

Enhanced Temporal Non-linearities in Human Object-related Occipito-temporal Cortex

Roy Mukamel¹, Michal Harel¹, Talma Hendler^{2,3} and Rafael Malach¹

¹Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel, ²Functional Brain Imaging Unit, Tel Aviv Sourasky Medical Center, Tel Aviv 64239, Israel and ³Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

To what extent does neural activation in human visual cortex follow the temporal dynamics of the optical retinal stimulus? Specifically, to what extent does stimulus evoked neural activation persist after stimulus termination? In the present study, we used functional magnetic resonance imaging (fMRI) to explore the resulting temporal non-linearities across the entire constellation of human visual areas. Gray-scale images of animals, houses and faces were presented at two different presentation rates — 1 and 4 Hz — and the fMRI signal was analyzed in retinotopic and in high order occipito-temporal visual areas. In early visual areas and the motion sensitive area MT/V5, a fourfold increase in stimulus presentation rate evoked a twofold increase in signal amplitude. However, in high order visual areas, signal amplitude increased only by 25%. A control experiment ruled out the possibility that this difference was due to signal saturation ('ceiling') effects. A likely explanation for the stronger non-linearities in occipito-temporal cortex is a persistent neuronal activation that continues well after stimulus termination in the 1 Hz condition. These persistent activations might serve as a short term (iconic) memory mechanism for preserving a trace of the stimulus even in its absence and for future integration with temporally correlated stimuli. Two alternative models of persistence (inhibitory and excitatory) are proposed to explain the data.

Keywords: fMRI, short-term memory, visual cortex

Introduction

A fundamental question concerning neural activity in high order visual areas is to what extent this activity may correlate with the physical properties of the visual stimulus. A number of recent studies have shown that often such activity departs from the optical features of the stimulus and correlates more tightly with the subject's perceptual state. For example, while early retinotopic cortex is sensitive to drastic changes in stimulus contrast, high order visual areas are invariant to such changes as long as stimulus recognition is not impaired (Avidan *et al.*, 2002a).

Similarly, holistic effects which go beyond the local physical properties of the stimulus have been amply documented. These effects, in which the perceptual state remains the same and the physical properties of the stimulus change, or vice versa include shape-based adaptation (Kourtzi and Kanwisher, 2001), completion effects of semi-occluded objects (Lerner *et al.*, 2002), illusory contours (Mendola *et al.*, 1999; Stanley and Rubin, 2003) and ambiguous figures such as the Rubin vase-face illusion (Hasson *et al.*, 2001; Andrews *et al.*, 2002). In all these experiments, fMRI activity in occipito-temporal cortex manifested a clear departure from the retinal stimulation and closer correlation to the perceptual state.

While this departure of brain activity from a tight correlation to the physical properties of the visual stimulus has been extensively investigated using static stimuli, the dynamic aspect of such transformations has received far less attention. We have previously shown, using a backward masking paradigm, that the relation between fMRI activation and exposure times of object images shows a high saturation non-linearity at exposure durations greater than 120 ms (Grill-Spector *et al.*, 2000). Such non-linearities were nicely correlated with subjects' recognition (Grill-Spector *et al.*, 2000; Bar *et al.*, 2001).

The dissociation of fMRI activation from linear dependence on the duration of visual stimulation is of particular interest since it could likely point to neuronal mechanisms that preserve the activity beyond the mere presence of the physical stimulus. These mechanisms could serve as a form of cortical short term memory. A variety of such memory effects have been demonstrated using different methodologies including single unit recordings, network modeling, neuroimaging and behavioral measurements of human recognition processes.

Using single unit recordings in IT cortex of the macaque, several groups have demonstrated that neuronal activation lasts beyond the actual retinal image presentation (Miyashita and Chang, 1988; Rolls and Tovee, 1994; Kovacs *et al.*, 1995; Yakovlev *et al.*, 1998; Tamura and Tanaka, 2001). These empirical results have also been reproduced in neural networks (e.g. Amit *et al.*, 1994). Finally, human psychophysical experiments have shown that when visual pictures are presented for a brief period (e.g. 110 ms), recognition performance depends on the interstimulus interval (ISI) and the type of the following stimulus (Potter, 1976; Intraub, 1980). These results suggest that stimulus evoked neural processing continues well after the optical stimulus has been removed. These processes might subserve the integration and linking of temporally related visual stimuli (Yakovlev *et al.*, 1998).

In the present study we have conducted a series of experiments specifically designed to map the neuroanatomical distribution of such putative, visually guided, memory-related processes. Keeping stimulus duration and epoch length fixed, we changed the presentation rate of visual stimuli within epochs. Under such conditions, brain regions in which no sustained memory processes occur during the ISI may be expected to have fMRI signals which are linearly correlated to stimulus presentation rate. On the other hand, the fMRI signal in brain regions manifesting sustained memory related processes during the ISI may be expected to increase non-linearly as a function of stimulus presentation rate. Figure 1 illustrates these two alternative neuronal responses and their predicted fMRI behavior in a graphic form.

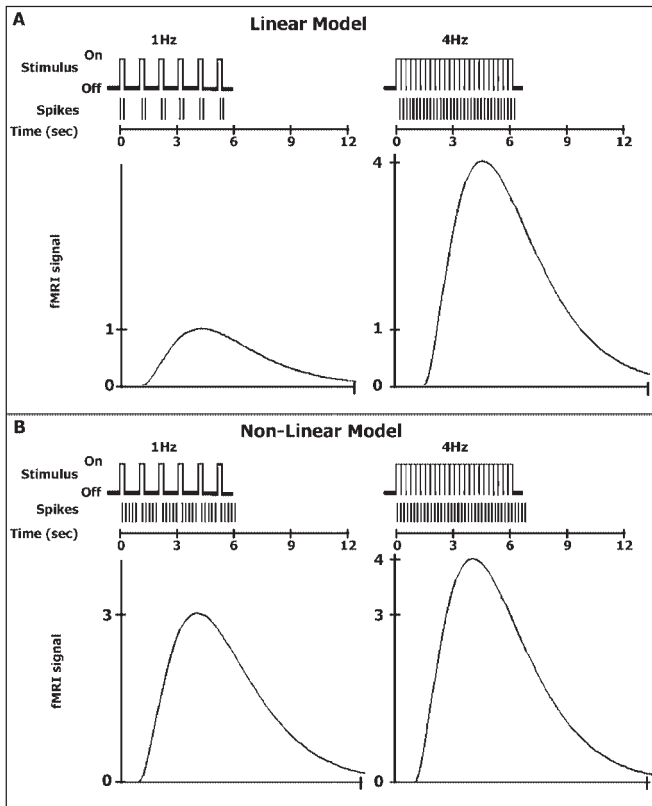


Figure 1. Persistent activity as the putative source of temporal non-linearity. The figure illustrates the linear versus non-linear correlation between stimulus presentation rate and fMRI signal. In the 1 Hz condition, six stimuli are presented for 250 ms followed by 750 ms blank and in the 4 Hz condition, 24 stimuli are presented in a consecutive manner. (A) In the linear model, the neural activity follows closely the stimulus presentation times resulting in a 1:4 ratio in the evoked fMRI signal during the 1 and 4 Hz conditions. (B) According to the non-linear model, during the 1 Hz condition, evoked neuronal activity persists in the absence of the stimulus (during the ISI), resulting in similar fMRI signals evoked by the 1 and 4 Hz conditions.

Previous studies have systematically changed stimulus presentation rate and measured neural activations in visual (Fox and Raichle, 1984; Mechelli *et al.*, 2000; Ozus *et al.*, 2001) and in auditory cortices (Binder *et al.*, 1994; Tanaka *et al.*, 2000). However, the above studies used either patterned flash stimuli or words which were not optimal stimuli for high order visual areas.

Our results show a consistent differential distribution of fMRI signal non-linearities across the visual cortex, with high order occipito-temporal cortex manifesting significantly higher levels of temporal non-linearities compared to early retinotopic areas and the motion sensitive area MT/V5. While the interpretation of such effects has to be cautious due to the sluggish nature of the fMRI response, a likely source for these enhanced non-linearities is increased persistent neuronal activity in human occipito-temporal cortex.

