# Functional MRI

Robert Turner and Karl Friston

**Contents**



## 1) INTRODUCTION

In the spring of 1991 the first of a group of new methods for mapping human brain function was successfully tested at the MRI Center of the Massachusetts General Hospital, in Boston. These methods use the supremely noninvasive technique of magnetic resonance imaging to create images of the brain that are sensitive to local changes in blood flow. In identification of cortical areas active for a particular brain task, they promise a spatial resolution of 2 mm, a temporal resolution of 1 s, and the opportunity to rescan a single subject as often as desired. As might have been expected, legions of brain researchers have since leapt upon magnetic resonance functional neuroimaging, or functional MRI (fMRI) as it is more frequently called, attempting to apply it to various studies in cognitive neuroscience. Whilst the underlying phenomenon is simple, in the sense that many MRI scanners were found to have the capability of performing fMRI studies once the technique had been discovered, quantification of the effect in relation to cortical electrical activity has been elusive, and for those unversed in MRI the origins of the functional signal require some explanation.

## 2) PRINCIPLES OF FMRI

### *2.1 Sources of Contrast in MRI*

The MRI signal in normal clinical use comes almost entirely from the protons of tissue water. The image intensity depends primarily on the density of these protons, of course, but can be profoundly affected by the local evironment of the water molecules. After excitation with a pulse of radiofrequency magnetic field ('rf pulse'), which changes the alignment of the proton magnetic moments from being parallel to the large static field produced by the MRI magnet, the protons recover their original alignment quite slowly, in several tenths of a

second to seconds. During this time the magnetization which has thus been created in the direction transverse to the steady magnetic field induces a signal voltage in an antenna surrounding the object or body being scanned. If the water protons are excited again before full recovery, a smaller signal is obtained. The rate of recovery, described by a 'longitudinal relaxation time' T1, depends on the type of tissue containing the relevant water molecules. For instance, in cerebrospinal fluid, which is close to pure water, protons relax in about 3 seconds, while in brain white matter, where water molecules interact strongly with lipid membranes and intracellular organelles, the recovery time is reduced to about 0.5 s.

Clearly, by varying the repeat time between rf pulses, the contrast between tissue regions of long and short T1 can be changed dramatically. But another type of relaxation also offers a means to distinguish tissue types. As we have seen, in order to observe an MRI signal the proton magnetization must be tipped away from the main field direction, to create a component of magnetization precessing about this axis, in the transverse plane. To produce maximum signal, the phase angle which the magnetization vector makes in the transverse plane must remain constant across the object, allowing the magnetization to add up constructively from each proton. However, small differences in the magnetic environment of each spin will cause them to precess at slightly different frequencies, resulting in a dispersion of phase angles and hence a signal which decreases with time. The decay of signal is generally roughly exponential, and can be characterized by a time constant T2, the 'transverse relaxation time'. In addition, if there are magnetic field variations across the object, which may be caused by the presence within the object of paramagnetic particles, such as an MRI contrast agent containing gadolinium, or simply by the spatial inhomogeneity of the object, there is a further phase dispersion causing a more rapid decay of the signal. This additional relaxation is denoted by T2'. Together the two effects result in a signal decaying with a time constant T2*, where $1/T2^* = 1/T2 + 1/T2'$.

*2.2 Bold Contrast fMRI*

Until the late 1980's the decay of signal associated with local magnetic field inhomogeneities, called T2* relaxation, was considered a nuisance and represented a limitation upon MR imaging. To mitigate this, either the so-called 'spin-echo' technique was used, in which a second 'refocussing' rf pulse following the initial excitation pulse removed the effects of dephasing, or the time between the rf excitation pulse and the acquisition of signal was reduced as much as possible, as in FLASH (Fast Low-Angle SHot imaging) (1).

It was only when it was realized that the presence of a paramagnetic substance in the bloodstream could act as a vascular marker, giving useful contrast, that sequences without a refocussing pulse, with a relatively long time between the excitation pulse and data acquisition (20-80 ms), began to be used. Initially the paramagnetic contrast agent was exogenous, a non-toxic compound of gadolinium introduced into the bloodstream via a leg vein. A fraction of a millimole of contrast agent per kilogram of body weight is sufficient to give a loss of signal from tissue surrounding cerebral blood vessels of perhaps 40% as a bolus of contrast agent passes through. The group at MGH pioneered the application of this approach to brain perfusion, using the Central Volume Theorem (2) to obtain local cerebral blood flow as the ratio of blood volume to mean transit time. Early

studies (3,4) examined the passage of contrast agent through the brains of rats and later dogs, making increasing use of an ultra-fast MRI imaging technique known as echo-planar imaging (EPI), invented by Sir Peter Mansfield in 1977 (5), which acquires a complete image in less than 100 ms, and thus allows 'snapshots' of the contrast agent distribution as it passes rapidly through the brain.

In a landmark paper published in 1991, Belliveau and collaborators finally applied the technique to functional activation studies in humans (6). The subjects were given visual ('photic') stimulation while a bolus of contrast agent was injected into a leg vein, and single slice images of the brain in the plane of the calcarine fissure were obtained at 0.75 s intervals to monitor the bolus passage. By integrating the time course of image intensity, estimates of relative blood volume were obtained, and compared (by image subtraction) with those obtained when the subjects were at rest in darkness. Consistent increases (up to 30%) of blood volume in primary V1 visual cortex were observed.

Subsequent developments rapidly overtook this pioneering work. Ogawa (7) and Turner(8) working independently, had shown in laboratory animals that similar changes of MRI image contrast extending around the blood vessels could be obtained simply by changing the oxygenation state of the blood. This observation arose from the fact (noted by Faraday, and measured by Pauling and Coryell (9)) that deoxyhemoglobin is more paramagnetic than oxyhemoglobin, which itself has almost exactly the same magnetic susceptibility as tissue. Thus deoxyhemoglobin can be seen as nature's own contrast agent. Interventions to the state of the brain which create an imbalance between oxygen uptake and blood flow will thus inevitably cause a change of MRI signal around the cortical vessels, if MR imaging sequences are used which are sensitive to magnetic field inhomogeneity.

