Neural Correlates of Thinking

Bearbeitet von
Eduard Kraft, Balázs Gulyás, Ernst Pöppel

ISBN 978 3 540 68042 0
Format (B x L): 15,5 x 23,5 cm
Gewicht: 619 g

Weitere Fachgebiete > Literatur, Sprache > Angewandte Sprachwissenschaft > Psycholinguistik, Neurolinguistik, Kognition

Zu Inhaltsverzeichnis

schnell und portofrei erhältlich bei

beck-shop.de

Die Online-Fachbuchhandlung beck-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.
Functional MRI Limitations and Aspirations

Peter A. Bandettini

Abstract Most would agree that knowing precisely what was happening in the brain during the act of thinking would help in our pursuit to understand what thinking really is. This chapter describes the basics, limits, and future directions of one of the more effective tools we have to observe the human brain while it is functioning – functional MRI. Functional MRI emerged in the early 1990s, and has since grown explosively in utility. In this chapter, an in-depth exploration is carried out of what limits functional MRI to a spatial resolution of millimeters and a temporal resolution of seconds. In addition, issues of how sensitive functional MRI is in detecting brain activity and how deeply we can interpret the signal changes are explored. Lastly, the chapter ends with a discussion on how imaging might be essential, or perhaps irrelevant, to understanding thinking.

1 Introduction

Before 1991, the thought that one could use magnetic resonance imaging (MRI) to map human brain activation noninvasively, rapidly, with full brain coverage, and with relatively high spatial and temporal resolution was pure fantasy. This fantasy became reality with a rapidity and decisiveness that surprised almost everyone, causing the neuroscience community to rapidly readjust itself as it embraced this new modality. Many researchers tailored many of their ongoing behavioral, electrophysiological, or other imaging modality studies to the MRI scanner environment. Since then, functional MRI (fMRI) has proven to be a powerful and robust technique.

Some argue that it is so easy to obtain eye-catching maps of “brain activation” that the quality of science performed with fMRI can be lacking at times.

P.A. Bandettini
Laboratory of Brain and Cognition, National Institute of Mental Health, Section on Functional Imaging Methods Building 10, Room 1D80B, 10 Center Dr. MSC 1148, Bethesda, MD 20892-1148, USA
bandettini@nih.gov
However, truly innovative experiments and unique, insightful findings have indeed been obtained using fMRI. The reality is that fMRI is an improving, powerful, but sometimes misused and overinterpreted technique that is the tool of choice for a growing community of researchers, clinicians, and thinkers. In this chapter, the limits and aspirations of fMRI are explored.

First, the basics of fMRI are introduced: the history, the physiologic processes behind the signal, and the current state of the field. This section is concluded with a list of some of the advances that have helped to define fMRI. Section 2 focuses on the limits of temporal and spatial resolution, sensitivity, and signal interpretation. The last section is an attempt to put fMRI in the context of understanding thinking – an objective of the Parmenides Foundation. In this last section, the question of “What really does fMRI contribute to our understanding of thinking?” is addressed.

2 Basics

2.1 History and Functional Contrast

The use of MRI to map brain activation in humans was introduced with a groundbreaking paper by Belliveau et al. (1991) in November of 1991 which described a technique involving two sequential bolus injections of the susceptibility contrast agent gadolinium-DTPA, to map blood volume during rest and activation. About the time that work was published, this approach was rendered obsolete (as far as functional activation imaging is concerned) by a completely noninvasive MRI-based technique utilizing endogenous functional contrast associated with localized changes in blood oxygenation during activation. Between the early spring and late fall of 1991, the first successful experiments in endogenous functional contrast fMRI were carried out. The findings of these first experiments were published within 2 weeks of each other in the early summer of 1992 (Bandettini et al. 1992; Kwong et al. 1992; Ogawa et al. 1992).

The mechanism of endogenous contrast by which these early results were based was pioneered in animal and phantom studies by Ogawa et al. (Ogawa and Lee 1990; Ogawa et al. 1990a,b), who coined the term “blood oxygenation level dependent” (BOLD), as well as by Turner et al. (1991), who further demonstrated this contrast in cat models.

The basics of the contrast mechanism are as follows. Hemoglobin is more paramagnetic (lower magnetic susceptibility) than tissue when deoxygenated, and has the same susceptibility as tissue when fully oxygenated. When it is deoxygenated and within a magnetic field, microscopic field inhomogeneities exist as a result of the different susceptibilities, leading to an attenuated magnetic resonance signal. During rest, venous blood is slightly deoxygenated. With activation, blood flow locally increases more than what is required by an increase in oxidative metabolic rate, causing an increase in blood oxygenation and a decrease in the amount of deoxygenated hemoglobin, therefore creating a small magnetic resonance signal increase.
Another noninvasive fMRI technique that emerged almost simultaneously with BOLD fMRI is known as arterial spin labeling (ASL) (Williams et al. 1992). The contrast in ASL arises from blood perfusion, independent of blood oxygenation. With ASL, blood is “labeled” with a radiofrequency (RF) pulse. This RF-labeled inflowing blood changes the signal in the plane being imaged as a function of blood perfusion in each voxel. ASL is unique in that maps of baseline and active-state perfusion may be made, whereas with BOLD contrast, only maps of changes can be obtained.

