in size thus contains 5.5 million neurons, 2.2–5.5×10^{10} synapses, 22 km of dendrites and 220 km of axons.

This 'large population view' is in contrast to the scope of the traditional microelectrode recordings. It would be nice if we could monitor every relevant neuron in the cortex during intracortical microelectrode recordings, but this is practically impossible. Instead, the typical electrophysiological measurements in behaving animals report only on the properties of most active large neurons that constitute a minority. The strong selection bias during extracellular recordings is partly due to practical limitations (for example, injury or simply size bias[62]) and partly to the physiological properties of neurons and/or the organizational principles of neural networks. In fact, many different types of electrical and optical measurements provide evidence that a substantial proportion of neurons, including the cortical pyramidal cells, might be silent[63]. Their silence might reflect unusually high input selectivity or the existence of decoding schemes relying on infrequent co-spiking of neuronal subsets. Most important for the comparison of neuroimaging and electrophysiology results is the fact that lack of measurable neuronal spiking may not necessarily imply lack of input and subthreshold processing.

A direct analogy between neuronal spiking as measured in animal experiments and the fMRI signal obtained in human recording is thus simply unrealistic and might often lead to incorrect conclusions. It is hardly surprising that most studies so far relying purely on BOLD fMRI have failed to reveal the actual neural properties of the studied area, at least those properties (for example, selectivity to various visual features) that were previously established in electrophysiological studies.

An example is cortical area V5 (or MT) that has been extensively studied in the context of motion processing and perception[64,65]. Electrophysiology has shown that the vast majority of the V5 neurons in monkeys are direction and speed selective. Neuroimaging localized the homologue of area V5 in humans as an area responding stronger to moving than to stationary stimuli. Later studies suggested that human V5 is sensitive to motion direction, and that it may be thought of as containing large populations of directionally selective units, just like its monkey homologue. The studies of directional specificity exploited the phenomenon of motion after-effect induced by motion adaptation. After prolonged exposure to a stimulus moving in one direction, subjects perceive a subsequent static stimulus to move in the opposite direction. It is assumed that motion after-effect is due to the fact that the balance of mutual inhibition (opponency) between detectors for opposite directions of movement is distorted after adaptation. The sensitivity of the detectors selective for the adapting direction is reduced, which in turn releases from inhibition the neurons selective for the opposite direction[66]. Using this phenomenon, human studies demonstrated that the fMRI response to a stationary stimulus was greater when the stimulus was preceded by a motion-after-effect-inducing, unidirectional adaptation, than when preceded by bidirectional adaptation[67]. Given the existing physiology data in the monkey V5, these findings were interpreted as demonstrating that the BOLD signal directly reflects direction-selective spiking activity of the area.

Yet, as I have indicated above, the BOLD signal is primarily affected by changes in excitation–inhibition balance, and this balance may be controlled by neuromodulation more than by the changes in spiking rate of a small set of neurons. In fact, the BOLD signal is strongly modulated by attention[68], and the results of the motion after-effect experiments could, in principle, be due to the fact that a stimulus with illusory motion automatically draws the attention of a subject more compared to a situation in which there is no motion after-effect. This hypothesis turned out to be correct, as a later study—in which balance in attentional load was accomplished by having the subjects perform a concurrent visual task—found no signal differences between the motion after-effect and no motion after-effect conditions[69].

A similar example pertains to the differences in neurophysiological and fMRI responses in the primary visual cortex during different perceptual states. It is known that physiological signals are in general stronger when stimuli are perceived as opposed to when they are not. Intriguingly, in some regions the BOLD response seems to reflect this even more sensitively than physiological measures like spikes and multi-unit activity[70]. An example is the pattern of fMRI activation changes in V1 during binocular rivalry (that is, the perceptual alternations experienced when the two eyes view different stimuli). This phenomenon has been studied extensively psychophysically and also over the last two decades in a series of electrophysiology studies

in monkeys[70]. These studies showed that only a small fraction of V1 cells modulate their spiking during the perceptual changes; neuroimaging, on the other hand, demonstrated fMRI-signal modulations that were nearly as large as those obtained during the physical alternation of stimuli[70]. The difference, once again, reflects the fact that neuromodulatory feedback from higher areas can be easily detected by means of fMRI, but not through the measurement of single-unit activity. Interestingly, measurements of subthreshold activity in another study of perceptual multistability revealed perception-related modulations in LFP, despite the unaltered spike rates[53]. Such clear spiking and BOLD signal mismatches appear even in simple experiments probing sensory processing. Simple stimuli, such as those used in the aforementioned studies, are most likely to generate a proportional enhancement in both the afferent and efferent activity of any sensory area. The activation of high-level association areas related to cognitive processing might be more sensitive or even dominated by feedback and neuromodulation, whose differential effect on spiking and haemodynamic responses is utterly unknown.

Conclusions and perspectives

The limitations of fMRI are not related to physics or poor engineering, and are unlikely to be resolved by increasing the sophistication and power of the scanners; they are instead due to the circuitry and
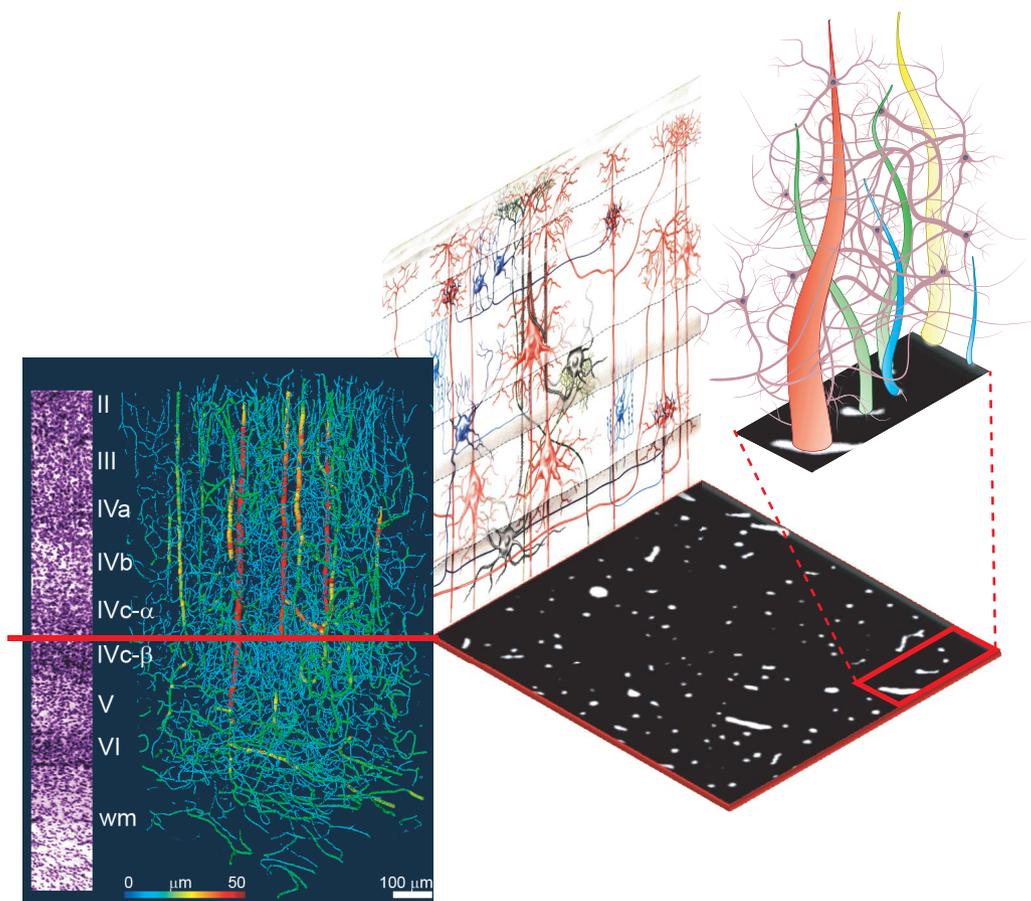
Figure 3 | Neural and vascular contents of a voxel. The left panel demonstrates the relative density of vessels in the visual cortex of monkeys. The dense vascular mesh is displayed by perfusing the tissue with barium sulphate and imaging it with synchrotron-based X-ray microtomography (courtesy B. Weber, MPI for Biological Cybernetics). The vessel diameter is colour coded. Cortical surface without pial vessels is displayed at the top; white matter at the bottom. At the left of the panel is a Nissl slice from the same area, showing the neural density for layers II through to the white matter (wm). Although the density of the vessels appears to be high in this three-dimensional representation, it is actually less the 3% (see section at the right; white spots are cross-sections of vessels). The average distance between the small vessels (capillaries) is about 50 μm. This is approximately the distance that oxygen molecules travel by diffusion within the limited transit time of the blood. The dense population of neurons, synapses and glia occupy the intervascular space, as depicted in the drawing at the top right—a hypothetical distribution of vascular and neural elements in a small section (red rectangle). The drawing in the background shows some of the typical neuronal types (for example, red, large pyramidal cell; dark blue, inhibitory basket cells; light blue, chandelier inhibitory neurons; and grey, stellate cells) and their processes.

functional organization of the brain, as well as to inappropriate experimental protocols that ignore this organization. The fMRI signal cannot easily differentiate between function-specific processing and neuromodulation, between bottom-up and top-down signals, and it may potentially confuse excitation and inhibition. The magnitude of the fMRI signal cannot be quantified to reflect accurately differences between brain regions, or between tasks within the same region. The origin of the latter problem is not due to our current inability to estimate accurately cerebral metabolic rate of oxygen ($CMRO_2$) from the BOLD signal, but to the fact that haemodynamic responses are sensitive to the size of the activated population, which may change as the sparsity of neural representations varies spatially and temporally. In cortical regions in which stimulus- or task-related perceptual or cognitive capacities are sparsely represented (for example, instantiated in the activity of a very small number of neurons), volume transmission (see Supplementary Information)—which probably underlies the altered states of motivation, attention, learning and memory—may dominate haemodynamic responses and make it impossible to deduce the exact role of the area in the task at hand. Neuromodulation is also likely to affect the ultimate spatiotemporal resolution of the signal.

This having been said, and despite its shortcomings, fMRI is currently the best tool we have for gaining insights into brain function and formulating interesting and eventually testable hypotheses, even though the plausibility of these hypotheses critically depends on used magnetic resonance technology, experimental protocol, statistical analysis and insightful modelling. Theories on the brain's functional organization (not just modelling of data) will probably be the best strategy for optimizing all of the above. Hypotheses formulated on the basis of fMRI experiments are unlikely to be analytically tested with fMRI itself in terms of neural mechanisms, and this is unlikely to change any time in the near future.

