NeuroImage 47 (2009) 1417–1424

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Investigating the consistency of brain activation using individual trial analysis of high-resolution fMRI in the human primary visual cortex

Ajay K. Nemani [a,b], Ian C. Atkinson [a,c,d], Keith R. Thulborn [a,d,c,*]

[a] Center for Magnetic Resonance Research, University of Illinois at Chicago, Chicago, IL 60612, USA
[b] Department of Physiology and Biophysics, University of Illinois at Chicago, Chicago, IL 60612, USA
[c] Department of Radiology, University of Illinois at Chicago, Chicago, IL 60612, USA
[d] Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL 60612, USA

ARTICLE INFO

Article history:
Received 27 January 2009
Revised 1 May 2009
Accepted 5 May 2009
Available online 14 May 2009

Keywords:
fMRI
BOLD
Individual trial
Brain function
Visual cortex

ABSTRACT

Conventional functional magnetic resonance imaging using blood oxygenation level dependent contrast requires signal averaging and statistical methods to detect activation. Signal averaging implicitly assumes that brain activation in response to a stimulus is reproducible on the scale of the imaging voxel. This assumption is examined in the absence of averaging by analyzing individual trials of individual voxels that approach the size of the functional unit, the cortical column, in the human primary visual cortex. In the absence of spatial and temporal averaging, even highly active voxels demonstrate inconsistent activation to the same repeated stimulus despite consistent behavioral responses. This observation implies a variable selection of suitable cortical columns from a population of available functional units to produce consistent perception of the stimulus. The implication of this observation for neuroplasticity and behavioral consistency at the level of functional units is discussed.

Introduction

Functional magnetic resonance imaging (fMRI) has always required signal averaging to detect brain function. This averaging has been typically implemented either temporally, by summation over repeated stimulus–response cycles, spatially, by blurring, image co-registration, transformation into a reference coordinate space or region-of-interest (ROI) analysis (Nieto-Castanon et al., 2003), or some combination thereof. Statistical analysis has been used to assign significance to these averaged data to reach a conclusion about the *general* pattern of activity (Duann et al., 2002; Friston, 1996; McKeown et al., 2002). Although tremendous strides in understanding brain function have been made with this approach, the use of temporal and spatial averaging precludes a detailed study of information processing at the level of the cortical column, the functional unit of the brain (Mountcastle, 1979; Obermayer and Blasdel, 1993). The implicit assumption underlying signal averaging is that, as a consistent stimulus evokes a consistent behavioral response, the underlying brain activation is also consistent. However, this assumption has never been examined *in vivo* on the temporospatial scale of individual responses by individual cortical columns. Measuring columnar responses to single cognitive events may establish if the consistency of perception exists at this level of brain organization.

Conventional fMRI with blood oxygenation level dependent (BOLD) contrast is based on the changes in magnetic susceptibility of hemoglobin within erythrocytes that accompany the hemodynamic response to neuronal activity to map the pattern of brain activation for a given cognitive task (Kwong et al., 1992; Ogawa et al., 1992). The measured signal change due to task-related neuronal activity is small compared to the background signal variation and, therefore, challenging to detect reliably. The primary motivation for using both temporal and spatial signal averaging is to achieve a sufficient BOLD contrast-to-noise ratio (CNR) for reliable detection. Corrections for head motion introduce spatial blurring by assuming a rigid brain that discretely moves from volume-to-volume despite continuous multi-planar imaging (Kim et al., 2008; Oakes et al., 2005). For intersubject analysis, co-registration algorithms require spatial filters in order to warp individual data to a common anatomical template (Brett et al., 2002; Lazar et al., 2002; Mazziotta et al., 2001; Talairach et al., 1993). Traditional statistical analysis techniques (e.g., Student's *t*-test) are robust for measurements of reproducible responses to repetitive stimuli. Such methods have produced useful summaries of average brain activation for individuals or groups of subjects.

Even with signal averaging, many studies have demonstrated intrinsic variability in trial-to-trial activity. This variability has been attributed to behavioral sources, including attention and anticipation (Drevets et al., 1995), and to task performance (Sapir et al., 2005).

* Corresponding author. Center for Magnetic Resonance Research, Room 1307, Outpatient Care Center, M/C 707, University of Illinois at Chicago, 1801 West Taylor Street Chicago IL 60612, USA. Fax: +1 312 355 3085.
E-mail address: kthulbor@uic.edu (K.R. Thulborn).

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.05.018

The inclusion of individual trial variability into the statistical model increases detection sensitivity (Donnet et al., 2006; Lu et al., 2005). Furthermore, Windischberger et al. (2002a) examined the possibility of inconsistent activation from trial-to-trial in individual voxels and found that the majority of active voxels have less than unity consistency. However, this result was limited by the use of the *t*-test, which is relatively insensitive at the single trial level and by spatial blurring due to motion correction and image co-registration.

While several studies have improved both the temporal and spatial resolution of fMRI, none has reported results without incorporating signal averaging in one form or another. Studies that have examined individual trial activity with high temporal resolution, such as time-resolved fMRI, used extensive spatial averaging in the form of an anatomical ROI over an entire gyrus (Richter et al., 2000). Conversely, studies that reached the columnar level of spatial resolution have only reported results averaged over multiple trials or with long stimulus durations (Zhao et al., 2005).

This report investigated the consistency of activation in human visual cortex from trial-to-trial at a level approaching that of the cortical column (Mountcastle, 1979; Obermayer and Blasdel, 1993). Results from high-resolution BOLD fMRI experiments of a simple, well-controlled visual paradigm with a short stimulus are presented with an analysis that allowed for a subset of highly stimulated voxels to be examined on the individual voxel and single trial level. These voxels demonstrated trial-to-trial activation with less than unity consistency that could not be explained by detection insensitivity, behavioral variability, or physiological fluctuations. These results were compared to those from the signal averaging method of fMRI analysis. The implication of these findings for understanding brain function, neuroplasticity, and behavior is discussed.

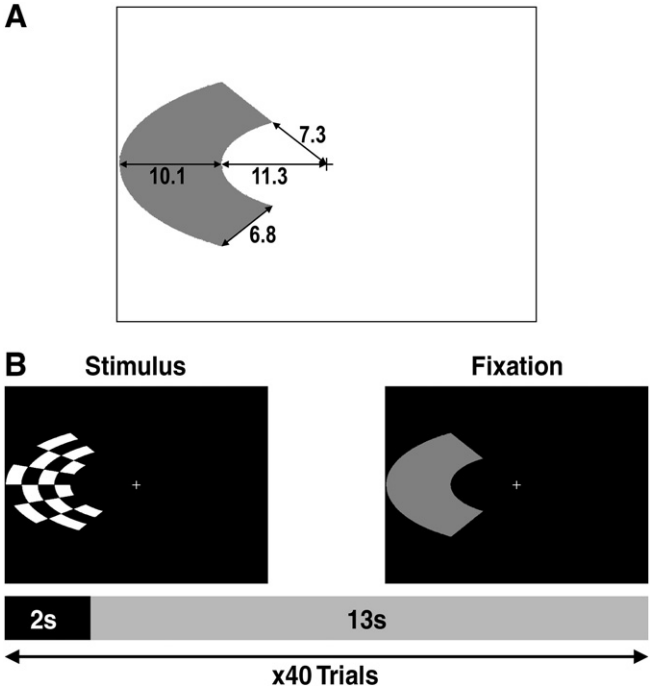

Fig. 1. Schematic diagram of the visual paradigm. (A) Locations and sizes of the stimulus components in degrees of visual angle. The central white fixation cross subtended an angle of 1.4°. (B) Stimulus timing for a single trial in the paradigm showing a left peripheral hemifield checkerboard of 2 s duration and an iso-luminescent fixation condition of 13 s. The subject fixated on the central white cross throughout the paradigm.

