Supporting Online Material for

Attention But Not Awareness Modulates the BOLD Signal in the Human V1 During Binocular Suppression

Masataka Watanabe,* Kang Cheng, Yusuke Murayama, Kenichi Ueno, Takeshi Asamizuya, Keiji Tanaka, Nikos Logothetis

*To whom correspondence should be addressed. E-mail: watanabe@tuebingen.mpg.de

Published 11 November 2011, Science 334, 829 (2011)
DOI: 10.1126/science.1203161

This PDF file includes:

Materials and Methods
SOM Text
Figs. S1 to S10
References
Supporting Online Material

Materials and methods

Subjects
Four healthy right-handed volunteers (two males and two females), aged 24 ± 5 years, participated in the experiment. All subjects had normal or corrected-to-normal visual acuity and normal stereo-depth perception. Two subjects were right-eye dominant and two were left-eye dominant. All subjects gave prior written informed consent before each experiment.

MRI acquisition and stimulus presentation
Functional images were acquired on a 3 Tesla TIM scanner located in Max Planck Institute Tuebingen, with a gradient echo planar imaging (EPI) pulse sequence and a 12-channel phased-array head coil. We collected 16 slices (slice thickness 1.9 mm) oriented perpendicular to individual subject’s calcarine sulcus with the first slice prescribed at the occipital pole, using an interleaved sequence and the following parameters: volume repetition time (TR) 1.87 s, echo time (TE) 39 ms, 96 x 96 matrix, and voxel size 1.3 x 1.3 x 1.9 mm³. A set of high-resolution (1 mm isotropic) T1-weighted 3D-MDEFT images was acquired from each subject as an anatomical reference.

All stimuli were programmed in Matlab in the PsychToolbox environment (S1-S2). They were presented at a resolution of 1280 x 1024 pixels and at a screen refresh rate of 60 Hz from a viewing distance of 82 cm. Stimuli were projected onto a translucent screen at the end of the scanner bore and viewed through a tilted mirror fixed to the head coil.

Prior to fixing the stimulus and task conditions as in the current paper, pilot experiments were conducted at RIKEN using a 4 Tesla whole-body MRI system equipped with a Magnex head gradient system.

Localizing targeted monocular region in individual subject’s V1
On a separate fMRI scan, the subject viewed either a disc of counter-phasing black and white checkerboard (6° in diameter; the center of the disc was 4° offset to the right from a centrally located fixation marker) or a dynamic Mondrian pattern (composed of an oscillating horizontal grating with multiple square patches whose positions were randomly refreshed at 10 Hz) placed in a ring-shaped region (inner diameter, 6°; outer diameter, 10°; the center of the ring was 4° offset to the right from the fixation marker). The subject was instructed to fixate on the central fixation marker (1° in size). To help maintain fixation and the level of arousal, the subject performed a letter detection task on the fixation marker. A pseudo-randomly ordered temporal sequence of white single characters altering at 10 Hz was presented on the fixation marker, and the subject was asked to report by button press immediately when they detected a target letter (f). The task performance was inspected online and analyzed offline; all subjects could perform the task with above 75% correct rate. The acquired EPI images were then analyzed with a boxcar regressor using SPM5 with a threshold of p < 0.05 (after family-wise error correction for multiple comparisons; see below
for detail) and the voxels in the targeted monocular region in V1, located mostly in the depth of the calcarine sulcus, were identified on the basis that they responded significantly to counter-phasing checkerboard but not to the dynamic Mondrian pattern (Figs 1c and 1d).

