Dynamics and non-linearities of the BOLD response at very short stimulus durations

B. Yeşilyurt¹, K. Uğurbil¹, ², K. Uludağ¹

¹Max-Planck-Institute for Biological Cybernetics, High-Field Magnetic Resonance Center Spemannstr. 41, Tübingen 72076, Germany.

²Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota Medical School, 2021 6th Street S.E., Minneapolis, MN 55455.

Keywords: BOLD, very short stimuli, nonlinearity, initial dip

Address for Correspondence: Barış Yeşilyurt
Max-Planck-Institute for Biological Cybernetics Spemannstr. 41 Tübingen 72076, Germany Phone: (+49 7071) 601 922 Email: baris.yesilyurt@tuebingen.mpg.de
ABSTRACT

In designing a functional imaging experiment or analyzing data, it is typically assumed that task duration and hemodynamic response are linearly related to each other. However, numerous human and animal studies have previously reported a deviation from linearity for short stimulus durations (< 4 seconds). Here, we investigated non-linearities of blood oxygenation level-dependent (BOLD) signals following visual stimulation of 5 to 1000 ms duration at two different luminance levels in human subjects. It was found that: i) a BOLD response to stimulus durations as short as 5 ms can be reliably detected; this stimulus duration is shorter than employed in any previous study investigating BOLD signal time courses; ii) the responses are more non-linear than in any other previous study: the BOLD response to 1000 ms stimulation is only twice as large as the BOLD response to 5 ms stimulation although 200 times more photons were projected onto the retina; iii) the degree of non-linearity depends on stimulus intensity, i.e. non-linearities have to be characterized not only by stimulus duration but also by stimulus features like luminance. These findings are especially of most practical importance in rapid event-related fMRI experimental designs. In addition, an ‘initial dip’ response -- thought to be generated by a rapid increase in cerebral metabolic rate of oxygen metabolism (CMRO\textsubscript{2}) relative to cerebral blood flow (CBF) -- was observed and shown to co-localize well with the positive BOLD response. Highly intense stimulation, better tolerated by human subjects for short stimulus durations, causes early CMRO\textsubscript{2} increase, and thus the experimental design utilized in this study is better for detecting the initial dip than standard fMRI designs. These results and those from other groups suggest that short stimulation combined with appropriate experimental designs allows neuronal events and interactions to be examined by BOLD signal analysis, despite its slow evolution.
INTRODUCTION

Blood oxygenation level-dependent (BOLD) [1] signal assessment using functional magnetic resonance imaging (fMRI) measures neuronal activity indirectly by recording combined changes in metabolism and hemodynamics accompanying neuronal activity [2-4]. Experiments simultaneously measuring fMRI or optical spectroscopy and intra-cortical recordings or electroencephalography document a close link between neuronal activity and the hemodynamic response [5-8], (for a review, see [9]). However, BOLD signal changes which take place over seconds are very slow compared to the underlying neuronal events which occur in milliseconds. Therefore, it has generally been accepted that the BOLD approach cannot temporally resolve alterations in neuronal activity in the range of milliseconds. However, using two-pulse stimulation of brief duration (310 µs in rats and 10 ms in humans) with different inter-stimulus intervals up to 1000 ms, Ogawa et al. showed that reduction of the neuronal response to the second pulse as assessed by somatosensory-evoked potential is in fact reflected by the amplitude of the BOLD signal [5]. That is, this experiment showed that using stimuli of very short duration, combined with an appropriate experimental design, the hemodynamic response can decode neuronal events and interactions on a millisecond time scale.

As a result of advances in MRI hardware and experimental approaches over the last decade, the duration of stimulus presentation evoking a measurable BOLD signal in humans has decreased. Thus, in one of the first fMRI studies, Blamire et al. reported a detectable BOLD response to visual stimulation lasting 2 s [10], whereas Bandettini and colleagues later demonstrated a BOLD response to stimulation of only 500 ms duration [11]. Robson et al., using a 100 ms auditory stimulus [12] and Fransson et al. using a 200 ms visual stimulus [13] have further decreased the lower limits of detectability in the auditory and in visual cortex, respectively. More recently, Pfeuffer et al. have also used a visual flash stimulus
of 100 ms duration to investigate the spatial dependence of non-linearity of the BOLD response [14]. Implicitly, many other studies, e.g. somatosensory stimulation using laser light [15], have also used very short stimulus duration. In most of these studies, however, the focus was not on exploring the hemodynamic response but on mapping function to brain areas.