Materials and methods

Subjects

Healthy human volunteers ($n=4$, ages 18–45 years, 100% male) with no history of psychiatric, neurological, or cardiac disease gave written informed consent to participate in this IRB approved protocol. Subjects had normal or corrected-to-normal (via contact lens) visual acuity. All subjects had participated in at least one fMRI session prior to this study, thereby minimizing effects from novelty and anxiety (Brennan et al., 1988).

Paradigm

The paradigm, illustrated in Fig. 1, used an event-related design of 41 trials, the first of which was used for eye movement calibration (see below). The remaining 40 trials contained a short stimulus (2 s duration) and long recovery time (13 s) for an interstimulus interval of 15 s. The stimulus consisted of a black background containing a black and white (100% contrast), semi-radial, left hemifield, alternating checkerboard with a central fixation cross present throughout the stimulus and recovery periods (Tootell et al., 1998). The frequency of checkerboard alternation was 10 Hz. The relative sizes and locations of the stimulus components are shown (Fig. 1A) in terms of visual angle. The unilateral peripheral checkerboard stimulus allowed the contralateral side of the visual cortex to be used as a control area that received no stimulation. An iso-luminescent gray background (50% contrast) replaced the checkerboard region during the baseline period (Fig. 1B) to minimize pupil contraction at the onset of the stimulus. Data from the first trial of each experiment were used for calibration of the eye tracking system and, therefore, were not used in subsequent analysis. Except for this calibration period, subjects were instructed to maintain fixation on the central cross throughout the scanning period. The total length of the paradigm was 615 s. The calibration was performed by having the subject follow a target (central fixation cross) with their eyes as it moved from the central fixation location to peripheral locations along the horizontal ($\pm 3.6°$, $\pm 7.2°$) direction. The exact angle was calculated from the position of the eye from the rear projection display screen. These distances span the distance from the fixation point to the peripheral checkerboard in the rest of the paradigm. The size of this known movement then allowed the visual angle of any eye movement to be measured.

The interstimulus interval of 15 s allowed the hemodynamic response to return to the baseline while maximizing the number of trials in each experiment (Bandettini and Cox, 2000). The definition of a short stimulus is important in the current context. Many studies have demonstrated robust, reproducible activation in the visual cortex with sustained (10+ s), high contrast stimulation (Tootell et al., 1998). While these results are useful for mapping the visual areas, extended stimulation may recruit more functional units and also enhance intravascular effects from draining veins, limiting spatial resolution. A short stimulus has the advantage of preserving high spatial resolution in the face of these effects (Goodyear and Menon, 2001).

Paradigms were projected onto a visor screen mounted approximately 200 mm in front of the subject's eyes via an fMRI synchronization control system (MRIχ Technologies, Bannockburn, IL). The screen was viewed via an angled mirror mounted above the subject's eyes.

Imaging

Functional and anatomical images were acquired using a clinical 3.0 Tesla whole-body MR scanner (HDx Signa, GE Healthcare, Waukesha, WI) with a water-cooled, high performance, clinical gradient set (model CRM, GE Healthcare, Waukesha, WI) in conjunction with a custom-built 4.5-inch transmit/receive surface coil. All MR procedures were carried out within the Food and Drug Administration guidelines (Zaremba, 2003). Head motion was minimized by tactile feedback from pads surrounding the subject's head and by visual

feedback provided by the visor apparatus (Thulborn, 1999). The slice prescription of functional scans was established as follows. A short, functional imaging study using a bilateral checkerboard in a block-design paradigm was performed (gradient-echo, echo-planar sequence, 8 slices, thickness/skip = 3/1 mm, FOV = 200 mm, matrix = 64×64, TR = 1000 ms, TE = 25 ms, flip angle = 60°, 150 volumes collected). Large blood vessels were identified with flow-sensitive axial and coronal 2D gradient-echo imaging (spoiled gradient-echo sequence (SPGR), 8 slices, co-incident with the fMRI slices; thickness/skip = 3/1 mm, FOV = 200 mm, matrix = 256×256, TR = 18 ms, TE = 5.5 ms, flip angle = 90°) that enhances through-plane flow against the T1 saturated in-plane static background. The slice with maximal BOLD activation, but away from large vessels, was chosen as the optimal location for further functional imaging.

High-resolution functional imaging was performed using T_2*-weighted gradient-echo, echo-planer imaging through the optimal axial slice identified above (thickness = 3 mm, FOV = 80 mm, matrix = 128×128, TR = 500 ms, TE = 35.5 ms, flip angle = 55°, 1230 volumes collected). These imaging parameters were chosen to maximize the signal-to-noise ratio (SNR) at submillimeter voxel in-plane dimensions with dense temporal sampling. After functional imaging, a high-resolution 3D SPGR was acquired for anatomical context (40 slices, thickness/skip = 1.5/0 mm, FOV = 120 mm, matrix = 256×256, TR = 10 ms, TE = 2 ms, flip angle = 25°).

Physiological and behavioral monitoring

Physiological data (cardiac and respiratory waveforms and peripheral arterial O_2 saturation) synchronized to the image acquisition were recorded using both scanner hardware (Signa HDx, GE Healthcare, Waukesha, WI) and fMRI hardware (MRIχ Technologies, Bannockburn, IL). Cardiac cycle and O_2 saturation were acquired via a pulse oximeter placed on the subject's non-dominant index finger. Respiration was measured by a tension transducer placed around the subject's chest.

As the experiment required subjects to fixate for an extended period of time, eye movement away from fixation was a potential source of behavioral variability. Eye movement was monitored using an infra-red (IR) camera via a dichroic mirror placed over the subject's right eye (MRIχ Technologies, Bannockburn, IL). These data were analyzed retrospectively frame-by-frame to track the corneal reflection. The calibration of sensitivity to horizontal eye movement was performed by having the subject saccade to visual targets projected onto the visor screen at fixed visual angles (±3.6°, ±7.2°) during the first trial (calibration trial) of each experiment.

Synchronized behavioral responses were captured with a fiber optic-based finger switch (MRIχ Technologies, Bannockburn, IL) placed in the dominant hand. For all experiments, the subjects were instructed to press the finger switch as quickly as possible at the onset of each visual stimulus.

Preliminary analysis

Data analysis was performed using customized software written in MATLAB (The Mathworks, Inc., Nantick, MA) and designed to avoid sacrificing spatial or temporal resolution. The following post acquisition procedures were employed.

