

# Two Visual Systems and Their Eye Movements: a Fixation-Based Event-Related Experiment with Ultrafast fMRI Reconciles Competing Views

DOI: 10.17691/stm2019.11.4.01

Received August 9, 2019



**B.M. Velichkovsky**, DSc, Professor, Corresponding Member of the Russian Academy of Sciences, Member of Academia Europaea (MAE), Chief Researcher<sup>1,2</sup>; Senior Professor<sup>3</sup>;

**A.N. Korosteleva**, Researcher<sup>1</sup>;

**S. Pannasch**, Dr. rer. nat., Dr. habil., Professor<sup>3</sup>;

**J.R. Helmert**, Dr. rer. nat., Leading Researcher<sup>3</sup>;

**V.A. Orlov**, Researcher<sup>1</sup>;

**M.G. Sharaev**, PhD, Leading Researcher<sup>4</sup>;

**B.B. Velichkovsky**, DSc, Professor, Department of Methodology of Psychology<sup>5</sup>;

**V.L. Ushakov**, PhD, Associate Professor, Leading Researcher<sup>1</sup>; Senior Researcher<sup>6</sup>

<sup>1</sup>National Research Center “Kurchatov Institute”, 1 Akademika Kurchatova Square, Moscow, 123182, Russia;

<sup>2</sup>Russian State University for the Humanities, 6 Miusskaya Square, Moscow, 125993, Russia;

<sup>3</sup>Technische Universität Dresden, 17 Zellescher Weg, Room A221, Dresden, 01069, Germany;

<sup>4</sup>Skolkovo Institute of Science and Technology, 30/1 Bolshoy Boulevard, Moscow, 121205, Russia;

<sup>5</sup>Lomonosov Moscow State University, 1 Leninskiye Gory St., Moscow, 119991, Russia;

<sup>6</sup>National Research Nuclear University MEPhI, 31 Kashirskoe Shosse, Moscow, 115409, Russia

Studies of active vision in naturalistic scenes show the existence of two classes of eye movements manifested in ambient and focal visual fixations. This finding seems to corroborate with the anatomical separation of two “streams” of visual processing related to localization (dorsal system) or to identification of objects (ventral system). Direct verification of this connection proved to be difficult due to an insufficient resolution of the conventional noninvasive brain-imaging methods. Another hypothesis recently attributed the same observation to the lateralization of global and local attention modes in the right and left hemispheres, correspondingly. Thus, there are two tentative explanations for the brain mechanisms of the same eye movement patterns in free image viewing. Our study aimed at resolution of this controversy.

**Materials and Methods.** 13 healthy subjects (age 21 to 31 years, right handed, 8 females) with normal or corrected to normal vision and without known history of neurological diseases participated in this experiment. Using a combination of ultrafast multi-band fMRI scanning with the fixation-based event-related (FIBER) paradigm of data collection, we measured the brain functional activity in its relation to tasks, a semantic category of the inspected object (houses or faces), brain regions as well as ambient and focal visual fixations during free viewing of complex images with an unprecedentedly high temporal and spatial resolution.

**Results.** The results unexpectedly showed that both competing hypotheses are confirmed. In line with our early proposal, ambient fixations were accompanied by activation of structures traditionally associated with the dorsal visual pathway, while focal fixations correlated with that of the ventral pathway. At the same time, the second hypothesis also proved to be correct: the activated structures of the dorsal pathway were localized in the right hemisphere and those of the ventral brain networks mainly — albeit not exclusively — in the left hemisphere.

**Conclusion.** The present study for the first time demonstrates pronounced lateralization of both basic brain mechanisms in charge of visual perception and eye movement control in free processing of complex images. This conclusion poses a number of further questions about a possible relation between two modes of active vision and other forms of asymmetries found at different levels of human brain organization.

**Corresponding author:** Boris M. Velichkovsky, e-mail: [boris.velichkovsky@tu-dresden.de](mailto:boris.velichkovsky@tu-dresden.de)

**Key words:** active vision; ambient vs. focal visual fixations; dorsal vs. ventral pathways of perception; lateralization of visual attention; fixation-based event-related (FIBER) paradigm; ultrafast multi-band fMRI scanning; hemodynamic response function; travelling wave.