Of course, fMRI is not the only methodology that has clear and serious limitations. Electrical measurements of brain activity, including invasive techniques with single or multiple electrodes, also fall short of affording real answers about network activity. Single-unit recordings and firing rates are better suited to the study of cellular properties than of neuronal assemblies, and field potentials share much of the ambiguity discussed in the context of the fMRI signal. None of the above techniques is a substitute for the others. Today, a multimodal approach is more necessary than ever for the study of the brain's function and dysfunction. Such an approach must include further improvements to MRI technology and its combination with other non-invasive techniques that directly assess the brain's electrical activity, but it also requires a profound understanding of the neural basis of haemodynamic responses and a tight coupling of human and animal experimentation that will allow us to fathom the homologies between humans and other primates that are amenable to invasive electrophysiological and pharmacological testing. Claims that computational methods and non-invasive neuroimaging (that is, excluding animal experimentation) should be sufficient to understand brain function and disorders are, in my opinion, naive and utterly incorrect. If we really wish to understand how our brain functions, we cannot afford to discard any relevant methodology, much less one providing direct information from the actual neural elements that underlie all our cognitive capacities.

1. Wandell, B. A., Brewer, A. A. & Dougherty, R. F. Visual field map clusters in human cortex. *Phil. Trans. R. Soc. Lond. B* **360**, 693–707 (2005).
   This paper provides a description of human visual field maps and the rationale generating and naming them.
2. Haacke, E. M. et al. Magnetic Resonance Imaging: Principles and Sequence Design (John Wiley & Son, New York, 1999).
3. Wood, M. L. & Wehrli, F. W. Principles of magnetic resonance imaging. In Magnetic Resonance Imaging 3rd edn (eds Stark, D. D. & Bradley, W.) 1–14 (Mosby, St Louis/Baltimore/Boston/London/Tokyo, 1999).
4. Buxton, R. B. Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques (Cambridge Univ. Press, Cambridge, UK, 2002).
5. Ogawa, S. & Lee, T. M. Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation. *Magn. Reson. Med.* **16**, 9–18 (1990).
6. Ogawa, S. et al. Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magn. Reson. Med.* **14**, 68–78 (1990).
7. Motter, B. C. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J. Neurophysiol.* **70**, 909–919 (1993).
8. Luck, S. J., Chelazzi, L., Hillyard, S. A. & Desimone, R. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.* **77**, 24–42 (1997).
9. Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. & Raichle, M. E. Positron emission tomographic studies of the processing of single words. *J. Cogn. Neurosci.* **1**, 153–170 (1989).
10. Friston, K. J. et al. The trouble with cognitive subtraction. *Neuroimage* **4**, 97–104 (1996).
11. Poeppel, D. A critical review of PET studies of phonological processing. *Brain Lang.* **55**, 317–351 (1996).
12. Buckner, R. L. et al. Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proc. Natl Acad. Sci. USA* **93**, 14878–14883 (1996).
13. Krekelberg, B., Boynton, G. M. & van Wezel, R. J. Adaptation: from single cells to BOLD signals. *Trends Neurosci.* **29**, 250–256 (2006).
    This paper discusses fMRI adaptation designs.
14. Friston, K. J. et al. Analysis of fMRI time-series revisited. *Neuroimage* **2**, 45–53 (1995).
15. Haynes, J. D. & Rees, G. Decoding mental states from brain activity in humans. *Nature Rev. Neurosci.* **7**, 523–534 (2006).
16. Braitenberg, V. & Schuez, A. Cortex: Statistics and Geometry of Neuronal Connectivity 2nd edn (Springer, Berlin, 1998).
17. Douglas, R. J. & Martin, K. A. Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* **27**, 419–451 (2004).
    This paper provides a review of cortical microcircuits.
18. Ullman, S. Sequence seeking and counter streams: A computational model for bidirectional information flow in the visual cortex. *Cereb. Cortex* **5**, 1–11 (1995).
19. Friston, K. A theory of cortical responses. *Phil. Trans. R. Soc. B* **360**, 815–836 (2005).
20. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
21. Douglas, R. J. & Martin, K. A. Mapping the matrix: the ways of neocortex. *Neuron* **56**, 226–238 (2007).
22. Freund, T. F. Interneuron diversity series: Rhythm and mood in perisomatic inhibition. *Trends Neurosci.* **26**, 489–495 (2003).
23. Markram, H. et al. Interneurons of the neocortical inhibitory system. *Nature Rev. Neurosci.* **5**, 793–807 (2004).
    This paper is a review on the various types of interneuron.
24. DeFelipe, J. Types of neurons, synaptic connections and chemical characteristics of cells immunoreactive for calbindin-D28K, parvalbumin and calretinin in the neocortex. *J. Chem. Neuroanat.* **14**, 1–19 (1997).
25. Szabadics, J. et al. Excitatory effect of GABAergic axo-axonic cells in cortical microcircuits. *Science* **311**, 233–235 (2006).
26. Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. & Suarez, H. H. Recurrent excitation in neocortical circuits. *Science* **269**, 981–985 (1995).
27. Shadlen, M. N. & Newsome, W. T. Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* **4**, 569–579 (1994).
28. Chance, F. S., Abbott, L. F. & Reyes, A. D. Gain modulation from background synaptic input. *Neuron* **35**, 773–782 (2002).
29. Brunel, N. & Wang, X. J. Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.* **11**, 63–85 (2001).
30. Steriade, M., Timofeev, I. & Grenier, F. Natural waking and sleep states: a view from inside neocortical neurons. *J. Neurophysiol.* **85**, 1969–1985 (2001).
31. McCormick, D. A., Shu, Y. S. & Hasenstaub, A. Balanced recurrent excitation and inhibition in local cortical networks. in Excitatory-Inhibitory Balance: Synapses, Circuits, Systems (ed. Hensch, T.) (Kluver Academic Press, New York, 2003).
32. Haider, B., Duque, A., Hasenstaub, A. R. & McCormick, D. A. Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J. Neurosci.* **26**, 4535–4545 (2006).
    This paper provides a demonstration of the regulation of excitation–inhibition balance changes in vivo.
33. Sherman, S. M. & Guillery, R. W. Exploring the Thalamus and its Role in Cortical Function 2nd edn (MIT Press, Cambridge, Massachusetts, 2006).
34. Crick, F. & Koch, C. Constraints on cortical and thalamic projections: the no-strong-loops hypothesis. *Nature* **391**, 245–250 (1998).
35. Jueptner, M. & Weiller, C. Review: does measurement of regional cerebral blood flow reflect synaptic activity? Implications for PET and fMRI. *Neuroimage* **2**, 148–156 (1995).
36. Sokoloff, L. et al. The [^{14}C]deoxyglucose method for the measurement of local cerebral glucose utilization: Theory, procedure and normal values in the conscious and anesthetized albino rat. *J. Neurochem.* **28**, 897–916 (1977).
37. Nudo, R. J. & Masterton, R. B. Stimulation-induced [^{14}C]2-deoxyglucose labeling of synaptic activity in the central auditory system. *J. Comp. Neurol.* **245**, 553–565 (1986).

38. Ackermann, R. F., Finch, D. M., Babb, T. L. & Engel, J. Jr. Increased glucose metabolism during long-duration recurrent inhibition of hippocampal pyramidal cells. *J. Neurosci.* **4**, 251–264 (1984).

39. Stefanovic, B., Warnking, J. M. & Pike, G. B. Hemodynamic and metabolic responses to neuronal inhibition. *Neuroimage* **22**, 771–778 (2004).

40. Shmuel, A. *et al.* Sustained negative BOLD, blood flow and oxygen consumption response and its coupling to the positive response in the human brain. *Neuron* **36**, 1195–1210 (2002).

41. Shmuel, A., Augath, M., Oeltermann, A. & Logothetis, N. K. Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. *Nature Neurosci.* **9**, 569–577 (2006).

42. Devor, A. *et al.* Suppressed neuronal activity and concurrent arteriolar vasoconstriction may explain negative blood oxygenation level-dependent signal. *J. Neurosci.* **27**, 4452–4459 (2007).

43. Attwell, D. & Gibb, A. Neuroenergetics and the kinetic design of excitatory synapses. *Nature Rev. Neurosci.* **6**, 841–849 (2005).

44. Payne, J. A., Rivera, C., Voipio, J. & Kaila, K. Cation-chloride co-transporters in neuronal communication, development and trauma. *Trends Neurosci.* **26**, 199–206 (2003).

45. McCormick, D. A., Connors, B. W., Lighthall, J. W. & Prince, D. A. Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *J. Neurophysiol.* **54**, 782–806 (1985).

46. Buzsaki, G., Geisler, C., Henze, D. A. & Wang, X. J. Interneuron diversity series: Circuit complexity and axon wiring economy of cortical interneurons. *Trends Neurosci.* **27**, 186–193 (2004).

47. Wang, Y., Gupta, A., Toledo-Rodriguez, M., Wu, C. Z. & Markram, H. Anatomical, physiological, molecular and circuit properties of nest basket cells in the developing somatosensory cortex. *Cereb. Cortex* **12**, 395–410 (2002).

48. Buzsaki, G., Kaila, K. & Raichle, M. Inhibition and brain work. *Neuron* **56**, 771–783 (2007).

49. Peppiatt, C. M. *et al.* Bidirectional control of CNS capillary diameter by pericytes. *Nature* **443**, 700–704 (2006).

50. Hamel, E. Perivascular nerves and the regulation of cerebrovascular tone. *J. Appl. Physiol.* **100**, 1059–1064 (2006).

51. Kayser, C. & Konig, P. Stimulus locking and feature selectivity prevail in complementary frequency ranges of V1 local field potentials. *Eur. J. Neurosci.* **19**, 485–489 (2004).

52. Liu, J. & Newsome, W. T. Local field potential in cortical area MT: Stimulus tuning and behavioral correlations. *J. Neurosci.* **26**, 7779–7790 (2006).

53. Wilke, M., Logothetis, N. K. & Leopold, D. A. Local field potential reflects perceptual suppression in monkey visual cortex. *Proc. Natl Acad. Sci. USA* **103**, 17507–17512 (2006).

54. Logothetis, N. K. *et al.* Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).

55. Goense, J. B. M. & Logothetis, N. K. Neurophysiology of the BOLD fMRI signal in awake monkeys. *Current Biol.* **18**, 631–640 (2008).