The distribution of reaction times was analyzed for each experiment as a measure of attention, variation in which has been shown to introduce variations in activation (O'Craven et al., 1997). If an outlier reaction time was found (>3 standard deviations), all data corresponding to that trial were rejected. This same procedure was applied for any deviations of the corneal reflection from central fixation. Since only the fixation cross (size = 1.4°) was present within the central portion of the paradigm, a shift in fixation on any part of the checkerboard would require at least 5.8° (horizontally) of movement in visual angle (Fig. 1A). Such a shift was readily detectable with the eye tracking system. These procedures ensured that the data were free from behavioral variations in attention and eye movement.

Overall head motion was estimated by two methods. Because only one slice location could be imaged per subject, SPM2 software (http://www.fil.ion.ucl.ac.uk/spm/) was unable to estimate the full set of six rigid body motion parameters (Friston et al., 1996). Therefore, a center-of-mass estimation was also performed to confirm the results from SPM2. If the standard deviation of motion was greater than 1/3 of the in-plane voxel resolution, the entire series was rejected. With experienced, trained subjects and adequate motion suppression (visual and tactile feedback), this criterion represented a strict but achievable standard (Thulborn, 1999). Therefore, the data were left uncorrected for motion, which avoided any spatial blurring involved with resampling and co-registering data. In addition to overall motion, if there was any single volume where the motion was greater than the in-plane voxel resolution, the data corresponding to that entire trial were rejected.

To increase the number of voxels with sufficient CNR for individual event analysis, the CNR of the image-space data was improved by processing the data with a blind signal estimation algorithm (Atkinson, 2008). This algorithm, which was designed to minimize temporal and spatial blurring, improves the CNR of fMRI data by removing noise without distorting the BOLD signal. This is achieved by approximating the optimal linear estimator for fMRI data using a hybrid hemodynamic response-wavelet subspace and a signal-preserving threshold. The hemodynamic response used by the algorithm was estimated from the time-averaged response of the most active voxels (identified with Bonferroni-corrected *t*-test, see Fig 3A). The other algorithm parameters were selected as suggested in Atkinson et al. (2008). This processing has been shown to improve BOLD CNR without introducing significant spatial and temporal distortions, and does not reduce trial-by-trial consistency in the analysis.

Cardiac and respiratory sources have been reported to influence fMRI data. These effects are largely confined to voxels on or near large vessels or cerebrospinal fluid (CSF) spaces (Dagli et al., 1999; Hu et al., 1995; Windischberger et al., 2002b).

Even though the voxels used in this study were away from large vessels, the data were corrected for physiological effects on a voxel-wise basis. Cardiac and respiratory physiological noise was removed using RETROICOR, a retrospective image-space correction algorithm for removing the influence of physiological noise from fMRI time-course data without blurring the data (Glover et al., 2000). Measurements of the cardiac and respiratory waveforms, made using the fMRI

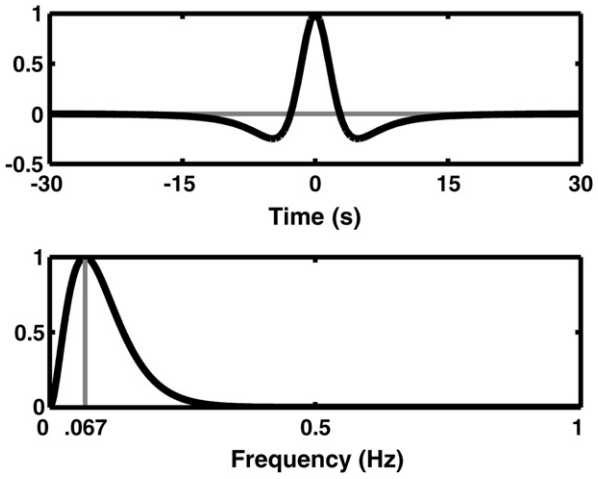

Fig. 2. The paradigm-analyzing wavelet, Ψ_p, in the time domain (top) and frequency domain (bottom). Note that Ψ_p forms a band-pass filter centered at the paradigm frequency.

presentation system, were used as linear regressors to remove signal contamination from cardiac and respiratory sources. All data were processed both with and without this algorithm without alteration of the consistency results.

After implementing the above procedures, the surviving data points were z-scored to remove the baseline offset and normalize the noise variance to unity. The mean and standard deviation used for the z-transform was derived from the baseline portion of the average BOLD response across all trials for each voxel. This baseline was defined as the last 4 s (8 samples) from each trial but this choice did not alter the result if a shorter period was used (3 s, 6 samples).

Student's t-test

A voxel-wise t-test was used to generate conventional activation maps by signal averaging across trials. Active condition data were defined as the most positive 3–4 s (6–8 samples) of the increased BOLD response centered around a latency of 5–6 s from the beginning of the stimuli. This was compared by t-test to the baseline data from the last 3–4 s (6–8 samples) of each trial. The number of samples used for active and baseline were matched and were manually chosen for each subject. The results were not sensitive to the choice of the number of samples and location of the center of the active and baseline conditions within the ranges given. The data were given a threshold ($p=0.05$) with a Bonferroni correction for multiple comparisons.

Specified resolution wavelet analysis (SRWA)

The determination of trial-to-trial consistency at high spatial and temporal resolution required that each voxel be examined for each trial. To that end, the following SRWA was adapted for analysis of each trial of each individual voxel time course (von Tscharner and Thulborn, 2001). A frequency-domain filter was constructed for efficient extraction of the major features of the hemodynamic response. A customized, non-orthogonal, pseudo-Paul-type wavelet was created from the mother wavelet, $\Psi_{s,m}(\nu)$, defined (in the frequency domain) by:

$$\Psi_{s\cdot m}(\nu) = \left(\frac{s}{m}\cdot\nu\right)^{m}\cdot \mathrm{e}^{\left(-\frac{s}{m}\cdot\nu+1\right)\cdot m} \tag{1}$$

where $s=$ scale, $m=$ mode, and $\nu=$ frequency. This wavelet definition does not have a closed form solution in the time domain. However, since a single realization of this wavelet was used as a band-pass filter, no such definition was necessary. The mode was chosen to ensure narrow temporal support of the wavelet ($m=2$), which prevents adjacent trials from influencing each other in the analysis. Since the central peak frequency of the wavelet is required to be the paradigm frequency ($\nu_c=\nu_p$), the scale was set to $s=2/\nu_c=30s$ for the paradigm frequency of 0.067 Hz (1/15 s) used in this study (Eq. (1) is maximized when $\nu=m/s$), leading to a paradigm-analyzing wavelet Ψ_p. Fig. 2 shows the time and frequency representations of Ψ_p. The temporal representation, $\Psi_p(t)$, resembles the conventional BOLD impulse response.

The time course of each voxel, $S(t)$, was first normalized by z-score to form $S_n(t)$ as described above. The wavelet transform for the voxel, $W(t)$, was calculated by convolving $\Psi_p(t)$ with $S_n(t)$. Since $W(t)$ was derived from the analyzing wavelet, which is centered on ν_p, $W(t)$ represented the strength of the contribution of ν_p to the signal at time t, with a specified temporal resolution $\sim s/2=1/\nu_p=15$ s. From $W(t)$, a set of peak parameters were extracted that represented the magnitude, latency, and width of the peak contribution of ν_p to the signal. $P_a(n)$, $P_p(n)$, and $P_w(n)$ were defined as the amplitude, phase (latency relative to trial onset), and full width at half maximum (FWHM) of the nth trial in $W(t)$, respectively. A peak vector $\boldsymbol{P}(n)=[P_a(n), P_p(n), P_w(n)]$ was defined which relates the extracted peak parameters to the activity of each trial in the voxel.