## Introduction

Studies of active vision in naturalistic environments show the existence of two classes of eye movements manifested in ambient and focal visual fixations [1–5]. This finding seems to corroborate with the known anatomical separation between two main cortical pathways (“streams”) of visual processing related primarily either to localization (dorsal system) or to identification of objects (ventral system). Direct verification of this connection proved to be difficult due to an insufficient resolution of the conventional noninvasive brain-imaging methods. Another hypothesis, recently proposed by Mills and colleagues [6], attributed the same behavioral observations to the lateralization of mechanisms for global and local modes of visual processing in the structures of right and left hemispheres, correspondingly. Thus, there are two tentative neurocognitive explanations for the same particularities of eye movement patterns in free image viewing.

Which of these explanations is the correct one? The answer to this question would only be possible when eye movements and brain activity can be measured with a sufficiently high temporal as well as spatial resolution. Such a measurement was not the case in previous research. For example, Mills and colleagues [6] used the functional transcranial Doppler ultrasound method with its exemplarily low temporal and spatial resolution. Therefore they relied on our finding that the exploration of complex visual scenes begins with an ambient mode of processing — short visual fixations and long-range saccades, which within a few seconds proceed to a more attentive, or focal mode of processing characterized by longer fixation and short-range saccades [3]. According to their results, this ambient-to-focal shift temporally overlaps with a shift of brain energy consumption from the right to the left hemisphere suggesting a right-hemispheric dominance in the early, ambient, phase of scene perception [6]<sup>1</sup>. Another option to access ambient and focal visual processing modes is to simply look at duration of visual fixations and the amplitudes of surrounding saccades whereby the length of the fixations is the single dominating variable [4, 5]. However, to use this option one would again need imaging tools with temporal resolution on a subsecond scale.

<sup>1</sup>Mills et al. [6] used two tasks, visual search and recognition, in their study. The interhemispheric shift of blood supply was apparent only in memory task. Unfortunately the construction of both tasks can be questioned — the targets of search tasks were too small to be detected and the recognition memory task included no memory tests at all.

An attempt to create a method to explore mechanisms behind ambient and focal visual processing has been made in the framework of EU NEST Project PERCEPT [7]. The idea was to modify the event-related paradigm where visual fixations would be considered as ‘events’ instead of physical stimulation, such as the presentation of visual or auditory stimuli. In its first application, the FIBER (fixation-based event-related) paradigm with functional magnetic resonance imaging (fMRI) scanning demonstrated in a free viewing experiment that fixations on different objects in different task contexts resulted in distinct cortical patterns of activation [8]. By using multivariate pattern analysis, the authors of this study showed that the blood-oxygen-level dependent (BOLD) signals revealed information about the task context of fixations and about the binary category of objects (houses vs. faces) being inspected during these fixations. There was only one follow-up study with the FIBER paradigm [9]. It aimed at exactly the question of brain mechanisms behind ambient and focal visual fixations but the obtained results were inconsistent with the previous findings. The authors described three nonlateralized regions of activation one of which was in the primary visual cortex while two others were located in the deep ventral parts of posterior cortex without relevance to known mechanisms of active vision (see, e.g., [10]). This inconsistency could be explained by the low temporal resolution of both eye tracking (50 Hz) and the standard fMRI scanning protocol with repetition time (TR) 2.0 s used in the experiments.

In a parallel development, there were efforts to speed up the fMRI with new schemes of parallel scanning. Employing these schemes generally showed that the BOLD responses may convey more finely graded timing information than usually believed however on the costs of a decline in spatial accuracy [11, 12]. We recently adopted an alternative approach to ultrafast scanning developed at the University of Minnesota, which is free from the detriment in spatial accuracy and offers whole-brain sampling resolutions of about 0.5 s<sup>2</sup>. Although this is still insufficient to directly describe eye movement dynamics, the technique is a major step on the way to revealing mechanisms of active vision especially when combined with the FIBER paradigm. In addition to the use of fast eye tracking, such a combination was the unique feature of the present study. Its aims were twofold. First, we wished to test whether previous FIBER results can be replicated after these changes of fMRI

<sup>2</sup>Multi-band Accelerated EPI Pulse Sequence, Release 016a from 19.12.2017 (curtesy Center for Magnetic Resonance Research, University of Minnesota).

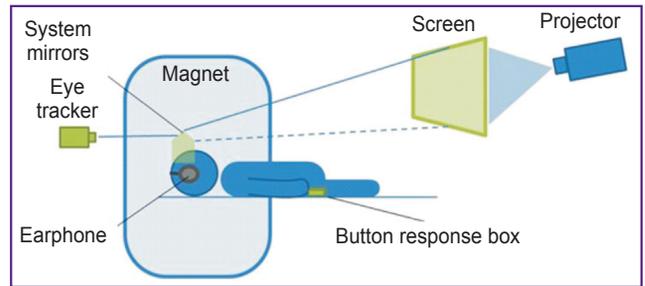
protocol<sup>3</sup>. The second and main goal of the study was the measurement of brain functional activity in relation to ambient and focal visual fixations. The division of ambient and focal fixations was operationalized on the basis of their durations as in some earlier experiments [4, 5].