Most of the aforementioned studies did not show the dynamic behavior of the BOLD response. The shortest stimulus duration in humans resulting in a BOLD time course is in the study of Savoy et al. which, however, was only presented in abstract form [16]. These investigators compared responses to visual stimulus durations of 34, 100 and 1000 ms (i.e. multiples of screen frame rate) and demonstrated a detectable BOLD response at all times. In addition, the responses were found to be temporally highly non-linear.

In most fMRI studies a linear, time-invariant transformation of stimulus design to BOLD response has been assumed, in order to simplify the analysis and interpretation of fMRI data. In humans, BOLD signal linearity in the visual cortex has been proposed to be valid for stimulus durations longer than ~4s [17-20]. Namely, the amplitude and the duration of the BOLD response for stimulus durations longer than ~4s were predictable from the response measured at ~4s. Although, response characteristics were possibly different across cortical areas and stimuli used. In contrast, the shorter the stimulus duration, the greater the observed non-linearity [12,14,18-23]. That is, below a certain threshold, the BOLD response to short duration stimuli cannot be used linearly to predict the response to long stimuli. Typically, the hemodynamic response to long stimulation is overestimated by responses to short stimulus durations.

The balloon and Windkessel models [24,25] suggest that vascular non-linearities stem from the mismatch between cerebral blood flow (CBF) and blood volume (CBV) due to delayed compliance of CBV and are minimized the shorter the stimulation is. Therefore, non-linear behavior of the BOLD signal for short
stimulus durations should mostly originate from neuronal activity which, however, has to be proven by simultaneous hemodynamic and electrophysiological measurements. In explaining the non-linearities seen in BOLD responses at short stimulus durations, Boynton et al. have theoretically stressed a possible role of neuronal adaptation effects [17]. That is, the neural activity is transiently greater at the beginning of stimulation and declines to a plateau after a few seconds [6]. Thus, short stimulus responses, implying no or little adaptation effect, would overestimate the longer stimulus responses, which possibly allow for more adaptation. In addition, neurons serve as non-linear amplifiers of physical stimuli for very short stimulus durations.

In the present study, we have explored the BOLD signal time course and non-linearities using ultrashort stimulus durations (5 milliseconds to 1 second) in humans. We show that a) BOLD responses to stimulus durations as short as 5 ms can be reliably detected in humans; b) BOLD responses to different stimulus durations are highly non-linear and c) non-linearities depend not only on stimulus duration but also on stimulus characteristics like luminance. Previous results on non-linearities suffer from this confounding factor which was not accounted for earlier. Additionally, the so-called ‘initial dip’ was found most reproducibly with 1000 ms high intensity stimulation. Furthermore, most voxels with the strongest initial dip have been found to be followed by a positive BOLD response and to overlap with the voxels with the strongest positive BOLD response. A preliminary description of these results has been previously presented in abstract form [26].

**MATERIALS AND METHODS**

**Experimental Set-Up and Design**

Custom-built non-magnetic white LED (light emitting diode) goggles were used for the stimulation of the entire visual field. The goggles were switched on and off in real-time by a data-acquisition card with DASYLAB as controlling software (Microstar Laboratories, Inc., Bellevue, WA, USA; Measurement Computing Co.
Middleboro, MA, USA). We measured the accuracy of the stimulus duration inside the MRI scanner using a 100 MHz oscilloscope (Tektronix, Inc., Beaverton, OR, USA) and found that the LED goggles reached predetermined signal intensity and also were switched off in less than 3 µs. We used eight different experimental paradigms with stimulus durations (SD) of 5, 50, 250, and 1000 ms and low and high stimulus intensities (SI) (high luminance ~ 1430 cd/m², low luminance ~ 440 cd/m² ~ 30% high luminance). Subjects were instructed to maintain fixation on the center of the goggles during the functional runs.

Following 45s of onset baseline period, eight visual stimuli of the same intensity and duration were presented about every 30s (Figure 1). The inter stimulus interval was randomized between 28 and 32s to avoid coherent summation of vascular oscillations. Each run ended with a baseline period of > 45s resulting in a total duration of each experimental run of 310s. 11 healthy human volunteers (age range 24-36, 3 females) participated in the study. Prior to scanning, informed written consent was obtained from the subjects according to the guidelines of the local ethics committee of University of Tübingen.