fMRI scans with and without binocular suppression

We adopted a modified version of continuous flash suppression (S3) to render the stimulus visibility. An “invisible stimulus” consisted of a target motion grating (1.5 cpd drifting at 5 Hz; 10° in diameter; the center of the grating was 4° offset to the right from the central fixation marker; target contrast, ranging from 15 to 45%, was adjusted for each subject to the highest value at which the robust invisibility of target in the binocular condition was achieved) presented to the subject’s non-dominant eye and a ring-shaped dynamic Mondrian pattern (identical to the one used in the localization scan described above) presented to the subject’s dominant eye. In contrast, for the “visible stimulus,” both the target motion grating and dynamic Mondrian pattern were presented to the non-dominant eye. Because these two stimulus components partially overlapped, they were temporally interleaved to induce a constant transparent percept: the target motion grating was presented in even frames while the dynamic Mondrian pattern was presented in odd frames at 60 Hz refresh rate. To match the total presentation duration of the two stimulus components in the invisible stimulus to that in the visible stimulus, the target motion grating and dynamic Mondrian pattern in the invisible stimulus were similarly interleaved.

A stimulus block consisted of 16 stimulus representations (500 ms stimulus on followed by 500 ms inter-stimulus interval). Based on the overwhelming stimulus visibility, two types of stimulus blocks were created. Regardless of attentional manipulations (see below), a “visible stimulus block” contained 14 visible stimuli and two invisible stimuli (as catch stimuli), the appearance order of which was randomly determined. An “invisible stimulus block,” on the other hand, contained 14 invisible stimuli and two randomly appeared visible stimuli.

Throughout an fMRI scan, the subject was instructed to constantly fixate on the central fixation marker, where a pseudo-randomly ordered temporal sequence of single characters altering at 10 Hz. The color at the fixation marker indicated the type of task in relation to the attending location. When the characters were in blue, signaling the “attention to non-target block,” the subject was required to perform a letter detection task reporting the appearance of a given character. When the characters appeared in red, signaling the “attention to target block,” however, the subject was asked, while maintaining central fixation, to direct their attention to the peripheral target and report whether the target motion grating was visible or not (two alternative forced choice with two separate button presses). Following each stimulus presentation block (16 s), a blank-display period (24 ± 4 s) served as a resting baseline condition, during which the fixation was in white and the subject was merely instructed to maintain fixation. The change of the fixation color from white to either blue or red began 1 s prior to the stimulus presentation block to inform the subject to get ready for the upcoming task.

In total, six to nine fMRI scans were conducted on each subject. A single scan contained all four types of stimulus blocks across two visibility conditions and two attentional conditions, that is, the “target visible with attention to target block,” “target
invisible with attention to target block,” “target visible with attention to non-target (fixation) block,” and “target invisible with attention to non-target block.” Within a scan, the appearance order for the four blocks were pseudo-randomly determined and a few seconds were inserted in between successive blocks.

fMRI analysis

fMRI data were preprocessed using SPM5 (Wellcome Dept of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk/spm). Preprocessing included image realignment and co-registration of functional and anatomical images. Functional images from the V1 localization scan were spatially smoothed with an isotropic Gaussian kernel of 20 mm full width at half-maximum (FWHM). A GLM analysis with a boxcar regressor was applied to calculate the statistical values to determine the targeted monocular region in V1. No spatial smoothing was applied on the functional images from the two-factorial-design experiment.

Supporting text and figures

Additional fMRI measurements with an MR compatible eye-tracker

To investigate the possible effects of eye movements, additional fMRI experiments were conducted on three subjects using an MR compatible eye-tracker (MReyetracking system: Resonance Technology Inc.), which has a sampling rate of 60 Hz, and an online analysis software package (Viewpoint: Arrington Research). These three subjects did not participate in the original experiment. All other aspects, including the stimulus design and fMRI data analysis, were identical to those described in the main experiment except for minor changes in scan parameters: volume repetition time (TR) 2.63 s, echo time (TE) 35 ms, 96 x 96 matrix, and voxel size 1.5 x 1.5 x 2.0 mm³. Fig. S1 shows X and Y eye-positions for all three subjects across the four blocks. Offline statistical analysis confirmed that there was no significant eye-position difference between different block conditions for all three subjects (one-way ANOVA, p = 0.32, 0.55, 0.48 for the mean and p = 0.57, 0.45, 0.64 for the standard deviation in the horizontal direction, and p =0.78, 0.46, 0.85 and p = 0.56, 0.61, 0.79 in the vertical direction). To further verify that larger V1 BOLD in attention to target blocks did not stem from horizontal eye position shifts towards the target, we also conducted a one-tail t-test between the two attention conditions, where visibility conditions were pooled together. For all three subjects, there was no significant shift towards the target (p = 0.24, 0.47, and 0.38, respectively). In agreement with the findings in the original experiment, we found no consistent effect of the visibility on BOLD responses, but a highly significant effect of attention regardless of visibility conditions (p < 0.001) for all three subjects studied (Fig. S2).