**Materials and Methods**

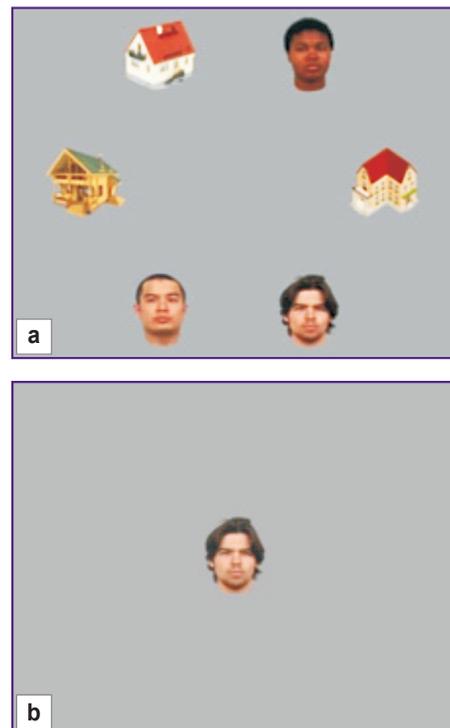
**Participants.** Overall 13 healthy subjects (age 21 to 31 years, right handed, 8 females) with normal or corrected to normal vision and without known history of neurological deceases participated in the experiment. Initially 15 subjects were scanned but 2 of them had to be excluded due to calibration issues of the eye tracker. Each participant gave informed consent prior to the experiment. Ethical approval for this study was provided by the local ethical committee of the National Research Center “Kurchatov Institute”.

**Stimuli and equipment.** To pass visual stimuli into an MRI scanner room during the functional research we used a projection system as schematically shown in Figure 1. Via a 45 degree tilted mirror, placed on top of the head coil, the subject was able to see the entire presentation display. The distance between the eyes and the screen was 75 cm. A second mirror relayed the image of the eye to the infrared camera of the eye tracker mounted at the head of the scanner bed. The paradigm was developed using NBS Presentation software. Parameters of presentation including the size of a translucent display were identical to those of the earlier FIBER study [8]. The display of 1024×768 pixels subtended a visual angle of 32×25 degrees for the entire screen. The house and face objects inside the stimulus display covered approximately 150×150 pixels. A typical stimulus configuration for free viewing conditions is shown in Figure 2. An Apple MacBook Pro (Apple, Cupertino, USA) was used to drive the stimulus display presentation.

For the eye tracking, a MRI-compatible eye tracker EyeLink 1000 Plus (SR Research, Canada) was used with a sampling frequency of 250 Hz. Visual fixations were recorded and categorized according to the type of object inspected and the task at hand (see below). Each subject was placed to Magnetom Verio 3T (Siemens, Germany) MRI scanner with 32-channel MRI head coil. During the experiment, we simultaneously recorded fMRI data by a combination of FIBER paradigm [8] and ultrafast Multi-band Accelerated EPI Pulse Sequence. The scanning process had two stages: capturing high-resolution anatomical data and recording functional data by a parallel scanning protocol with ultrafast EPI-sequence (TR=720 ms, TE=33 ms, 56 slices, slice thickness = 2 mm, spatial resolution in each slice = 2×2 mm).



**Figure 1. Schematic view of the experiment with simultaneous registration of BOLD-signal and eye movements**

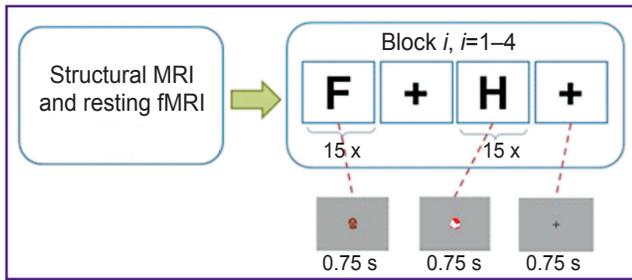


**Figure 2. Typical stimulus displays in the free viewing condition:**

(a) multiobject presentation; (b) a control test-object

**Procedure.** During the scanning session, we collected functional data in two experimental paradigms: “localizer” and “multiobject viewing”. The localizer part was a standard experiment with passive observation (block paradigm). Subjects viewed pictures of houses or faces in the center of the screen while BOLD-activity was tested in search of cortical areas predominately involved in perception of houses or faces [8, 14–16]. Figure 3 shows a scheme of the experiment for functional localization. The experiment consisted of 4 blocks, each including 2 parts. The first part of a block contained 15 stimuli with pictures of faces and a fixation cross. The second part of a block contained 15 stimuli with pictures of houses and a fixation cross. The time of each stimulus was 750 ms, fixation cross time — 10 s, blank interval —

<sup>3</sup>For a preliminary report on the method, see [13].