This development culminated in the work of Kwong et al (10) and Ogawa et al (11), who succeeded in showing that the change in deoxyhaemoglobin in human visual cortex, while the subject viewed a bright light, was sufficient to cause measurable changes in gradient-echo MRI images of a slice passing through the calcarine fissure. The technique was dubbed 'Blood Oxygenation Level Dependent Contrast" (BOLD). Thus the way was opened to functional mapping studies of the human brain without use of contrast agent, no radiation dose, and with the high spatial resolution of MRI.

It is significant that a rise in signal can be observed during visual stimulus, indicating a relative decrease in the concentration of paramagnetic deoxyhemoglobin. This confirms earlier work using PET (12) in which the measured rise in oxygen uptake rate during visual stimulus was found to be much smaller than the rise in blood flow. Earlier observation during open skull surgery (e.g. (13)) had also demonstrated that the blood leaving active cortical regions is brighter red, i.e. more oxygenated, than normal, as a result of this mismatch between demand and supply.

In theory, the change of signal can be affected by changes in blood arterial oxygenation, blood volume, blood flow, hematocrit, tissue oxygen uptake and possibly blood velocity. It increases with field strength (14). Changes in blood flow apparently dominate, via dilution of deoxyhaemoglobin, but detailed modeling and

experimental investigation are in progress (15, 16, 17, 18). Recently a number of studies (19,20) have shown that a significant fraction of the water molecules responsible for the change in signal observed in BOLD contrast are in fact inside unresolvable venules or veins draining the active cortex. This represents a limitation on the spatial resolution, which will be discussed further below.

Since these pioneering studies, many laboratories have explored a diversity of brain tasks using BOLD contrast fMRI.

BOLD methods for functional brain mapping suffer from certain problems which can give false positive results, or obscure the effect altogether. It must be understood that the signal changes of interest are usually not much larger than the thermal and physiological noise which inevitably accompany MR images of living tissue, so that equipment stability is of great importance. But two other issues have been the subject of controversy in the literature. These relate to the precision of localization of neural activity, and to artifactual image differences associated with subject motion, which will not be discussed here.

*2.2.1 Localization*

It is not yet clear how well the increase in delivery of blood to brain tissue is spatially matched to the increased metabolic demand caused by neural electrochemical activity. The early qualitative observations of Roy and Sherrington did not involve microscopy of the cortex, and techniques such as $^{15}$O-water Positron Emission Tomography (PET), which rely on co-localization of flow and neural activity increases, have a spatial resolution of no better than 5 mm. On this scale it is evident that there is an excellent spatial match between flow changes and electrical activity. Single-electrode studies show a good correlation between activity and flow, as measured with iodo-antipyrine (21) in animal models, when cortical areas a few mm across are considered. But it is obvious that at a sufficiently fine scale the match must fail, since blood occupies only about 2% of the cortical volume.

Cerebral perfusion is controlled largely by the smooth muscle walls of the smaller arterioles, which dilate or contract according to the demand from tissue which they supply. However, the entire arterial circulation is capable of vasoactivity, and quite large arteries can expand to increase blood flow if required, as in the case of hypercapnia. This implies that it is possible for local increases in neural metabolism to be accompanied by an increase in blood flow over a wider area of cortex. A mechanism for this has been investigated by Duling and collaborators (22), where the arteriolar endothelium has been shown capable of propagating for 2 mm or further an electrotonic spread of hyperpolarization leading to vasodilation. This has been suggested as a mechanism for co-ordination of vascular control of blood flow over the arteriolar distance scale of a few millimeters, but it could also give rise to surplus perfusion.

It is possible thus to interpret the data from visual transcranial studies made by Grinvald and co-workers (23), using visual stimulation of lightly anesthetized cats or monkeys. Changes in blood oxygenation are indicated by the differential absorption of light of a specific wavelength in the red (610 nm). In these experiments (24, 25) a

two-stage change in oxygenation resulting from monocular visual stimulation can be observed. Initially, the relevant ocular dominance columns show a relative loss of signal, indicating an increased capillary deoxyhemoglobin concentration, resulting from greater oxygen extraction that can be associated directly with increased neural activity in the same tissue. However, after a period of one or two seconds, along an extensive network of small sub-pial venules, there is a large increase of intensity of reflected light which overlies the active cortical tissue and extends perhaps 2-3 mm beyond it, along the larger draining vessels, which indicates that the blood in these vesels has become hyperoxygenated.

This finding is economically explained by the suggestion that the increase of blood flow, triggered by the neural activity, takes place in arterioles which supply wider cortical regions than those actually active. If the messenger which transmits the vasodilatory signal travels by diffusion from active tissue to nearby arterioles, such a degree of spatial mismatch is quite possible.

The implication of this hypothesis is that BOLD contrast fMRI, and indeed arterial spin tagging (AST), to be discussed in the next section, cannot normally be expected to provide a spatial resolution within the cortex of better than about 2 mm, the distance by which the BOLD signal, and the increase of blood flow, extends beyond the area of activated cortical neurons. This resolution would be obtained for tasks which differentially activate small areas of cortex (up to say 25 $mm^2$, say). For subcortical structures the spatial regulation of blood flow may be more precise. Where complementary tasks are selected for which two cortical areas share identical arterial supply and venous drainage vessels, such as adjacent ocular dominance columns, it is possible that suitably designed differential experiments, using high resolution single-slice fMRI, may resolve activity down to a scale of 0.5 mm or better. This is feasible because the activated neuronal tissue has a greater oxygen extraction from the local capillary bed than non-activated tissue, even when the more poorly localised increase in blood flow leads to increased oxygenation in both areas. Thus BOLD signal will increase in both areas, but less in the activated area (26). Model calculations show that for reasons of increased BOLD contrast sensitivity, and greater differential sensitivity to small vessel effects, it is desirable to work at high field strengths if such high spatail resolution is to be obtained.