Details of other techniques exist which allow noninvasive assessment of activation-induced changes in blood volume (Lu et al. 2003) and oxidative metabolic rate (Davis et al. 1998; Hoge et al. 1999), temperature (Yablonskiy et al. 2000a,b), and diffusion coefficient (Le Bihan et al. 2006) have since been published. Currently, efforts are being made to develop a successful approach to directly imaging neuronal activity using MRI (Bandettini et al. 2005). So far, no results have been convincingly demonstrated, as models suggest that the sensitivity required is an order of magnitude higher than what is currently available when imaging humans.

BOLD fMRI is the brain activation mapping method of choice for almost all neuroscientists because it is easiest to implement and the functional contrast to noise (defined as the signal change magnitude divided by the background fluctuation magnitude – ranging from 2 to about 6 with BOLD contrast) is generally a factor of 2–4 higher than for the other MRI-based methods. The need for sensitivity and ease of use outweighs, most of the time, the advantages in specificity, quantitation, or baseline information inherent to ASL.

Picking up in 1992, only a handful of laboratories could perform fMRI because it required not only an MRI scanner but also the capability of performing high-speed MRI – known as echo planar imaging (EPI). EPI is a technique by which an entire image (or “plane”) is collected with the use of a single RF pulse and a single subsequent signal “echo” – hence the name “echo planar imaging.” Collecting an entire image in 30 ms (as with EPI) “freezes” physiologic processes that contribute to nonrepeatable artifacts in slower MRI methods, leading to a significantly higher temporal stability – critical for fMRI. Until about 1996, the hardware for performing EPI was not available on clinical systems. Now, practically every standard MRI scanner is equipped to perform EPI.

After about 1996, with rapid proliferation of EPI-capable MRI scanners incorporating whole-body gradients, the standard platform for fMRI reached a plateau that is still mostly in use today. In addition, the number of people able to perform fMRI increased dramatically. Figure 1 shows the increase in the number of articles and reviews published (using the Scopus search engine) dealing with fMRI.

This standard platform pulse sequence typically used is gradient-echo EPI: echo time 40 ms, matrix size $64 \times 64$, field of view 24 cm, slice thickness 4 mm (voxel dimension of about $4 \times 4 \times 4$ mm). For studies incorporating spatial normalization and multisubject spatial averaging, going to any higher resolution gives no gains and a loss in sensitivity. (In fact, perhaps the optimal matrix size to use when performing multisubject averaging is $32 \times 32$ since sensitivity is increased, and the acquisition resolution approximately matches the resolution that spatial smoothing,
Fig. 1 The number of articles or reviews published per year, obtained from the Scopus search engine. The production of papers in functional magnetic resonance imaging (fMRI) shows no signs of slowing down.

normalization, and multisubject averaging reduces functional images to.) Typically, whole brain volume coverage using a repetition time (TR) of 2 s is performed. Time series are collected, lasting on the order of 5–8 min. A typical experiment involves the collection of about seven time series per subject scanning session. Multisubject studies usually settle on assessing about 12 such sessions. Regarding hardware, in about the year 2002, the “standard” field strength increased from 1.5 to 3 T, thus improving sensitivity. Currently, a new standard in RF coils has begun to take hold. In the past, the standard was the use of a quadrature RF coil. Currently, the trend is to use multireceiver coils, ranging from eight channels to 32 channels. The use of these coils translates into either an increase in sensitivity (smaller coils have greater sensitivity at the expense of less coverage – hence the need for more RF coils) or an increase in resolution (and/or speed) that comes with a novel strategy known as “sensitivity encoding.”

Typical paradigm design methods are either “box-car,” involving steady-state activation periods for 10 s or more, or more commonly “even-related” designs enjoying the flexibility inherent to brief activation periods interspersed within the time series. For postprocessing, SPM is the most common processing software program, but platforms such as Brain Voyager, FSL, and AFNI are almost as common (information regarding these platforms is at the end of the chapter). The fundamental concept in all of functional imaging creation is the statistical comparison of what is expected to happen in the hemodynamic response, as defined by a “reference” function or a “regressor,” with the data, on a voxelwise basis.
2.2 Advances

An approximately chronological list of a few of the many significant developments in fMRI is shown below. Neither this list nor the references associated with each topic are comprehensive. The goal is to provide a quick perspective of some of the highlights over the past 15 or so years and to give a sense that fMRI is a method that is very much in the hands of the users, as they drive many of the most innovative advances.