**Data Acquisition**

All measurements were performed on a 3T Siemens Trio (Erlangen, Germany) scanner using an 8-channel head coil. Anatomical high-resolution T1-weighted MPRAGE images (TR=2s; TE=4.38 ms; FA=90° FOV=256 mm; voxel-size=1x1x1m³) were acquired. Functional measurements were made using T2*-weighted gradient echo planar imaging sequence with TR = 1s, TE = 40 ms and FA = 90°. A total of 12 slices with slice thickness of 3 mm and in-plane resolution of 3 x 3 mm² were obtained.

**Data Analysis**

Data analysis was performed using FSL software (http://www.fmrib.ox.ac.uk/fsl) [27] and self-written MATLAB (The MathWorks, Natick, MA, USA) routines. Time
courses of all functional volumes were motion-corrected with MCFLIRT [28] using the mean volume of the corresponding run as reference. Additionally, all runs were registered to the run with SD = 1000 ms and SI = High, which had the highest signal to noise (SNR) ratio. Data were high-pass filtered (cutoff was 50s) to remove slow varying drifts. No additional detrending or spatial smoothing was applied.

In order to determine optimal hemodynamic response functions for statistical analysis, the FSL toolkit FLOBS [29] was utilized. Briefly, randomly sampling each delay and height parameter from a range specified by the user, the toolkit generates various hemodynamic response functions and uses principal component analysis to create an optimal basis set that maximally spans the space of the generated samples. Using FLOBS, three basis functions were generated (Figure 2a). Statistical maps were created finding the best linear combination of these basis functions for each voxel. With this approach, differently shaped hemodynamic response functions (i.e. with and without initial dip; different amplitudes of post-stimulus undershoot) can be fitted with approximately equal statistical accuracy, i.e. with this procedure the bias towards a hemodynamic response function with a pre-defined shape was reduced.

Based on the functional map of 1000 ms and a high intensity stimulation experiment, a region-of-interest (ROI) containing the 50 most activated voxels (voxels with the highest F values) was generated. The BOLD signal time courses in this ROI were averaged. Because functional runs of all experiments are co-registered, choosing the same ROI ensures that the same voxels are encompassed in all functional runs. Stimulus-locked single trial responses for each subject and experimental run were determined and averaged across subjects without any BOLD signal amplitude normalization. The response amplitudes were calculated from the peak of BOLD response intensity. The response integral was defined as the product of the amplitude and the full-width-
at-half-maximum (FWHM). Alternatively, the integral of the response was defined as the sum of positive values (see Figure 2b).

Initial dip was most reliably detected for SD = 1000 ms and high stimulus intensity (see results). To compare the spatial distribution of the initial dip response with the spatial distribution of the positive BOLD response under these conditions, functional maps were created individually for both responses. For the generation of the initial dip statistical map, a Gaussian function ($\sigma = 0.8s.$ and mean lag = 2s) and its first derivative were used in a general linear model (GLM) as explanatory variables. For improving SNR and thereby detectability of initial dip responses, a spatial smoothing filter of 5 mm was applied. Using a standard gamma variate function with common parameters and its first derivative as explanatory variables, the same procedure was executed for generating statistical maps of positive responses. On both statistical maps, z-values were chosen to yield 50 voxels (those which have the greatest z-values). Again, independent ROIs containing these voxels were generated both for the initial dip and positive responses, for further analysis.

**RESULTS & DISCUSSION**

Figure 3 shows a representative activation map (SD = 1000 ms, SI = High) superimposed on an anatomical image showing that early visual areas around the calcarine sulcus are activated by the stimulus. The average BOLD responses (n = 11, 50 voxels) for SD = 5, 50, 250 and 1000 ms, and for SI = low and high are shown in Figure 4a and b. Reliable BOLD responses in the early visual cortex were detected in all subjects under all stimulus conditions. Average BOLD responses exhibit an ‘initial dip’ followed by a positive response peaking between 5 and 7s. A small or negligible post-stimulus undershoot is observed. The initial dip and positive BOLD response have larger amplitudes for high stimulus intensity. The amplitudes of BOLD signal intensity changes elicited by the 5 ms stimulus were $1.07 \pm 0.39\%$ at high stimulus intensity and $0.92 \pm 0.24\%$ at low
intensity; those elicited by 1000 ms stimulus duration were 2.36 ± 0.64% and 1.66 ± 0.27%, respectively (see Table 1a).