The effective resolution is likely to be significantly poorer than 2 mm if a large area of cortex (several $cm^2$) is activated, when draining veins quite distant from the active neurons may also experience a considerable change of flow and oxygenation (13).

*2.2.2 Inflow Effects*

A number of MRI studies (27,28), using high resolution FLASH-based imaging techniques, have shown activated structures with very much the appearance of sulcal vessels of perhaps 0.5 mm diameter, for a simple finger-tapping task. Such an appearance is not found with EPI studies, nor with some FLASH fMRI experiments. Much depends on the choice of imaging sequence parameters. Use of FLASH methods with a relatively large flip angle, of $30^o$ or greater, makes the image very sensitive to changes in flow velocity. The geometry of the venous circulation (29,30) implies that activation-related changes in flow velocity can propagate

relatively unchanged from the capillary bed through to the larger veins. By contrast, oxygenation changes originating from a small fraction of the area of a large vein's watershed become rapidly diluted downstream by blood which has not changed its state of oxygenation. Researchers continuing to use FLASH methods (e.g. 31) now prefer to use a very low flip angle, $10^{o}$ or less, which drastically reduces the image sensitivity to flow changes. However, to ensure confidence, it is still important to design experiments such that areas of differential neuronal activation are small.

Use of EPI methods, with a TR as long as 3 s, ensures that flow effects are minimal, since the nuclear spins are nearly fully relaxed by the time of the next rf pulse. With such MRI sequences, sensitivity to BOLD activation in larger vessels may be reduced by suppressing signal from more rapidly flowing blood. This can be done by applying so-called 'diffusion gradients' (19,20) which cause rapid dephasing of spins in flowing blood, or by using an asymmetric spin-echo sequence (32). By these means a high degree of confidence can be maintained that the areas of increased MR signal indeed represent increased neural activity, although a quantitative relationship has not yet been established.

*2.3 Arterial Spin Tagging*

Meanwhile, another MRI method for noninvasive monitoring of blood flow has been developed. This approach, begun at Pittsburgh in 1990 (33), uses a blood labelling technique akin to diffusible tracer studies. In this case, the label is applied using a preparatory radio-frequency pulse, before the MRI data acquisition sequence, which tags the spins of protons in arterial blood flowing into the brain region of interest. The label is evanescent, disappearing with the relaxation time constant T1(typically 1-2 s), so that it is not possible to build up a large concentration of labelled spins in brain tissue. However, the label, which consists of inverting the nuclear spins, freely diffuses through the capillary walls into the tissue and reduces the MR signal there by up to 2%, sufficient to be measured with reasonable accuracy if care is taken (33). Variations on this approach, now known as arterial spin tagging (AST) have been developed, in which the spins in the slice to be imaged are labelled, while unlabelled arterial spins move into the slice and change the image intensity proportionally to local CBF. Potentially AST methods can be quantified to give absolute values for local cerebral perfusion, although a number of important correction terms need to be evaluated.

AST has more recently been combined with EPI to measure relative changes in blood flow during activation. EPISTAR (Echo-Planar MR Imaging and Signal Targeting with Alternating Radio frequency) (34) involves inversion of the MR spins in a thick slab of the brain/head which is physically close to the section of interest. The effect of labelled spins, carried by the blood into the imaging slice, is observed as a signal change within the slice, and blood flow changes can be monitored by comparing images with and without activation. This method enables haemodynamics to be investigated by varying the delay between inversion and detection. The change of signal related to brain activity observable with this technique depends on the relaxation time T1, both directly (34), and indirectly since the magnetic labelling of the blood decays with the T1 of blood. T1 increases with magnetic field, and thus the sensitivity of EPISTAR should also increase with field strength.

*2.4 Imaging Techniques*

For human brain activation studies it is highly desirable to obtain image data quickly. There are several reasons for this, apart from the obvious need to avoid experiments lasting many hours. The first is that many perceptual and cognitive tasks of interest can be continued only for a few minutes without habituation, fatigue, or boredom. Secondly, the spatial resolution is generally on the order of 1-2 mm, which means that effective head immobilization is essential. The longer a subject is lying in an uncomfortable position inside the MRI magnet, the greater the chance of large movement. Thirdly, it is important to sample the activation state of the whole brain as synchronously as possible. Since MRI usually obtains image data one slice at a time, and 20-30 slices are necessary to cover the entire brain, this implies acquisition of a slice in a time very short compared with the haemodynamic response time of the cerebral vasculature of 6-8 s. The only successful MRI technique capable of this speed with reasonable spatial resolution and good signal: noise ratio is Echo-Planar Imaging (35), as previously mentioned.

It has to be mentioned that compared with slower MRI techniques, the spatial resolution of EPI is somewhat impaired (to about 2 mm), while the temporal resolution is improved to about 100 ms. Furthermore, the rate of information capture, that is, the signal to noise per unit time, is highest for EPI and its variants (such as single-shot GRASE (36), and single-shot spiral EPI (37)). However, EPI images (especially the gradient-echo version of EPI) suffer more than the other techniques from distortion and signal loss arising from magnetic field inhomogeneities in the brain, which often derive from the natural susceptibility differences between brain and air and are thus not amenable to correction by improved shimming (adjustment of the static field homogeneity with a set of correction coils).