- **Parametric manipulation** of brain activation demonstrated that BOLD contrast approximately followed the level of brain activation: visual system (Kwong et al. 1992), auditory system (Binder et al. 1994), and motor system (Rao et al. 1996).
- **Event-related fMRI** was first demonstrated (Blamire et al. 1992). The application of event-related fMRI to cognitive activation was shown (Buckner et al. 1996; McCarthy et al. 1997). Development of mixed event-related and block designs was put forward: (Donaldson et al. 2003). Paradigms were demonstrated in which the activation timing of multiple brain systems was orthogonal, allowing multiple conditions to be cleanly extracted from a single run (Courtney et al. 1997).
- **High spatial resolution** maps were created: For spatial resolution ocular dominance columns (Cheng et al. 2001; Menon et al. 1997; Yacoub et al. 2006, 2007) and cortical layer activation maps (Logothetis et al. 2002) were created. Figure 2 illustrates graphically the spatial scales of brain organization that are able to be imaged with fMRI. First able to be imaged was the large-scale organization (i.e. V1, V2, etc...), then the ocular dominance column scale, then the layer-specific scale, and then the smallest scale – the orientation column scale.

![FMRI-Accessible Scales of Visual Cortex Organization](http://www.thebrain.mcgill.ca)

Fig. 2. The spatial scales of cortical organization that are accessible to fMRI measures: functional units such as V1 *(left)*, cortical columns such as ocular dominance columns, cortical layers, and orientation columns *(right)*. (Obtained from http://www.thebrain.mcgill.ca)
High temporal resolution fMRI developed: relative onset timings from milliseconds to hundreds of milliseconds were extracted (Bellgowan et al. 2003; Henson et al. 2002; Menon et al. 1998; Ogawa et al. 2000).

The development of “deconvolution” methods allowed for rapid presentation of event-related stimuli (Dale and Buckner 1997).

Early BOLD contrast models were put forward (Buxton and Frank 1997; Ogawa et al. 1993). More sophisticated models were published that more fully integrated the latest data on hemodynamic and metabolic changes (Buxton et al. 2004).

The use of continuous variation of visual stimuli parameters as a function of time was proven a powerful method for fMRI-based retinotopy. (Deyoe et al. 1994; Engel et al. 1994; Sereno et al. 1995).

The development of “clustered volume” acquisition was put forth as a method to avoid scanner acoustic noise artifacts (Amaro et al. 2002; Edmister et al. 1999).

The findings of functionally related resting state correlations (Biswal et al. 1995) and regions that consistently show deactivation (Binder et al. 1999; Raichle et al. 2001) were described. This exploration of resting state connectivity has currently emerged as a major new research area in fMRI (Raichle and Snyder 2007). The very recent, explosive growth of this area of fMRI is illustrated in Fig. 3.

Observation of the pre-undershoot in fMRI (Hennig et al. 1997; Hu et al. 1997; Menon et al. 1995) and correlation with optical imaging (Malonek and Grinvald 1996) was reported. This is still highly controversial as the effect is very subtle.

Fig. 3 The number of articles published per year which discuss resting state fluctuations in fMRI. While this effect was first discovered in 1995, this aspect of fMRI took off dramatically in 2006 and 2007.
and the hypothesized mechanisms producing it (rapid change in blood volume or $\Delta$CMR$O_2$) remain hotly debated.

- Structural equation modeling was developed in the context of fMRI time series analysis (Buchel and Friston 1998).
- Simultaneous use of fMRI and direct electrophysiological recording in nonhuman primate brains during visual stimulation has elucidated the relationship between fMRI and BOLD contrast (Logothetis et al. 2001), suggesting that BOLD contrast is more correlated with synaptic activity (local field potentials) than with spiking activity. Simultaneous electrophysiological recordings in animal models revealed a correlation between negative signal changes and decreased neuronal activity (Shmuel et al. 2002). Simultaneous electrophysiological recordings in animal models provided evidence that inhibitory input could cause an increase in cerebral blood flow (Matheiesen et al. 1998).
- A technique known as vascular space occupancy (VASO) has emerged as a way to noninvasively map blood volume changes (Lu et al. 2003). Recently, this technique has also come under some scrutiny.
- Extraction of information at high spatial frequencies within regions of activation was first demonstrated (Haxby et al. 2001). This approach, which focuses on information extraction rather than mapping (Kriegeskorte et al. 2006), has developed rapidly as many groups are making efforts at “brain reading” rather than brain mapping (Cox and Savoy 2003; Haynes et al. 2005; Haynes and Rees 2005a,b; Kriegeskorte 2007a,b). The field of “fMRI decoding” has also become a major direction of research in fMRI (Haynes et al. 2004; Haynes and Rees 2005a; Kamitani and Tong 2005; Kay et al. 2008).
- Several papers have been published showing the potential for direct neuronal current imaging with MRI (Bandettini et al. 2005; Buracas et al. 2008; Cassara et al. 2008; Kraus et al. 2008; Mandelkow et al. 2007; Matlachov et al. 2007; Park and Lee 2007; Park et al. 2004; Parkes et al. 2007; Singh and Sungkarat 2005; Truong et al. 2008; Xiong et al. 2003). No paper has been published showing convincing neuronal current activation maps in humans.
- While real-time fMRI has been in existence since at least 1995 (Cox et al. 1995), a paper was recently published (DeCharms et al. 2005) demonstrating that real-time feedback of brain activation to subjects experiencing chronic pain not only allowed them to modulate their activation, but that this resulted in their perception of the pain – opening up a potentially rich area of real time fMRI based therapy.
- The uses of parallel imaging and high field strength have been the major developments on the technical side of fMRI (Bellgowan et al. 2006; De Zwart et al. 2004; Moeller et al. 2006; Preibisch et al. 2008). Parallel imaging allows for higher resolution, or a more rapid acquisition, or higher sensitivity. Higher field allows for higher resolution and higher sensitivity. Interestingly, the progression of field strength used for humans has been linear since clinical human MRI began around 1984. Figure 4 shows this progression. Scanners with field strengths of 11.7 T have been proposed for 2011, keeping the trend linear.
3 Limitations and Aspirations