Next to test for possible effects on V1 BOLD from non-significant differences in eye-positions and eye-movements, we conducted the following analysis. For each condition, we divided the trials into two pools depending on the mean horizontal eye-position shifts toward the target. The V1 BOLD time courses showed no consistent trend on the trial-by-trial eye-position bias, and the differences between the two trial pools, if any, were always smaller
than the difference between the two attentional conditions (Fig. S3). Similar analysis was also conducted on the standard deviation of eye-positions. Again, we saw no consistent trend, and the differences between the two trial pools were also smaller than the difference between the two attentional conditions (Fig. S4). These results indicate that any effect of eye positions in the trial-to-trial variability is negligible compared to other potential sources of the variability, thus providing evidence that eye-positions and eye-movements played only an insignificant role in the reported elevation of BOLD response in V1 by top-down attention.

Finally we analyzed saccades to test for the possibility that increased saccades towards the target in attention to target conditions may have led to the observed increase of BOLD signal in V1. We adopted a saccade-detection algorithm proposed by Martinez-Conde and colleagues (S4). Signal from the eye-tracker was first up-sampled to 1 KHz and then smoothed by a boxcar filter of ± 25 ms to reduce noise (up-sampling was required for applying the filter). An eye-velocity threshold of 4 °/s was then applied to detect saccadic eye movements. The lower limit of saccadic distance was set to 0.3° in accordance to the results of eye-tracker evaluation experiments described in the next section. First, in line with the results from the average eye-position analysis, we did not find any significant increase in the number of saccades toward the target in “attention to target” conditions compared to “attention to fixation” conditions (p values of one-sided t-test: 0.66, 0.58, 0.61 for subjects 1, 2 and 3, respectively). Second, to test for possible effects of non-significant differences in saccadic eye-movements on BOLD signal in V1, we conducted the following analysis. For each condition, we divided the trials into two pools depending on the ratio between saccades landing towards and away from the visual target relative to the center of fixation. Time courses of BOLD responses in V1 did not exhibit any consistent trend that depended on the trial-by-trial saccadic bias, and the differences between the two trial pools, if any, were always smaller than the difference between the two attentional conditions (Fig. S5). These results indicate that any effect of saccadic bias towards the target is negligible compared to other potential sources of the variability, thus providing evidence that saccadic eye-movements played at most an insignificant role in the reported, top-down attention evoked, elevation of the BOLD response in V1.

Performance of the eye-tracker system

We evaluated the performance of the eye-tracker by conducting two additional eye-movement experiments inside the scanner with the three subjects who participated in the combined fMRI and eye-tracking experiment described above. The first was aimed to evaluate the sensitivity of the eye-tracker for detecting eye position shifts. The subjects performed a character detection task as in the main experiment (the central fixation task) while the position of characters shifted horizontally every 2 s. Fonts smaller than those used in the original experiment (0.22° in size) were used for precision measurements. The results from a representative subject are shown in Fig. S6. We performed a Student’s t-test between eye positions obtained immediately before the fixation shift and those obtained 1 s after the shift (duration, 1 s) and found that the difference in measured eye positions was statistically significant (p < 0.01) for all fixation shifts greater than or equal to 0.13° (Fig. S6). These results indicate that the spatial resolution of the eye-tracker is high enough to detect the
general position bias or the difference in deviations within the size of the characters used in the main experiment. We thus contend that our conclusion drawn above that there were no statistical differences in the average and standard deviation of eye positions between experimental conditions is solid.