Materials and Methods

The data presented in this paper were acquired from 15 different subjects. All subjects who participated in experiments 1 or 2 also conducted the meridian mapping experiment (see below) in order to define the borders of their retinotopic regions. All subjects had normal or corrected to normal vision and provided written informed consent to participate in the experiments. The Tel-Aviv Sourasky Medical Center approved the experimental protocol.

Magnetic Resonance Protocol

Subjects were scanned in a 1.5 T Signa Horizon LX 8.25 GE scanner equipped with a quadrature surface coil that covered the posterior brain regions (Nova Medical, Wake-field, MA). Blood oxygenation level dependent (BOLD) contrast was obtained with a gradient-echo echo-planar imaging (EPI) sequence (T_R 3000 ms; T_E 55 ms; flip angle 90°; field of view 24 × 24 cm; matrix size 80 × 80). The scanned volume included 17 nearly axial slices of 4 mm thickness and 1 mm gap. T_1 -weighted high resolution (0.93 × 0.93 mm in-plane) anatomical images were acquired at the same orientation as the functional images in order to allow accurate alignment with a three-dimensional spoiled gradient echo (SPGR) sequence (0.93 × 0.93 × 1.2 mm).

Data Analysis

fMRI data were analyzed with the BrainVoyager 4.9 software package (R. Goebel, Brain Innovation, Maastricht, Netherlands). The functional images were superimposed on two-dimensional anatomical images and incorporated into the three-dimensional data sets through trilinear interpolation. The complete data set was transformed into Talairach space (Talairach and Tournoux, 1988). Preprocessing of functional scans included three-dimensional motion correction and filtering out of low frequencies up to five cycles per experiment. Anatomical data from the 3D SPGR sequence were used to segment white/gray matter and reconstruct the cortical surface.

Activation Maps and Statistics

Statistical analysis was based on the General Linear Model (Friston *et al.*, 1995). Each experimental condition was modeled with a boxcar regressor (taking into account a 3 s hemodynamic lag). The signal time course of each individual voxel was then modeled as a linear combination of the different regressors (best fit using least-squares). The regressor coefficients were then used for *t*-test analysis comparing between different experimental conditions. The multi subject maps were obtained by *z*-normalizing the time course of individual subjects and using a random effect procedure (Friston *et al.*, 1999). Significance levels of maps were calculated taking into account the minimum cluster size and the probability threshold of a false detection of any given cluster (Forman *et al.*, 1995). This was accomplished by a Monte Carlo simulation [AlphaSim by B. Douglas Ward, a software module in Cox (1996)], using the combination of individual voxel probability thresholding and minimum cluster size of six contiguous voxels. The probability of a false positive detection per image was determined from the frequency count of clusters of the same size within the entire cortical surface (not including white matter and sub nuclei).

Stimuli

Stimuli were generated on a PC, projected onto a tangent screen positioned in front of the subject's forehead and viewed through a tilted mirror. In each experiment, the order of experimental conditions was counter balanced. A small fixation cross was present in the center of all visual stimuli and subjects were instructed to fixate during the experiments. Each experiment started and ended in a 21 s period in which a blank gray image was presented.

Experiment 1: Temporal Summation – House/Face

We used images of faces and houses to test sensitivity to stimulus presentation rate in face-related and house-related areas. Eight subjects participated in this experiment. Four experimental conditions were used – Houses 4 Hz, Houses 1 Hz, Faces 4 Hz and Faces 1 Hz – and each condition was repeated eight times during the experiment. In the Houses epochs, stimuli consisted of gray-scale photographs of houses and in the Faces epochs, stimuli consisted of gray-scale photographs of frontal views of faces. Each epoch lasted 6 s and was followed by a blank gray image lasting another 6 s before the next epoch began (Fig. 2A). Images in each epoch were presented for 250 ms in order to ensure image perception and avoid eye movements. In the 4 Hz conditions, images were presented consecutively (i.e. a total of 24 images per epoch). In the 1 Hz conditions, each image presentation was followed by a blank gray image lasting 750 ms (i.e. a total of six images per epoch; Fig. 2B). The images in the 1 Hz conditions were a subset of the images presented in the 4 Hz

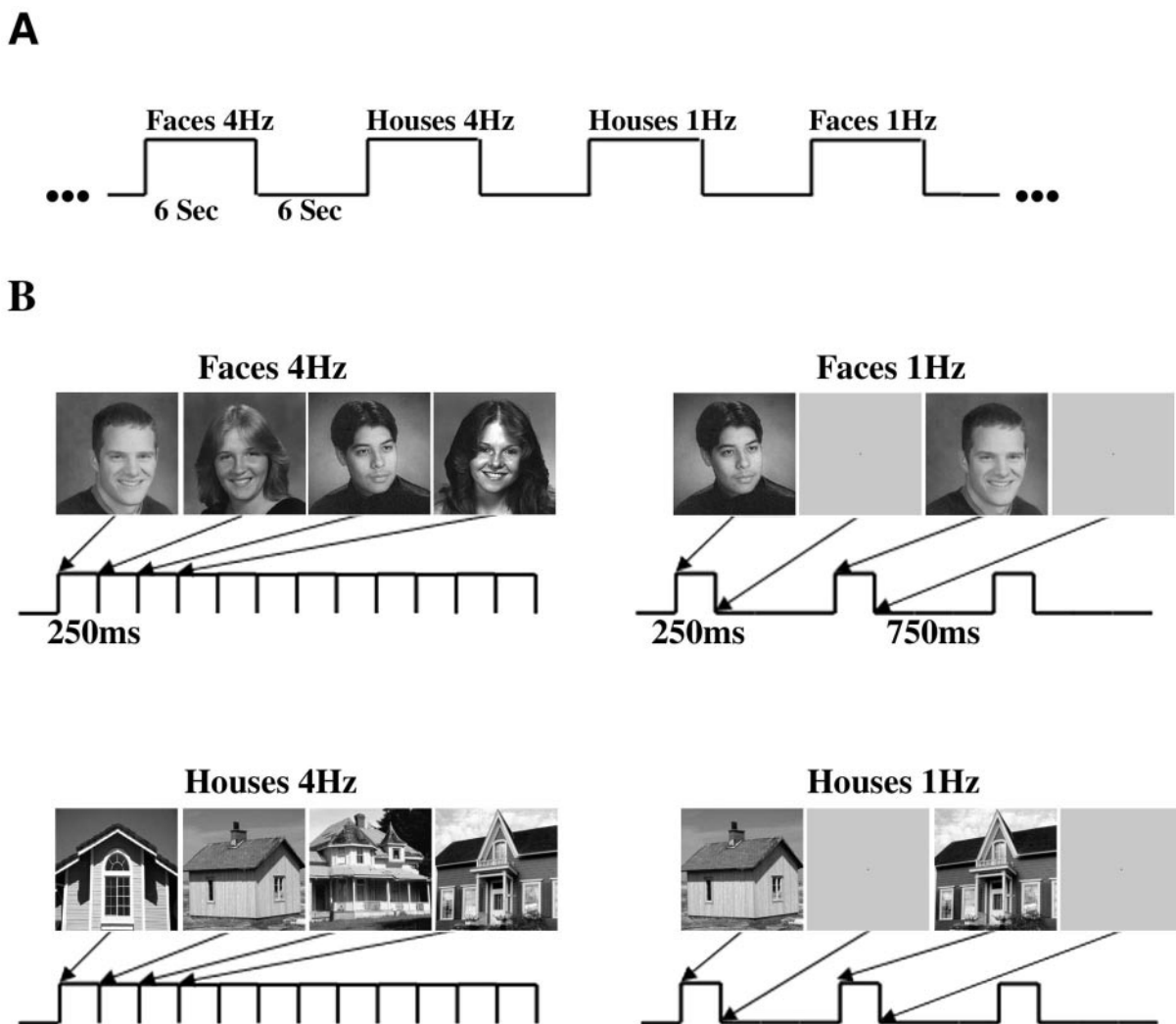


Figure 2. Experiment 1. (A) A scheme of the experiment's design. Visual epochs lasted 6 s and were separated by a blank gray image lasting 6 s. The order of the four experimental conditions was counter balanced. (B) Example of the four experimental conditions. Each image was presented for 250 ms. In the 4 Hz conditions, stimuli were presented in a consecutive manner (24 images in 6 s) and in the 1 Hz conditions each stimulus presentation was followed by 750 ms of a blank gray image (total of six images in 6 s). In each epoch, stimuli consisted of gray-scale images of either faces or houses.

conditions, and the order of the 1 Hz and 4 Hz conditions was counterbalanced. Stimuli covered 12° of the visual field and subjects' task was to covertly determine if the pictures are of a house or a face.