In order to probe the response amplitude linearity of FWHM as well as of the response integral, we normalized all response values to the value of the experiment with SD = 1000 ms at both stimulus intensities (Figure 5 a-c). The response amplitude and integral exhibited a statistically significant monotonic increase (between successive points, p < 0.05). The amplitude of the BOLD signal intensity to the 5 ms stimulus was about 47% of the amplitude of the response to a 1000 ms stimulus at high intensity and 55% at low intensity. The non-linearity of the integral showed a similar pattern to that of the amplitude. Note that the amplitude of the BOLD response for 1000 ms stimulation is only twice as large as for 5 ms stimulation, although the stimulus duration is 200 times longer.

The BOLD response typically exhibits a temporal nonlinearity, such that an appropriately shifted and added response to a brief stimulus over-predicts the true response to an extended stimulus [14,17,20,22,23,30-32]. This temporal nonlinearity is most pronounced when the brief stimulus is less than about 4s and the extended stimulus is longer than 6s. Usually, the degree of non-linearity differs across cortical areas. It has been suggested that non-linearity is less prevalent in the primary visual cortex than other primary cortical areas. Soltysik et al. showed that the peak amplitude of the BOLD response to a 16s stimulus is about 4, 3 and 2 times smaller than predicted on the basis if a 1s stimulus in auditory, motor and visual cortices, respectively. Even within the same cortical area, spatially heterogeneous non-linearities have been reported by Birn et al [22]. They found a 4-fold larger response amplitude and even 6-fold larger response area to 250 ms visual stimulation than predicted from a linear model, ascertained by extrapolation of the responses linearly from a 20s stimulus response. Moreover, the degree and behavior of the non-linearities are also field strength-dependent [14]. Pfeuffer et al. reported a deviation of ~250% in
response area at 200 ms stimulus duration, when predicted using the convolution product of stimulus duration and a Gamma function defined by Cohen et. al. [33], at 4T, whereas the deviation was only about 40% at 7T. In all previous studies using short stimulus durations in humans, non-linearities were much lower than observed in the current study.

In addition, it is apparent that non-linearities depend not only on stimulus duration but also on stimulus characteristics like luminance. That is, non-linearities in BOLD responses must be characterized as a function not only of stimulus duration, as it is usually done, but also of stimulus characteristics.

What is the origin of the non-linearities observed for short stimulus durations? BOLD signal non-linearities with respect to stimulus duration might arise 1) from the transition of the physical stimulus detected at the retina to neuronal signaling in the early visual cortex, and 2) at the vascular level. Regarding the first possibility, previous studies have reported non-linearities during stimulus information transfer from retina to primary visual cortex. Already in the retina the responses are broadened in time compared to the physical stimulus, as already shown in turtle and monkey photoreceptor cells [34],[35]. Further evidence for nonlinearities regarding the temporal dynamics of cell responses has been found in retinal ganglion cells. Electrophysiological measurements on cat ganglion cells revealed that decreasing the stimulus duration decreases the duration of the ganglion cell response if the presented stimulus is longer than about 32 ms; further decreases in stimulus duration only reduce the number of evoked spikes at less than 32 ms. That is, there is minimum response duration of 50-70 ms independent of the brevity of the light flash [36]. Levick et al. reported that increasing stimulus intensity serves to increase the magnitude of the response but had little or no effect on its duration. The point to be made is that there is definite minimum ganglion cell response duration regardless of the duration of the stimulus. However, the response amplitude is still able to encode (non-linearly) the duration and the intensity of stimulus.
Boynton et al have suggested that the source of nonlinearity seen in brief stimuli might be due to a neural adaptation effect [17] because the neural activity is transiently greater at the beginning of stimulation and reaches a plateau after a few seconds. The transient overshoot of neuronal activity leads to over-prediction of the neuronal and, as a consequence, also of the hemodynamic response to longer stimuli. The initial overshoot of neuronal activity is a common observation in electrophysiological studies. For example, Logothetis et al., measuring the BOLD signal, local field potentials (LFP), multi unit activity (MUA) and single unit activity (SUA) simultaneously showed that there is a transient high LFP signal at the beginning of visual stimulation [6,37].