The second experiment was aimed to examine the detectability of saccades. Subjects were asked to visually track a moving dot on the screen. A single trial consisted of a dot (0.03° in diameter) appearing at the center of the screen for 752 ms, followed by an instant jump to the next position, where it stayed for another 752 ms. The dot then disappeared from the screen, and following an inter-trial interval of 256 ms, a new trial was commenced. We adopted the previously mentioned saccade-detection algorithm proposed by Martinez-Conde and colleagues (S4). The result from a representative subject is shown in Fig. S7. The standard deviation of the measured saccadic distance from the three subjects was (horizontal x vertical) 0.18° x 0.19°, 0.17° x 0.19° and 0.18° x 0.20° (horizontal x vertical), respectively. In addition, we plotted the saccadic distance and maximum velocity to further confirm the performance of the eye-tracker in detecting saccadic eye-movements (Fig. S8). It is known that the saccadic distance and peak velocity exhibit a linear relationship in distance larger than 0.1 - 0.2° in idealistic eye-movement recordings (S5). Thus, the lower range of the saccadic distance showing a linear relationship can be used as a lower bound for reliable saccade detection. In our eye-tracking data, we confirmed a linear relationship down to 0.3° in the saccadic distance in all three subjects. Together with the standard-deviation results, we adopted 0.3° as the lower distance limit for saccadic detection for the data analysis of the combined fMRI and eye-tracker experiment.

Psychophysics experiment for testing the occurrence of inattentional blindness

Additional psychophysical experiments were performed to test for the possible inattentional blindness in the stimulus condition used in fMRI measurements. Eight healthy volunteers (three males and five females), aged 26 ± 3 years, participated in the experiment. None of them participated in the main fMRI experiment. Identical task and stimulus conditions were used to mimic the subject’s state during fMRI measurements as much as possible. Subjects received full instructions for the task and underwent practice trials until they were comfortable with the task and their performance reached certain criteria (80% for the central fixation task and 90% for the target detection task). The actual run consisted of 4 blocks in the order of 1) invisible-central fixation task, 2) visible-target detection task, 3) invisible-target detection task, and 4) visible-central fixation task. We aborted the fourth block right after the first 500 ms presentation of the visible stimulus, and replaced the stimulus with an open circle corresponding to the monocular region of the target grating. Immediately after this, the subjects were asked the following question: “Did you notice anything inside the circle?” If the answer was “yes,” they were then asked the next question: “Can you describe any feature of the visual object you have seen?” All eight subjects answered “yes” to the first question, but the answer to the second question varied. Three subjects described only its color (monochrome), two subjects described its color and that something was moving, and the remaining three subjects guessed that it was identical to the visible target used in target detection blocks. From these observations, we conclude that
inattentional blindness is less likely to happen in our stimulus conditions, most likely due to the fact that a simple motion-grating appears abruptly in a total blank field only 1° away from the center of fixation.

**Multivariate pattern analysis of visual awareness**

We have conducted a multivariate pattern analysis (MVPA) to investigate the possibility of dissociable activation patterns between visible and invisible stimuli. Voxels in the V1 ROI defined in the main analysis were analyzed using established MVPA methods (S6-S7) as follows. Single trial/single voxel BOLD responses were transformed into scalar values by averaging the values of percent signal changes between 5 and 21 seconds after the stimulus onset, during which trial-averaged BOLD time courses show a typical positive response. Single trial response vectors of N dimension, where N is the number of voxels within the ROI, were constructed with the calculated average responses and were used for training and testing a support vector machine (SVM) classifier. We adopted a 10-fold cross-validation method, where 90% of trial response vectors were used to train the classifier for predicting whether the remaining responses belonged to “visible” or “invisible” trials. Two attentional conditions, that is, attention directed to the target and attention directed to the central fixation, were analyzed separately. The performance of trained classifiers in distinguishing visible and invisible trials in both attentional conditions was close to the chance level for all subjects, and in no case was the performance significantly higher than the performance with surrogate data (Fig. S9). Here, surrogate data was constructed using original BOLD time courses of voxels in the same V1 ROI, but their labels of “visible” and “invisible” trials were randomly shuffled. These results provide further support to our claim that BOLD signal in V1 is not modulated by the change in awareness.