3.1 Temporal Resolution

An echo planar image has an acquisition window that is about 20–30 ms in duration. In general, about 15 echo planar images can be collected in 1 s. For volume collection, typically consisting of 30 slices, a TR of 2 s is therefore required. It is also possible to collect one image (as opposed to multiple images in a volume) at a rate of 15 images per second (TR = 1,000 ms/15 images = 66.7 ms per image). Relative to the limits in temporal resolution imposed by the sluggishness and variability of the hemodynamic response in fMRI, the image acquisition rate is quite fast.

The dynamics, location, and magnitude of the signal are highly influenced by the vasculature as it is sampled in each voxel. If a voxel happens to capture large vessel effects, the magnitude of the signal may be large (up to an order of magnitude larger than capillary effects), the timing a bit more delayed than average (up to 4 s delayed from capillary effects), and the location of the signal somewhat distal (up to 1 cm) from the true region of activation. The problem of variable vasculature and hemodynamic coupling in fMRI remains to some extent at all field strengths and poses significant limits on the depth and range of questions that can be addressed using fMRI.
On average, the fMRI signal begins to increase approximately 2 s after neuronal activity begins, and plateaus in the “on” state after about 7–10 s. A “pre-undershoot” in the signal is sometimes observed at about 0.2–1.0 s and a post-undershoot is much more commonly observed, lasting up to 1 min. These more subtle dynamics are still not fully understood, but are likely due to temporal mismatches among the hemodynamic factors which most influence the signal: flow, blood volume, or CMRO$_2$ (Buxton and Frank 1997; Buxton et al. 2004).

The hemodynamic response can be thought of as behaving like a low-pass filter for neuronal activity (Bandettini 1999; Kim et al. 1997; Richter et al. 1997). At on/off frequencies of 6 s on/6 s off (0.08 Hz), BOLD responses begin to be attenuated relative to longer on/off times. At on/off frequencies of 2 s on/2 s off (0.25 Hz) the BOLD response is almost completely attenuated. Even though BOLD signal is attenuated by these rapid on/off responses, activity of very brief duration can be observed. Activity durations as low as 16 ms have been shown to cause robust BOLD signal changes, suggesting that there is no apparent limit to the briefness of detectible activation (Birn and Bandettini 2005). When repeated experiments are performed, the hemodynamic response in each voxel shows variability of only about 100 ms (Bandettini 1999).

A strong desire of those who use functional brain imaging is to determine the precise timing of activation between different regions of the brain – either relative to the stimulus or input or relative to each other. The temporal resolution required for this type of assessment is on the order of at least tens of milliseconds. With BOLD contrast, the latency of the hemodynamic response has a range of 4 s across voxels owing to spatial variations in underlying hemodynamics or neurovascular coupling dynamics from voxel to voxel even within the same region of activity (Bandettini 1999). If a voxel contains mostly larger venous vessels, the response is typically more delayed than if the voxel captures predominantly capillaries. This observation is only approximate. The precise reasons for latency variations have still not been completely determined.

Methods have been proposed to alleviate this problem of spatial heterogeneity of the latency of the hemodynamic response. The most direct is to try to identify and remove larger vessel effects by thresholding based on the percentage signal change or temporal fluctuation characteristics. The accuracy of these methods remains undetermined and is likely to be low since high percentage signal changes may occur quite proximal to an active area (and therefore should not be eliminated), and draining veins may have low fluctuations (and therefore be missed, while they should be eliminated).

Another solution is to use pulse sequences sensitive only to capillary effects. ASL techniques are more sensitive to capillaries, but the practical limitations of lower functional contrast to noise and longer interimage waiting time (owing to the additionally required time to excite the inflowing blood and to wait for it to arrive in the plane – about 1.5 s) make this unworkable for most studies.

An alternative strategy to push temporal resolution is to focus on localized changes in latency and width of the hemodynamic response with task timing changes. As mentioned, within a voxel, the hemodynamic response, while exhibiting a delay of 4 s from the mean, still only shows a variation (with repeated, identical
trials) on the order of 100 ms, allowing significantly more accurate assessment if activation timing were to be varied within a region (Bellgowan et al. 2003; Henson et al. 2002).