**Figure 3. Scheme representation of the localizer experiment**

500 ms, total trial time — 11.25 s. Structural MRI and resting state BOLD-activity were registered individually with closed eyes during 10 min preceding each localizer experiment. Overall length of the experiment was about 15 min.

In the multiobject experiment, subjects inspected a display containing spatially-fixed circular arrangement of 3 house and 3 face objects with the instruction to look at houses (house task, HT), at faces (face task, FT), or at houses and faces (all task, AT) with the purpose of a later recognition. A control test-object was shown after each multiobject stimulus presentation. Subjects had to answer by pressing “YES” and “NO” buttons to indicate whether the object was shown among the objects on the previous stimulus display (akin in Sternberg short-term recognition memory test). Random guessing baseline was 50%.

Figure 4 shows a scheme of the experiment with multiobject viewing. The experiment consisted of 4 blocks. Each block consisted of 4 screens.

*Screen 1:* Instruction (10 s). Subjects were given 3 types of tasks: “Look at faces” (FT), “Look at houses” (HT), or “Look at all objects” (AT).

*Screen 2:* Stimulus display consisting of 3 images of houses and 3 of faces, placed randomly in a circle at equal distances in 6 fixed positions (Figure 2 (a)). Stimulus duration varied pseudorandomly from 8 to

18 s. Due to the varying duration of the presentation, the blocks also varied in their total duration.

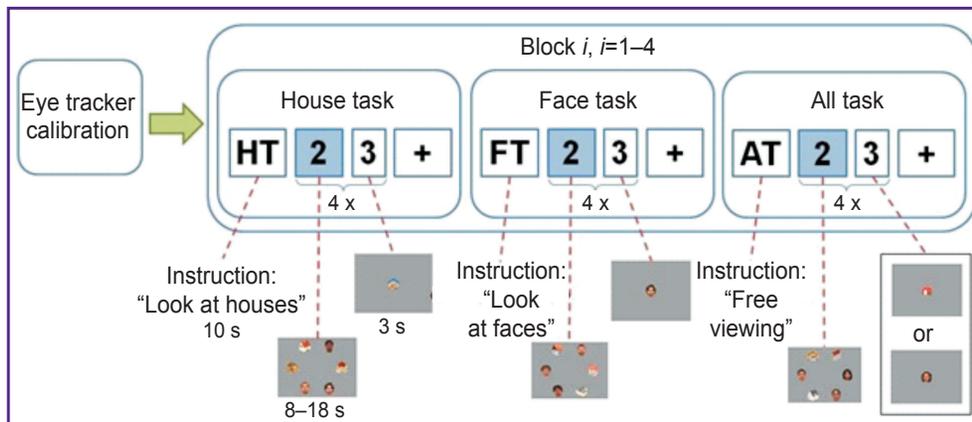
*Screen 3:* Control test-stimulus (3 s). The stimulus was a single house or face image (depending on the task) which was placed at the center of display.

*Screen 4:* Fixation cross (10 s).

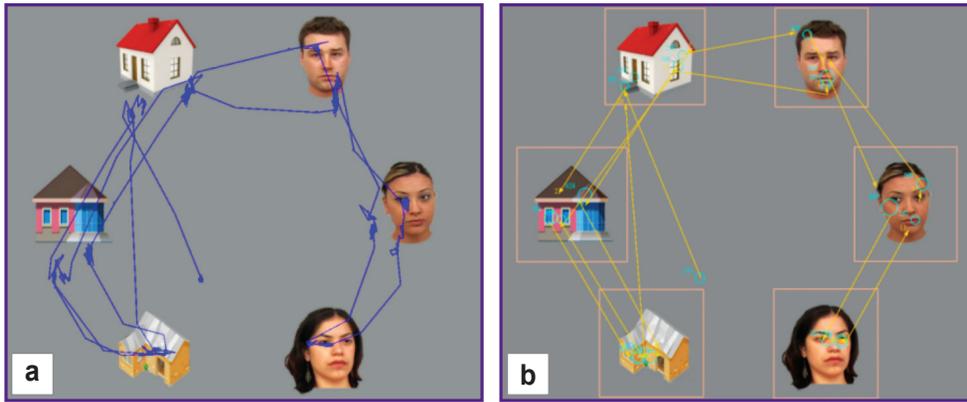
Each block included 3 groups of presentation cyclically repeated 4 times and differed by the type of task. Numbers 2 and 3 correspond to “Screen 2” and “Screen 3” in a block. Each block ended with the fixation cross. Eye tracking was employed only during “Screen 2”. This experiment was preceded by standard 9 points eye tracker calibration and lasted for about 20 min.