Experiment 2: Temporal Summation – Animals

Seven subjects participated in this experiment. Two experimental conditions were used – 4 Hz and 1 Hz (Fig. 5A). Visual stimuli consisted of gray-scale images of various animals including birds, mammals and fish which covered 12° of the visual field. Timing within epochs was the same as in the 4 and 1 Hz conditions of experiment 1 (Fig. 5B). Each experimental condition was repeated 13 times during the entire experiment. The subjects' task was to covertly name the animals in the pictures.

Experiment 3: Signal Saturation Control

Four subjects participated in this experiment. Four experimental conditions were used – epoch length was either 4 or 6 s and stimuli were presented at a rate of either 4 or 1 Hz. The 6 s epochs (4 and 1 Hz) were similar to experiment 2 (Temporal Summation – Animals). In the 4 s epochs, stimulus presentation manner was similar to the 6 s epochs with a total of 16 images presented in the 4 Hz epochs and four images presented in the 1 Hz epochs. Each experimental condition repeated four times during the entire experiment.

Experiment 4: Meridian Mapping

In order to delineate the borders of early retinotopic areas, we mapped the representation of the vertical and horizontal visual field meridians in all 15 subjects who participated in the other experiments (Engel *et al.*, 1994; Sereno *et al.*, 1995; DeYoe *et al.*, 1996). Stimuli consisted of triangular wedges, compensating for the expanded foveal representation. Wedges were either vertical (pointing up or down), or horizontal (pointing left or right) and consisted of either gray-scale natural images or black and white objects defined from textures. Subjects were instructed to fixate on a small central cross during the whole experiment. Each epoch lasted 18 s and was followed by a 6 s blank. During each epoch, pictures of the same orientation (up, down, left, or right) were presented for 250 ms in a consecutive manner. A total of four cycles per condition were shown.

Experiment 5: MT/V5 Localizer

Six of the eight subjects who participated in the House/Face temporal summation experiment (experiment 1) also participated in the MT/V5 Localizer experiment which was conducted in order to delineate the borders of the human motion sensitive area MT/V5 (Tootell *et al.*, 1995). Visual stimuli consisted of low contrast (1.5%), thin white rings, surrounding a small fixation cross, embedded in a black background (1.5 cycles/degree; duty cycle = 0.2). The experiment consisted of two conditions: Moving and Stationary. In the Moving

condition the rings expanded for 2 s and contracted for 2 s. In the Stationary condition the rings were presented every 3 s. Each trial lasted 18 s and was followed by a blank gray image for 6 s. Each condition was repeated eight times during the experiment. For each subject, area MT/V5 was defined by contrasting the Moving with the Stationary conditions.

Results

In order to test the sensitivity of different visual areas to stimulus presentation rate, we presented gray-scale images for 250 ms each, at a rate of either 1 or 4 Hz (Fig. 2A). In experiment 1, the visual stimuli we used consisted of either house images or face images (Fig. 2B).

First, we defined the different regions of interest according to their stimulus selectivity. It has been previously shown that the posterior fusiform gyrus (pFs, termed the fusiform face area, FFA), inferior occipital gyrus (IOG) and superior temporal sulcus (STS) are preferentially activated by face images compared with house images and various other objects (Kanwisher *et al.*, 1997; Ishai *et al.*, 1999; Hasson *et al.*, 2003). On the other hand the collateral sulcus (CoS) and transverse occipital sulcus (TOS) have preferential activation to house images compared with face images and various other object images (Epstein and Kanwisher, 1998; Maguire *et al.*, 2001; Hasson *et al.*, 2003). Using face and house stimuli in experiment 1 enabled us to determine the exact location of these regions in individual subjects before testing their sensitivity to stimulus presentation rate.

Depicted in Figure 3A is the map contrasting the Houses 4 Hz and Faces 4 Hz conditions on a flattened brain format. The 1 Hz conditions were excluded from the statistical step of defining the activated regions in order not to favorably bias regions with strong signals during the 1 Hz condition. Green voxels represent regions with stronger activation for the Houses 4 Hz condition compared with the Faces 4 Hz condition and red voxels represent regions with stronger activation to the Faces 4 Hz condition compared with the Houses 4 Hz condition. Both in the single subject (Fig. 3A, left) and the multi subject, random effect, averaged map (Fig. 3A, right) preferential activation to houses is clearly evident in the CoS and TOS and preferential activation to faces is clearly evident in the pFs, IOG and STS.

In order to map the level of sensitivity to stimulus rate across cortical regions, we created a two color activation map contrasting the 4 Hz conditions (both Houses and Faces) with the 1 Hz conditions (both Houses and Faces; Fig. 3B). This map was created by coloring all significantly activated regions ($P < 10^{-3}$ for $n = 1$ and $P < 0.03$ for $n = 8$) in blue or green according to the relative strength of the fMRI signal during the 4 and 1 Hz conditions. Brain regions with strong activation during the 4 Hz conditions compared with the 1 Hz conditions (i.e. closer to linear relationship with stimulus presentation rate) were colored green and brain regions with fMRI activations of similar strength during both conditions (i.e. having a highly non-linear response) were colored blue. Notice that early retinotopic cortex showed high bias to green colors (higher linearity), while high order visual areas show a higher tendency to blue colors (higher non-linearity) both in the single subject (Fig. 3B, left) and in the multi subject average map (Fig. 3B, right).

Another way to present the difference between retinotopic cortex and high order visual areas is depicted in Figure 3C.

Here we mapped regions showing significant preferential activation to the 4 Hz conditions compared with the 1 Hz condition in blue. Notice that these regions concentrate in early retinotopic regions and do not extend to high order visual areas. This restriction to more posterior regions was not due to lack of visual responsiveness in more anterior regions. This fact is demonstrated by coloring the visually responsive voxels (at the same statistical threshold as the 4 versus 1 Hz condition map) in red. As can be seen in the map of a single subject (Fig. 3C, left), and in the multi subject average map (Fig. 3C, right), the red voxels extended more anteriorly and encompassed high order visual areas, which were not revealed in the 4 versus 1 Hz contrast. Thus, the difference in fMRI signal between the 4 and 1 Hz conditions in high order visual areas was not statistically significant for these regions to be included in the 4 versus 1 Hz contrast test (blue voxels) although these regions did have significant visual activation at the same statistical threshold.

To obtain a quantitative estimate of the difference in activations, we sampled the time course of voxels in retinotopic and high order visual areas. High order visual areas were defined in each subject by using a selectivity test (Houses versus Faces and Faces versus Houses) in the 4 Hz conditions. The 1 Hz conditions were not used to define the regions of interest since we did not want to favorably bias regions with strong activations in these conditions. An example of such a map in a single subject is depicted in the left panel of Figure 3A. For the retinotopic regions and area MT/V5, we used a test of all visually active voxels (compared to rest) and sampled the time course in areas V1, V2, V3 and MT/V5. The extent and borders of retinotopic regions in each subject were defined by mapping the horizontal and vertical meridians in the Meridian Mapping experiment (see Materials and Methods for details) and area MT/V5 was defined by contrasting the Moving versus Stationary conditions in the MT/V5 Localizer experiment (experiment 5 in Materials and Methods).

The average time courses of retinotopic and high order visual areas are depicted in Figure 4A. Time courses are locked to stimulus presentation time (taking into account the hemodynamic lag) and averaged across eight subjects. Notice that while there was a large difference in signal amplitude between the 4 and 1 Hz conditions in retinotopic regions (V1, V2, V3) and area MT/V5, signal amplitudes in high order visual areas were quite similar for both the preferred (faces in pFs, STS, IOG and houses in CoS, TOS) and non-preferred stimuli. The difference in signal amplitude between the Houses 4 Hz conditions and the Faces 4 Hz conditions in high order visual areas was expected since these regions were defined by their selectivity in the 4 Hz conditions. However, signal amplitudes during the 1 Hz conditions were similar to the 4 Hz conditions although they were not included in the statistical step of defining the regions of interest.