56. Mathiesen, C., Caesar, K., Akgoren, N. & Lauritzen, M. Modification of activity-dependent increases of cerebral blood flow by excitatory synaptic activity and spikes in rat cerebellar cortex. *J. Physiol.* **512**, 555–566 (1998).

57. Viswanathan, A. & Freeman, R. D. Neurometabolic coupling in cerebral cortex reflects synaptic more than spiking activity. *Nature Neurosci.* **10**, 1308–1312 (2007).
This paper provides a demonstration of the coupling between $CMRO_2$ and the LFP.

58. Rauch, A., Rainer, G. & Logothetis, N. K. The effect of a serotonin-induced dissociation between spiking and perisynaptic activity on BOLD functional MRI. *Proc. Natl Acad. Sci. USA* **105**, 6759–6764 (2008).

59. Rockel, A. J., Hiorns, R. W. & Powell, T. P. The basic uniformity in structure of the neocortex. *Brain* **103**, 221–244 (1980).

60. Cragg, B. G. The density of synapses and neurones in the motor and visual areas of the cerebral cortex. *J. Anat.* **101**, 639–654 (1967).

61. Schuz, A. & Demianenko, G. P. Constancy and variability in cortical structure. A study on synapses and dendritic spines in hedgehog and monkey. *J. Hirnforsch.* **36**, 113–122 (1995).

62. Logothetis, N. K. & Wandell, B. A. Interpreting the BOLD signal. *Annu. Rev. Physiol.* **66**, 735–769 (2004).

63. Shoham, S., O'Connor, D. H. & Segev, R. How silent is the brain: is there a ''dark matter'' problem in neuroscience? *J. Comp. Physiol. A* **192**, 777–784 (2006).

64. Born, R. T. & Bradley, D. C. Structure and function of visual area MT. *Annu. Rev. Neurosci.* **28**, 157–189 (2005).

65. Zeki, S. Thirty years of a very special visual area, area V5. *J. Physiol.* **557**, 1–2 (2004).

66. Mather, G., Verstraten, F. A. & Anstis, S. M. *The Motion Aftereffect: a Modern Perspective* (MIT Press, Cambridge, Massachusetts, 1998).

67. Tootell, R. B. H. *et al.* Visual motion aftereffect in human cortical area MT revealed by functional magnetic-resonance-imaging. *Nature* **375**, 139–141 (1995).

68. Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L. & Petersen, S. E. Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* **248**, 1556–1559 (1990).

69. Huk, A. C., Ress, D. & Heeger, D. J. Neuronal basis of the motion aftereffect reconsidered. *Neuron* **32**, 161–172 (2001).

70. Blake, R. & Logothetis, N. K. Visual competition. *Nature Rev. Neurosci.* **3**, 13–21 (2002).

71. Douglas, R. J., Martin, K. A. C. & Whitteridge, D. A canonical microcircuit for neocortex. *Neural Comput.* **1**, 480–488 (1989).

72. Bartels, A., Logothetis, N. K. & Moutoussis, K. FMRI and its interpretations: An illustration on directional sensitivity in area V5/MT. *Trends Neurosci.* (in the press).

What we can do and what we cannot do with fMRI

Nikos K. Logothetis[1]

[1]Max Planck Institute for Biological Cybernetics, 72076 Tuebingen, Germany, and Imaging Science and Biomedical Engineering, University of Manchester, Manchester M13 9PL, UK.

MAGNETIC RESONANCE IMAGING

MRI Principles

There are many textbooks that thoroughly cover the theoretical and practical aspects of MRI (e.g.[38,108]) and its functional variants[11,48,69]. This brief overview is only given to provide an understandable definition of the terms and parameters commonly used in MR and fMR imaging.

Many nuclei possess a quantum mechanical property called spin, which can be most easily thought of as a tiny spinning magnet. For the purpose of this introduction we consider only spin-½ nuclei like the water and fat protons in biological tissues that dominate the signals measured in conventional MRI. Normally spins are randomly oriented; when an external magnetic field B_0 is applied, however, the spins orient themselves according to the magnetic field, i.e. they start to precess around the axis of the magnetic field (think of the motion of a gyroscope or spinning top). Because the parallel alignment with the external field is energetically more favorable than the antiparallel state, there will be a net excess of spins aligned with the field; i.e. the spin system can be represented by a *magnetization vector*. The difference between the number of parallel and antiparallel spins follows a Boltzmann distribution and increases with field strength. This difference in population is small, on the order of one in 10,000 000 spins at 1.5T, and NMR (*nuclear magnetic resonance*) is actually considered an insensitive technique. The spins precess according to the *Larmor frequency* (ω), which is related to the field through the *gyromagnetic ratio* (γ) (i.e. ω = γ * B_0).

When a radiofrequency (RF) field of amplitude B_1 rotating synchronously with the precessing spins is applied, the magnetization vector rotates away from its initial equilibrium position (aligned along the z-direction) by 90 degrees into the transverse (i.e. xy) plane, i.e. the longitudinal magnetization is converted to transverse magnetization. This happens only when the carrier frequency of the RF pulse is equal to the Larmor frequency, hence the term *magnetic resonance*. While on the transverse plane, the magnetization can be detected by an RF receiver coil.

The application of B_1 not only equalizes the populations of spins in the two energy levels, but also introduces phase coherence among the spins. Coherence decreases quickly as the magnetic moments move out of phase as a result of their mutual interaction. The transverse magnetization, in other words, is short-lived; it decays exponentially as a result of processes known as *relaxation*. There are different kinds of relaxation processes known as T_1, T_2, *or* T_2*, each reflecting different interactions of the spins with their environment or with other spins, and each specified by its time constant (T_i) or its inverse, the relaxation rate ($R_i = 1/T_i$). These relaxation rates differ depending on the properties of the tissue, and these differences are the basis of image contrast.

In biological and clinical sciences, most MRI signals are derived from the hydrogen nuclei of water, as the latter is the most abundant component (70-90%) of living tissues. The gyromagnetic ratio of protons is 42.58 MHz/Tesla, which means that the typical clinical magnet of 1.5T has a resonance frequency of approximately 64MHz, while a 7T magnet has a frequency of about 300 MHz. Field strength and frequency can be used

interchangeably to characterize the strength of a scanner (the former term is preferred in clinical research and fMRI and the latter in chemistry).

To generate an MR image, say an anatomical image of the brain, one needs a contrast mechanism that can separate different brain parts from one another (e.g. gray from white matter or from cerebrospinal fluids), and a mechanism by virtue of which that contrast can be calculated for each single volume element, commonly termed as *MRI voxel*.

MRI contrast can be generated from a number of different quantities, including regional *spin density*, *water diffusion*, or the *relaxation times* mentioned above. Most MR images in biomedicine rely on a clever selection of acquisition parameters that make the image differentially sensitive (also called *weighted*) to the relaxation process. It is impressive indeed to see how the choice of these simple parameters affects the contrast between the various tissues of the brain. For example, the usual anatomical images of the brain are most often T_1-weighted images, exploiting the differences in T_1 of gray and white matter.

Just as relaxation-rate differences at different sites might be used to generate *anatomical contrast*, changes over time in one or more relaxation rates at a single site might be exploited to image changes in the physical-chemical state. This very simple principle underlies the *functional contrast* in fMRI, which often uses T_2^*-weighted imaging to detect changes in relaxation times (T_2^*) brought about by differences in blood oxygenation between test and control epochs.

Spatial localization, on the other hand, is achieved with the use of smaller magnetic field gradients that are superimposed on the homogenous magnetic field of the scanner and by subsequently exploiting the aforementioned Larmor relationship. According to the latter, the positions of protons within a voxel at different positions along the gradient field are encoded by their differences in resonance frequency.

Conventional MR imaging does not measure the signal decay immediately after the RF pulse. Instead, the decaying signal is "recovered" to some extent by *refocusing* the dephasing spins with additional RF pulses or by appropriately changing the magnetic field gradients used for spatial localization. The time between the peak of the RF excitation pulse and that of the

recovered signal, i.e. the echo, is termed *echo time (TE)*. Together with the *repetition time (TR)* these two parameters characterize the contrast and quality of an MR image. The combination of different types of RF and gradient pulses used to obtain the desired MR image is known as a *pulse sequence*. There are very many different pulse sequences, but two are most often used in cognitive applications: *gradient-recalled echo (GRE or GE)* and *spin echo (SE)*. They differ in the way they refocus the dephasing spines, whereby the former increases the coherence of precession of spins via gradient reversals, and the latter by means of RF pulses.

The type of information obtained from the usual MR images depends on the parameters just described. Pulse sequences whose TR is much longer than the longitudinal relaxation times T_1 and whose TE is much shorter than T_2 are typically used to obtain spin-density images. The usual anatomical images are obtained with pulse sequences having a TR value less or equal to T_1 and a TE value that is much shorter than the transverse relaxation time T_2. Using TR values that are much longer than T_1 and TE values that are equal to or larger than T_2 or T_2^*, on the other hand, produces T_2- and T_2^*-weighted images, respectively, such as those used in functional imaging. Nonetheless, when high temporal resolution is of importance, the TR in functional imaging can be even shorter than T_1.

Finally, an image can be acquired by using multiple excitation pulses, but also following a single RF pulse with a method known as *echo planar imaging or EPI*[85]. Two types of EPI are most commonly used: *spin-echo (SE-EPI)* and *gradient-echo (GE-EPI)* echoplanar imaging. They differ in the way they refocus the spins prior to data acquisition, and they can be fine-tuned to yield the T_2- and T_2^*-weighted images used in functional neuroimaging.

Functional MRI

Fortuitously, the physical and chemical events underlying the relaxation-rate changes can reflect alterations of neural activity in a reliable way. Relaxation rates may be affected by spin motion (e.g. blood flow), diffusion, or by changes in field homogeneity. Correspondingly, functional activations in the brain can be detected with MRI via direct measurements of *tissue perfusion*[17], *blood*

oxygenation[73], *blood volume*[8], or reportedly even *water diffusion*[58].

The first demonstration of the capacity of MRI to map cortical activation through hemodynamics was accomplished using the clinically well-established intravascular contrast agent Magnevist. When the blood-brain barrier (BBB) is intact, this lanthanide-gadolinium chelate, Gd-DTPA, is confined to the intravascular space, and this compartmentalization is the basis of the observed magnetic susceptibility effects in the surrounding tissues. Following an intravenous bolus injection, Magnevist produces significant signal changes during first-pass cerebral transit. Rapid imaging techniques detecting susceptibility differences in tissues (e.g. gradient-echo echo-planar imaging GE-EPI) can detect signal decreases during the passage of the intravascular agent through the capillary bed, and accurate estimates of the *regional cerebral blood volume (rCBV)* can be made by integrating over the first passage of the agent[84].