**Data processing.** For all subjects in the multiobject experiment, we acquired saccades and fixations data that were afterwards regularized using EyeLink Data Viewer (SR Research, Canada) software. Pictures of eyetracks preceding and following such regularization are shown in Figure 5. Each fixation was additionally marked depending on the category of the foveated object and actual instruction (type of task). Some fixations were excluded from the further modelling of hemodynamic responses (see below) according to the following criteria: 1) fixations shorter than 80 ms, 2) fixations longer than 1000 ms, and 3) fixations that were not clearly attributable either to houses or faces, e.g. located at the center of display presentation of the 6 objects. In our analysis of brain mechanisms behind ambient and focal visual processes, we considered a larger set of fixations by also including fixations not clearly attributable either to houses or faces. We further differentiated ambient and focal fixations on the basis of fixation length parameter using for decision approximately the median-threshold of 280 ms as it was done earlier [5]. To ensure that subjects followed the instruction, we automatically registered correctness of their responses in the recognition-memory test.

BOLD-data from localizer and multiobject experiments for each subject were preprocessed using SPM8 software (Wellcome Trust Centre for Neuroimaging,



**Figure 4. Scheme representation of the multiobject experiment with house (HT), face (FT), and free viewing (AT) conditions (see text for explanation)**



**Figure 5. An example of eye tracking data in free viewing task (AT):**  
 (a) raw eyetrack; (b) regularized eyetrack

London, UK) in Matlab R2018a (Mathworks, Natick, MA). Preprocessing consisted of realignment to correct for subject movements, co-registration to align all functional data to subject’s anatomical volume, normalization to convert all images to Montreal Neurological Institute (MNI) space and spatial smoothing with a Gaussian kernel of 8 mm (full width at half maximum).

From the data of the localizer experiment, two regions of interest (ROIs) were defined which were predominately involved in perception of houses or faces. In fMRI data analysis of multiobject experiment, we reconstructed the hemodynamic responses to study activity patterns in both ROIs revealed in the localizer experiment and in early visual cortex areas. The latter were defined as spherical regions of 10 mm radius in the visual cortex of each hemisphere around location with MNI coordinates  $-34 -71 -16$  in the left hemisphere and  $28 -71 -16$  in the right hemisphere.

Differences in the BOLD responses evoked by each fixation category (fixations on houses or on faces) were investigated by modelling their associated hemodynamic responses. At the single subject level, a model was defined using both the onsets and the durations of fixations of both categories corrected for a typical delay of the BOLD hemodynamic response function (HRF). These models were estimated in SPM8 (Restricted Maximum Likelihood estimation) using the informed basis set represented by HRF amplitude, derivative and dispersion [8, 17]. In the final reconstruction of brain structures involved in ambient and focal visual processing, we used individual resting-state data as baseline.

**Results**

**Eye movements and memory performance.** During the multiobject experiment, we registered nearly 80,000 visual fixations. Their proportion according to task at hand and exclusion criteria is shown in Table 1. Most of the excluded fixations belonged to the group of “not

clearly attributable either to houses or faces”. At the end of our study, they were again taken into account in the analysis of brain mechanisms behind ambient and focal processing, the proportion of considered fixations then increased to more than 92%. The data summarized in Table 1 show that subjects generally followed task instruction by looking predominately at the objects of appropriate category. In the group of “wrong” fixations, we found 3 times more fixations on faces during the HT than on houses in the FT. This result evidences a strong bias for fixating faces compared to houses. A similar tendency can be seen even in the balanced AT condition.