One other unique method for probing very rapid neuronal interactions was pioneered by Ogawa et al. (2000). In this approach, paired pulses of stimuli either activate the same region or activate two different regions that are connected by inhibitory or excitatory synaptic input. The time between the stimuli pairs is modulated and the amplitude modulation of the second response is observed. This method has demonstrated a 50-ms optimal inhibitory timing between left and right forepaw in rat – as indicated by maximal reduction of the second BOLD response at that timing, and also has shown a 100-ms optimal inhibition in human visual cortex in the same manner. The precise neuronal mechanisms behind these findings are still not fully understood, but the method itself is potentially quite useful for probing the timing and connectivity between either excitatory or inhibitory processing nodes in the brain using fMRI.

A summary of current temporal limits of fMRI as well as speculations on improvements are given below.

Temporal limits:
- Able to detect transient activity as short as 16 ms.
- Able to create a functional image within 20 s for more robust activation (more an issue of sensitivity, but can be considered practical temporal resolution).
- Able to detect differences in brain activation timing across regions of no less than 4–6 s because across voxels the temporal lag varies by 4 s.
- Able to detect modulations in brain activation timing within the same voxel or region of the brain that are no smaller than 100 ms.
- Able to detect inhibitory interaction, with no temporal limit, between connected and interacting nodes of activity in the brain, provided that these interactions are able to cause detectible modulations in BOLD signal in the area being affected.

Temporal aspirations:
- Calibration methods (identification and removal of draining vein effects) are being developed to reduce physiologic fluctuations such that the temporal resolution would be most influenced by temporal signal to noise, perhaps pushing all temporal resolutions to about 50 ms.
- Neuronal current imaging aspires to bypass hemodynamics altogether, aiming to detect neuronal transients on the order of 10 ms. This will take a breakthrough to be accomplished, since, so far, no robust results at all have been observed using neuronal current imaging in humans.

3.2 Spatial Resolution

Single-shot (one excitation RF pulse per imaging plane) imaging is easily the most robust, stable, and common imaging procedure in fMRI. The primary drawback
is that the spatial resolution and overall image quality of echo planar images are significantly less than those of clinical anatomical scans. The upper in-plane resolution of standard single-shot EPI is about 2 mm$^2$. One of the most promising developments in fMRI scanner technology has been the use of multiple parallel RF receive coils to help spatially encode the data, thus allowing for much higher resolution with a single excitation pulse. This approach can allow functional image resolutions of about 1 mm$^3$.

Nevertheless, because the voxel volume directly determines functional signal to noise, the signal to noise of these high-resolution images is considerably lower than the signal to noise of lower-resolution images, requiring functional imaging to be performed at 3 T or preferably higher (because of greater image signal to noise, and larger functional contrast at higher field strengths) to produce useful data in a workable amount of time (Murphy et al. 2007).

A fundamentally important caveat to improving spatial resolution that is worth mentioning at this point is that most fMRI studies involve spatial smoothing, spatial normalization, and multisubject averaging – effectively reducing the spatial resolution to, at best, 10 mm$^3$ and completely nullifying any advantages of collecting data at high resolution, and at high field for that matter. The main purpose of high-resolution studies is single-subject assessment that is not subsequently spatially smoothed, transformed into a normalized space, and averaged with 20+ other brains. Individual subject assessment is a growing area of fMRI as methods are being developed to extract ever more subtle and useful information on a subjectwise basis (Kriegeskorte and Bandettini 2007a,b). There is no compelling reason to perform fMRI at resolution higher than 4 mm$^3$ or lower if multisubject averaging is part of the processing path.

As with temporal resolution limits, the spatial resolution limits are predominantly determined not by limits in the acquisition method but by the relatively wide spatial spread of oxygenation and perfusion changes that accompany focal brain activation. This “hemodynamic point spread function” has been empirically determined to be on the order of 3.5 mm (Engel et al. 1997). At 7 T, more sensitive to microvessels, Shmuel et al. (2007) found the point spread function to be about 2.3 mm. Interestingly, they also found that the point spread function was narrower for BOLD signal change obtained during the third (1.52 mm) and fourth (1.99 mm) seconds of stimulation.

The approaches for dealing with the hemodynamic smoothing function are similarly applicable as those for temporal resolution limits. The primary effort has been to eliminate large vessel effects while preserving sufficient functional contrast for creating functional images in a reasonable amount of time. Approaches to eliminate large vessel effects have generally included:

1. The use of spin-echo imaging. A spin-echo image (as opposed to the more typically used gradient-echo image) is more sensitive to small susceptibility variations but is still sensitive to red blood cells in large vessels (Jochimsen et al. 2004) – except at high fields where intravascular signal is almost completely gone because of T2 shortening of deoxygenated blood.
The use of high field. At high field, there is slightly more sensitivity to small vessels and less to large vessel intravascular signal since both $T2^*$ and $T2$ of blood decrease.

The use of diffusion weighting (otherwise known as “velocity nulling” in this context). Diffusion weighting removes rapidly flowing large vessel intravascular signal but not extravascular effects.

The use of ASL. Imaging capillary perfusion completely bypasses the large vessel problem. This method has a contrast to noise that is about a factor of 2–4 lower than $T2^*$ BOLD signal except for long duration activation paradigms where the BOLD baseline tends to drift, whereas the ASL baseline is steady.