It is important to note, that the statistical analysis of retinotopic and high order cortex was deliberately conservative as follows: in retinotopic regions we used a visual selectivity test which biases for voxels with strong activations also during the 1 Hz conditions. Therefore, the large difference in signal amplitude between the 4 and 1 Hz conditions is found in spite of the statistical test. On the other hand, in high order visual areas we used a selectivity test involving only the 4 Hz conditions (biasing for strong activations in the 4 Hz conditions) however the activations during the 1 Hz conditions are still high.

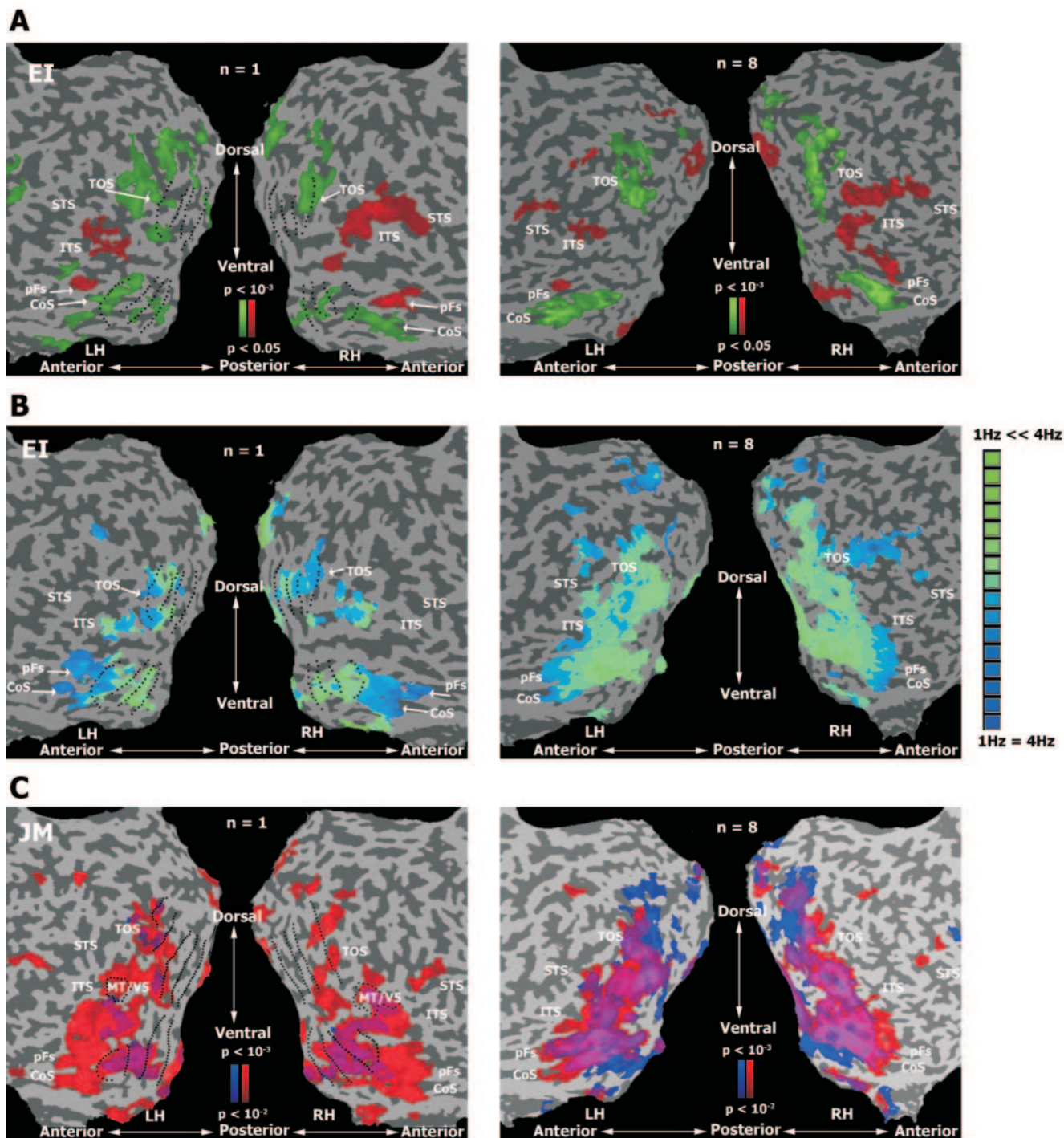


Figure 3. Experiment 1 — activation maps. All activation maps are presented on a flattened format, cut along the calcarine sulcus. Retinotopic borders of single subjects were mapped in a separate experiment (see Meridian Mapping experiment in Materials and Methods) and are plotted in black dotted lines. (A) Selectivity maps for houses and faces. Green voxels represent regions with stronger activation during the Houses 4 Hz condition compared with the Faces 4 Hz condition and red voxels represent regions with stronger activation during the Faces 4 Hz condition compared with the Houses 4 Hz condition. (B) Two color map representing the relative contribution of the 4 Hz (both Houses and Faces conditions) versus 1 Hz conditions. Visual activation ($P < 10^{-3}$ for $n = 1$ and $P < 0.03$ for $n = 8$) was colored according to the relative signal strength during the 4 and 1 Hz conditions. Green voxels represent 4 Hz \gg 1 Hz activations and blue voxels represent 4 Hz = 1 Hz activations. Note that high order object-selective cortex showed a tendency for equal 4 Hz and 1 Hz activations. (C) Superposition of a 4 Hz versus 1 Hz contrast map onto a visual activation map. Red voxels represent regions with visual activation, while blue voxels represent regions with stronger activation during the 4 Hz condition compared with the 1 Hz condition (both Houses and Faces) at the same threshold. Purple voxels represent the overlap between both maps. Note that the visual activation (red voxels) extends further anteriorly beyond the 4 versus 1 Hz contrast map (blue voxels). STS, superior temporal sulcus; pFs, posterior fusiform gyrus; CoS, collateral sulcus; ITS, inferior temporal sulcus; TOS, transverse occipital sulcus; LH, left hemisphere; RH, right hemisphere.

For each subject we also measured the maximal average percent signal change during the 1 Hz conditions and in the 4 Hz conditions and calculated their ratio (1 Hz/4 Hz). A low

ratio signifies a large difference in signal amplitude between the two conditions and a ratio of 1 indicates no difference in signal amplitude between the two conditions. The average