In their seminal study, Belliveau and his colleagues at Harvard University investigated the visual cortex of human subjects looking at a flickering visual pattern that was known to produce large hemodynamic changes. Regional CBV maps of the visual cortex in resting and activated states were obtained[8]; when these were superimposed on high-resolution anatomical images, the activated cortical regions were precisely delineated, demonstrating the potential of the technique to map brain areas involved in sensory stimulation, and potentially in cognitive processing as well. Shortly after this *CBV-fMRI* experiment three groups simultaneously and independently published the first functional maps of the human brain based on an endogenous intravascular contrast agent, the *deoxyhemoglobin (dHb)* of the blood[5,57,77].

The functional images utilized the *blood-oxygen-level-dependent (BOLD)* contrast, which depicts differences in blood oxygenation. This phenomenon was discovered by *Seiji Ogawa*, who noticed that the contrast of very high resolution brain images ($65\times65\times700\mu m^3$) acquired with GE pulse sequences depicts numerous dark lines of varying thickness. These lines were invisible when SE sequences were used, and they turned out to be signal dropouts from blood vessels[74]. Shortly thereafter, this effect was demonstrated in the cat brain during the course of experimentally induced anoxia[100].

Most neuroimaging today relies on signal variations induced by such local changes in field homogeneity. In biological samples inhomogeneities are induced by different materials that respond differently to the scanner field. Molecules like water weakly counteract the applied magnetic field and are known as *diamagnetic*; in contrast, molecules like gadolinium compounds slightly enhance the magnetic field and are called *paramagnetic*. The proximity of paramagnetic substances decreases T_2^*. An inhomogeneous distribution of paramagnetic materials introduces a field distortion into an MRI image known as *susceptibility artifact*.

Almost all non-invasive fMRI is possible because the hemoglobin of the blood - which is sequestered in the blood red cells and thus inhomogeneously distributed in the tissues - is diamagnetic when it is oxygenated and paramagnetic when it is reduced, an effect first described by Linus Pauling and his student Coryell[79], who used a simple apparatus consisting of a wire, a glass tube, a balance, and a small electromagnet! The wire was suspended from one arm of the balance and held the tube that was placed between the electromagnet poles. They filled it with cattle blood, balanced it, and measured its weight. Passing an electric current through the coils changed the apparent weight, and this change was proportional to the amount of dHb in the blood.

With this simple but brilliant method, Pauli and Coryell reported that the magnetic susceptibility of fully oxygenated arterial blood differs by as much as 20% from that of fully deoxygenated venous blood. In fact, oxygenated blood is diamagnetic, minimally affecting the magnetic field just like the water and most of the macromolecules of the tissues, while deoxygenated blood is paramagnetic, producing a clearly measurable, additive magnetic field. The physical-chemical explanation of this behavior is straightforward: hemoglobin consists of two pairs of polypeptide chains (globins), each of which is attached to a complex of protoporphyrin and iron (heme group). In dHb the iron (Fe^{2+}) is in a paramagnetic high-spin state, as four of its six outer electrons are unpaired and act as a paramagnetic agent. When oxygenated, the heme iron changes to a low-spin state by receiving the oxygen's electrons, and Hb becomes diamagnetic.

Studies revealed via NMR that blood T_2 is indeed directly affected by its oxygenation[96], suggesting that T_2 changes alone could be used for functional imaging. In

the meantime, of course, this is a standard method. In a 1.5T clinical magnet, a change in oxygenation from 60% to 70% changes the net intravascular T_2 by 40% in venules and 15% in capillaries. I note here in passing that venules and capillaries have an equal share (each 40%) of any particular voxel's blood volume, but venules are approximately twice as magnetic.

The usual signal increases reported in BOLD fMRI experiments are due to the fact that neural activation induces a regional increase in *cerebral blood flow (CBF)* and glucose utilization that is always larger than the *oxygen consumption rate (CMRO2)*[25,26], since oxygen uptake is diffusion-limited. The net effect of neural excitation is thus a seemingly paradoxical drop in the deoxyhemoglobin concentration, which in turn increases the signal strength.

It is now clear that the BOLD signal depends on the $CMRO_2$ as well as on CBF and cerebral blood volume (CBV), thereby representing a complex response controlled by several parameters[76]. Despite this complexity, much progress has been made toward quantitatively elucidating various aspects of the BOLD signal and the way it relates to the hemodynamic and metabolic changes occurring in response to changes in neuronal activity[54,65].

Spatial Specificity

The primary biological factor determining the specificity of the fMRI signal is the density and architecture of the brain's microvasculature. Intracortical vessels vary in their degree of cortical penetration, ranging from those having a short course, i.e. branching close to the pial surface, to those penetrating the entire width of the cortical sheet and covering a large territory. Their diameter ranges from 20μm to 250μm, and their branching pattern depends on the cortical layer[20]. Successive branching of arteries ultimately creates the capillary bed, a network of capillaries whose thin walls and huge combined surface area allows the ready exchange of oxygen, energy substrates and metabolic wastes. The inter-capillary mesh size (ca. 50μm in cortex) is probably dictated by the diffusion coefficient of oxygen. The human brain has no arteriovenous anastomoses[20], i.e. blood vessels that branch and reconnect, so an increase of arterial flow most likely always leads to a commensurate increase in perfusion, the flow of arterial blood into the capillaries, which in turn enhances delivery of oxygen and nutrients in the parenchyma. An increase in neural activity will evidently result in oxygenation changes in all vessels.

A recent study has also provided the first detailed results on the microvascular system of the macaque primate cortex using an array of anatomical techniques that included corrosion casts, immunohistochemistry, and cytochrome oxidase (COX) staining[106]. Detailed measurements of regional vascular length density, volume fraction, and surface density revealed a similar vascularization in different visual areas. In the lower cortical layers a positive correlation between the vascular and cell density was found, but this relationship was very weak in the upper layers. Interestingly, the vascular density was instead strongly correlated with the steady-state metabolic demand as measured by COX activity, suggesting that the number of neurons and synapses determines an upper limit to an area's integrative capacity; vascularization actually reflects the neural activity of those subpopulations that represent a "default" mode of brain steady state.

FMRI specificity depends on the extent to which the methodology reflects primarily the oxygenation changes in the capillary bed and the surrounding tissue rather than in large vessels (e.g. 100 μm and larger). A significant contribution of the latter, which depends on the area of cortex activated[99], can generate BOLD contrast downstream of the tissue with high metabolic activity, resulting in a mislocalization of activation patterns and an erroneous estimation of their extent. Unless the draining vein is remote from the site of activation this is a minor problem for many cognitive studies using voxels of several millimeters linear dimension, but it is of critical importance for any fMRI study with high spatial resolution.

Strategies for improving specificity have been convincingly demonstrated, although they have yet to be fully implemented in most neuroscientific applications. For example, the contribution of vessels of different sizes towards the BOLD signal depends on the strength of the magnetic field and the MR-pulse sequence employed[75]. High magnetic fields de-emphasize the contribution of large vessels and increase the specificity of the fMRI signal. The reasons for the field benefit are twofold. First, local field inhomogeneities are less forgiving at high field: T_2^* relaxation times become shorter as the field strength increases. This short T_2^*

decreases the visibility of signals from larger vessels. Second, extravascular BOLD increases with field strength more rapidly for small vessels than it does for large vessels. This means that the higher the field the more signal is obtained proportionately from the capillary bed and the parenchyma, which is the actual site of activation. The effects of field strength on specificity have been convincingly demonstrated in animal studies, in which precise cortical maps of columnar organization could be obtained[101].

For any given field strength, signal specificity can also be increased by using pulse sequences that are less sensitive to signals from within and around large vessels. The GE-EPI used by the vast majority of cognitive studies is sensitive to (almost) all vessel sizes[76,107]. In contrast, SE-EPI favors signals from small vessels and parenchyma[95], especially at high field when the signal from venous blood becomes very small.

MRI signal losses can originate from static field inhomogeneities like local variations in field near a large vessel, which can be effectively recovered by spin-echo methods, where an RF pulse refocuses the spins that have previously dephased. Signal loss due to dynamic processes, on the other hand, cannot be refocused. This occurs when spins diffuse in a local field gradient (for instance the field gradient produced by a capillary); because of their displacement, they experience a different field before and after the refocusing pulse, resulting in incomplete refocusing.

This recovery of the MRI signal from static inhomogeneities (near large vessels) accomplished by SE-EPI weights the extravascular contribution toward the microvasculature[107]. Last but not least, because signal changes in SE-EPI originate from the average water diffusion, they show apparent T_2, i.e. water diffusion sensitive transverse relaxation, rather than T_2^* dependency, and are therefore less affected by intravascular effects[10,76,107]. Two drawbacks of SE-EPI, however, are decreased sensitivity and imaging speed compared with GE-EPI.

Other fMRI methods that are more specific to capillary signal than GE-BOLD are CBV and CBF methods (e.g. [19,111]). In CBV methods T_2^* of blood is reduced by the injection of iron-containing contrast agent, which eliminates all signal from large and medium-sized vessels. In CBF methods the arterial blood is tagged.

After an appropriate waiting period the tagged blood arrives in the capillary bed, so if the proper waiting period is chosen, most of the signal will be arising from the capillary bed.

Spatial Resolution

The spatial resolution of MRI can be reliably determined by calculating the *point spread function (PSF)* of the imaging method, which is a collective measure of all factors contributing to image blurring, including intrinsic resolution of the imaging system and any necessary spatial smoothing during or following the imaging procedure. The PSF of an imaging method is determined by having a delta function (i.e. a point source with very high intensity) as input and by determining the width of the resulting spot in the final image. When this concept is applied to the transfer function of the whole fMRI procedure, a point source of 100% contrast is assumed as visual input. The size of the activation is then determined by the neural and vascular transfer function, i.e. the representation of the point source on, say, the primary visual cortex, by imaging factors, i.e. resolution, blurring and post-processing/smoothing, and by neurovascular factors, i.e. vascular smearing and the spatial extent of the hemodynamic response.