Lower detection limit (LDL)

The investigation of whether voxels showed variable responses in which some trials were active and other trials were inactive demanded that both Type I and Type II errors be given balanced consideration at the individual trial level. This was achieved for each voxel by determining the minimum required BOLD signal amplitude for which both Type I and Type II error could be controlled at a specified error rate ($\alpha=0.05$) using the same peak threshold vector to discriminate between active and inactive trials.

The null and alternate distributions of the peak vector were estimated using the stationary bootstrap technique (Politis and Romano, 1994). Many studies have found fMRI data to be non-white or non-Gaussian, although the origins of this noise have not been adequately settled (Gautama and Van Hulle, 2005; Lund et al., 2006; Zarahn et al., 1997). The stationary bootstrap was used as a means to establish the statistical significance of $\boldsymbol{P}(n)$ while respecting any autocorrelation in the data. To avoid bias from signal-related variance (Friman and Westin, 2005), activation was first removed from the normalized time-series, $S_n(t)$, by means of a first order deconvolution with the cycle-averaged response to yield a designaled time-series, $S_d(t)$. Specifically, the trial-averaged BOLD signal was fitted to each individual trial and then subtracted to yield a response-free baseline time-series suitable for resampling. The optimal bootstrap block length, β, was estimated from the autocorrelation spectrum of $S_d(t)$ (Politis, 2003; Politis and White, 2004). Multiple ($N=50$) surrogate time courses were generated by random circular sampling from $S_d(t)$ in blocks of length b, where b is a sample from a geometric distribution with mean β (Politis and Romano, 1994). After bias correction, these surrogate time courses represented synthetic realizations of $S_d(t)$, with similar statistical properties. Half of the 2000 synthesized events (50 surrogate time courses with 40 events each) were randomly selected to remain unchanged (inactive synthetic events). Synthetic BOLD activation was added to the remaining 1000 events (active synthetic events). The shape of the synthetic signal was estimated from the original normalized time-series by fitting a gamma-variate BOLD signal model to the averaged data (Ollinger et al., 2001).

SRWA was applied to each surrogate time course, each of which contained a mixture of inactive and active synthetic events. The null and alternate distributions of $\boldsymbol{P}(n)$ were estimated from the peak parameters extracted from the inactive and active synthetic events, respectively. The null distribution was used to select the peak parameter thresholds to control Type I errors at the specified rate of $\alpha=0.05$. The amplitude of the BOLD signal added to the active synthetic events was iteratively adjusted using the Nealder–Mead simplex procedure such that the peak parameter thresholds found to control Type I errors at a rate of α also controlled Type II errors at a rate of α. This amplitude was defined as the LDL and represented the minimum contrast required for control of both Type I and Type II errors to a rate of $\alpha=0.05$. The null and alternate distributions were estimated for each iteration of synthetic BOLD contrast.

Individual trial analysis

All voxels with a BOLD contrast averaged over all 40 trials of less than the LDL were eliminated from individual trial analysis. Each of the 40 trials in voxels meeting the LDL was classified as active if $P_a(n)$, $P_p(n)$, and $P_w(n)$ all met or exceeded the corresponding peak parameter threshold. The trial-by-trial consistency of a voxel was defined as the percentage of active trials over the course of the experiment (Windischberger et al., 2002a). The visual paradigm was expected to activate only voxels in the right visual cortex. Individual trial analysis was applied to comparable non-active (by conventional