**A hybrid conceptual model providing a possible interpretation to our results**

Numerous psychophysical studies have demonstrated that attention and awareness are behaviorally “dissociable.” Top-down attention directed toward perceptually invisible stimuli leads to negative aftereffects (S8) or priming (S9), while the gist of scenes can be perceived despite the near absence of attention (S10-S13). Although it is true that previous studies, as well as our own study, demonstrating “awareness without attention,” indicate that the reportable content is limited compared to conditions where attention is directed toward the object. We may summarize that, in terms of behavior, attention and awareness are “dissociable yet coupled.” Thus, it is likely that the neural mechanisms of attention and awareness are also “dissociable yet coupled.” The crucial question is, where is V1 situated in this “dissociable yet coupled” neural architecture for attention and awareness?

Although we cannot generalize our findings to different measurement modalities and it is possible that future work may reveal modulation of V1 activity by contents of awareness under attentional control, we hereafter construct a conceptual model under the hypothesis that V1 is not a part of the neural correlate of visual awareness.

Before describing our conceptual model, let us address the key aspects of our reasoning to provide an overview:
a) Both attention and awareness are not a single entity in terms of neural functions.

b) Awareness can be divided into “access consciousness,” where the contents are verbally reportable, and “phenomenal consciousness,” in which an object is outside of phenomenal consciousness if it cannot be perceived with the aid of directed attention (for review, see (S14)).

c) According to Knudsen, attention is divided into “working memory,” “competitive selection,” “sensitivity control” and “salience filters” (S15).

d) Awareness and attention have much high-level neural functions in common regarding “competitive selection/working memory” and “access consciousness.”

e) Meanwhile, “sensitivity control” has been psychophysically shown to work upon unconscious contents of neural representations (S8-S9).

f) d) and e) explain the “dissociable yet coupled” nature of attention and awareness, and our results may be straightforwardly interpreted by placing V1 in a position that it receives “sensitivity control,” but is not a part of the neural representation of “phenomenal consciousness.”

We next introduce a hybrid conceptual model (Fig. S10) based on a model by Knudsen (S15) and that by Lamme (S14). In order to avoid confusion, we have changed the descriptions of Lamme’s model from “attended/unattended” to “selected/non-selected.” In addition, for the sake of simplicity, we have omitted the aspect of the gaze control in Knudsen’s model.

The key point of our hybrid model is that the “sensitivity control” is applied to both conscious and unconscious neural representations. Here, “sensitivity control” is defined as modulation of the sensitivity of neural circuits to improve the quality of information. This is verified by psychophysical studies demonstrating that top-down attention directed toward perceptually invisible stimuli leads to negative aftereffects (S8) or priming (S9). To improve the quality of the available information, it should be beneficial for “sensitivity control” to tap into sensory neural circuits as early as possible.