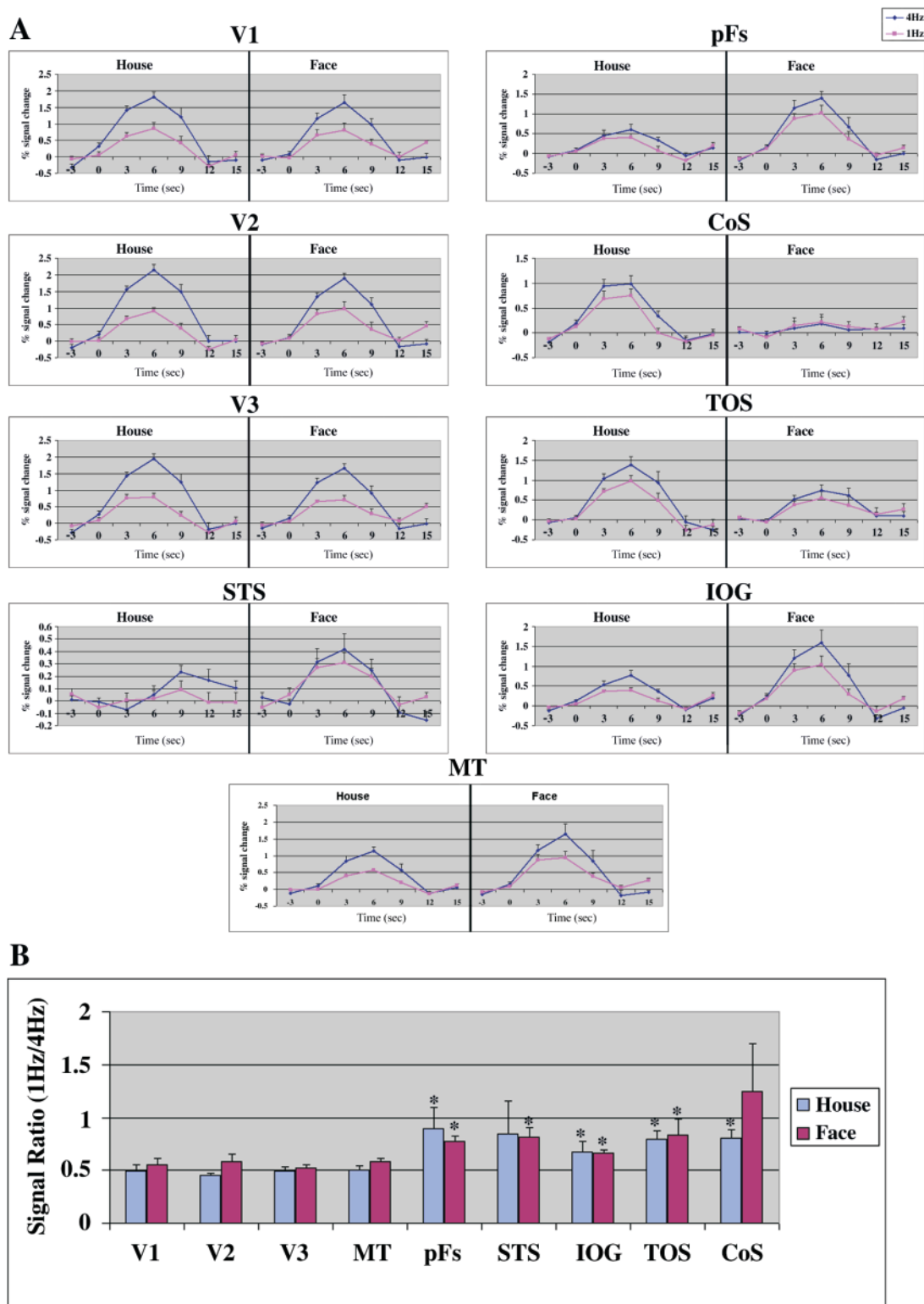


Figure 4. Experiment 1 — time course analysis. (A) Average signal time courses ($n = 8$ in visual areas and $n = 6$ in area MT/V5) across retinotopic and high order visual areas. The signals are time locked to the time of stimulus presentation (0 s) after taking into account the hemodynamic lag. Error bars indicate \pm standard error of the mean (SEM). Voxels in retinotopic regions (V1, V2, V3) and area MT/V5 were sampled using a visual activation test and the borders between regions were defined from the Meridian Mapping and MT/V5 Localizer experiments (see Materials and Methods). Voxels in high order areas were defined using an object selectivity test in the 4 Hz conditions (i.e. Houses 4 Hz versus Faces 4 Hz or Faces 4 Hz versus Houses 4 Hz). All activation clusters used to obtain the time courses had a significance of at least $P < 0.05$. (B) Average linearity ratio (calculated as maximal percentage signal change during the 1 Hz divided by the maximal percentage signal change during the 4 Hz conditions) in the Houses and Faces epochs across regions. A ratio of 1 signifies equal activation in the two conditions. Asterisks denote significance levels (paired t -test) of at least $P < 0.05$ between the high order and all retinotopic regions. The only exceptions were the linearity indices for faces in STS and IOG which were significant compared with V1 and V3 but failed to reach significance level with V2 ($P = 0.06$ and $P = 0.1$, respectively). Note that occipito-temporal cortex showed significantly higher non-linearity (larger ratio) compared to retinotopic areas and area MT/V5.

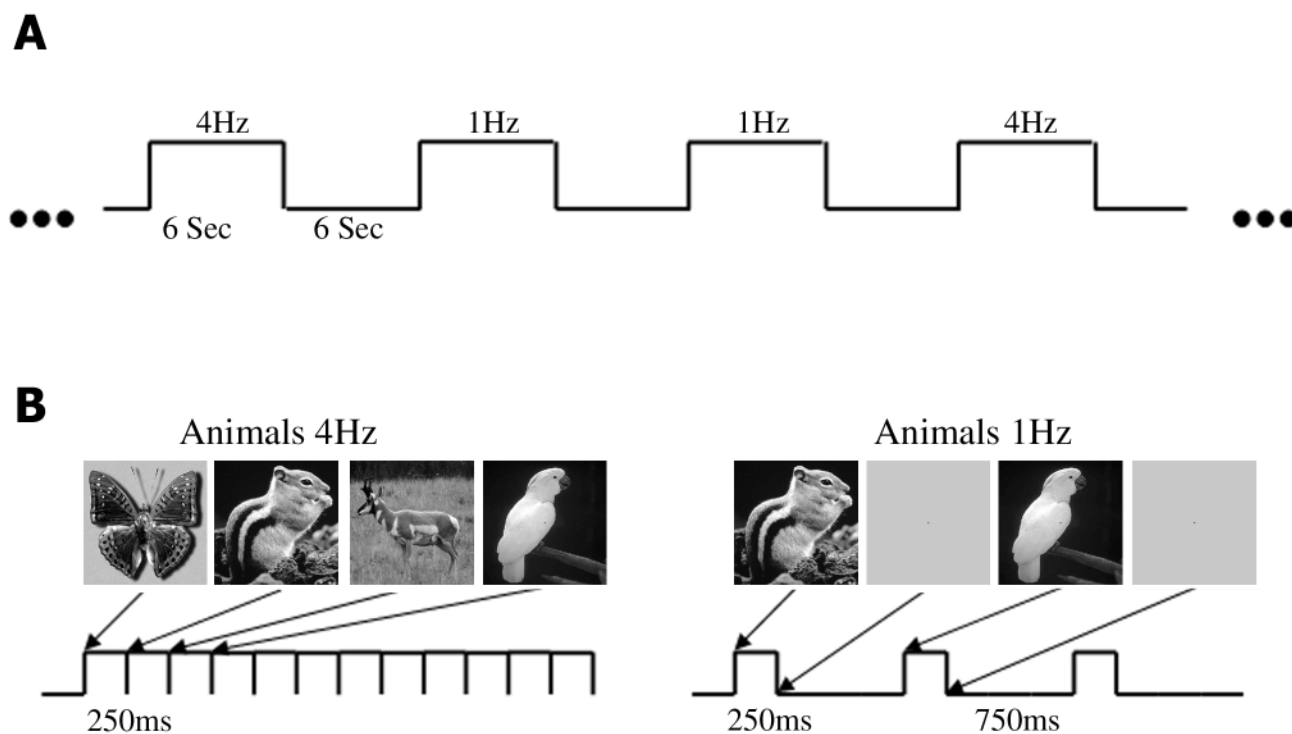


Figure 5. Experiment 2 — a scheme of the experiment's design. Stimuli consisted of gray-scale images of animals presented at either 4 or 1 Hz. Timing and epoch lengths were similar to those used in experiment 1 (see Fig. 2).

ratio of eight subjects in the different regions is depicted in Figure 4B. Ratios in retinotopic cortex were significantly smaller than in high order visual areas ($P < 0.05$ paired t -test between V1, V2 and V3 versus high order areas). It must be noted that the difference between high order and early visual areas may be even larger and that these numbers represent a lower bound on the non-linearities since in each region our scheme of voxel selection was biased against the non-linearity outcome.

While house images and face images evoke neural activity in more focused, anatomically distinct regions (e.g. collateral sulcus and fusiform gyrus), animal images evoke a more distributed pattern of activation. Therefore, in order to test whether the observed non-linearity is specific to house and face related regions, or more general in high order visual cortex, we used in experiment 2 gray-scale images of various animals and the same presentation rates (1 and 4 Hz) as in experiment 1 (Fig. 5). Similar to experiment 1, we created the two color map showing the relative contributions of the 4 and 1 Hz conditions (Fig. 6A). Green voxels represent brain regions with stronger fMRI signal during the 4 Hz conditions compared with the 1 Hz conditions and blue voxels represent brain regions with similar activations during both conditions. As in the previous experiment, retinotopic regions are mostly green, while high order visual areas are mostly blue — both in the single subject (Fig. 6A, left) and in the multi subject average map (Fig. 6A, right).

As in experiment 1, we mapped brain regions in which the fMRI signal was stronger during the 4 Hz condition compared with the 1 Hz condition (blue voxels in Fig. 6B). These regions included mostly retinotopic areas and some more anterior non-retinotopic regions. However, mapping all visually active voxels at the same statistical threshold included more anterior regions in high order visual cortex (red voxels in Fig. 6B). The

red voxels, outside the overlapping region, represent visually responsive voxels in which the difference in fMRI signal between the 4 Hz conditions and 1 Hz conditions was not statistically significant although these regions did have statistically significant visual activations.