*2.5 BOLD and AST fMRI Compared*

The difference in susceptibility between fully oxygenated and deoxygenated blood is small (about $0.02 \times 10^{-6}$ cgs units), and thus image intensity changes in BOLD contrast studies are generally small, less than 15% at a magnetic field of 2 T even when the haemoglobin oxygen saturation is reduced to 20% in acute hypoxia (38). In brain activation studies at 1.5 T (10) the signal change is no more than 2-4%. Experiments (39) and model calculations (16,17,18) have shown that the change in T2* relaxation rate associated with this signal loss increases with the static magnetic field of the MRI scanner, so that the change of signal is typically about three times larger at 4 T than at the more commonly used field strength of 1.5 T, for the same echo time and sequence type. Given the improved signal: noise ratio at higher field (roughly proportional to field strength) there is a strong case for the use of high magnetic fields for this type of functional MRI. At 4T, signal changes of up to 30% have been observed for visual stimulation, with typical single-shot background noise of 0.5% or less, giving z-values of 60 or more for some pixels. Of course, for more subtle brain activations in association cortex, intensity changes are generally much smaller, from 2-8% at this field strength, and proportionately lower at the more commonly used field of 1.5 T. At the lower field strength it is often necessary to perform averaging over images to obtain significant results, although single-subject studies remain eminently feasible.

By contrast with BOLD, AST methods characteristically have a poorer contrast to noise ratio. Here the change in signal depends on the relaxation time T1, which increases only weakly with field strength. At 1.5 T the contrast::noise ratio is about one third of that obtainable with BOLD, and signal averaging is always required. One benefit of the EPISTAR method is that the signal from static tissue in each slice is nulled in the absence of perfusion (33), so that the large background 'anatomical' signal plays no part in functional studies. This means that the large subtraction errors in functional maps resulting from head movement and consequent image misregistration, which bedevil BOLD studies, are not found.

A major disadvantage of AST methods in general, however, is the considerable difficulty in obtaining multislice data. This arises from the fact that the nuclear spins of the water protons carry the label in the blood, and that once these spins have moved into a slice to be imaged, they are disturbed by the rf pulse needed in image acquisition, so that the blood must be re-labelled for each image. Since it takes 1-2 seconds for the label to move into the capillary bed, during which time no other rf pulse can be allowed to affect it, it is not feasible to obtain volume data at the rate of 10 slices per second, as is feasible with BOLD EPI methods.

Comparing areas in BOLD and AST functional images apparently activated by a simple brain stimulus, there is a very large measure of agreement. It can be shown that AST images (when obtained correctly) selectively show areas of increased capillary flow, and changes in arterial and venous flow are not represented. BOLD images show primarily oxygenation changes in small venules overlying neurally active cortical tissue, since these contain the majority of the blood experiencing local changes in oxygenation (24,25). On the scale of the spatial resolution of EPI, the activated areas seen on BOLD and AST images overlap, with residual differences ascribable to the lower sensitivity of AST, as discussed earlier. Higher resolution BOLD images, at high magnetic field which selectively enhances contributions from smaller cortical vessels, may (as discussed above) show yet more local oxygenation changes such as those associated with differential oxygen demand from adjacent cortical columns. AST is unlikely to be able to provide such fine discrimination, since it appears that the arteriolar control of blood flow lacks sufficient precision to selectively supply such small structures, except in specialized cortex such as rat whisker barrels (40).

### *2.6 Comparisons with PET Scanning*

A number of groups have compared results obtained with fMRI and PET methods, for the same subjects performing the same tasks. When the poorer spatial resolution of PET is taken into account, the spatial localization of the two methods is in good agreement (41) and a parametric study, in which finger pressure was varied, showed that changes in fMRI signal, averaged over a group of six volunteers, are proportional to changes in PET measurements of rCBF (42).

Recent work (43,44) has compared PET with fMRI in more detail. Here a single subject, listening to the same auditory stimulation protocol, was scanned first by fMRI and then by PET methods. In order to ensure the similarity of the auditory stimulus, the noise of the MRI scanner was recorded and mixed into the stimulus presented to the subject during PET scanning (which is of course normally silent). The stimulus consisted of English bisyllabic words, presented at varying rates from one every 20 s up to 90 words per minute. Data analysis was performing using Statistical Parametric Mapping (4). Functional imaging data were realigned, spatially normalised and smoothed, and statistically analyzed. Within the spatial resolution of PET (5 mm), the activated areas (Figure 1) were in agreement between the two imaging modalities. When homologous areas of activation in auditory cortex had been identified in both datasets, the signal changes in the voxels of peak activation in these regions were identified and plotted as a function of stimulus presentation rate (Figure 2)

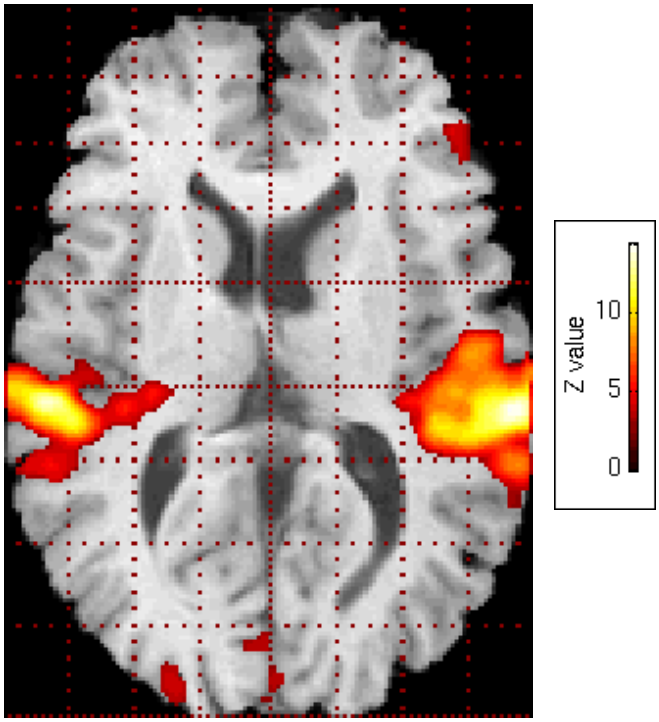


Figure 1. Areas of brain in single subject activated by word-listening task (BOLD fMRI data)