The combination of approaches 1 and 3 could work in theory but, in practice, there is no functional signal left owing to signal-to-noise limitations. The combination of approaches 1 and 2 has been used successfully for mapping ocular dominance columns.

Another approach to increasing functional spatial resolution is calibration of the hemodynamic factors which influence BOLD signal change. Spatial calibration methods have been proposed involving hypercapnia (Bandettini and Wong 1997; Cohen et al. 2004; Thomason et al. 2007), CO$_2$ stress (Chiarelli et al. 2007; Handwerker et al. 2007), and, recently, even the resting state fluctuation data (Birn et al. 2008). The general idea in calibration is to create a map of the “potential magnitude of BOLD” by giving a global hemodynamic stress that is evenly distributed throughout the brain (this maps resembles closely a map of gray and white matter combined with a venous angiogram), then to divide activation-induced signal changes, on a voxelwise basis, by this map of signal change to a global stress.

In spite of the limitations in spatial resolution, ocular dominance column (~1 mm$^3$) (Cheng et al. 2001; Goodyear and Menon 2001), cortical layer (less than 0.5 mm$^3$) (Logothetis et al. 2002), and orientation column (Yacoub et al. 2006) delineation have been achieved. An ongoing issue with regard to the upper resolution of fMRI is whether or not fine delineation necessarily translates to accurate delineation (Kriegeskorte and Bandettini 2007a,b) – meaning that detailed activation maps may not be precisely registered with underlying function. This remains to be demonstrated using a gold standard, but compelling data suggesting fine and accurate delineation have been presented in an animal model in which optical imaging data were compared with hemodynamic changes as measured by fMRI (Fukuda et al. 2006; Moon et al. 2007a,b).

Lastly, a method involving neuronal adaptation paradigms may be able to selectively image neuronal populations on a subvoxel scale. This approach has been termed “fMR-adaptation” (Grill-Spector and Malach 2001), and relies on the relatively rapid adaptation and recovery properties of specific neuronal pools, and the reflection of these properties in fMRI signal, to characterize and differentiate sub-voxel populations of neurons that are sensitive to subtle differences in stimulus or general paradigm properties. Grill-Spector and Malach described this method as a paradigm that proceeds in two stages: first, a neuronal population is adapted by repeated presentation of a single stimulus; second, a property of the stimulus is
varied and the recovery from adaptation (manifest as an increase in fMRI signal) is assessed. If the signal remains adapted, it indicates that the neurons are invariant to that attribute. However, if the fMRI signal recovers from the adapted state, it implies that the neurons are sensitive to the property that was varied.

A summary of current spatial limits of fMRI as well as speculation on improvements is given below.

Spatial limits:
- At 3 T, only enough sensitivity to practically achieve 1.5-mm$^3$ resolution. The functional point spread function is about 3.5 mm.
- At 7 T, enough sensitivity to practically achieve 0.5-mm$^3$ resolution. The functional point spread function can be has high as 1.5 mm.
- At 7 T, using spin-echo sequences, the smallest resolved functional unit was orientation columns (on the order of 0.5-mm width).
- With fMR-adaptation paradigms, the highest resolution is unknown, as very small pools of neurons within voxels may be selectively modulated.
- With calibration methods using a global hemodynamic stress, it is speculated that the functional point spread function can be reduced to 1.5 mm at all field strengths.

Spatial aspirations:
- It’s not clear that there is a need to go to a resolution higher than 0.5 mm. With methods for improving sensitivity (RF coil addition being most promising), and selectivity to capillaries, and neuronal subpopulations, the functional resolution may be readily approaching 0.5 mm. It is not clear how abundant functional units smaller than 0.5 mm are.
- Averaging or pooling of multi-subject high-resolution data remains a challenge. I believe that normalization algorithms have room for improvement – as more information about the principles of brain variability may be incorporated. It’s hard to speculate on how high a resolution multisubject averaging may achieve. It may reach 2 mm as algorithms become more focused on particular brain structures.

### 3.3 Sensitivity

Currently, the functional contrast to noise ratio in fMRI is about 4:1 at typical resolutions at 3 T. This means that the functional signal change is approximately 4 times larger than the underlying noise levels. Increases in sensitivity can directly translate into being able to extract more subtle functional information either in space or in time, and as shown in Fig. 5 can translate into creating a usable functional image in significantly less time – extremely important in a clinical setting. In general, it is perhaps the most desired commodity in fMRI, as most researchers are willing to sacrifice temporal resolution, spatial resolution, and higher specificity in order to maximize sensitivity.
Sensitivity to brain activation related hemodynamic changes can be increased in the following ways:

1. Increasing image signal-to-noise ratio – either by reducing RF coil size (and adding more for whole brain coverage), or increasing field strength (which improves both signal change magnitude and signal-to-noise ratio).
2. Increasing the activation-induced signal change – by increasing field strength, or, in animal studies, using an intravascular paramagnetic contrast agent (Smirnakis et al. 2007).
3. Better modeling and accounting for variations in activation-induced signal changes, as much sensitivity is lost when the temporal models do not match the signal change dynamics exactly. The hemodynamic response varies considerably in shape, latency, and with across regions, individuals, and voxels.
4. Better modeling of and accounting for the noise – mostly the physiologic noise (which includes respiration, cardiac, respiration changes, etc.). This noise is additionally problematic since it sets an upper bound on the temporal signal-to-noise ratio (Bodurka et al. 2007; Triantafyllou et al. 2005). This upper bound is about 100:1 (1% fluctuations).