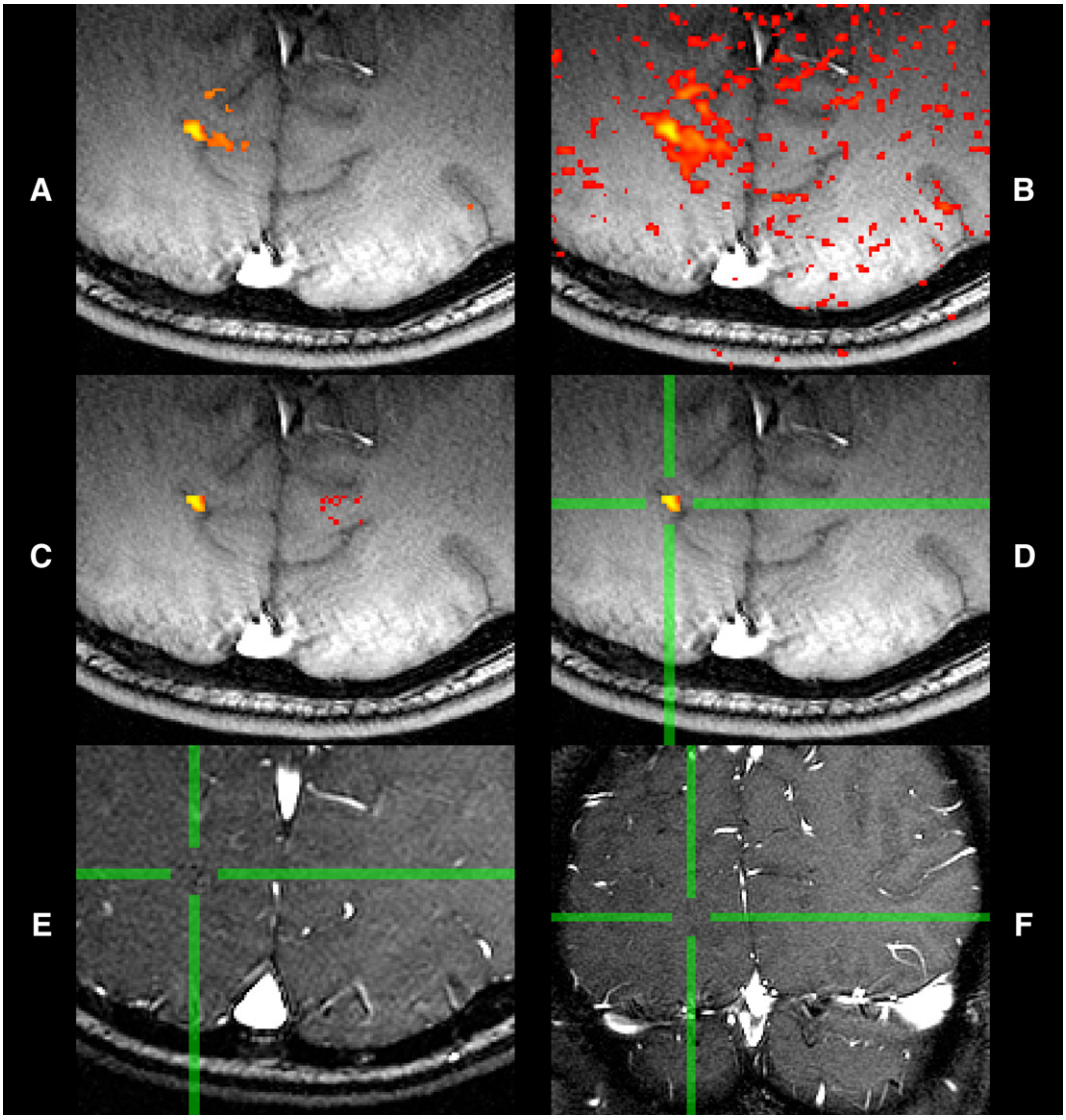

Fig. 3. Conventional activation maps calculated with a *t*-test for a 40-trial event-related left peripheral hemifield flashing checkerboard in which all trials were averaged. (A) Activation map with a Bonferroni-corrected threshold ($t=4.46$, $p=4.99\times10^{-6}$). (B) Activation map with an uncorrected threshold ($t=1.65$, $p=0.05$). (C) Voxels chosen for individual trial analysis are shown as the highly stimulated voxels (yellow) in the right calcarine gray matter and the non-active control voxels (red) in the contralateral left cortical gray matter. Non-active voxels were chosen to have similar means and variances as the highly stimulated voxels. (D) Highly stimulated voxels chosen for individual trial analysis with green guidelines. (E) Axial flow-sensitive image co-registered with (D) showing the same green guidelines. (F) Coronal flow-sensitive image co-registered with (D) showing the same green guidelines.

t-test analysis) control voxels from the contralateral visual cortex. These inactive voxels were randomly chosen from a pool of voxels in the left visual cortex with similar mean signal intensity and variance as the contralateral active voxels during baseline periods.

Results

Fig. 3 shows representative conventional activation maps based on the *t*-test that were created at two statistical thresholds (Figs. 3A and B) to identify candidate active voxels for individual trial analysis (Figs. 3C and D) and to confirm that these voxels were not on or adjacent to large vessels as identified by flow-sensitive imaging (Figs. 3E and F).

The largest peak from each trial, extracted using the SRWA in conjunction with LDL thresholding, is shown in Fig. 4A for the time course of a representative voxel (only the first 10 of 40 trials of this representative voxel are shown for illustration). The amplitude and phase (FWHM not shown) of the peak extracted from each trial were then superimposed over the averaged BOLD response over all 40 trials (Fig. 4B, all 40 peaks shown from the same voxel as Fig. 4A). While most peaks were distributed near the apex of the trial-averaged BOLD signal, some trials showed peaks (marked by arrows) that were apparently outside of this distribution, suggesting the possibility that these trials may represent inactive events.

Very few voxels survived the LDL criterion: only 11–13 highly stimulated voxels on the single image available per subject contained sufficient activity among active trials to meet the LDL. Two representative voxels meeting the LDL (as well as an inactive control voxel from the contralateral side of the brain) are shown in Fig. 5 as raster diagrams to illustrate trial-specific activation. Rasters for both the original normalized time-series and associated wavelet-transformed time-series are shown.

The classification of individual trial activity shown in Fig. 5 revealed that these voxels showed variable responses across trials despite a consistent stimulus and behavioral response. These three voxels demonstrated highly consistent (90%, same voxel as shown in Fig. 4), moderately consistent (60%), and non-active (7.5%) behavior.

The patterns of trial-to-trial activation across voxels meeting the LDL as well as comparable control voxels (Fig. 3C) for all four subjects

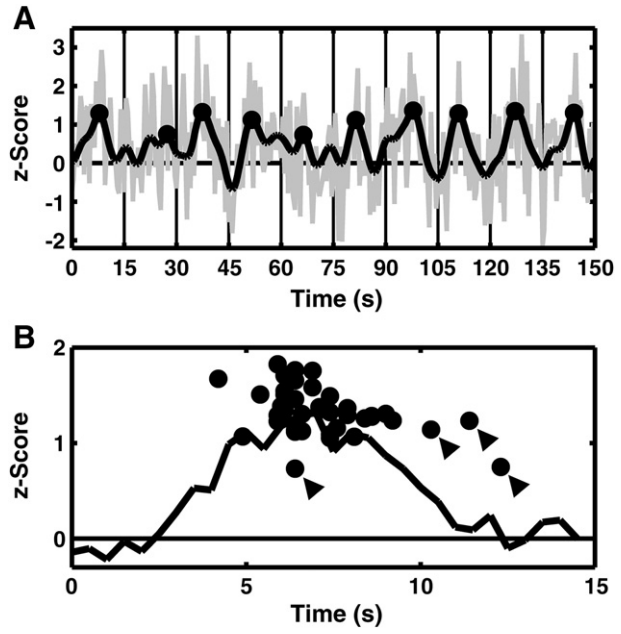

Fig. 4. Specified resolution wavelet analysis of a representative active voxel. (A) Wavelet-transformed time-series (black curved line) superimposed over the normalized time-series (gray line) for a representative active voxel from the right primary visual cortex. Data corresponding to the first 10 of 40 trials of the visual stimulus paradigm are shown, each giving rise to an extracted peak (black circles). (B) Distribution of all 40 extracted peaks (black circles) collapsed to a single trial superimposed over the averaged BOLD response (black curved line). While most of the peaks from this active voxel are distributed near the apex of the BOLD signal, four trials (marked by arrows) appeared to distribute differently, suggesting that these peaks may represent inactive trials.

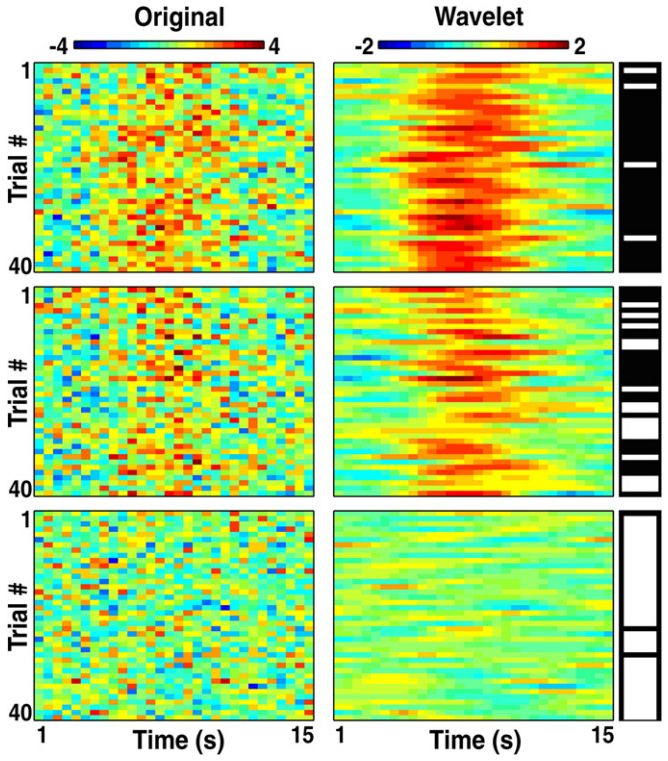

Fig. 5. Raster diagrams of representative voxels of the original (left column) and wavelet-transformed (middle column) time-course data for highly active (top, 90% consistency), moderately active (middle, 60% consistency) and non-active (bottom, 7.5% consistency) voxels. The classification of each trial as active (black) or inactive (white) based on the LDL is shown (right column). The original and wavelet color scales correspond to the *z*-score for each time point and applies for all three voxels.