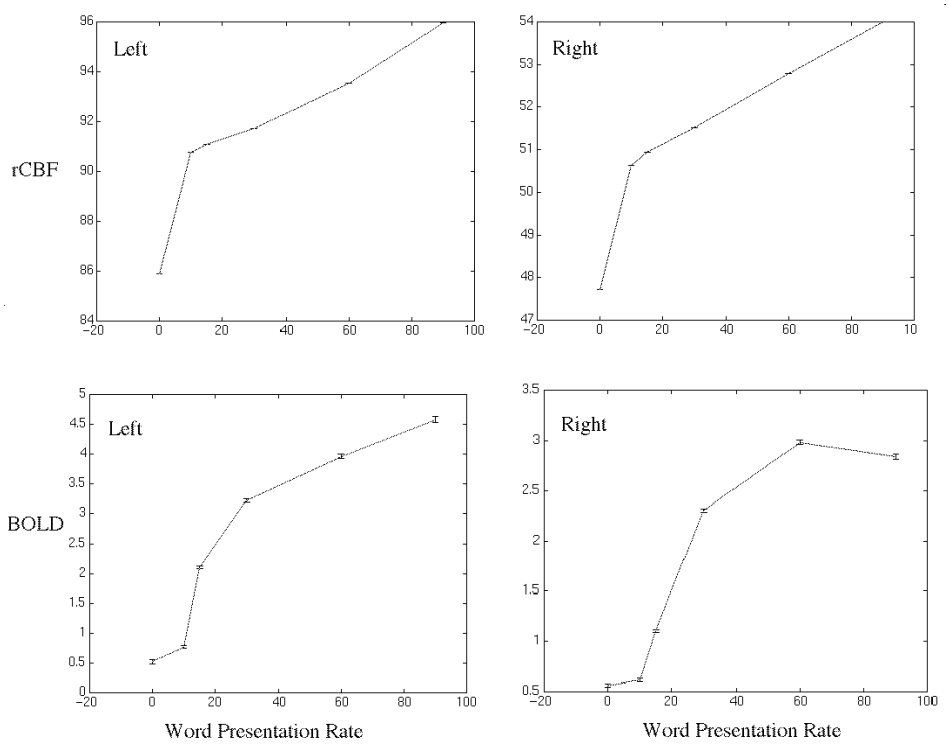


Figure 2. Comparison of PET (rCBF) and fMRI temporal cortical signal increase with word presentation rate

The PET rCBF values show an approximately linear relationship between word presentation rate and rCBF, except for the 'rest' condition where only echoplanar noise was presented. However, the relationship between BOLD contrast signal change and word presentation rate is clearly non-linear, showing a saturation effect with increasing word presentation rate consistent with previous data (45). The PET data show that this departure from linearity is not due to nonlinearity in rCBF changes.

BOLD contrast, as we have seen, is due to a mismatch between flow and metabolism, not only spatially but also temporally. The rate dependence of these data can be explained in terms of a hemodynamic 'refractoriness', in which a prior stimulus modulates the response to a subsequent stimulus, if it occurs within a second or so. This modulation represents an interaction, over time, between the responses to successive stimuli and results in reduced responsiveness at high stimulus frequencies.

This nonlinearity of the BOLD response with respect to blood flow can be modelled elegantly using the so-called 'Volterra integral' formulation (44). This decomposes the response into a linear component and a nonlinear term which quantifies the refractory behaviour of the activation following a given stimulus. For the auditory stimulus under consideration, the BOLD respose to the presentation of a single word is shown in Figure 3.

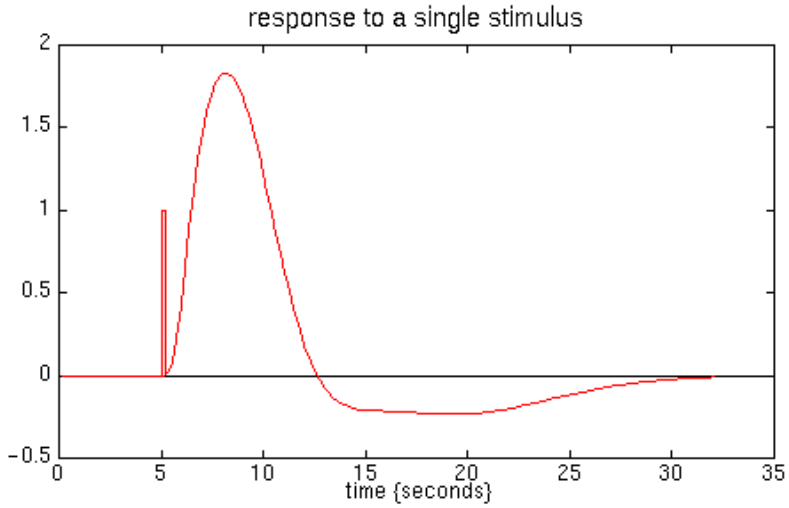


Figure 3. BOLD response in auditory cortex to presentation of a single word. Note that full recovery to baseline does not take place until about 30 seconds after stimulus onset.

The Volterra formulation allows one to predict the effect of applying multiple stimuli closely spaced in time. Consider the response to a pair of stimuli, separated by a second, in relation to the responses to each presented in isolation. Figure 4 shows these responses to the stimuli (upper panel: together - solid line and separately - broken lines). To assess the impact of the first stimulus, on the response to the second, we can subtract the response to the first stimulus from the response to both. This gives the response to the second stimulus in the context of the first (solid line in the lower panel of Figure 4). It can be seen that this response is attenuated markedly, with an augmented undershoot, in relation to the response obtained when the stimulus is presented in isolation (broken line). In short, the response to a stimulus is compromised or modulated by preceding stimuli to give a nonlinear 'refractoriness' that depends on the inter-stimulus interval or rate.