To sample the signal time course in retinotopic areas we used a contrast of visually responsive voxels relative to baseline. Similar to experiment 1, the extent and borders of areas V1, V2 and V3 in each subject were defined externally from the Meridian Mapping experiment (see Materials and Methods). High order visual areas were defined using the same visual responsiveness contrast. In this experiment we sampled the pFs and IOG in high order visual areas since they showed particularly robust activation to the animal stimuli. As in experiment 1, retinotopic areas had much stronger activation during the rapid stimulus presentation (4 Hz conditions) compared to the low rate, while high order visual areas had similar signal amplitudes during the rapid and slow presentation rates.

We obtained a quantitative estimate of this difference by calculating a linearity index for each subject (maximal average percent signal change during the 1 Hz conditions divided by the maximal average percentage signal change during the 4 Hz conditions). The average ratio of seven subjects in the different visual areas is depicted in Figure 7. Ratios in pFs and IOG were significantly larger than ratios in retinotopic cortex (paired t -test $P < 0.01$).

While a likely source for the observed non-linearity is neuronal, a possible alternative for this interpretation could be a hemodynamic saturation or 'ceiling' effect. It could be argued that in areas manifesting similar signal activations in both the 4 and 1 Hz conditions we have reached the maximum hemodynamic range and therefore see no difference between

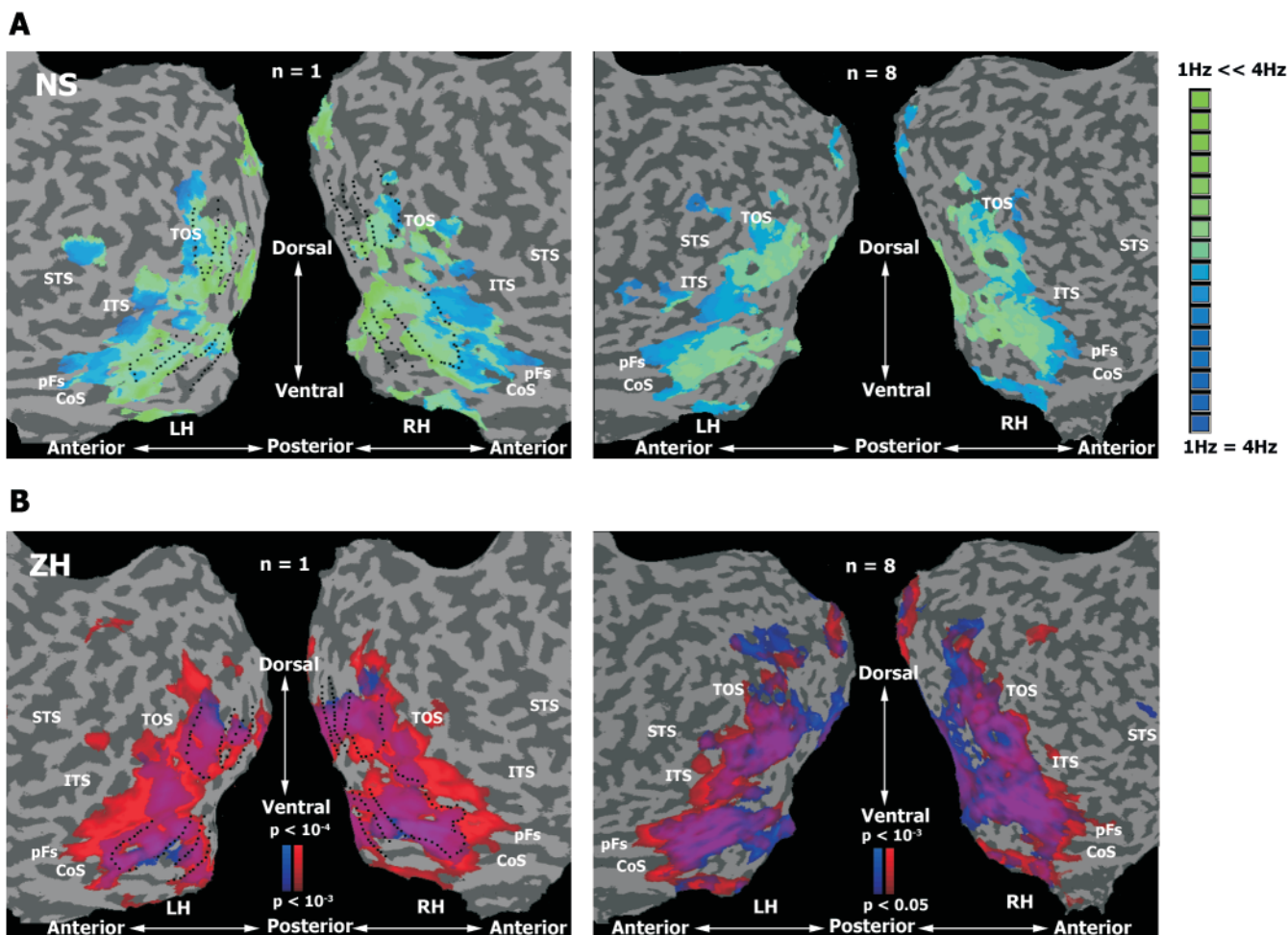


Figure 6. Experiment 2 — activation maps. (A) Two color map contrasting the 4 Hz with the 1 Hz condition. Visual activation was mapped [$P < 10^{-3}$ in both $n = 1$ (left panel) and $n = 8$ (right panel)] and color coded according to the same conventions as the map in Figure 3B. (B) Superposition of 4 versus 1 Hz contrast onto visual activation. As in Figure 3C, the red voxels represent the visual activation map and the blue voxels represent the 4 Hz versus 1 Hz conditions activation map. Purple voxels represent the overlap between both maps. Abbreviations as in Figure 4. Note that here too, occipito-temporal cortex showed a higher similarity between the 4 and 1 Hz activations.

conditions, even though the underlying neuronal activation might be different. While this option is possible, it is unlikely since high order visual areas (manifesting similar activations for both conditions) actually tended to have smaller signal amplitudes relative to retinotopic cortex (see time courses in Fig. 4A). Therefore, if anything, a hemodynamic ‘ceiling’ effect is more likely to occur in retinotopic cortex. This might explain the fact that the linearity ratios in retinotopic cortex were closer to 0.5 rather than the value of 0.25 – which would have been expected from a purely linear system when increasing presentation rate from 1 to 4 Hz. An alternative explanation for the somewhat non-linear responses in V1 might have been the use of stimuli, such as gray-scale object images, which may not be optimal for these areas compared to checkerboard stimuli (Boynton *et al.*, 1996; Dale and Buckner, 1997).

Additional evidence against a ceiling effect is the fact that non-optimal stimuli (e.g. Houses in the fusiform gyrus) which evoked smaller activations than optimal stimuli (and therefore were well within the dynamic range) had similar linearity index values as optimal stimuli (Fig. 4B).

However, it may still be possible that despite the small signal amplitudes in high order visual areas, the dynamic range of

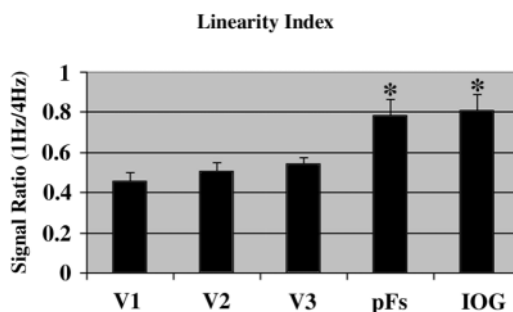


Figure 7. Experiment 2 — time course analysis. Average linearity ratio in the different regions. Regions of interest and borders between areas were defined similar to experiment 1 (see Fig. 4). Note the significantly higher linearity in retinotopic versus occipito-temporal cortex.

hemodynamic signals in these areas was also smaller. This could lead to saturated signals during the 4 Hz conditions and therefore result in similar hemodynamic signals during the rapid and slow presentation rates.