Brian Wandell and colleagues were the first to quantify the spatial specificity of the GE BOLD response by estimating its PSF with an elegant paradigm that exploits the retinotopic organization of the primary visual cortex[24]. They found that at 1.5T the width of the BOLD PSF at half its maximal height (i.e. full-width-half-maximum or FWHM) is 3.5 mm. Similar values, 3.9 mm for GE-BOLD and 3.4 mm for SE-BOLD, have been reported at 3T[78]. Predictably, the PSF of GE-BOLD is narrower at 7T, and it can be as small as 2 mm when signals from macrovessels are masked out[87]. Methods such as SE-BOLD, CBV imaging at high fields, and CBF fMRI, which stress oxygenation changes in the capillary bed and parenchyma, further reduce PSF width to 1 mm or less[101]. It should be noted that such measurements only set an upper bound for the width of PSF, because it may also be limited by factors unrelated to imaging, such as organization of cortex or subject behavior. In the visual system, for instance, the PSF will necessarily be affected by eye movements, as well as by intracortical connectivity patterns (e.g. horizontal

connections in early visual cortices) that may spread local activation horizontally[34].

Alternatively to PSF estimation, the limits of the fMRI spatial resolution can be assessed by resolving functional subunits of the brain that have already been shown by invasive methods. The cortical columns of neocortex are prominent examples of such structurally and functionally specialized subunits, whose organization has been studied extensively in the early sensory areas of animals[70]. Not surprisingly, visualization of such structures by means of BOLD, CBF, or CBV imaging has been thought of as the ultimate test for the resolving power of the method and has been attempted ever since the advent of fMRI.

In humans the ocular dominance columns have a mean width of roughly 1 mm[1]. Convincing maps of these columns in normal human subjects were reported only a few years ago using high-field (4T) fMRI that permitted a resolution of $0.47\times0.47mm^2$ in-plane for a slice thickness of 3mm[16]. However, successful mapping relied on explicit masking out of regions having non-specific large vessels, a procedure that cannot be easily applied in other cortical areas with less well-defined neural and vascular architecture. An alternative to this area-specific macrovessel exclusion is the application of subtraction designs on data collected at very high magnetic field (7T) and by using the SE-BOLD pulse sequence[110]. The combination of field strength and SE-BOLD enabled the unconstrained mapping of human ocular dominance columns with an in-plane spatial resolution of $0.5\times0.5mm^2$ and a slice thickness of 3mm. In animals, i.e. cats, orientation columns (~1 mm diameter) have also been shown using GE-BOLD, CBV and CBF-based methods (e.g.[19,111]).

Leaving aside certain limitations imposed by the scanner hardware, spatially resolved MRI/fMRI usually suffers from substantial drops in signal-to-noise (SNR) and contrast-to-noise (CNR) ratios that occur as the voxel size decreases. In a single-channel acquisition system small surface coils are often used to substantially improve SNR and increase resolution. Such coils increase SNR both because they can be placed very close to the region of interest and because they reduce the *sensitive volume*, i.e. the volume from within which signals are received, thus decreasing noise by eliminating signals originating in irrelevant tissues. In animal studies, for instance, resolutions as high as

$0.25\times0.25mm^2$ with a slice thickness of 2mm can be achieved with either GE or SE BOLD using either implantable or surface coils[31]. Yet, in this methodology there is a clear tradeoff between resolution and coverage, the latter being indeed one of the greatest virtues of fMRI and one that cannot be easily given up. But this can be recovered to some extent by parallel imaging.

Temporal Resolution

High spatial resolution is needed to resolve anatomical and functional detail, but fast scanning is needed to achieve a detailed description of successive neural events. Temporal resolution is constrained by both technical and physiological – mostly vascular – factors. Technically, temporal resolution is determined by the sampling rate of the volumes. With conventional acquisition methods the fastest technique with adequate signal-to-noise is single-shot echo planar imaging (EPI)[85], provided that the scanner has strong and fast gradients and that a number of other parameters have been sufficiently optimized to avoid image distortions. The total time needed for the acquisition of an entire image with single-shot EPI can be as short as 40ms. The drawback is that single shot EPI is very sensitive to field inhomogeneities and other artifacts (e.g. T2* blurring). Many of the problems of single-shot EPI can be overcome by using multi-shot (segmented) EPI, which collects data over multiple excitations (shots), although this incurs the risk of significant motion artifact. Currently, using BOLD fMRI with segmented GE-EPI one can collect single slices at a sampling rate of less than 100ms, and volumes of multiple slices at a sampling rate of 1-3 seconds. High sampling rates (e.g. 100 ms or less), however, require special caution to avoid activation-independent signal increases, e.g. inflow effects[21].

Developments and Perspectives

Temporal resolution, as well as SNR and CNR problems in spatially resolved imaging can be dealt with to some extent by using multiple MR-signal acquisition channels, each collecting data from a different small RF coil[47,109]. Parallel imaging with such coil arrays can tremendously improve the nominal resolution of MRI and fMRI, in particular if superconducting (and hence very low noise) coils are used. In fact, within the last 8 years, parallel

imaging methods have become commercially available and are now in broad clinical use.

Parallel MRI (pMRI) works by taking advantage of the spatial sensitivity information inherent in an array of multiple receiver surface coils to speed up the time-consuming spatial encoding normally performed by switching magnetic field gradients.

In the MRI introduction above I briefly indicated that spatial localization in MRI is achieved by using smaller magnetic field gradients that are superimposed on the homogenous magnetic field of the scanner. This strategy exploits the Larmor relationship and encodes spatial information in one direction (e.g. x or *readout* direction) in terms of frequency. The usual two-dimensional image is obtained by encoding the other direction (e.g. y or *phase-encoding* direction) into the phase of MR signals. Phase- and frequency-encoding directions are perpendicular to each other. While the phase-encoding gradient (PEG) is on, the Larmor frequency becomes linearly proportional to the position along the phase-encoding direction. The phase shift that accumulates by the time the PEG is turned off characterizes each "line" (or row) of the two dimensional image. Accordingly, in a MR image of 128 lines - each 128 voxels long - 128 acquisitions will be required to obtain the complete image. In contrast, with pMRI only a fraction of the phase-encoding steps have to be acquired, resulting in substantially faster image acquisition yet still maintaining full spatial resolution and image contrast.

This decreased acquisition time can obviously be used to improve spatial resolution as well. In addition, pMRI diminishes susceptibility-caused artifacts by reducing the echo train length of single- and multi-shot pulse sequences. Currently, the best known pMRI methods are the so-called *simultaneous acquisition of spatial harmonics (SMASH)*[89], *sensitivity encoding (SENSE)*[81], and *generalized autocalibrating partially parallel acquisitions (GRAPPA* [35]. For an overview on the advantages and disadvantages of these methods see the review by Blaimer and colleagues[9].

COGNITIVE AND NEURAL SCIENCES

On Pure Insertion

Serial subtraction designs rely strictly on *pure insertion,* a concept introduced in the context of reaction times

(RT). In the 1960s, the work of Donders[18] had proposed that the time between stimulus and response is occupied by a train of successive processes, or stages, each component of which begins only when the preceding one has ended. For example, the visual search for an item defined by the conjunction of two or more stimulus dimensions (e.g. shape and color) is linearly (additively) related to the number of distractors in the visual field [98]. Empirical demonstrations of similar additive relationships between increments in cognitive load and reaction times were later the basis of Sternberg's proposal [93]. Sternberg improved and extended Donders' method to better study various stages of information processing. The main feature of his *additive-factor* method was the search for non-interacting effects of experimental factors on mean RT. He suggested that stage durations may be additive without being stochastically independent, and that the effect of a new experimental factor can be localized among a set of already established stages. In other words, according to this pure insertion hypothesis, a single cognitive process can be inserted into a task without affecting the remainder. This assumption is the basis of a large number of experimental designs and analysis methods in cognitive fMRI.

But the idea of pure insertion in the context of brain activation studies is actually misleading and all too often simply not tenable. Even if an experimental design could satisfy this assumption at the cognitive level, the assumption would be condemned to fail utterly at the level of its neuronal instantiation due to the highly non-linear nature of most brain processes[2]. This fact was elegantly demonstrated by Karl Friston and colleagues[28] in their critical assessment of cognitive subtraction as a conceptual framework in brain activation studies.

Neurotransmission and Neuromodulation

Sensory information reaches cortex via fast, mostly glutamatergic or aspartergic synapses, typically producing direct excitatory effects on the postsynaptic neurons via the AMPA and NMDA receptors. The dynamics of activation are strongly affected by recurrent inhibition mediated by GABAergic interneurons. Together, the glutamatergic and GABAergic neurons are responsible for a major part of neurotransmission, which in turns affects the regional cerebral blood flow (CBF). Yet, CBF is also considerably affected by the overall

regulation of cortical dynamics and cell excitability mediated by a number of other neurochemicals, including acetylcholine, norepinephrine, serotonin, dopamine, and various peptides. Although such chemicals were initially defined as classical neurotransmitters in peripheral systems, in cortex they appear to have an important neuromodulatory role. Neurotransmission and neuromodulation are often discriminated on the basis of the activated receptors, with neurotransmitters targeting receptors that are coupled directly to ion channels and neuromodulation affecting receptors coupled to channels via second messenger pathways.

Neuromodulatory innervation originates from various nuclei in the brainstem and basal forebrain and spreads out in a rather diffuse fashion to broad cortical regions[41]. Well studied examples include the noradrenergic ascending system primarily originating in the locus ceruleus (LC), the serotonergic system in nuclei near the midline and the raphe regions of the pons and upper brainstem, the cholinergic system innervating neocortex from the nucleus basalis of Meynert, and the dopaminergic system from the ventral tegmental area (VTA).

In relation to neurovascular coupling, it is important to note that neurotransmission and neuromodulation may often involve different types of interneuronal communication with potentially different spatial specificity and spatiotemporal resolution. Neuromodulation is slow and diffuse, and while it may be induced by sensory stimulation, it is not necessarily stimulus-specific[46].

The lack of strict specificity is also evident in the different types of interneuronal communication that characterize neuromodulation. Many axons of the neuromodulatory nuclei, which typically run long distances through various cortical structures, have certain characteristic *non-junctional varicosities* with high densities of small vesicles and absence of synaptic specializations (for a review see[112]). Such varicosities are often located far from any recognizable postsynaptic density. In fact, their serial reconstruction of processes originating in noradrenergic, serotonergic, cholinergic, and dopaminergic nuclei has demonstrated that about 80% of their varicosities do not really match the local postsynaptic densities. It thus appears that there exists a frequent spatial uncoupling between release sites for

transmitters and their respective receptors, implying that the transmitter may in principle be poured out in the extracellular space and affect tissue volumes containing relevant receptors over distances much larger than those of the typical synaptic cleft.