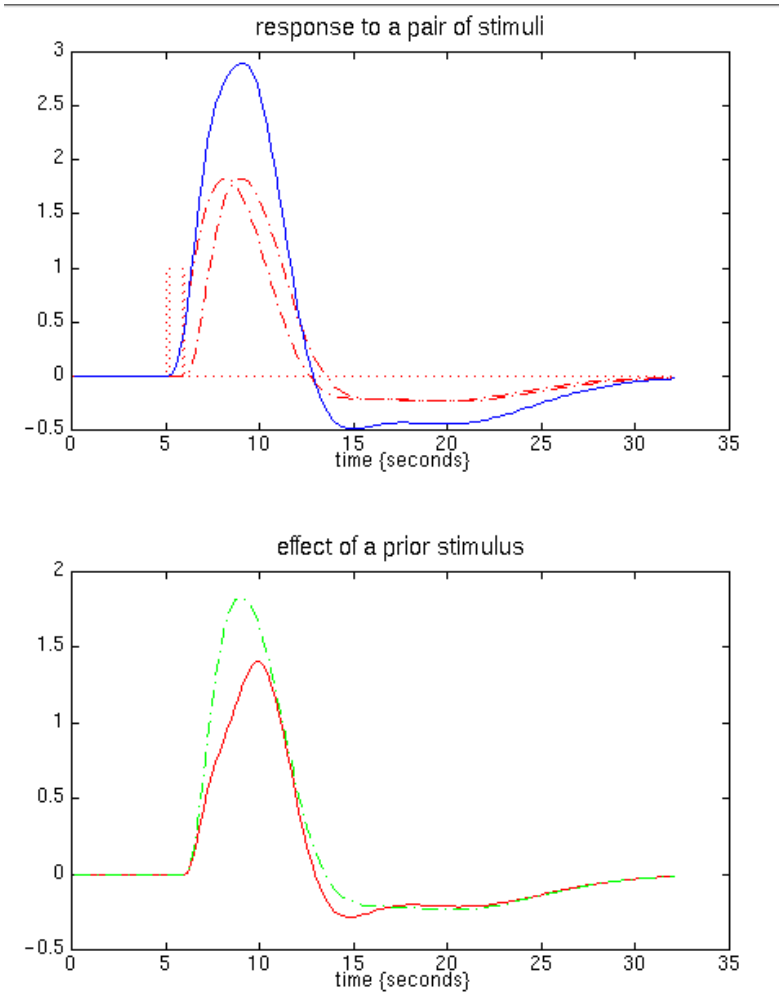


Figure 4. (Top) Predicted BOLD response (solid line) in auditory cortex to a closely-spaced pair of auditory stimuli (broken lines). (Bottom) Effective response to the second auditory stimulus in the context of the first (solid line), compared with response to isolated single auditory stimulus (broken line).

*2.6.1 Neurobiological mechanisms*

We have assumed that underlying neuronal activity is equivalent to word presentation rate. This assumption allows one to assign all the nonlinearities observed in the response to the mapping between neuronal activity and hemodynamic response. It is of course possible that much of the implicit refractoriness can be explained directly in neuronal terms. There is no way of distinguishing between these explanations using the data obtained. The reason we have assumed that neuronal activity (i.e. neuronal discharge rate) is proportional to word presentation rate is that previous observations (46) show a proportional relationship between blood flow and word presentation rate, and blood flow is generally regarded as an index of pre-synaptic activity. Although neuronal adaptation, facilitation and refractoriness certainly contribute to our results, we can make the following (very heuristic) argument. If we assume that each stimulus, or event, evokes roughly the same degree of spike

activity in a neuronal population (providing that these events are at least several hundred milliseconds apart) then spike activity and blood flow will increase in proportion to stimulus frequency. The demonstration of highly significant nonlinear components in the hemodynamic response based on BOLD effects therefore suggests a nonlinear relationship between flow and oxygen extraction fraction. This nonlinear flow-dependency is fully expected (47) and represents a sufficient explanation for the observed nonlinearities.

### *2.6.2 Nonlinear Modelling vs Linear Modelling*

What are the implications of these findings for experimental design and analysis? It is important to note that there is a fundamental distinction between positing a universal nonlinear hemodynamic response function, that can account for varying responses to stimuli presented at different rates, and a series of rate-dependent, linear hemodynamic response functions.

Which is the most appropriate analysis to use? This question is only posed in parametric experimental designs (48) when some experimental parameter is varied (for example rate of stimulus presentation, response rate, duration of task etc.). The alternatives are to model a nonlinear response in terms of a parametrically varying input, or treat each level of the parameter manipulated as a separate condition. In some circumstances only the nonlinear option may be viable, for example if the parameter is changing continuously and does not conform to a series of discrete levels. In general, however, the results that obtain from the two approaches would be similar and the question reduces to one of implementational expediency and simplicity in describing the results. Although it is possible to implement the nonlinear analysis above with existing tools (e.g SPM96) the construction of the design matrices is complicated and a simple linear analysis may be quite sufficient for most purposes.

In terms of experimental design the nonlinear effects above mean that it is possible to drive the brain 'too hard' with very high stimulus presentation rates. The analyses of the present study suggest that the optimum presentation rate, for words, is about one per second. Further discussion of fMRI experimental design will be found in the next chapter.

## *2.7 Typical fMRI Experimental Parameters*

In a typical FMRI experiment, a subject lies on the bed of an MRI scanner while viewing a screen illuminated by an LCD video projector, or listens to auditory input via headphones, or performs some other form of cognitive task, while a sequence of images is obtained. When EPI is used, commonly 3-10 images per second are acquired, for a period of 5-10 minutes. The subject's head is kept still using foam pads, or one of a number of proprietary methods for head immobilization. Ideally, multislice images are obtained, giving upwards of 50 consecutive images of the entire brain during the experimental run. In-plane resolution is typically 2-3 mm, with a field of view of 16-24 cm, and slice thickness is between 3 and 10 mm. With a repeat time of 3 s for any particular slice, effects on the image caused by cardiac and respiratory pulsations are relatively small compared

with functionally-related changes, and are mostly confined to large vessels and CSF (49,50). However, head movement of even as little as 0.5 mm can cause an apparent change of signal in a given voxel of as much as 40%.