To further examine this possibility, we conducted experiment 3. In this experiment, stimuli were similar to experiment 2 (i.e. gray-scale images of animals); however, two epoch

lengths of either 6 or 4 s were used (Fig. 8A). We reasoned that if the non-linearity was due to the high signal amplitude during the 6 s epochs, then lowering this amplitude by shortening epoch length should uncover the linear response. Thus, if the activations to the 4 and 1 Hz conditions were similar due to saturation, we would expect an increased difference between them when the epoch duration was shortened.

As can be seen in Figure 8B, this was not the case. As expected, there was an increase in signal amplitude with increased epoch length. However, the average signal amplitude in the 4 Hz condition was similar to that in the 1 Hz condition both in the short and long epochs although the signal in the short epochs had further room for increase. This conclusion was substantiated by calculating the linearity ratio (1 Hz/4 Hz; Fig. 8C). A ceiling effect would predict small ratio values for the short epochs and larger values for the long epochs. As can be seen, the average ratios were similar in magnitude to those we obtained in experiment 2 and there was no significant increase in ratio values with longer epoch durations.

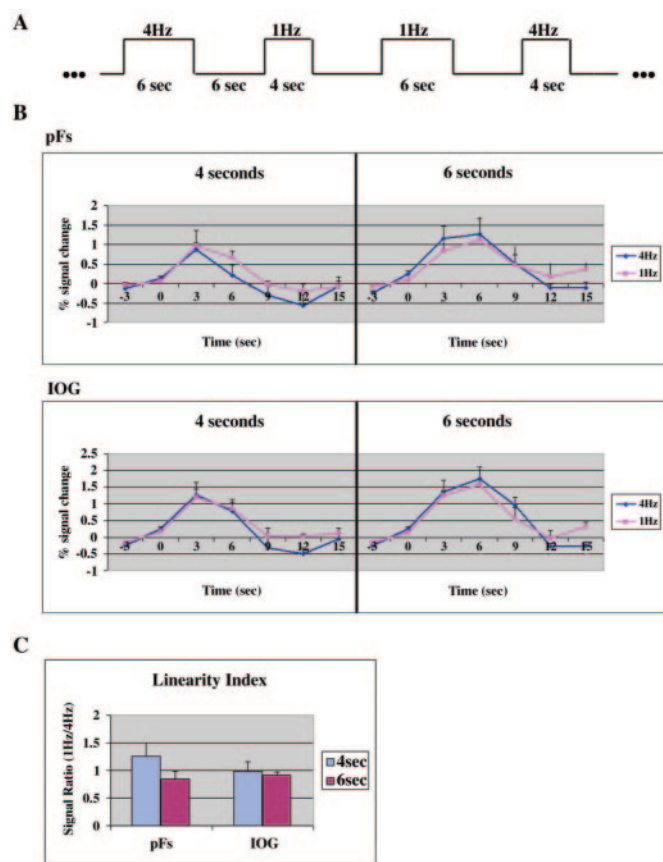


Figure 8. Control experiment — saturation effect. (A) A scheme of the experimental design of the experiment aimed to examine the possibility of a saturation effect as the source of the signal non-linearity. Four experimental conditions were used. The 4 and 1 Hz conditions were similar to experiment 2, but with the exception that two epoch lengths of either 4 or 6 s were used. In the 4 s epochs, 16 images were presented in the 4 Hz conditions and four images were presented in the 1 Hz conditions. (B) Average signal time courses ($n = 4$) in the 4 and 6 s epochs. The signals were time locked to stimulus onset (0 s) after taking into account the hemodynamic lag. (C) Average linearity ratio ($n = 4$) of the maximum signal strength during the 1 and 4 Hz conditions for the 4 and 6 s epochs. Note the similar level of non-linearity in the two epoch lengths, despite the lower activation amplitude in the short epoch durations.

Discussion

The main finding of the present study is the enhanced temporal non-linearity observed in human, high order occipito-temporal visual cortex, compared to early retinotopic visual areas. More specifically, a fourfold increase in stimulus presentation rate resulted in a significant increase in fMRI signal in all visual areas. While in early visual areas activation increased twofold, in occipito-temporal cortex activation increased by ~25%. Thus, we found that activation level in early visual cortex follows more closely the presentation rate (and image duration) of visual images, while high order visual cortex is significantly less affected by an increase in image presentation rate and duration. This conclusion is of course valid only for the presentation rates (4 and 1 Hz) employed in the present study.

What could be the source of this enhanced temporal non-linearity? An obvious component of the non-linearity could be saturation of the hemodynamic response which is not of neuronal origin. While this component cannot be ruled out completely, particularly for early retinotopic cortex, the robust, albeit lower signals in high order visual areas, as well as our control experiment using short and long stimulus durations make this possibility unlikely. In addition, the fact that the degree of non-linear activation for optimal and non-optimal stimuli was the same, implies that a saturated hemodynamic response could not be a major reason for our results. Furthermore, the fact that the non-linearity changes were matched topographically to the borders of the independently defined object-related cortex (Fig. 3) and area MT/V5 (Fig. 4) further supports the notion that at least part of the differential levels of non-linearities may be attributed to the unique neuronal functional processing occurring in different high order occipito-temporal cortex.

Considering specific neuronal mechanisms that could account for the temporal non-linearity, a likely mechanism involves sustained neural activity. Suppose that for the stimulus duration we used (250 ms), the stimulus evoked neural activity in human high order visual areas persists for ~800 ms. In this case, during the 1 Hz conditions we used, the neural activity would have continued throughout the ISI, leading to similar levels of activation during the 1 and 4 Hz conditions and thus strong non-linearity (see Fig. 1B). In contrast, stimulus-evoked activity which follows precisely the timing of the physical stimulus, should terminate during the 750 ms ISI in the 1 Hz conditions and will thus lead to a fourfold signal decrease in the 1 versus 4 Hz activations (see Fig. 1A).

A wealth of electrophysiological studies in experimental animals show sustained neural activity in macaque IT cortex (Miyashita and Chang, 1988; Rolls and Tovee, 1994; Kovacs *et al.*, 1995; Yakovlev *et al.*, 1998; Keyser *et al.*, 2001; Tamura and Tanaka, 2001). In the Miyashita study (Miyashita and Chang, 1988) this sustained increased activity even lasted for 16 s. However, in all these studies a delayed match to sample paradigm was used in which the animal explicitly had to retain the sample image in working memory. In our experiments, the subjects' task was simple object naming and did not involve any explicit memory component.

Human neuroimaging studies (Courtney *et al.*, 1997; Haxby *et al.*, 2000) used a delayed match to sample paradigm using face images and found strong delayed activity in prefrontal cortex and only weak delayed activity in extrastriate cortex. Using an *n*-back paradigm, it has been shown that the fMRI

signal in the fusiform gyrus increases directly with memory load (Druzgal and D'Esposito, 2001). In a recent study, Ferber and colleagues report greater sustained activity in LOC compared to area MT/V5 by using stimuli of objects defined from motion (Ferber *et al.*, 2003). Our results agree with their data and extend them to static object images. Furthermore, our data agree with their finding that MT/V5 shows different persistence properties relative to adjacent object areas and that in this regard area MT/V5 appears functionally closer to early visual cortex.

Finally, further evidence for sustained neural activity comes from human behavioral studies. When human subjects were presented with pictures for 110 ms, followed by a long ISI, their performance in a later recognition test was similar to that for pictures presented for longer durations. However, as the ISI was shortened, recognition performance dropped (Intraub, 1980) suggesting that additional processing is done during the ISI.

While this persistent activation mechanism appears most likely, it should be noted that our fMRI results are also compatible with an alternative, persistent effect which involves an inhibitory mechanism. Similar to the sustained activity mechanism, here the suggestion is that the stimulus evoked activity is followed by a refractory period (e.g. of 800 ms) during which the presentation of additional stimuli does not evoke neural activation (e.g. Ogawa *et al.*, 2000). In this case, the fMRI signal during the 4 Hz conditions in our experiment would be reduced since many of the pictures will be presented during the refractory period. Similar to the persistent activation, such an inhibitory effect will also tend to 'equalize' the signal amplitudes during the 4 and 1 Hz conditions and thus produce a strong temporal non-linearity.

The existence of cells manifesting sustained inhibition has been shown in macaque IT (Tamura and Tanaka, 2001). It should be emphasized that the suggestions of excitatory or inhibitory mechanisms are not exclusive of each other and one can envision, for example, that persistent excitation or inhibition will occur depending on the type of visual stimuli presented (Keysers and Perrett, 2002).