On the basis of such observations, two different communication frames have been suggested: *wiring* and *volume transmission*[3]. Wiring transmission implies the existence of specialized communication channels within the neuronal and/or glia cell network, and includes the well-studied synaptic transmission and gap junctions. Volume transmission, on the other hand, is characterized by signal diffusion in a three-dimensional fashion within the brain's extracellular fluid. While such nonjunctional volume transmission is the rule in the peripheral autonomic nervous system, it has also been proposed for the CNS[104]. In the cerebral cortex of monkeys, for example, acetylcholine can reach both pyramidal and nonpyramidal neurons via volume transmission; Cortical noradrenaline innervation is mostly nonjunctional, and the innervation of the forebrain by the dorsal and median raphe nuclei is characterized by spaced varicosities distributed across all areas of the cerebral cortex[45]. The activity of the diffuse ascending systems may thus influence the activation level of large cell populations, and it may do so in a manner less specific to the function in question.

Neural Signals

In order to understand the relationship between the BOLD signal and its underlying neural events it is necessary to comprehend the nature of the neurophysiological signals commonly reported in animal studies, as these are the telltale signs of the underlying neuronal processes (see also reviews[62,63,65]). Physiological studies at the systems and behavioral level in anesthetized or conscious animals typically report extracellular recordings. While the mechanisms at work during the monitoring of transmembrane electrical events with intracellularly placed electrodes are reasonably well understood, the interpretation of different types of extracellular recordings, in particular those reporting on the activity of neural masses, proved to be difficult and it still requires discussion.

Signals Spread in the Extracellular Microenvironment, Which Acts As a Volume Conductor

Neurons are considered to be embedded in an extracellular medium that acts as a *volume conductor*[66] (for a detailed review on field potentials see[27]; for background in biophysics see textbooks[4,49]). The extracellular microenvironment consists of narrow gaps between cellular processes an average of about 200 Å wide. These spaces form a complex three-dimensional mosaic filled with extracellular fluid, and they account for about 12 to 25% of the brain's volume. Currents and ions spread in this space, and as theoretical reasoning suggests, spread mostly in the extracellular fluid between the cells, but not through them[83]. Hence, the resistance depends on the spatial layout of neurons and glia, resulting in an intricately shaped conductive medium that, in principle, can carry electrical signals over large distances. Confirmation comes from direct measurements of current flow and studies of the diffusion of ions in this microenvironment[42,72,82,102,103]. Given these intricate properties of the extracellular medium it is important to know how this medium conducts the variety of signals generated by subthreshold and spiking activity.

The Volume Conductor Is Quasi-Static, Ohmic and Tangentially Isotropic

Within the physiological frequency range (0 to about 5 kHz), the inductive, magnetic, and propagative (wave) effects of the bioelectrical signals in the extra-cellular space can be neglected[67,83], permitting the description of a current electrode as a simple static point source. Moreover, any description of the volume conductor is further simplified by the fact that the propagation of signals is independent of their frequency[64]. For many years the volume conductor has been described as having a capacitive and anisotropic nature, implying a frequency- and direction-dependent signal propagation (e.g. see review[65]). Yet recent intracranial measurements using a novel variant of the four-point technique[86] showed that tissue impedance is actually frequency-independent, allowing the description of cortex as an ohmic resistor. As such, the tissue can be described by its specific electrical resistance (or resistivity). Finally, the resistive properties of the gray matter are largely isotropic, i.e. its resistivity is the same along each direction[64]. These results have important implications for the origin and spatial summation of different types of neural signals (see below).

Dipoles and Dipole-Layers in the Volume Conductor

What do extracellularly placed electrodes actually measure? When a neuron discharges, it undergoes an increase in conductivity over the excitable membrane regions, i.e. usually at the axon hillock and/or soma. A current flowing into the cell across the regionally increased conductance will flow along the core of the cell, and then exit at various regions of adjacent, inactive membrane to ultimately return to the site of current entry by way of diverse paths through the volume conductor. The active regions of the membrane at any given time point are considered to act as a current *sink*, while the inactive ones as a *source* for the active regions. During the generation of the action potential the neuron can be considered an *electric dipole* as the dendrites are positive with respect to the soma. The aforementioned non-zero resistivity of the volume conductor is the reason that the current flow in the extracellular fluid during the activity of neurons generates measurable gradients of potential.

These *extracellular field potentials (EFP)* add up linearly and algebraically throughout the volume conductor (the principle of electric superposition), representing the weighted sum of all sinks and sources along multiple cells. The superposition principle sets some constraints in the interpretation of the *mean extracellular potential (mEFP)* commonly measured with electrodes. Specifically, the mEFP critically depends on the local geometrical arrangements of neurons. For cells with diametrically opposite orientations, currents of equal magnitude but opposite polarity will generate potentials that tend to cancel each other. As a result, information is potentially lost that cannot be recovered without additional knowledge (e.g. anatomical or intracellularly obtained single-unit data). For certain geometrical arrangements, the way the current flows and the position of sinks and sources along the membranes can be calculated. For others they cannot.

Open, Closed and Open-Closed Fields

Three arrangements produce characteristic mEFP: the *open*, the *closed*, and the *open-closed fields*. The first is encountered when the neurons are organized in a laminar array, with their dendrites facing in one direction and the

somata in the other; typical examples of open field are the neocortex, the cerebellum and the hippocampus. Simultaneous activation of the dendrites of neurons in this arrangement generates strong dipole layers whose field potentials are easily captured by electrodes. The second arrangement produces dipoles with spherical symmetry; the polarity of a measurement within the sphere depends on its location within the spherical dipole, while measurements from outside would record zero potential. Lastly, the third arrangement yields a mixture, whereby measurements within the spherical dipole sense combination of the two fields, while those outside the sphere would simply record changes in the open field component.

Single and Multiple Unit Activity vs. Local Field Potentials

When a microelectrode with a small tip is placed close to the soma or axon of a neuron, then the measured mEFP directly reports the spike traffic of that neuron and frequently that of its immediate neighbors as well. The recordings then reflect *single unit activity (SUA).* Tetrodes placed close (within 50 μm) to pyramidal neurons in hippocampus provide accurate information on a number of their parameters such as latency, amplitude and shape of action potentials, as has been demonstrated by means of simultaneous intracellular recordings[40,43]. Single spike monitoring has the best possible spatial and temporal resolution but it provides information mainly on single receptive fields, with no access to subthreshold integrative processes or the associational operations taking place at a given site. Moreover, it is biased toward certain cell types (cf.[94]) and sizes[97]. The measured spikes mostly represent only very small neural populations of large cells, which in cortex are by and large the principal cells (e.g. pyramidal cells in cerebral cortex and Purkinje neurons in cerebellar cortex). Recording from interneurons (e.g. inhibitory cells) is often very difficult both because of their size and because their response is often found to be uncorrelated to the stimulus or behavior state of the animal.

Nonetheless, more than SUA can be measured with microelectrodes. In fact, two different signal types can be easily extracted from the mEFP by using traditional Fourier band separation or the wavelet approach. A high-pass filter cutoff of approximately 1000Hz can be used to obtain *multiple-unit spiking activity (MUA)*[29], and a low-pass filter cutoff of ca. 300Hz to obtain the so-called *local field potentials (LFPs).* MUA has been often calculated by using lower frequency cutoffs (300-800Hz); yet lower cutoffs make far-field potentials noticeable and may occasionally (e.g. by averaging) lead in erroneous MUA interpretations[59]. A large number of experiments have presented data indicating that such a band separation does indeed underlie different neural events. MUA-range activity reflects the variations in the magnitude of extracellular spike potentials. LFPs, on the other hand, represent mostly slow events reflecting cooperative activity in neural populations[62,63,65].

New insights into the generation of LFPs, as well as into intracortical processing in general, have come from the study of inhibitory networks in hippocampus[14,53,56]. These studies provided evidence of the existence of other types of slow activity unrelated to synaptic events, including voltage-dependent membrane oscillations (e.g.[52]) and spike afterpotentials. To be more specific, the soma-dendritic spikes in the neurons of the central nervous system are generally followed by afterpotentials, a brief delayed depolarization, the *afterdepolarization*, and a longer lasting *afterhyperpolarization*, which are thought to play an important role in the control of excitation-to-frequency transduction (e.g.[32,37,39]). Afterpotentials, which were shown to be generated by calcium-activated potassium currents (e.g.[15,39,44,55,105]) have a duration on the order of 10s of milliseconds and most likely contribute to the generation of the LFP signals[12,13]. The delta waves of the EEG signal, for example, are not necessarily a result of synaptic activity, but rather reflect the summation of long-lasting afterhyperpolarizations of layer V pyramidal neurons; the suppression of delta waves during neocortical arousal is mainly due to blockade of this hyperpolarization by cholinergic input.

Band Limited Power (BLP) Signals in the LFP Range

Traditionally, low-frequency signal modulations are classified in a number of specific frequency bands initially introduced in the EEG literature (e.g.[22,23,80]). Rhythmic EEG is subdivided into frequency bands known as *delta* (δ, 0-4Hz), *theta* (θ, 4-8Hz), alpha (α, 8-12Hz), *beta* (β, 12-24Hz), and *gamma* (γ, 24-40/80Hz) which are typically characterized by different amplitudes[6,61,90,91].

This classification was based on the strong correlation of each band with a distinct behavioral state. The oscillatory activity in these bands is associated with the thalamocortical loops and is modulated by the ascending network system and basal forebrain[90,92]. But rhythmic activity related to some of the EEG frequency bands has also been reported for the spiking principal neurons. Many pyramidal neurons in layer 5 of the neocortex show prolonged, 5- to 12-Hertz rhythmic firing patterns due to intrinsic membrane properties such as sodium conductance, which is essential for rhythmicity, and calcium-dependent conductance, which strongly modifies it[88]. Although synaptic networks of intrinsically rhythmic neurons may still be the origin of the synchronized cortical oscillations, spiking activity in this case will be tightly correlated with the LFPs and will contribute to the modulation of their amplitude.

The definition of the above bands is empirical, and often the range of a particular band is determined arbitrarily, varying, in the case of the gamma band, for example, from 20-50, 20-70, 24-60, or 24-90Hz. An alternative band separation is that based on information theory. Recently the information carried by individual BLP in the LFP or MUA range was calculated in extracellular recordings during the presentation of movies with natural images[7]. The most informative LFP frequency ranges were *1-8Hz* and *60-100Hz*. LFPs in the range of 20-60Hz carried very little information about the stimulus, although they shared strong trial-to-trial correlations, indicating that they might be influenced by a common source, such as diffuse neuromodulatory input. The upper range of the latter band is may often have transition characteristics between the two nearby regions. LFPs in the range of *12-40Hz* are distinct and most likely are the best reflection of neuromodulatory input.