The FWHM of the responses became wider with increasing stimulus duration as well as with higher intensity (Figure 5c). At high intensity, the width of the 5 ms and 1000 ms response was ~4s and ~4.8s, respectively, about a 1s difference; and at low intensity these values were ~3.5s and ~4.5s, respectively, the same difference (see Table 1b). This increase in FWHM is particularly prominent between 5 and 250 ms. Using a canonical gamma-variate impulse hemodynamic response function, the FWHM of the BOLD responses are predicted to be ~ 100 ms broader for a 1000 ms short stimulus compared to 5 ms stimulation. However, the broadening experimentally found is ~ 800 ms for high and low stimulus intensity. Interestingly, FWHM for high intensity stimulation was larger than for low intensity stimulation for the same stimulus durations. That is, both findings argue that the broadening of FWHM is due to sustained neuronal activity.

Regarding the second possibility mentioned above, by simultaneously recording BOLD and visual evoked potentials (VEP) Janz et al. showed that neural adaptation alone cannot cause the observed nonlinear behavior of the BOLD response [37]. They suggested a further nonlinear coupling of the changes in CBF, CBV, and cerebral metabolic rate of oxygen metabolism (CMRO₂)
underlying the BOLD response. Furthermore, Gu et al. have found in their simultaneous CBF, CBV and BOLD measurements that the BOLD signal showed a stronger degree of nonlinearity. That is, there are additional nonlinear effects in the signal transduction step from CBV/CBF to BOLD.

Miller et al. have simultaneously measured the CBF and BOLD response in humans [23]. Both responses were non-linearly related to stimulus duration. However, the BOLD response exhibited greater deviation from linearity than CBF responses. Based on these results, using the balloon model, they suggested that the steps from stimulus to neuronal activity and from CBF to BOLD response are nonlinear but the step from neural activity to CBF is linear. Friston et al. suggested that a large part of the nonlinearity of the BOLD response arises from the transformation of a CBF response to the BOLD signal change [30]. In a subsequent study they showed that the Volterra kernel characterization of experimentally observed nonlinearities could be accounted for with the balloon model and a linear transformation up to the CBF response, again supporting the idea that the primary nonlinearity is in the transformation from the CBF to the BOLD response [38].

The balloon and Windkessel models suggest that the BOLD signal non-linearity increases with stimulus duration [25,39,40]. As proposed by these models, venous structure changes closely follow the time course of CBF during stimulation. However, they revert to baseline more slowly than CBF after stimulation due to the lack of active smooth muscle control of venules and veins. The longer the stimulus duration, the larger the mismatch of expansion and reversion time constants of venous CBV and the larger the discrepancy from a linear model.

In summary, the following picture is obtained for the non-linearities observed in BOLD response characteristics for short stimulus durations: At the retina, the duration of the neuronal activity projecting to the visual cortex via LGN and other
subcortical pathways is broadened with respect to the physical stimulus duration. This broadening is relatively more pronounced the shorter the stimulation is. As a consequence, although 200 times more photons reach the retina during the 1000 ms than 5 ms stimulation, the input to the visual cortex differs much less and therefore only a ~50% smaller BOLD response is observed for the latter. As discussed above, further non-linearities might arise from the translation of the neuronal activity into the hemodynamic response and into the BOLD signal although this would only add a small fraction to the overall observed non-linearities.

All average BOLD responses exhibited a significant initial dip prior to the larger signal increase with maximal deviation about 2s after stimulus onset (t-test; p<0.05) (spatial distribution of the initial dip compared to the positive BOLD response, see below). The so-called ‘initial dip’ shortly after stimulus onset is thought to be caused by an early increase in CMRO₂ compared to CBF (for an overview, see [41,42]) in optical imaging [43]. The effect is small and not always present, but it has stirred interest because it may be better localized to the area of increased metabolism (that is, the CBF and CBV increases may cover a wider area).

For a 1000 ms stimulus duration at high and low intensity the initial dip values were -0.50 ± 0.26% and -0.31 ± 0.13%, respectively. In Figure 6, it is shown that the region where the initial dip occurs overlaps largely with the region of the maximal positive response, displayed on an average amplitude map of one subject in the time interval 0-2s and 6-8s. (here we used clustering to eliminate single voxel large amplitude occurrences, merely for display purposes).