Sensitivity limits:
- Even though image signal-to-noise ratio can be as high as 800:1, we are currently limited to a temporal signal-to-noise ratio of about 100:1 across all field strengths (Bodurka et al. 2007). This limit is most strongly influenced by physiologic...
The relationship between temporal SNR (y-axis) and image SNR (x-axis). If no physiologic fluctuations exist, the relationship is a straight line – they would equal each other as shown in the phantom data. In reality, the temporal SNR in gray matter plateaus around 100:1, as image SNR continues to increase. The vertical lines indicate what is the “optimal” resolutions to scan at – where temporal and spatial SNRs are highest and closest to each other. If one collects a time series of images in which the temporal SNR is greater than 100, that person is throwing away signal that could be translated into perhaps higher resolution. WM white matter, GM gray matter, CSF cerebrospinal fluid. (Adapted from Bodurka et al. 2007.)

fluctuations that occur over time. A graphic illustration of this effect is shown in Fig. 6, in which the temporal signal-to-noise ratio in gray matter is shown to plateau, even as image signal-to-noise ratio continues to increase.

- The functional contrast-to-noise ratio is about 4:1 at 3 T and up to 5:1 at 7 T.

Sensitivity aspirations:

- The goal is to achieve a temporal signal-to-noise ratio that matches the image signal-to-noise ratio above 100:1. While many physiologic fluctuations can be accounted for, to account for all nonneuronal fluctuations would require a much higher sampling rate and better spatial and temporal modeling of the noise. A temporal signal-to-noise ratio of 200:1 and a functional contrast-to-noise ratio of 10:1 is likely to be achievable relatively soon.
- Processing methods that take into account “patterns” of activation rather than individual voxels as independent measures show considerable promise not only for fMRI decoding efforts but also for increasing sensitivity (Kriegeskorte and Bandettini 2007b; Kriegeskorte et al. 2006). These methods are still in their infancy, and their potential is yet to be fully realized. Fundamentally, this will be a paradigm shift in fMRI since, instead of looking for 1-cm “blobs” of activation, we will start looking for unique patterns of activation, across several spatial scales – down to individual voxelwise patterns.
3.4 Interpretation

A fundamental goal in fMRI is to be able to infer precisely where, when, and how much neuronal activity is taking place in the brain on the basis of the measured BOLD signal. This goal is problematic since BOLD changes depend on variables other than neuronal activity itself, including hemodynamic coupling and volume in each voxel. The hemodynamics vary from voxel to voxel, so even if studies demonstrate that within a region there is a relationship between neuronal activity and BOLD signal, this does not get any closer, in practice, to being able to say that “neuronal activity is x in this voxel.” To do this, spatial calibration (voxelwise calibration) of the hemodynamic response is necessary.

Progress has been made in at least confirming the BOLD signal change is a reliable and a high enough fidelity measure of neuronal activation to be widely used and depended on. Strategies for characterizing the relationship between neuronal activity and BOLD signal changes have included (1) animal models and the simultaneous use of other measures of neuronal activity such as multunit electrodes or more precise measures of hemodynamic changes, such as optical imaging; (2) parametric modulation of magnitude or timing of activation in humans with corresponding measurement of fMRI signal changes; (3) simultaneous measures of neuronal activity (implanted electrode or electroencephalography, EEG) and fMRI signal changes; (4) nonsimultaneous measures of neuronal activity (magnetoencephalography, MEG; EEG) and fMRI signal changes; and (5) modeling of the hemodynamic response and comparison with precise activation magnitude, timing, or pharmacological manipulations.

Primary findings of these efforts have been BOLD signal changes appear to be driven by synaptic activity, as indicated by field potentials, rather than spiking itself (Viswanathan and Freeman 2007; Logothetis et al. 2001; Niessing et al. 2005), and MEG coherence changes in the gamma frequency range correlate spatially with fMRI signal changes (Muthukumaraswamy and Singh 2008; Singh et al. 2002).

Interpretation limits:

- Cannot differentiate inhibitory from excitatory activity.
- BOLD signal change is not a quantitative measure. Hemodynamic factors (baseline blood volume, neurovascular coupling) influence location, magnitude, and dynamics.
- Interpretation aspirations.
- Work has been progressing rapidly in the area of calibration. It will likely be possible to perform a spatial calibration of BOLD signal using the resting state fluctuation data (accompanied by a measurement of breathing depth using a chest strap). This will not only increase spatial specificity but will also reduce intrasubject variability and increase statistical power when averaging multisubject data.
- Convergent evidence from multimodal studies will continue to increase the confidence in the fidelity of the relationship between BOLD signal and underlying neuronal activity. This will have direct impact on the clinical applications of fMRI.
4 What About Thinking?