Spike power, on the other hand, was informative only at frequencies *below 12 Hz* (frequency of bursts). Positive signal correlations were found between LFPs (60-100Hz) and spikes, as well as between signals within this LFP range, suggesting that the 60-100Hz range of LFPs and the spikes might be generated within the same network.

Spatial Summation of Neuronal Signals

The summation range for the fast MUA has been studied by a number of investigators. Electrodes with exposed tips of approximately 100μm (impedance from 40-120kΩ), for example, were estimated to record from a sphere with a radius of 50-350μm[33,36,60], whereby the activity from each point within the sphere is weighted by a factor depending on the distance of the point from the tip of the electrode [71].

LFPs reflect a weighted average of synchronized dendro-somatic components of the synaptic signals of a neural population within 0.5-3 millimeters of the electrode tip[51,68]. The upper limits of the spatial extent of LFP summation were indirectly calculated by computing the phase coherence of LFPs as a function of inter-electrode distance in experiments with simultaneous multiple-electrode recordings[50].

Recently the spatial summation of different BLPs in the LFP range was estimated by *reverse correlation* and *coherence analysis*. Reverse correlation indicated that the site-RF sizes (square root of the area) are about 6.3 mm (delta-theta), 4.2 mm (20-60Hz), 2.3 mm (gamma), 1.9 mm (MUA), and 1.5 mm (SUA). The half-maximum of coherence-to-distance functions, on the other hand, show a coupling region of 2.9mm (2-8Hz), 2.38mm (8-15Hz), 1.94mm (20-60Hz), and 1.48mm for MUA[30]. These findings suggest that the spatial summation is in the range of a typical voxel size in high resolution imaging for both the LFP band as well as the MUA bands. The stronger coupling of different bands of LFP to BOLD can therefore not be explained by a difference in spatial summation.

REFERENCE LIST

1. D. L. Adams, L. C. Sincich, and J. C. Horton, Complete pattern of ocular dominance columns in human primary visual cortex, 27(39), 10391 (2007).
2. A Aertsen and H Preissl, Dynamics of Activity and Connectivity in Physiological Neuronal Networks,in *Nonlinear Dynamics and Neuronal Networks,* edited by H Schuster (VCH Verlag, Weinheim, 1991), pp.281-301.
3. L. F. Agnati *et al.*, Intercellular communication in the brain: wiring versus volume transmission, Neuroscience 69(3), 711 (1995).
4. D. J. Aidley, *the physiology of excitable cells*, 3 ed. (Cambridge University Press, Cambridge, 1989).
5. Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S., and Hyde, J. S., Time course EPI of human brain function during task activation, Magnetic Resonance in Medicine 25(2), 390-397 (1992).
6. E. Basar, *eeg-brain dynamics: relation between eeg and brain evoked potentials* (Elsevier/North Holland Biomedical Press, Amsterdam, New York, Oxford, 1980).
7. Belitski, A., Gretton, A., Magri, C., Murayama, Y., Montemurro, M. A., Logothetis, N. K., and Panzeri, S., Low-

frequency local field potentials and spikes in primary visual cortex convey independent visual information, Journal of Neuroscience (in press) (2008).

8. Belliveau, J. W., Kennedy, D. N., McKinstry, R. C., Buchbinder, B. R., Weisskoff, R. M., Cohen, M. S., Vevea, J. M., Brady, T. J., and Rosen, B. R., Functional mapping of the human visual cortex by magnetic resonance imaging, Science 254, 716-719 (1991).

9. M. Blaimer, *et al.*, SMASH, SENSE, PILS, GRAPPA: how to choose the optimal method, Top. Magn Reson. Imaging 15(4), 223 (2004).

10. Boxerman, J. L., Hamberg, L. M., Rosen, B. R., and Weisskoff, R. M., MR contrast due to intravascular magnetic susceptibility perturbations, Magnetic Resonance in Medicine 34(4), 555-566 (1995).

11. R. B. Buxton, *introduction to functional magnetic resonance imaging: principles and techniques* (Cambridge University Press, Cambridge, UK., 2002).

12. Buzsaki, G., Theta oscillations in the hippocampus, Neuron 33(3), 325-340 (1-31-2002).

13. G. Buzsaki, *et al.*, Nucleus basalis and thalamic control of neocortical activity in the freely moving rat, 8(11), 4007 (1988).

14. G. Buzsaki and J. J. Chrobak, Temporal structure in spatially organized neuronal ensembles: a role for interneuronal networks. [61 refs], 5(4), 504 (1995).

15. S. H. Chandler, *et al.*, Electrophysiological properties of guinea pig trigeminal motoneurons recorded in vitro, Journal of Neurophysiology. 71(1), 129 (1994).

16. Cheng, K., Waggoner, R. A., and Tanaka, K., Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging, Neuron 32(2), 359-374 (2001).

17. Detre, J. A., Leigh, J. S., Williams, D. S., Koretsky, A. P., Detre, J. A., Leigh, J. S., Williams, D. S., and Koretsky, A. P., Perfusion imaging, Magnetic Resonance in Medicine 23(1), 37-45 (1992).

18. F. C. Donders, On the speed of mental processes, Acta Psychol. (Amst) 30, 412 (1969).

19. T. Q. Duong, *et al.*, Localized cerebral blood flow response at submillimeter columnar resolution, Proc. Natl. Acad. Sci. U. S A 98(19), 10904 (2001).

20. Duvernoy, H. M., Delon, S., and Vannson, J. L., Cortical blood vessels of the human brain, Brain Research Bulletin 7(5), 519-579 (1981).

21. J. H. Duyn, *et al.*, Inflow Versus Deoxyhemoglobin Effects in Bold Functional Mri Using Gradient Echoes at 1.5 T, NMR Biomed. 7(1-2), 83 (1994).

22. R. Elul, The physiological interpretation of amplitude histograms of the EEG, 27(7), 703 (1969).

23. R. Elul, The genesis of the EEG. [100 refs], 15, 227 (1971).

24. S. A. Engel, G. H. Glover, and B. A. Wandell, Retinotopic organization in human visual cortex and the spatial precision of functional MRI, Cereb. Cortex 7(2), 181 (1997).

25. Fox, P. T. and Raichle, M. E., Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects, Proc.Natl.Acad.Sci.U.S.A 83(4), 1140-1144 (1986).

26. Fox, P. T., Raichle, M. E., Mintun, M. A., and Dence, C., Nonoxidative glucose consumption during focal physiologic neural activity, Science 241(4864), 462-464 (7-22-1988).

27. W. J. Freeman, *mass action in the nervous system* (Academic Press, New York, 1975).

28. Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S., and Dolan, R. J., The trouble with cognitive subtraction, Neuroimage. 4(2), 97-104 (1996).

29. A. Gail, H. J. Brinksmeyer, and R. Eckhorn, Perception-related modulations of local field potential power and coherence in primary visual cortex of awake monkey during binocular rivalry, 14(3), 300 (2004).

30. J. B. Goense and N. K. Logothetis, Neurophysiology of the BOLD fMRI Signal in Awake Monkeys, Curr. Biol. 18(9), 631 (2008).

31. J. B. Goense, A. C. Zappe, and N. K. Logothetis, High-resolution fMRI of macaque V1, Magn Reson. Imaging (2007).

32. R. Granit, D Kernell, and R. S Smith, Delayed depolarization and the repetitive response to intracellular stimulation of mammalian motoneurones, 168, 890 (1963).

33. C. M. Gray, *et al.*, Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex, 63, 43 (1995).

34. Grinvald, A., Lieke, E. E., Frostig, R. D., and Hildesheim, R., Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex, Journal of Neuroscience 14(5:Pt 1), t-68 (1994).

35. M. A. Griswold, *et al.*, Generalized autocalibrating partially parallel acquisitions (GRAPPA), Magn Reson. Med. 47(6), 1202 (2002).

36. F. S. Grover and J. S. Buchwald, Correlation of cell size with amplitude of background fast activity in specific brain nuclei, 33(1), 160 (1970).

37. B. Gustafsson, Afterpotentials and transduction properties in different types of central neurones, 122(1), 17 (1984).

38. E. M. Haacke, *et al.*, *magnetic resonance imaging: principles and sequence design* (Wiley-Liss: John Wiley & Son, Inc, New York, 1999).

39. Y. Harada and T. Takahashi, The calcium component of the action potential in spinal motoneurones of the rat, 335, 89 (1983).

40. K. D. Harris, *et al.*, Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements, 84(1), 401 (2000).

41. Hasselmo, M. E., Neuromodulation and cortical function: modeling the physiological basis of behavior, Behav.Brain Res. 67(1), 1-27 (1995).

42. J. W. Havstad, Ph.D. Stanford University, 1976.

43. D. A. Henze, *et al.*, Intracellular features predicted by extracellular recordings in the hippocampus in vivo, 84(1), 390 (2000).

44. H. Higashi, *et al.*, Ionic mechanisms underlying the depolarizing and hyperpolarizing afterpotentials of single spike in guinea-pig cingulate cortical neurons, 55(1), 129 (1993).

45. J. P. Hornung, The human raphe nuclei and the serotonergic system, J Chem. Neuroanat. 26(4), 331 (2003).

46. Hurley, L. M., Devilbiss, D. M., and Waterhouse, B. D., A matter of focus: monoaminergic modulation of stimulus coding in mammalian sensory networks, Curr.Opin.Neurobiol. 14(4), 488-495 (2004).

47. Hyde, J. S., Froncisz, W., Jesmanowicz, A., and Kneeland, J. B., Planar-pair local coils for high-resolution magnetic resonance imaging, particularly of the temporomandibular joint, Med.Phys. 13(1), 1-7 (1986).

48. P. Jezzard, P. M. Matthews, and S. M Smith, *functional magnetic resonance imaging: an introduction to methods* (Oxford University Press, Oxford New York, 2002).

49. D. Johnston and S. M. Wu, *foundations of cellular neurophysiology* (MIT Press, Cambridge, Massachusetts, 1995).

50. E. Juergens*, et al.*, Restricted Coupling Range of Fast Oscillations in Striate Cortex of Awake Monkey,in *Brain and Evolution,*(Thieme, Berlin, New York, 1996), p.418.

51. E. Juergens, A. Guettler, and R. Eckhorn, Visual stimulation elicits locked and induced gamma oscillations in monkey intracortical- and EEG-potentials, but not in human EEG, 129(2), 247 (1999).

52. Kamondi, A., Acsady, L., Wang, X. J., and Buzsaki, G., Theta oscillations in somata and dendrites of hippocampal pyramidal cells in vivo: activity-dependent phase-precession of action potentials, Hippocampus 8(3), 244-261 (1998).