In Figure 7a, the distribution of z-values of the 50 voxels over all subjects with largest z-values for the initial dip is shown (because the initial dip was detected most reliably for SI = high and SD = 1000 ms, we present the results only for this). For each subject, the z-values for the positive BOLD response for the 50
most significant initial dip voxels were calculated. The dashed line indicates the z-value corresponding to $p = 0.01$ resulting in 82% of voxels being statistically significant at $p<0.01$. 91% of these voxels exhibit a significant positive response at $p<0.01$ (Figure 7b), showing that in most of these voxels with an initial dip, a significant positive BOLD response was also observed. For comparison, the $z$-values for the voxels with the largest positive response are shown in Figure 7c. In addition, it was found that 51% of the voxels with the greatest positive response $z$-values also have the greatest initial dip $z$-values.

Because the largest BOLD responses should be in venous structures, only a partial overlap of the greatest initial dip and positive BOLD responses is expected. A second alternative interpretation for the partial overlap is that the low SNR of the initial dip yields false positives, thereby reducing the overlap with the positive response. However, this question cannot be resolved conclusively at the spatial resolution used in human studies. In Figure 8, the average time course ($n = 11$, 50 voxels) for ROIs based on the initial dip, positive response and overlap of both are shown. For all of these ROIs, an initial dip followed by a positive response and a small post-stimulus undershoot are observed. The ratio of the maximum positive response to the maximum magnitude of the initial dip is $\sim 3$ for the initial dip-ROI and $\sim 10$ for the positive response-ROI.

The early rise of CMRO$_2$ is more easily detectable following strong stimulation. Because high intensity stimulation is better tolerated by human subjects (and conscious animals) the shorter the stimulus duration is, we could present visual stimulation at higher intensity than typically employed in visual stimulation experiments. That is, the transient uncoupling between CBF and CMRO$_2$ will be more pronounced by an ultrashort or short high intensity stimulation paradigm (up to few seconds), suggesting that such a paradigm could be ideal for detection of the initial dip as demonstrated by these results.
Only a small or even negligible post-stimulus undershoot was observed. However, preliminary data in an ongoing follow-up study show that the BOLD response can exhibit a post-stimulus undershoot also for ultrashort stimulation. Whether this discrepancy is due to physiological inter-subject variability or to signal-to-noise characteristics of the different scanners used remains to be investigated.

In humans, the 5 ms stimulus is shorter than employed in any previous study examining the characteristics of the BOLD response. The results presented in the present study suggest that there is no detectable minimal hemodynamic response down to 5 ms stimulus duration. The possibility of a minimal hemodynamic response was suggested by Logothetis et al. and Pfeuffer et al. [6,14]. In the former study, it was found that if normalized BOLD response was plotted as a function of normalized local field potentials (LFP) or multi-unit activity (MUA), at the limit of zero LFP or MUA, the BOLD signal was still ~50%. In the latter study, the BOLD response was found to be linear with stimulus duration with, however, also a non-zero intercept at the limit of zero stimulus duration. The data presented in the current study are inconclusive regarding the existence of a minimal hemodynamic response, because stimulation durations shorter than 5 ms were not studied. Therefore, we are currently performing a study using even shorter stimulus duration down to 0.1 ms and measuring both the BOLD signal and visual evoked potentials (VEP) to explore whether a minimal hemodynamic response exists.

**CONCLUSION**

In this study, non-linearity of the BOLD response to ultrashort visual stimulation (5 – 1000 ms duration) was investigated in human subjects. The BOLD responses were found not to be linear time-invariant transformations of stimulus duration as is usually assumed in fMRI analysis. Practically, this deviation from linearity has the most impact in rapid event-related experiments. In addition, a
hemodynamic response was reliably detected to visual stimulus duration as short as 5 ms, shorter than in any other study examining the characteristics of the BOLD response. Using even shorter stimulus duration, future studies will address whether a minimal hemodynamic response exists or not. Averaged BOLD signal time courses exhibited a significant initial dip response. The location of detected initial responses overlapped largely with positive responses, showing that the largest early \( \text{CMRO}_2 \) increase occurs at the same location as the local increase in CBF, at least at the spatial resolution used in human fMRI studies. Vascular non-linearities are minimized at very short stimulus duration and therefore the major contribution to the observed non-linearities is neuronal in origin. Further experiments simultaneously measuring hemodynamic and neuronal responses are needed to show their relative contributions to the observed non-linearities.
REFERENCES


[26] Uludag K, Yesilyurt B, Ugurbil K. Hemodynamics and nonlinearities of BOLD response to ultrashort visual stimulation (5ms - 1s). 2006; Seattle, USA.