Another interesting question is whether the observed non-linearity was manifested only for the preferred stimuli, or whether it was a more general phenomenon. In the current study, we used gray-scale images of houses, faces and animals and measured the linearity index in brain regions selective for these stimuli. Our results show that although non-preferred stimuli (e.g. houses in the pFs) had lower signals (Fig. 4A), their linearity index was quite similar to that of the preferred stimulus (Fig. 4B).

The fact that both preferred and non-preferred stimuli showed similar temporal non-linearities is interesting and is compatible with our earlier suggestion that weak signals might stem from activation by highly selective, albeit few, neurons and does not necessarily imply the lack of optimal representational circuitry (Avidan *et al.*, 2002b). Furthermore, the fact that non-preferred stimuli (producing low amplitude fMRI signals) had similar linearity indices as the preferred stimuli, strengthens the hypothesis that the non-linearities we observed are of neural origin and not a mere hemodynamic saturation 'ceiling' effect.

Conclusion

The stronger temporal non-linearities we found in high order visual areas relative to retinotopic areas, along with the electrophysiological and neuroimaging literature, point to the existence of short-term persistence effects in high order visual areas. These effects can serve as a mechanism for short term 'iconic' memory, integrating stimuli which are coupled in time. Although we cannot distinguish between the excitation and inhibition mechanisms using the fMRI technique, both mechanisms point to the existence of continued neural processing in high order visual areas even when the evoking stimulus is no longer in view.

Notes

This study was funded by ISF 644/99, Israel Academy grant 8009 and the Dominic Center. The authors thank I. Goldberg for running the experiments and Y. Nir, U. Hasson and S. Gilaie-Dotan for comments on the manuscript.

Address correspondence to R. Malach, Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel. Email: rafi.malach@weizmann.ac.il.

References

- Amit DJ, Brunel N, Tsodyks MV (1994) Correlations of cortical hebbian reverberations: theory versus experiment. *J Neurosci* 14:6435–6445.
- Andrews TJ, Schluppeck D, Homfray D, Matthews P, Blakemore C (2002) Activity in the fusiform gyrus predicts conscious perception of Rubin's vase-face illusion. *Neuroimage* 17:890–901.
- Avidan G, Harel M, Hendler T, Ben-Bashat D, Zohary E, Malach R (2002a) Contrast sensitivity in human visual areas and its relationship to object recognition. *J Neurophysiol* 87:3102–3116.
- Avidan G, Hasson U, Hendler T, Zohary U, Malach R (2002b) Analysis of the neuronal selectivity underlying low fMRI signals. *Curr Biol* 12:964–972.
- Bar M, Tootell RB, Schacter D, Greve D, Fischl B, Mendola J, Rosen B, Dale A (2001) Cortical mechanisms specific to explicit visual object recognition. *Neuron* 29:529–535.
- Binder JR, Rao SM, Hammeke TA, Frost JA, Bandettini PA, Hyde JS (1994) Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex. *Brain Res Cogn Brain Res* 2:31–38.
- Boynton GA, Engel SA, Glover G, Heeger D (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207–4221.
- Courtney SM, Ungerleider LG, Keil K, Haxby JV (1997) Transient and sustained activity in a distributed neural system for human working memory. *Nature* 386:608–611.
- Cox R (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29:162–173.
- Dale AM, Buckner RL (1997) Selective averaging of rapidly presented individual trials using fMRI. *Hum Brain Mapp* 5:329–340.
- DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, Neitz J (1996) Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci USA* 93:2382–2386.
- Druzgal TJ, D'Esposito M (2001) Activity in fusiform face area modulated as a function of working memory load. *Brain Res Cogn Brain Res* 10:355–364.
- Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN (1994) fMRI of human visual cortex. *Nature* 369:525.
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601.
- Ferber S, Humphrey GK, Vilis T (2003) The lateral occipital complex subserves the perceptual persistence of motion-defined groupings. *Cereb Cortex* 13:716–721.

- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic-resonance-imaging (fMRI) – use of a cluster-size threshold. *Magn Reson Med* 33:636–647.
- Fox P, Raichle ME (1984) Stimulus rate dependence of regional cerebral blood flow in human striate cortex, demonstrated by positron emission tomography. *J Neurophysiol* 51:1109–1120.
- Friston J, Holmes A, Worsley K, Poline J, Frith C, Frackowiak R (1995) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2:189–210.
- Friston KJ, Holmes AP, Price CJ, Buchel C, Worsley KJ (1999) Multi-subject fMRI studies and conjunction analyses. *Neuroimage* 10:385–396.
- Grill-Spector K, Kushnir T, Hendler T, Malach R (2000) The dynamics of object-selective activation correlate with recognition performance in humans. *Nat Neurosci* 3:837–843.
- Hasson U, Hendler T, Ben Bashat D, Malach R (2001) Vase or face? A neural correlate of shape-selective grouping processes in the human brain. *J Cogn Neurosci* 13:744–753.
- Hasson U, Harel M, Levy I, Malach R (2003) Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron* 37:1027–1041.
- Haxby JV, Petit L, Ungerleider LG, Courtney SM (2000) Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *Neuroimage* 11:380–391.
- Intraub H (1980) Presentation rate and representation of briefly glimpsed pictures in memory. *J Exp Psychol: Hum Learn Mem* 6:1–12.
- Ishai A, Ungerleider LG, Martin A, Schouten HL, Haxby JV (1999) Distributed representation of objects in the human ventral visual pathway. *Proc Natl Acad Sci USA* 96:9379–9384.
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311.
- Keysers C, Perrett DI (2002) Visual masking and RSVP reveal neural competition. *Trends Cogn Sci* 6:120–125.
- Keysers C, Xiao DK, Foldiak P, Perrett DI (2001) The speed of sight. *J Cogn Neurosci* 13:90–101.
- Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. *Science* 293:1506–1509.
- Kovacs G, Vogels R, Orban GA (1995) Cortical correlate of pattern backward masking. *Proc Natl Acad Sci USA* 92:5587–5591.
- Lerner Y, Hendler T, Malach R (2002) Object-completion effects in the human lateral occipital complex. *Cereb Cortex* 12:163–177.
- Maguire EA, Frith CD, Cipolotti L (2001) Distinct neural systems for the encoding and recognition of topography and faces. *Neuroimage* 13:743–750.
- Mechelli A, Friston KJ, Price CJ (2000) The effects of presentation rate during word and pseudoword reading: a comparison of PET and fMRI. *J Cogn Neurosci* 12:145–156.
- Mendola JD, Dale AM, Fischl B, Liu AK, Tootell RB (1999) The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J Neurosci* 19:8560–8572.
- Miyashita Y, Chang HS (1988) Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331:68–70.
- Ogawa S, Lee T, Stepnoski R, Chen W, Zhu X, Ugurbil K (2000) An approach to probe some neural systems interaction by functional MRI at neural time scale down to milliseconds. *Proc Natl Acad Sci USA* 97:11026–11031.
- Ozuz B, Liu H-L, Chen L, Iyer MB, Fox PT, Gao J-H (2001) Rate dependence of human visual cortical response due to brief stimulation: an event-related fMRI study. *Magn Reson Imaging* 19:21–25.
- Potter MC (1976) Short-term conceptual memory for pictures. *J Exp Psychol: Hum Learn Mem* 2:509–521.
- Rolls ET, Tovee MJ (1994) Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proc R Soc Lond B Biol Sci* 257:9–15.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889–893.
- Stanley DA, Rubi, N (2003) fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. *Neuron* 37:323–331.
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. New York: Thieme.
- Tamura H, Tanaka K (2001) Visual response properties of cells in the ventral and dorsal parts of the macaque inferotemporal cortex. *Cereb Cortex* 11:384–399.
- Tanaka H, Fujita N, Watanabe Y, Hirabuki N, Takanashi M, Oshiro Y, Nakamura H (2000) Effects of stimulus rate on the auditory cortex using fMRI with 'sparse' temporal sampling. *Neuroreport* 11:2045–2049.
- Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 15:3215–3230.
- Yakovlev V, Fusi S, Berman E, Zohary E (1998) Inter-trial neuronal activity in inferior temporal cortex: a putative vehicle to generate long-term visual associations. *Nat Neurosci* 1:310–317.