In order to demonstrate directly that the subjects performed the required tasks, we examined their scores on the Sternberg recognition-memory task. Average performance was 85% ( $\pm 10\%$  SD) correct for both the HT and FT and 63% ( $\pm 12\%$  SD) for the AT. The difference in performance is not surprising given that the AT condition requires remembering double the number of items. One can conclude from the data that subjects

**Table 1**  
**Mean number — and respective percents — of visual fixations used in fMRI analysis of multiobject experiment (data for one “averaged” subject from the group of 13 participants)**

Task	Foveated object	Code	Mean number of fixations	%	
All	House	H-AT	730	11.9	
House	House	H-HT	1370	22.3	
Face	House	H-FT	40	0.7	
All	Face	F-AT	960	15.6	
House	Face	F-HT	130	2.1	
Face	Face	F-FT	1395	22.7	
Excluded fixations			N	1520	24.7
Total			6145	100	

performed the tasks appropriately. These scores were not used in the further fMRI analyses.

In view of a large number of earlier studies, the time course of relation between fixation lengths and saccade amplitudes during the first several seconds after presentation of complex images is of particular interest. Figure 6 shows the results of such analysis. Contrary to the bulk of previous reports [3, 6, 8, 18], our data show no trends of growing fixation durations and falling saccadic amplitude. This unusual result could be related to the full spatial predictability of the multiobject displays in the present study. It also provides the reason for our selection of fixation length as the critical parameter to distinguish between ambient and focal modes of visual processing.

**fMRI data.** The localizer experiments led us to the description of the same two bilateral areas

predominately involved in perception of faces or houses as in pioneering works by Kanwisher and colleagues [14, 16] and in the referent FIBER study [8] namely fusiform face area (FFA) and parahippocampal place area (PPA). Their MNI centers of coordinates were as follows: FFA (left: -40 -50 -19; right: 42 -49 -17); PPA (left: -24 -49 -9; right: 26 -47 -10). The four ROIs were defined from the T-maps of the localizer experiment, using MarsBar [19] and uncorrected p-values of 0.01. Together with earlier defined regions of primary visual cortex, this resulted in 6 (or 3 bilaterally located) ROIs which were used in modelling hemodynamic response depending on the category of the foveated objects and the task at hand. The results of such modelling averaged across subjects and hemisphere localization are presented in Figure 7.

In the last phase of our analysis, we looked for

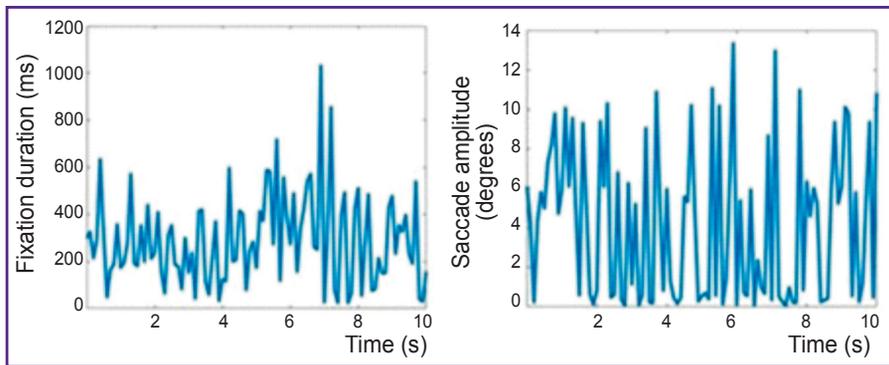


Figure 6. Fixation duration (left) and saccadic amplitude (right) over stimulus presentation time with a running average across 50 ms both