are shown in Fig. 6. Although some strongly stimulated voxels showed high consistency, all failed to reach 100%. The mean consistencies of all voxels meeting the LDL were: 68 ± 12%, 49 ± 6.6%, 47 ± 6.0%, and 47 ± 5.8% for subjects 1 through 4, respectively. When the consistency over the first half of the experiment was compared to consistency over the second half, there were no significant differences (67 ± 14% vs. 69 ± 16% ($p = 0.76$), 43 ± 9% vs. 55 ± 9% ($p = 0.08$), 44 ± 7% vs. 47 ± 11% ($p = 0.71$), and 44 ± 10% vs. 47 ± 9% ($p = 0.59$) for subjects 1 through 4, respectively). The associated control voxels from contralateral visual cortex showed low false positive rates corroborated by a consistency across all subjects of 5.4 ± 1.9%. This false positive rate was expected based on the specified Type I and Type II error rates of $\alpha = 0.05$. The difference in average consistency between subjects can be attributed to the limited extent of coverage that was constrained by the scanner at high resolution to a single slice through the visual cortex. The low-resolution block-design paradigm using a similar checkerboard stimulus provided multiple slices for greater coverage and showed activity extending well beyond the high-resolution imaging slice (data not shown). Nonetheless, the voxels used for this individual trial analysis covered by the high-resolution experiment represented the most strongly active voxels identified by conventional average-based analyses. This is illustrated by comparing the selected voxels (Fig. 3C) with a conventional *t*-map that used conservative Bonferroni thresholding (Fig. 3A).

A further analysis of the active and inactive trials based on the LDL criterion is shown in Fig. 7 in which the averaged hemodynamic response of all active trials and all inactive trials are plotted for each subject. The active trial BOLD response (dark gray line) showed an improvement over the conventional BOLD response (black line) representing an average across all trials. More importantly, the inactive trial BOLD response (light gray line) showed no significant signal ($p = 0.63, 0.31, 0.44, 0.51$ for subjects 1 through 4, respectively). Despite the large number of trials contributing to these averages (active vs. inactive trial totals = 352 vs. 168, 255 vs. 265, 201 vs. 228, and 219 vs. 249 for subjects 1 through 4, respectively), a bimodal distribution of responses was identified. Active trial averages showed significant improvement in contrast over conventional averaging (increases of 48%, 80%, 96%, and 92% for subjects 1 through 4,

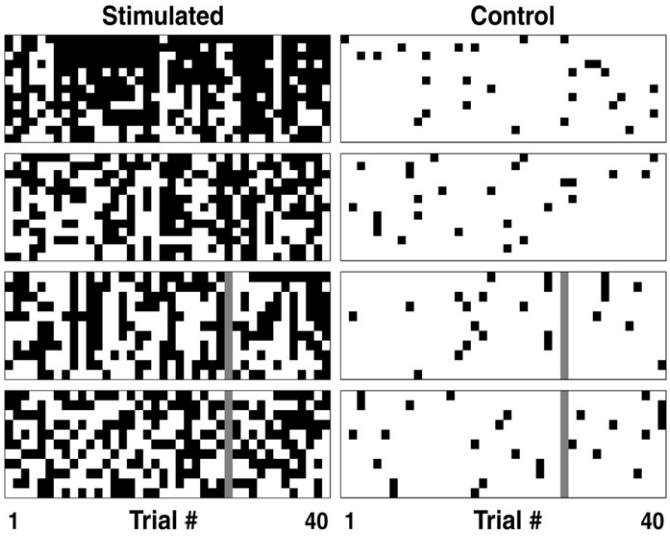

Fig. 6. Trial-based activity for all stimulated voxels meeting the LDL threshold (left column) and an equivalent number of control voxels (right column) for all 4 subjects. Each trial is shown as active (black) or inactive (white) for each subject. The gray lines show the individual trials rejected from the analysis due to behavioral or physiological aberrations. The consistencies for the stimulated voxels in each subject were 68 ± 12%, 49 ± 6.6%, 47 ± 6.0%, and 47 ± 5.8%. The consistency for the control voxels across all subjects was 5.4 ± 1.9%.

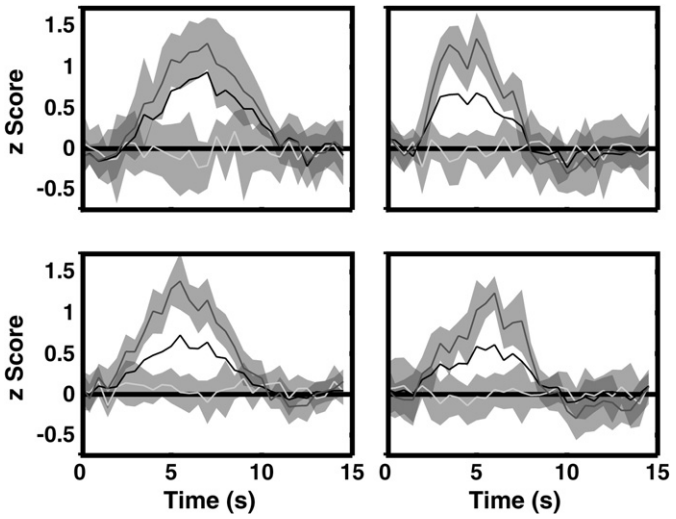

Fig. 7. The trial-averaged BOLD signal is shown for all trials (black line), active trials as classified by the LDL threshold (dark grey line) and inactive trials as classified by the LDL threshold (light grey line) for each of the four subjects. The grey highlight shows the standard deviation of the averaged data. Averaging all trials produced the conventional BOLD response that peaks at approximately 6 s. Averaging only active trials produced a stronger BOLD response despite fewer trials (352/520, 255/520, 201/429, and 219/468 trials for subjects 1 to 4, respectively). In contrast, averaging only inactive trials did not show any significant BOLD signal, consistent with these trials representing true negative responses ($p = 0.63$, 0.31, 0.44, 0.51 for subjects 1 to 4, respectively).

respectively), and inactive trial averaging showed no BOLD contrast. This supports the assertion that the LDL threshold has correctly classified (to an error of $\alpha = 0.05$) true positive and true negative trials for this population of voxels with high intrinsic single trial activity.

Although head motion can be a potential source of variability, that was not the case for these data which were chosen to have head motion of less than 1/3 of an in-plane voxel dimension (i.e., motion less than 208 μm). The standard deviation of translational head motion estimated by SPM2 was 72, 72, 53, and 82 μm while the same measure using center-of-mass was 93, 69, 122, and 110 μm for subjects 1 through 4, respectively. No data were rejected due to head motion across all four subjects and the data were left uncorrected for head motion.

Voxels selected for individual trial analysis (Fig. 3D) were shown to be away from large vessels by the absence of flow enhancement in co-registered axial and coronal flow-sensitive images (Figs. 3E and F, respectively). Consequently, the RETROICOR processing resulted in negligible alteration of active voxel time courses (data not shown).

Eye movement and finger switch response times were used to measure behavioral performance to eliminate variation in attention as a confound in brain activation on a trial-to-trial basis (O'Craven et al., 1997).