53. A. Kandel and G. Buzsaki, Cellular-synaptic generation of sleep spindles, spike-and-wave discharges, and evoked thalamocortical responses in the neocortex of the rat, 17(17), 6783 (1997).

54. Kim, S. G. and Ugurbil, K., Comparison of blood oxygenation and cerebral blood flow effects in fMRI: estimation of relative oxygen consumption change, Magnetic Resonance in Medicine 38(1), 59-65 (1997).

55. M. Kobayashi*, et al.*, Role of calcium conductances on spike afterpotentials in rat trigeminal motoneurons, Journal of Neurophysiology. 77(6), 3273 (1997).

56. B. Kocsis, A. Bragin, and G. Buzsaki, Interdependence of multiple theta generators in the hippocampus: a partial coherence analysis, 19(14), 6200 (1999).

57. Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., and Turner, R., Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation, Proc.Natl.Acad.Sci.USA 89(12), 5675-5679 (6-15-1992).

58. Le Bihan, D., The 'wet mind': water and functional neuroimaging, Physics in Medicine and Biology 52(7), R57-R90 (4-7-2007).

59. A. D. Legatt, J. Arezzo, and H. G. Vaughan, Jr., Averaged multiple unit activity as an estimate of phasic changes in local neuronal activity: effects of volume-conducted potentials, J. Neurosci. Methods 2(2), 203 (1980).

60. A. D. Legatt, J. Arezzo, and H. G. J. Vaughan, Averaged multiple unit activity as an estimate of phasic changes in local neuronal activity: effects of volume-conducted potentials, 2(2), 203 (1980).

61. D. B. Lindsley and J. D. Wicke, The Electroencephalogram: Autonomous Electrical Activity in Man and Animals,in *Electroencephalography and Human Brain Potentials,* edited by R. F Thomson and M. M. Patterson (Academic Press, New York, 1974), pp.3-83.

62. N. K. Logothetis, The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. [270 refs], 357(1424), 1003 (2002).

63. N. K. Logothetis, The underpinnings of the BOLD functional magnetic resonance imaging signal. [125 refs], 23(10), 3963 (2003).

64. N. K. Logothetis, C. Kayser, and A. Oeltermann, In vivo measurement of cortical impedance spectrum in monkeys: implications for signal propagation, 55(5), 809 (2007).

65. Logothetis, N. K. and Wandell, B. A., Interpreting the BOLD signal, Annual Review of Physiology 66, 735-769 (2004).

66. Lorente de Nó, R., Analysis of the distribution of action currents of nerve in volume conductors, Studies from the Rockefeller Institute Medical Res 132(A Study of Nerve Physiology), 384-477 (1947).

67. U. Mitzdorf, Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. [324 refs], 65(1), 37 (1985).

68. U. Mitzdorf, Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex, 33(1-2), 33 (1987).

69. C. T. Moonen and P. A. Bandettini, *functional mri* (Springer Verlag, Berlin, 1999).

70. V. B. Mountcastle, Modality and topographic properties of single neurons of cat's somatic sensory cortex, 20(4), 408 (1957).

71. C. Nicholson and R. Llinas, Field potentials in the alligator cerebellum and theory of their relationship to Purkinje cell dendritic spikes, 34(4), 509 (1971).

72. Nicholson, P. W., Specific Impedance of Cerebral White Matter, Experimental Neurology 13, 386-401 (1965).

73. Ogawa, S. and Lee, T. M., Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation, Magnetic Resonance in Medicine 16(1), 9-18 (1990).

74. S. Ogawa*, et al.*, Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields, Magn Reson. Med 14(1), 68 (1990).

75. Ogawa, S., Menon, R. S., Kim, S. G., and Ugurbil, K., On the characteristics of functional magnetic resonance imaging of the brain, Annual.Review of Biophysics & Biomolecular.Structure 27, 447-474 (1998).

76. S. Ogawa*, et al.*, Functional Brain Mapping by Blood Oxygenation Level-Dependent Contrast Magnetic-Resonance-Imaging - A Comparison of Signal Characteristics with A Biophysical Model, 64(3), 803 (1993).

77. Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., and Ugurbil, K., Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging, Proceedings of the National Academy of Sciences of the United States of America 89(13), 5951-5955 (7-1-1992).

78. Parkes, L. M., Schwarzbach, J. V., Bouts, A. A., Deckers, R. H., Pullens, P., Kerskens, C. M., and Norris, D. G., Quantifying the spatial resolution of the gradient echo and spin echo BOLD response at 3 Tesla, Magn Reson.Med. 54(6), 1465-1472 (2005).

79. Pauling, L and Coryell, C, The magnetic properties and structure of hemogblogin, Proc Natl Acad Sci U S A 22, 210-216 (1936).

80. T. A. Pedley and R. D. Traub, Physiological Basis of the EEG,in *Current Practice of Clinical Electroencephalography,* 2 ed. edited by D. D Daly and T. A. Pedley (Raven Press, New York, 1990), pp.107-137.

81. K. P. Pruessmann*, et al.*, SENSE: Sensitivity encoding for fast MRI, Magn. Reson. Med. 42(5), 952 (1999).

82. J. B. Ranck, Analysis of Specific Impedance of Rabbit Cerebral Cortex, 7(2), 153 (1963).

83. D. A. Robinson, The Electric Properties of Metal Microelectrodes, 56(6), 1065 (1968).

84. B. R. Rosen, *et al.*, Susceptibility contrast imaging of cerebral blood volume: human experience, Magn Reson. Med 22(2), 293 (1991).

85. F. Schmitt, M. K. Stehling, and R. Turner, *echo-planar imaging: theory, technique and application* (Springer, Berlin, 1998).

86. H. P. Schwan and C. D. Ferris, 4-Electrode Null Techniques for Impedance Measurement with High Resolution, 39(4), 481-& (1968).

87. A. Shmuel, *et al.*, Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla, 35(2), 539 (2007).

88. L. R. Silva, Y. Amitai, and B. W. Connors, Intrinsic oscillations of neocortex generated by layer 5 pyramidal neurons, 251(4992), 432 (1991).

89. Sodickson, D. K. and Manning, W. J., Simultaneous acquisition of spatial harmonics (SMASH): Fast imaging with radiofrequency coil arrays, Magnetic Resonance in Medicine 38(4), 591-603 (1997).

90. M. Steriade, Alertness, Quiet Sleep, Dreaming,in *Cerebral Cortex,*(Plenum Press, New York, London, 1991), pp.279-357.

91. M. Steriade and J. Hobson, Neuronal activity during the sleep-waking cycle, 6(3-4), 155 (1976).

92. M. Steriade, D. A. McCormick, and T. J. Sejnowski, Thalamocortical oscillations in the sleeping and aroused brain. [70 refs], 262(5134), 679 (1993).

93. S. Sternberg, Discovery of Processing Stages - Extensions of Donders Method, Acta Psychol. 30, 276-+ (1969).

94. Stone, J., Sampling properties of microelectrodes assessed in the cat's retina, J Neurophysiol 36, 1071-1079 (1973).

95. K. R. Thulborn, *et al.*, High-resolution echo-planar fMRI of human visual cortex at 3.0 tesla, NMR Biomed. 10(4-5), 183 (1997).

96. Thulborn, K. R., Waterton, J. C., Matthews, P. M., and Radda, G. K., Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field, Biochimica et Biophysica Acta 714(2), 265-270 (2-2-1982).

97. A. L. Towe and G. W. Harding, Extracellular microelectrode sampling bias, 29(2), 366 (1970).

98. A. M. Treisman and G. Gelade, A feature-integration theory of attention, Cognit. Psychol. 12(1), 97 (1980).

99. R. Turner, How much cortex can a vein drain? Downstream dilution of activation-related cerebral blood oxygenation changes, 16(4), 1062 (2002).

100. R. Turner, *et al.*, Echo-planar time course MRI of cat brain oxygenation changes, Magn Reson. Med 22(1), 159 (1991).

101. Ugurbil, K., Adriany, G., Andersen, P., Chen, W., Garwood, M., Gruetter, R., Henry, P. G., Kim, S. G., Lieu, H., Tkac, I., Vaughan, T., Van de Moortele, P. F., Yacoub, E., and Zhu, X. H., Ultrahigh field magnetic resonance imaging and spectroscopy, Magnetic Resonance Imaging 21(10), 1263-1281 (2003).

102. Van Harreveld, A., Murphy, T, and Nobel, K. W., Specific impedance of rabbit's cortical tissue, Am.J.Physiol. 205, 203-207 (1963).

103. Van Harreveld, A. and Ochs, S., Cerebral Impedance Changes After Circulatory Arrest, Am.J.Physiol 187, 180-192 (1956).

104. Vizi, E. S., Kiss, J. P., and Lendvai, B., Nonsynaptic communication in the central nervous system, Neurochemistry International 45(4), 443-451 (2004).

105. K. Walton and B. P. Fulton, Ionic mechanisms underlying the firing properties of rat neonatal motoneurons studied in vitro, 19(3), 669 (1986).

106. B. Weber, *et al.*, The Microvascular System of the Striate and Extrastriate Visual Cortex of the Macaque, Cerebral Cortex, Doi:10. 1093/Cercor/Bhm259 (2008).

107. Weisskoff, R. M., Zuo, C. S., Boxerman, J. L., and Rosen, B. R., Microscopic susceptibility variation and transverse relaxation: theory and experiment, Magnetic Resonance in Medicine 31(6), 601-610 (1994).

108. M. L. Wood and F. W. Wehrli, Principles of Magnetic Resonance Imaging,in *Magnetic Resonance Imaging.,* 3 ed. edited by D. D. Stark and W Bradley (Mosby, St. Louis, Baltimore, Boston, London, Tokyo, 1999), pp.1-14.

109. S. M. Wright, R. L. Magin, and J. R. Kelton, Arrays of Mutually Coupled Receiver Coils - Theory and Application, Magn. Reson. Med. 17(1), 252 (1991).

110. E. Yacoub, *et al.*, Robust Detection of Ocular Dominance Columns in Humans Using Hahn Spin Echo BOLD Functional MRI at 7 Tesla,in 2006)

111. F. Q. Zhao, *et al.*, Spatial speciricity of cerebral blood volume-weighted fMRI responses at columnar resolution, 27(2), 416 (2005).

112. M. Zoli, *et al.*, Volume transmission in the CNS and its relevance for neuropsychopharmacology, Trends Pharmacol. Sci. 20(4), 142 (1999).