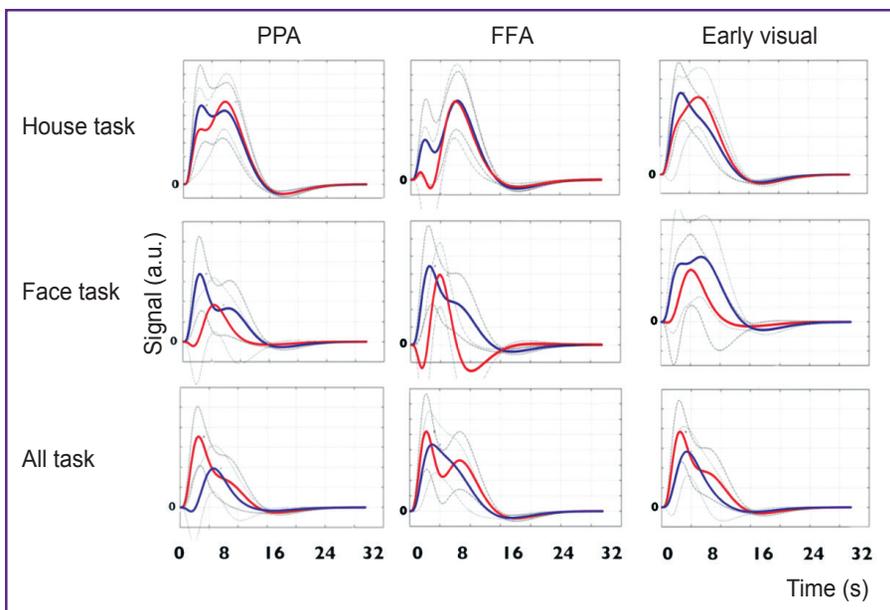
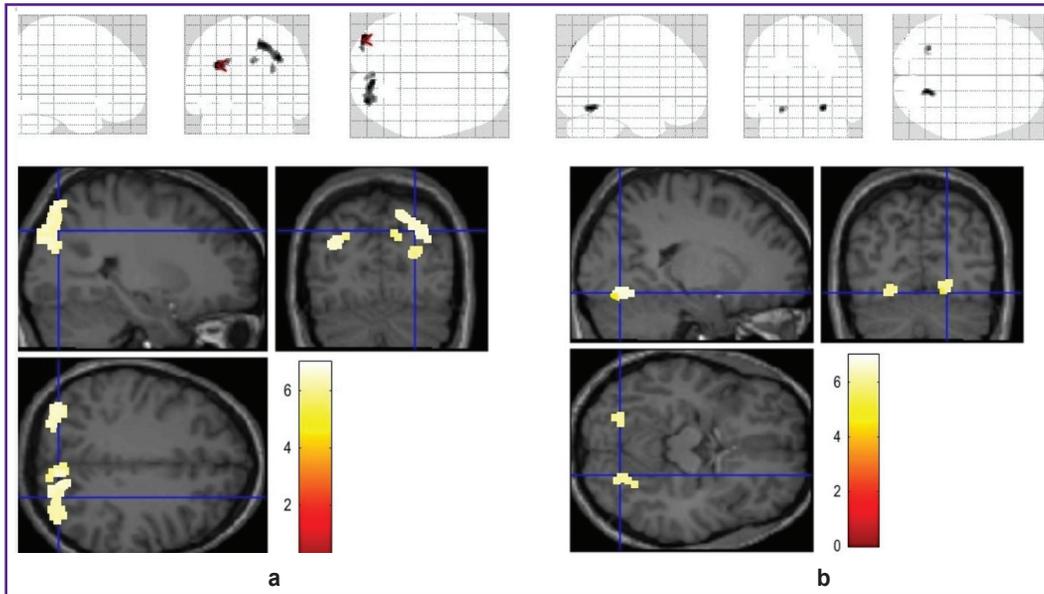


Figure 7. Modelled hemodynamic responses to house fixations (red) and face fixations (blue) with standard error of the mean (gray) in three task conditions for three ROIs: PPA, FFA, and early visual cortex (averaged models for all subjects and both hemispheres)



**Figure 8. Brain activity maps in the multiobject free viewing experiment:**  
 (a) ambient mode of visual processing (short fixations); (b) focal mode of visual processing (long fixations)

**Table 2**  
**Brain regions showing a significant increase in FIBER whole-brain fMRI activity associated with ambient and focal visual fixations (all significant changes to the resting-state baseline data,  $p < 0.0001$ , uncorrected)**

Brain structures	Center of clusters	Number of voxels
<i>Ambient modus (short fixations)</i>		
Occipital medial l.	-34 -88 30	295
Occipital medial r.	32 -84 36	293
Occipital superior l.	-24 -90 32	53
Occipital superior r.	18 -82 48	316
Parietal superior r.	24 -82 42	40
Cuneus r.	4 -78 32	272
<i>Focal modus (long fixations)</i>		
Fusiform l.	-26 -62 -16	60
Cerebellum VI r.	20 -64 -14	109

changes of the whole-brain functional activity in relation to ambient and focal visual modi of visual processing operationalized on the basis of fixation durations. Eye movement information for FIBER computing was used irrespective of fixating or not fixating particular objects. The contrasts with the resting state baseline data were drawn using a more stringent level of significance criterion than before:  $p$ -values  $< 0.0001$  (uncorrected). Our analyses revealed several regions in posterior parts of the brain that were activated in comparison to the resting state. No inhibited regions were discovered. The MNI coordinates of identified regions with numbers

of voxels involved are listed in the Table 2. The same results are presented in a graphical form in Figure 8.