What about thinking? A major theme in this book is the quest to understand thinking. The question that most reading this chapter will want to know the answer to is: “What can fMRI, or more generally, neuroimaging, contribute to our pursuit of an understanding of thinking?” Does it really help to be able to look into the brain? To borrow an analogy, can one really truly understand how computers work by opening up a computer chassis and probing the components with a heat gun? Can identifying the when, where, and how much in the brain provide enough information so that we can begin, from this information, to derive principles of thinking? Even if we had a perfect picture at infinite spatial and temporal resolution of what was actually happening in the brain during thought, would we even then begin to understand thinking? Does it really matter what the limits of fMRI are with regard to answering questions about thinking?

It seems apparent that to truly understand the brain, a much wider context (physical and evolutionary factors) needs to be considered. Thinking itself might someday be deconstructed into simple algorithms that can be carried out within different media other than brains. Perhaps a simple model of interacting layers of neuronal networks may emerge as being able to explain thought (Hawkins and Blakeslee 2004). It is my feeling that because thinking is a subjective process, it tends to be shrouded in mystery, and potentially elevated to a status, either correctly or incorrectly, that defies understanding.

fMRI has been an extremely effective tool with regard to deepening our understanding of specific aspects of thinking. Specific regions, networks, dynamics, and patterns have been revealed as they are associated with, among many other processes, learning (Cabeza and Nyberg 2000; Peissig and Tarr 2007; Posnanski and Riera 2006), working memory (Wager and Smith 2003), emotional processing (Beauregard 2007; Saxe and Kanwisher 2003; Vogeley et al. 2001), deception (Langleben et al. 2006; Phillips 2004), social interaction (Montague et al. 2002), humor (Berns 2004; Moran et al. 2004; Watson et al. 2007; Wild et al. 2006), and introspection (Goldberg et al. 2006).

At the end of the day, we might be able to then say that x network, on x spatial scale, is directly related to say, theory of mind, willed action, and humor. So fMRI reveals the functions of specific processing modules. Does this really tell us anything that will help our understanding of thinking? Do we need to know what modules overlap in function or how large they are or where they are located in the brain? Does this information really matter? What spatial scale in the brain is the most critical for the understanding of thinking? While all of our tools are able to probe many different spatial scales, there are also many which have not been investigated yet. Does this matter?

I believe that more will be understood about thinking once we can integrate data across all temporal and spatial scales and use this information to construct testable...
models of thinking. The fMRI specific spatial scale of millimeters and a temporal resolution, in most cases, of seconds is a relatively narrow temporal/spatial niche to be studying how humans think.

The human thinking process has evolved as a strategy for human survival in very specific context. One might say that how we see the world and, therefore, how we think about the world, is highly tuned to the physical and social environment in which we evolved. fMRI can tell us how the brain works on a very specific temporal and spatial scale. It can certainly contribute to but not provide the whole story of how we think. To answer this we need to draw upon not only the vast array of imaging and behavioral measures, but also on data-driven models of how functional units in the brain interact across spatial and temporal scales to create such emergent activity such that human beings can see, hear, feel, move, react, solve problems, create, learn, and introspect.

5 Further Information

MRI and fMRI basics:
- http://www.simplyphysics.com/MAIN.HTM
- http://defiant.ssc.uwo.ca/Jody_web/fmri4dummies.htm

Processing software:
- http://www.bic.mni.mcgill.ca/software/: from the Brain Imaging Center at McGill University
- http://grommit.lrdc.pitt.edu/fiswidgets/: a Java graphical user interface for a number of neuroimaging analysis packages
- http://www.mayo.edu/bir/Software/Analyze/Analyze.html: from the Mayo Clinic
- http://www.brainvoyager.com/: a commercial product from Brain Innovation (Rainer Goebel)
- http://www.math.mcgill.ca/keith/fmristat/: a set of useful MATLAB programs (Keith Worsley)
- http://www.fmrib.ox.ac.uk/fsl/: a comprehensive set of analysis programs (Steve Smith, Oxford University)

Books:
• Functional MRI (Eds. Chrit Moonen and Peter A. Bandettini), Springer, Berlin, 1999
• Functional MRI: Basic Principles and Clinical Applications (Eds. Scott H. Faro, and Feroze B. Mohamed), Springer, Berlin, 2005

fMRI course Web sites:
• http://www.nmr.mgh.harvard.edu/fmrivfp/
• http://www.firc.mcw.edu/course/

References


Goldberg I, Ullman S, Malach R (in press) Neuronal correlates of “free will” are associated with regional specialization in the human intrinsic/default network. Consciousness Cogn
Goodyear BG, Menon RS (2001) Brief visual stimulation allows mapping of ocular dominance in visual cortex using fMRI. Hum Brain Mapp 14:210–217
Haynes JD, Rees G (2005a) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat Neurosci 8:686–691


Xiong J, Fox PT, Gao JH (2003) Directly mapping magnetic field effects of neuronal activity by magnetic resonance imaging. Hum Brain Mapp 20:41–49