## 3) APPLICATIONS OF FMRI

Since fMRI methods are entirely noninvasive, multiple studies on single subjects may be performed, and large representative groups of volunteers may be recruited so that population studies can be performed. Furthermore, by comparison with other nonsurgical functional brain mapping techniques, such as PET, EEG, magnetic source imaging, and near-infrared spectroscopic imaging, fMRI has excellent spatial resolution and quite good temporal resolution. These features have lent themselves to a variety of novel cognitive neuroscience experiments with human subjects, and suggest others which may have considerable clinical benefit in cases of brain pathology. For a detailed review, see (51).

### *3.1 Longitudinal Studies*

Fundamental questions exist regarding neural plasticity, latency, and memory storage mechanisms, that make it of the greatest interest to follow changes in cortical organization associated with learning and memory, in child development and in recovery from brain damage. While all of these areas can be studied with fMRI, so far only one study has been published (52), which describes the process of acquiring a simple motor skill over a period of several weeks of training. Functional MRI maps were obtained while the subject was performing a learned sequence of finger-to-thumb oppositions paced at two taps per second, and compared with maps obtained during performance of an unpractised control sequence containing the same component movements. Each of the six subjects was randomly assigned one of the two sequences to be practiced for ten to twenty minutes each day. The other sequence served as a control, to be performed only during scanning.

Subjects were scanned once a week using EPI for 4-6 consecutive weeks in a 4 Tesla MRI system. Each study consisted of 4 experimental runs of 64 images per slice, each experimental run lasting 4.5 minutes. The general paradigm for each run was rest-X-rest-Y-rest, where X and Y were the trained or untrained task. At the start of the study, no statistically significant difference were seen between the size of regions in M1 above threshold for the trained or untrained sequence. By week 4, however, the MRI data showed a highly significant increase in the area activated in M1 for the practiced sequence versus the unpracticed sequence, perfectly matching observed improvement in performance.

Given the reasonable assumption that the spatial extent of significant changes in fMRI signal is an indication of the area of neuronal activity, the implications are intriguing. Apparently, in the process of learning a motor task a particular network of neurons in M1 (among other regions) becomes specialized specifically for that task. The spatial extent of the network, presumably in this case remaining within the borders of the hand representation,

correlates with the skill in performance of the task. This study can be regarded as a demonstration of neuronal plasticity in humans, comparable to Merzenich's single-electrode work in monkeys.

Future studies will examine changes in cortical activity associated with recovery of function following stroke or brain trauma, cortical reorganization following limb amputation or destruction of peripheral nerves, 'ghost limb' syndrome, and selective loss of cortical activity in Alzheimer's patients or simply in old age.

*3.2 Population Studies*

Comparisons of patterns of cortical activation for a set of common tasks between well-characterized groups are still few in number for fMRI, but it already clear that such studies can easily be done. Among the more successful are studies of language lateralization among epileptic patients and normal subjects (53) and comparison of areas involved in reading English and American Sign Language sentences in prelingually deaf and normally hearing subjects (54). None of these studies used particularly powerful methods of statistical analysis, and the results would certainly be strengthened if SPM techniques had been employed. Future possibilities include cross-cultural studies of cognition (starting with the cortical areas involved with reading differing types of script), correlations of speed in non-verbal reasoning with area and location of cortical activity, study of acquisition of second languages, bilingualism, and so forth.

*3.3 Event-Related fMRI*

Recent extensions to SPM and fMRI acquisition schemes (54,55) have enabled the development of event-related fMRI (see next chapter for more detail), where the local cerebral haemodynamic response to single events can be studied. An example is shown in Figure 3. Such events can comprise a planned protocol, where odd-ball or go/no-go paradigms are now feasible, and the differential time courses relating to the different events can be observed in relevant cortical areas. In addition, the noninvasive nature of fMRI makes it ideal for study of unpredictable events, in which the subject has only to signal that such an event has occurred. Typical subjects might include sufferers from Tourette's syndrome, in whom the cortical areas responsible for causing or inhibiting tics could be studied, schizophrenics, for whom the cortical localization of auditory or visual hallucinations could be elucidated, tremor patients, epileptics, and normal subjects dreaming in the MRI magnet. Such studies may require much machine time, but the rewards are likely to be considerable. The statistical tools of event-related fMRI will be equally applicable in these studies.

## REFERENCES

1) Haase A, Frahm J, Matthaei D, Hänicke W, Merboldt K-D. J. Magn. Reson. 67, 258 (1986)
2) Stewart GN. J. Physiol. 15, 1 (1894)
3) Villringer A et al, Proc. Society of Magnetic Resonance in Medicine, p 21 (1986)
4) Belliveau JW et al, Proc. Society of Magnetic Resonance in Medicine, p 222 (1988)
5) Mansfield P J. Phys. C 10: L55-L58 (1977)

6) Belliveau, J.W.et al. Science 254: 716-719 (1991)

7) Ogawa S, Lee T-M, Nayak AS, Glynn P. Magn. Reson. Med. 14, 68 (1990)

8) Turner R, Le Bihan D, Moonen CTW, Despres D, Frank J. Magn. Reson. Med. 22, 159 (1991)

9) Pauling L and Coryell CD. Proc Natl Acad Sci USA 22: 210-216 (1936)

10) Ogawa S, Tank DW, Menon R, Ellerman JM, Kim S-G, Merkle H and Ugurbil K. Proc. Natl. Acad. Sci. USA, 89:5951-5955 (1992)

11) Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R, Cheng H-M, Brady TJ, Rosen BR . Proc Natl Acad Sci USA 89: 5675-5679 (1992)

12) Fox PT and Raichle ME. Proc. Natl. Acad. Sci. USA 83: 1140-4 (1986)

13) Penfield W . Ann Intern Med 7: 303-310 (1933)

14) Turner R, Jezzard P, Wen H, Kwong KK, Le Bihan D, Zeffiro T, Balaban RS. . Magn. Reson. Med. 29: 277-281 (1993)

15) Fisel CR, Ackerman JL, Buxton RB, Garrido L, Belliveau JW, Rosen BR, Brady TJ. Magn. Reson. Med. 17: 336-347 (1991)

16) Ogawa S, Menon RS, Tank DW, Kim SG, Merkle H, Ellerman JM and Ugurbil K. Biophys. J. 64: 803-812 (1993)

17) Weisskoff RM, Zuo CS, Boxerman JL and Rosen BR. Magn. Reson. Med., 31: 601-610 (1994).