On the other hand, the neural representation of awareness is a totally different issue. For example, area V1 possesses information that does not enter our conscious vision: spatiotemporal jitter of visual inputs by microsaccades (S16), raw wavelength information (without color constancy correction) (S17), high spatial frequency information exceeding visual acuity (S18), high temporal frequency flickering luminance stimuli (S19), anti-correlated random dot stereograms (S20), etc. The actual question is whether “V1 as a whole” is outside of conscious neural representation, or put it more precisely, whether neural activity in V1 is not modulated by the contents of awareness. Our experimental results, by observing the BOLD signal, indicate that this may be the case. Assigning entire V1 as “neural representations (unconscious)” in our hybrid model provides a possible interpretation as why we see attentional modulation (due to sensitivity control), but not awareness modulation.
Fig. S1. Eye movement traces from the additional experiment. Peri-stimulus trial average (red traces) and standard deviation (green traces) of the eye position relative to the fixation marker are plotted for horizontal (X: positive values indicate rightward shift) and vertical (Y: positive values indicate upward shift) movements independently. Gray bar indicates the stimulus presentation period.
Fig. S2. Time courses of averaged BOLD responses and 95% confidence intervals in the targeted monocular region of the additional experiment. The three time courses in each panel (for each of three subjects) were obtained from target visible with attention to target block (magenta), target invisible with attention to target block (cyan), target visible with attention to non-target location block (red) and target invisible with attention to non-target block (blue), respectively.
Fig. S3. Time courses of averaged BOLD responses and 95% confidence intervals in the targeted monocular region of the additional experiment where trials were divided into two pools according to the mean horizontal eye positions. The eight time courses for each of three subjects (visibility conditions are separately presented in the two columns) were obtained from target visible with attention to target block (magenta), target invisible with attention to target block (cyan), target visible with attention to non-target location block (red) and target invisible with attention to non-target block (blue), respectively. Average time courses with the confidence interval coded in gray were obtained from the trials with the eye-position bias toward the target, while those with the confidence interval in lighter colors were obtained from the trials with the bias away from the target.
Fig. S4. Time courses of averaged BOLD responses and 95% confidence intervals in the targeted monocular region of the additional experiment where trials were divided into two pools according to the standard deviation of horizontal eye position. The eight time courses for each of three subjects (visibility conditions are separately presented in the two columns) were obtained from target visible with attention to target block (magenta), target invisible with attention to target block (cyan), target visible with attention to non-target location block (red) and target invisible with attention to non-target block (blue), respectively. Average time courses with the confidence interval coded in gray were obtained from the trials with smaller standard deviations of eye-positions, while those with the confidence interval in lighter colors were obtained from the trials with larger standard deviations.
Fig. S5. Time courses of averaged BOLD responses and 95% confidence intervals in the targeted monocular region of the additional experiment where trials were divided into two pools according to the ratio of saccades made toward and away from the target. The eight time courses for each of three subjects (visibility conditions are separately presented in the two columns) were obtained from target visible with attention to target block (magenta), target invisible with attention to target block (cyan), target visible with attention to non-target location block (red) and target invisible with attention to non-target block (blue), respectively. Average time courses with the confidence interval coded in gray were obtained from the trials with larger ratios of saccades towards the target, while those with the confidence interval in lighter colors were obtained from the trials with smaller ratios.
**Fig. S6.** Measured eye positions (mean ± standard deviation) following the shifts of fixation characters. The number at the top of each panel indicates the distance of the horizontal shift of fixation, which took place at time zero. The number of trials per condition (30 trials) approximately matches that in the main experiment.
Fig. S7. Measured saccadic distance, plotted against the actual shift of fixation. Each dot denotes a single saccade. A, horizontal distance, and B, vertical distance.

Fig. S8. Peak eye velocity versus saccadic distance. Data were collected from the moving-dot tracking experiment.
Fig. S9. Results of multivariate pattern analysis. Filled red square, blue dots, and the open square denote the actual classification ratio, classification ratios for individual surrogate data and the mean classification ratio of all surrogate data, respectively. Each pair of black horizontal bars denotes the 95% confidence interval based on the analysis of surrogate data. For each subject, the results from the “attention to target” condition are plotted on the left, and those from the “attention to fixation” condition are plotted on the right.
Fig. S10. A hybrid conceptual model based on the models by Knudsen (S15) and Lamme (S14) provides a possible interpretation for our experimental findings. Oval compartments are taken from Knudsen’s model, whereas rectangle compartments are taken from Lamme’s model (in order to avoid confusion, descriptions of Lamme’s model have been changed from “attended/unattended” to “selected/non-selected.”). Processes that contribute to attention are shown in red. Sensitivity control, which modulates the sensitivity of neural circuits to improve the quality of information, is directed toward both conscious and unconscious neural representations. Full conscious report requires information to be maintained and processed in working memory.

Supporting online material references