Representative eye movement data of the corneal reflection are displayed in Fig. 8. The eye movement data showed that the standard deviation about the center was $0.44° \pm 0.18°$ ($n = 4$), which was considerably smaller than the size of the fixation cross (1.4°). Subjects were able to maintain focus on the paradigm for the entire duration of the experiment so that visual stimuli were consistently delivered only to a specific area of the left visual field. No data were rejected due to failure to properly fixate throughout the experiment.

Reaction times to press a finger switch in response to the appearance of the visual stimulus over all 40 trials were 345 ± 45 ms, 270 ± 20 ms, 297 ± 17 ms, and 337 ± 100 ms for subjects 1 through 4, respectively, yielding an overall reaction time across all subjects of 314 ± 65 ms ($n = 4$). There was no statistically significant difference in reaction time between the first and seconds halves of the experiment (338 ± 44 ms vs. 355 ± 46 ms ($p = 0.27$), 269 ± 18 ms vs. 270 ± 22 ms ($p = 0.88$), 297 ± 17 ms vs. 296 ± 17 ms ($p = 0.95$), and 330 ± 104 ms vs. 344 ± 98 ms ($p = 0.67$) for subjects 1 through 4, respectively). Across all four subjects (160 trials), only data from two trials were rejected due to outlier reaction times (>3 standard deviations from that subject's mean reaction time). The data from these trials were not used in any analysis performed in this study, as reflected by the gray bars in the trial-to-trial activity diagram (Fig. 6).

Discussion

Individual responses in the primary visual cortex to individual visual stimuli have been investigated at a structural scale approaching the functional unit level. Although only limited coverage of the visual cortex could be achieved, voxels typically labeled as highly active with conventional analyses using signal averaging were shown to activate with less than 100% consistency despite consistent perception of the stimulus. This observation implies that a variable subpopulation of responsive units is selected on a stimulus-to-stimulus basis from a larger pool of available functional units in the visual cortex. Although the nature of the selection process remains speculative, this model suggests a mechanism by which the visual cortex maintains the flexibility to processes multiple visual stimuli in overlapping visual fields.

Individual trial analysis at high spatial resolution was achieved through a combination of strategies to ensure the observed results were valid. Decreased image SNR with high-resolution imaging approaching the size of single functional units was balanced by increasing BOLD contrast from reduced partial voluming of active and inactive tissue (Thulborn et al., 1997). Sensitivity was further enhanced by high field (3.0 T) imaging in combination with a small surface coil placed over the occipital lobe. A denoising algorithm was applied to enhance low CNR signals while avoiding significant spatial or temporal blurring (Atkinson et al., 2008). SRWA was used in combination with LDL thresholding to detect responses in individual voxels without averaging.

The assumptions behind the analysis were conservative. The wavelet-based analysis technique was developed to characterize the magnitude, latency, and width of the BOLD response for each trial based on *a priori* knowledge of the frequency of stimulus presentation. The corresponding amplitude, phase, and FWHM of these responses

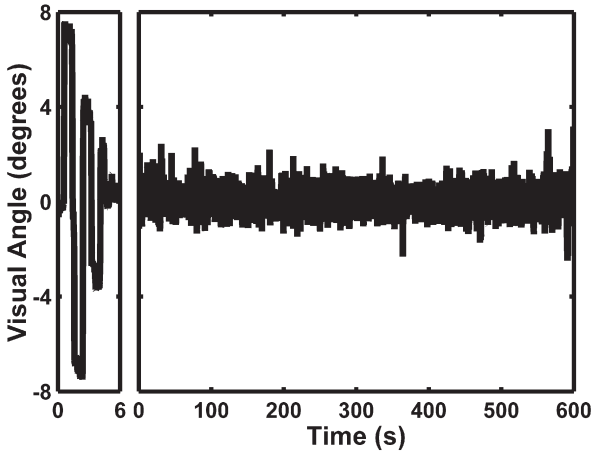

Fig. 8. Representative eye movement data of the corneal reflection tracked during the initial calibration period (left) and the 40 trials of the paradigm (right). The calibration shows the corneal reflection moving to 4 different locations across the left and right visual fields ($\pm 3.6°$, $\pm 7.2°$). The corneal reflection in the *x* direction is shown (the *y* direction results are similar). The data were low pass-filtered (cutoff frequency = 0.1 Hz) to remove fluctuations corresponding not to eye movement, but to very small head motion apparent only at the resolution of the video camera. The standard deviation about the baseline for this subject was 0.47°. Similar results were observed for the other three subjects (0.39°, 0.49°, 0.52°, respectively).

were used to classify individual trials as active or inactive based on a threshold that balanced Type I and Type II errors. This threshold was estimated for each individual voxel from non-parametric distributions based on simulations of null and signal-added surrogate data produced by stationary bootstrapping methods applied to the voxel time-series. The majority of voxels were rejected for individual trial classification due to the balanced error rate exceeding the specified error threshold of $\alpha = 0.05$. Only a small number of highly stimulated voxels (11–13 voxels per subject, $n = 4$) contained sufficient CNR to meet the LDL threshold and hence classify individual trials as active or inactive with less than $\alpha = 0.05$ Type I and Type II errors. This classification was confirmed by demonstrating that the average signal over inactive trials showed no significant response, eliminating false negatives as a source of misclassification. In addition, signal averaging either temporally (averaging over multiple trials) or spatially (averaging over multiple voxels) yielded conventional activation results.

These results imply that a subpopulation of cortical columns is selected from a larger population of available cortical columns that can respond to a given stimulus. The selection of cortical columns for a repeated stimulus varies from stimulus to stimulus, while still producing a consistent perception of the stimulus. Such redundancy may represent a form of neuroplasticity that allows flexible parallel processing of simultaneous stimuli. This model of BOLD activation at the scale of functional units reduces to the reproducible results produced by standard fMRI at lower resolution where the responses of multiple functional units are averaged together spatially or temporally.

References

Atkinson, I.C., Kamalabadi, F., Jones, D.L., Thulborn, K.R., 2008. "Blind Estimation for Localized Low Contrast-to-Noise Ratio BOLD Signals". IEEE Journal of Selected Topics in Signal Processing 2, 879–890.

Bandettini, P.A., Cox, R.W., 2000. Event-related fMRI contrast when using constant interstimulus interval: theory and experiment. Magn. Reson. Med. 43, 540–548.

Brennan, S.C., Redd, W.H., Jacobsen, P.B., Schorr, O., Heelan, R.T., Sze, G.K., Krol, G., Peters, B.E., Morrissey, J.K., 1988. Anxiety and panic during magnetic resonance scans. Lancet 2, 512.

Brett, M., Johnsrude, I.S., Owen, A.M., 2002. The problem of functional localization in the human brain. Nat. Rev. Neurosci. 3, 243–249.

Dagli, M.S., Ingeholm, J.E., Haxby, J.V., 1999. Localization of cardiac-induced signal change in fMRI. Neuroimage 9, 407–415.

Donnet, S., Lavielle, M., Poline, J.B., 2006. Are fMRI event-related response constant in time? A model selection answer. Neuroimage 31, 1169–1176.