18) Kennan RP, Zhong J and Gore JC, Magn. Reson. Med., 31:9-21 (1994).

19) Boxerman JL, Bandettini PA, Kwong KK, Baker JR, Davis TL, Rosen BR, Weisskoff RM.. Magn Reson Med, 34, 4-10 (1995)

20) Menon RS, Hu X, Adriany G, Andersen P, Ogawa S and Ugurbil K. Proc. Soc. Magn. Reson. (1994) 2: 622 (1994)

21) Hossmann K-A, Ueki M, Kocher M and Linn F. (1991) Multiparametric imaging of coupling between functional activity, blood flow and metabolism uder physiological and pathophysiologic conditions, in Brain Work and Mental Activation, Alfred Benzon Symposium 31, (Lassen NA, Ingvar DH, Raichle ME and Friberg L, eds) pp 158-173. Copenhagen: Munksgaard

22) Duling B, Matsuki T, Segal S. Conduction in the resistance-vessel wall: contributions to vasomotor tone and vascular communication. In: Bevan JA ed. The Resistance Vasculature. Totowa, NJ: Humana Press, 1991: 193-215

23) Frostig RD, Lieke EE, Ts'o DY, Grinvald A. Proc Natl Acad. Sci USA 87: 6082-6086 (1990).

24) Turner R and Grinvald A. (1994). Proc. Soc. Magn. Reson. (1994) 1: 430

25) Malonek D and Grinvald A. Science 272:551-4 (1995)

26) Menon RS, Ogawa S, Strupp JP, Ugurbil K. Proc Soc Magn. Reson, 3rd Meeting, p 163 (1995)

27) Lai S, Hopkins AL, Haacke EM, Li D, Wasserman BA, Buckley P, Friedman L, Meltzer H, Hedera P, Friedland R. Magn. Reson. Med. 30: 387-392 (1993).

28) Segebarth C, Belle V, Delon C, Massarelli R, Decety J, Le Bas JF, Decorps M, Benabid AL. NeuroReport 5, 813-816 (1994).

29) Henkelman RM, Neil JJ, Xiang Q-S. Magn. Reson. Med. 32: 464-469 (1994).

30) Wiederman MP Circ. Res. 12: 375 (1963).

31) Frahm J, Merboldt K-D, Hänicke W, Kleinschmidt A, Boecker H. NMR in Biomedicine 7: 45-53 (1994).

32) Baker JR, Hoppel BE, Stern CE, Kwong KK, Weisskoff RM, Rosen BR. Proc. Soc. Magn. Reson. Med.(1993) 3: 1400

33) Williams DS, Detre JA, Leigh JS, Koretsky AP. Proc. Natl. Acad. Sci. USA 89: 212-216 (1992)

34) Edelman RR et al. Radiology 192: 513-520 (1994)

35) Stehling MK, Turner R, Mansfield P. Science 254: 43-50 (1991)

36) Feinberg DA and Oshio K. Radiology 181: 597 (1991)

37) Ahn CB, Kim JH and Cho ZH. IEEE Trans. Med. Imag. MI-5:2 (1986)

38) Jezzard P et al NMR in Biomedicine, 7:35-44 (1994)

39) Turner R, Jezzard P, Wen H, Kwong KK, Le Bihan D, Zeffiro T, Balaban RS. Magn Reson Med **29**, 281-283 (1993)

40) Cox, SB, Woolsey TA, Rovainen CM. J. Cereb. Blood Flow Metab. 13, 899 (1993)

41) Paulesu E et al, Neuroimaging Clinics of N. America, 5, 207 (1995)

42) Dettmers C, Connelly A, Stephan KM, Turner R, Friston KF, Frackowiak RSJ, Gadian DG. NeuroImage ; 4: 194-200 (1996)

43) Rees G et al (1997) submitted

44) Friston KJ, Josephs O, Rees G and Turner R, Nonlinear event-related responses in fMRI (1997) submitted.

45) Binder J et al Cog Brain Res 2, 31-38 (1994)

46) Price C et al Neurosci Lett 145, 179-92 (1992)

47) Buxton RB and Frank LR. J. Cereb. Blood. Flow. Metab. - in press (1997)

48) Friston KJ Holmes AP Worsley KJ Poline J-B Frith CD and Frackowiak RSJ. Human Brain Mapping 2:189-210 (1995)

49) Noll, DC, Schneider W, Cohen JD. Proc. Soc. Magn. Reson.. Med. 3: 1407 (1993)

50) Le TH, Hu X. Magn. Reson. Med. 35: 290 (1996)

51) Turner R. Sem. Neurosci. 7: 179-194 (1995)

52) Karni A, Meyer G, Jezzard P, Adams MM, Turner R and Ungerleider LG. Nature 377: 155-158 (1995)

53) Hertz-Pannier L et al. Proc. Soc. Magn. Reson. (1994) 1: 326

54) Bavelier D, Corina D, Jezzard P, Clark V, Karni A, Padmanhaban S, Rauschecker J, Turner R, Neville H. Brain and Cognition 32: 165-167 (1996)