FIGURE CAPTIONS

FIGURE 1  Experimental design schematic. Stimulation starts after an onset baseline period of 45 seconds. Four very short stimulus durations (SD=5, 50, 250, 1000 ms) and two different stimulus intensities (SI=low and high) were used.
Inter stimulus interval (ISI) were randomized between 28 to 32 seconds in each run. Each run ended with a cessation baseline period > 45 seconds.

**FIGURE 2**  a) Three basis functions were generated using FLOBS. Each voxel’s intensity time course was then cross correlated with the best linear fit of these basis functions to construct statistical maps. b) Representative averaged time course (Figure created for illustration purposes). Response amplitude and initial dip were calculated from the peak value of the response and the early decrease of BOLD signal intensity respectively. The integral was calculated by the product amplitude multiplied by full width at half maximum (alternatively, integral borders are shown in the figure as well).

**FIGURE 3** Representative activation map (z > 8) superimposed on the anatomical image (SI = high, SD = 1000 ms).

**FIGURE 4** Averaged BOLD signal intensity time courses (averaged over 8 trials, 50 voxels and 11 subjects) for a) low stimulus intensity and b) high stimulus intensity. All responses exhibit an initial dip. The positive BOLD responses increase with stimulus duration.

**FIGURE 5** Normalized a) response amplitudes, b) full width at half maximum (FWHM) values and c) response integrals. Error bars indicate the standard error. Amplitude and Integral of the BOLD response are highly non-linear with respect to stimulus duration. FWHM decreases for SD = 5 ms showing a neuronal contribution to the width of the BOLD response even at these short stimulus durations.

**FIGURE 6** Representative thresholded and clustered amplitude maps for the time intervals 0-2s (left) and 6-8s (right) overlaid on an anatomical image. Initial dip and positive BOLD response maps largely overlap.
FIGURE 7  Distribution of z-values using voxels (50 voxels for each subject, N = 11) based on: a) largest initial dip (sign of z-values inverted) b) positive response z-scores using the same voxels as in a); c) positive responses. Dashed line corresponds to p = 0.01.

FIGURE 8  Averaged time courses of 50 voxels (N = 11) based on: Black curve: most significant positive response; Blue curve: most significant initial dip; Red curve: overlap of both.

TABLE  a) Amplitude of average BOLD responses in [%], b) FWHM of the average BOLD responses in [s] for all stimulus durations and low and high stimulus intensities (8 trials, 50 voxels, N = 11).
Figure 4

Low

BOLD (%)

1000 ms
250 ms
50 ms
5 ms

High

BOLD (%)

1000 ms
250 ms
50 ms
5 ms

(time (s))

(time (s))
Figure 5

(a) Amplitude vs. Stimulus Duration (ms)
(b) Integral vs. Stimulus Duration (ms)
(c) FWHM vs. Stimulus Duration (ms)
Figure 6

0-2 sec

6-8 sec
Figure 7
Figure 8

[BOLD (%)]

- Positive Resp.
- Initial Dip Resp.
- Positive + Initial Dip Resp.

(time (s))
### TABLE 1

**a) Amplitude of the BOLD responses in [%]**

<table>
<thead>
<tr>
<th></th>
<th>5 ms</th>
<th>50 ms</th>
<th>250 ms</th>
<th>1000 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low Intensity</strong></td>
<td>0.92 ± 0.24</td>
<td>1.19 ± 0.32</td>
<td>1.39 ± 0.23</td>
<td>1.66 ± 0.27</td>
</tr>
<tr>
<td><strong>High Intensity</strong></td>
<td>1.07 ± 0.40</td>
<td>1.27 ± 0.40</td>
<td>1.60 ± 0.45</td>
<td>2.36 ± 0.64</td>
</tr>
</tbody>
</table>

**b) Width of the BOLD responses in [s]**

<table>
<thead>
<tr>
<th></th>
<th>5 ms</th>
<th>50 ms</th>
<th>250 ms</th>
<th>1000 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low Intensity</strong></td>
<td>3.5 ± 0.8</td>
<td>4.1 ± 1.0</td>
<td>4.3 ± 0.9</td>
<td>4.5 ± 0.9</td>
</tr>
<tr>
<td><strong>High Intensity</strong></td>
<td>4.0 ± 0.7</td>
<td>4.2 ± 1.0</td>
<td>4.6 ± 1.4</td>
<td>4.8 ± 1.3</td>
</tr>
</tbody>
</table>