Drevets, W.C., Burton, H., Videen, T.O., Snyder, A.Z., Simpson Jr., J.R., Raichle, M.E., 1995. Blood flow changes in human somatosensory cortex during anticipated stimulation. Nature 373, 249–252.

Duann, J.R., Jung, T.P., Kuo, W.J., Yeh, T.C., Makeig, S., Hsieh, J.C., Sejnowski, T.J., 2002. Single-trial variability in event-related BOLD signals. Neuroimage 15, 823–835.

Friman, O., Westin, C.F., 2005. Resampling fMRI time series. Neuroimage 25, 859–867.

Friston, K.J., 1996. Statistical parametric mapping and other analyses of functional imaging data. In: Toga, A.W., Mazziotta, J.C. (Eds.), Brain Mapping: The Methods. InAcademic Press, San Diego, pp. 363–396.

Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fMRI time-series. Magn. Reson. Med. 35, 346–355.

Gautama, T., Van Hulle, M.M., 2005. Estimating the global order of the fMRI noise model. Neuroimage 26, 1211–1217.

Glover, G.H., Li, T.Q., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn. Reson. Med. 44, 162–167.

Goodyear, B.G., Menon, R.S., 2001. Brief visual stimulation allows mapping of ocular dominance in visual cortex using fMRI. Hum. Brain Mapp. 14, 210–217.

Hu, X., Le, T.H., Parrish, T., Erhard, P., 1995. Retrospective estimation and correction of physiological fluctuation in functional MRI. Magn. Reson. Med. 34, 201–212.

Kim, B., Yeo, D.T., Bhagalia, R., 2008. Comprehensive mathematical simulation of functional magnetic resonance imaging time series including motion-related image distortion and spin saturation effect. Magn. Reson. Imaging 26, 147–159.

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

Lazar, N.A., Luna, B., Sweeney, J.A., Eddy, W.F., 2002. Combining brains: a survey of methods for statistical pooling of information. Neuroimage 16, 538–550.

Lu, Y., Jiang, T., Zang, Y., 2005. Single-trial variable model for event-related fMRI data analysis. IEEE Trans. Med. Imaging 24, 236–245.

Lund, T.E., Madsen, K.H., Sidaros, K., Luo, W.L., Nichols, T.E., 2006. Non-white noise in fMRI: does modelling have an impact? Neuroimage 29, 54–66.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., 2001. A probabilistic atlas and reference system for the human brain: international consortium for brain mapping (ICBM). Philos. Trans. R Soc. Lond. B Biol. Sci. 356, 1293–1322.

McKeown, M.J., Varadarajan, V., Huettel, S., McCarthy, G., 2002. Deterministic and stochastic features of fMRI data: implications for analysis of event-related experiments. J. Neurosci. Methods 118, 103–113.

Mountcastle, V.B., 1979. An organizing principle for cerebral function: the unit module and the distributed system. In: Schmitt, F.O., Worden, F.G. (Eds.), The Neurosciences: Fourth Study Program. InMIT Press, Cambridge, Mass, pp. 21–42.

Nieto-Castanon, A., Ghosh, S.S., Tourville, J.A., Guenther, F.H., 2003. Region of interest based analysis of functional imaging data. Neuroimage 19, 1303–1316.

O'Craven, K.M., Rosen, B.R., Kwong, K.K., Treisman, A., Savoy, R.L., 1997. Voluntary attention modulates fMRI activity in human MT-MST. Neuron 18, 591–598.

Oakes, T.R., Johnstone, T., Ores Walsh, K.S., Greischar, L.L., Alexander, A.L., Fox, A.S., Davidson, R.J., 2005. Comparison of fMRI motion correction software tools. Neuroimage 28, 529–543.

Obermayer, K., Blasdel, G.G., 1993. Geometry of orientation and ocular dominance columns in monkey striate cortex. J. Neurosci. 13, 4114–4129.

Ogawa, S., Tank, D.W., Menon, R., Ellermann, J.M., Kim, S.G., Merkle, H., Ugurbil, K., 1992. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc. Natl. Acad. Sci. U. S. A. 89, 5951–5955.

Ollinger, J.M., Corbetta, M., Shulman, G.L., 2001. Separating processes within a trial in event-related functional MRI. Neuroimage 13, 218–229.

Politis, D.N., 2003. Adaptive bandwidth choice. J. Nonparam. Stat. 15, 517–533.

Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. J. Amer. Stat. Assoc. 89, 1303–1313.

Politis, D.N., White, H., 2004. Automatic block-length selection for the dependent bootstrap. Econ. Rev. 23, 53–70.

Richter, W., Somorjai, R., Summers, R., Jarmasz, M., Menon, R.S., Gati, J.S., Georgopoulos, A.P., Tegeler, C., Ugurbil, K., Kim, S.G., 2000. Motor area activity during mental rotation studied by time-resolved single-trial fMRI. J. Cogn. Neurosci. 12, 310–320.

Sapir, A., d'Avossa, G., McAvoy, M., Shulman, G.L., Corbetta, M., 2005. Brain signals for spatial attention predict performance in a motion discrimination task. Proc. Natl. Acad. Sci. U. S. A. 102, 17810–17815.

Talairach, J., Tournoux, P., Missir, O., 1993. Referentially Oriented Cerebral MRI Anatomy : An Atlas of Stereotaxic Anatomical Correlations for Gray and White Matter. Thieme Medical Publishers, New York.

Thulborn, K.R., 1999. Visual feedback to stabilize head position for fMRI. Magn. Reson. Med. 41, 1039–1043.

Thulborn, K.R., Chang, S.Y., Shen, G.X., Voyvodic, J.T., 1997. High-resolution echo-planar fMRI of human visual cortex at 3.0 Tesla. NMR Biomed. 10, 183–190.

Tootell, R.B., Hadjikhani, N.K., Vanduffel, W., Liu, A.K., Mendola, J.D., Sereno, M.I., Dale, A. M., 1998. Functional analysis of primary visual cortex (V1) in humans. Proc. Natl. Acad. Sci. U. S. A. 95, 811–817.

von Tscharner, V., Thulborn, K.R., 2001. Specified-resolution wavelet analysis of activation patterns from BOLD contrast fMRI. IEEE Trans. Med. Imaging 20, 704–714.

Windischberger, C., Lamm, C., Bauer, H., Moser, E., 2002a. Consistency of inter-trial activation using single-trial fMRI: assessment of regional differences. Brain Res. Cogn. Brain Res. 13, 129–138.

Windischberger, C., Langenberger, H., Sycha, T., Tschernko, E.M., Fuchsjager-Mayerl, G., Schmetterer, L., Moser, E., 2002b. On the origin of respiratory artifacts in BOLD-EPI of the human brain. Magn. Reson. Imaging 20, 575–582.

Zarahn, E., Aguirre, G.K., D'Esposito, M., 1997. Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. Neuroimage 5, 179–197.

Zaremba, L.A., 2003. Guidance for magnetic resonance diagnostic devices—criteria for significant risk investigations, F.a.D.A. United States Center for Devices and Radiological Health, Department of Health and Human Services, ed.

Zhao, F., Wang, P., Hendrich, K., Kim, S.G., 2005. Spatial specificity of cerebral blood volume-weighted fMRI responses at columnar resolution. Neuroimage 27, 416–424.