### Discussion

In the present study, we followed two interrelated goals. First, we aimed at a substantial improvement of temporal resolution in BOLD-signal measurement up to a level sufficient for multimodal fMRI and eye tracking experiments. Our second goal was to describe brain functional activity in relation to ambient and focal fixations in free viewing complex visual displays. Using eye movement information as parameters for the FIBER fMRI analysis combined with ultrafast scanning protocol, we confirmed reports by Marsman, Renken, Velichkovsky, et al. [8] that the fMRI signal contains information to identify the task and executed and the type of object inspected by the subject. Moreover, in some respects we got a more elaborated pattern of results than in the previous study. This can be explained by a number of details that multiplied sensitivity of our method: a 5 times higher temporal resolution of the eye tracking (250 vs. 50 Hz), the whole-brain fMRI scanning within 720 ms (instead of 2 s) and the use of 32-channels MRI head coil (instead of one with 8 channels)<sup>4</sup>.

First of all, it can be seen that brain processing of both categories of objects — houses and faces — is rather dissimilar. Data on eye movements (see Table 1)

<sup>4</sup>Recently, one can observe a re-birth of interest to the FIBER type of analysis however up to date only in reading studies and without employment of ultrafast scanning protocols [20, 21].

show a strong bias to fixate faces and not houses while modelling of hemodynamic responses in the FT as well as in the HT conditions (see Figure 7) demonstrates a faster increase of reaction to faces in comparison to houses. This special character of face processing is widely discussed in the literature, for instance, in EEG event-related potentials studies that reliably demonstrated a face-specific reaction as early as 170 ms after stimulus presentation [22].

Similar arguments were once defined to consider the FFA as a prototypical module of brain cognitive organization [16]. However other aspects of results seem to challenge the modular interpretation of face processing, which should be mandatory task-independent and associated with a particular brain structure [23, 24]. One of the discrepancies is that similar hemodynamic responses to face images can be observed also in the PPA and, rather counter-intuitively, in the early visual cortex. One would need additional hypotheses to understand this broad similarity such as the assumption that hemodynamic responses have a travelling-wave nature [25, 26]. The next inconsistency is more difficult to explain. As one can see, the AT condition reverses the relations between the hemodynamic responses to faces and houses so that now the latter start to produce faster and stronger reactions at all studied brain locations including the early visual cortex. Such a reversal under the influence of a task is impossible in strictly modular architecture working as a mosaic of “cognitive reflexes”. In our view, the reversal can be related to subject’s attempts of counterbalancing the prioritization of faces in the AT condition that explicitly requires equal attention to both object categories.

If indeed the BOLD-signals may “travel” across different brain locations, one would expect the observed broad similarity of hemodynamic responses in dependence on the task at hand at intervals of several seconds. In contrast, the core specialization of particular brain structures would become less evident in such long time segments. Thus, for a sharp spatial localization one needs a maximum of temporal resolution. In the second part of our study, we applied ultrafast multi-band scanning protocol together with FIBER analysis to reveal brain mechanisms behind ambient and focal visual fixations as events on a subsecond scale.

There are currently two alternative explanations for these brain mechanisms: one stresses the division of dorsal and ventral pathways in perception [4], another emphasizes interhemispheric differences in global and local modes of processing [6]. The results of this analysis summarized in Table 2 and Figure 8 show that both hypotheses are to some extent confirmed. In line with our early proposal, ambient fixations were accompanied by an activation of structures associated with the dorsal visual pathway, while focal fixations correlated with that of the ventral pathway. At the same time, the second hypothesis is also proved to be correct: the activated structures of the dorsal pathway were localized in

the right hemisphere and those of the ventral stream mainly — albeit not exclusively — in the left hemisphere.

More specifically, short fixations associated with ambient mode of processing correlate with bilateral activation of early visual cortex (BA17 left and right) and a number of structures that increasingly demonstrate rightward lateralization. The asymmetry is already present in the superior occipital gyrus (BA19): 53 activated voxels on the left vs. 316 on the right side. Area 19 is considered to be the differentiation point of the two visual streams [27] and probably receives inputs from the retina via the superior colliculus and pulvinar. Furthermore, only right-hemisphere lateralized parts of the cuneus and the superior parietal lobule are activated. The latter is considered as the core region of the dorsal stream involved with multimodal functions of the parietal lobe in general, such as reaching and grasping behaviour [28]. In contrast, long visual fixations associated with focal mode of processing are connected to activation of the fusiform gyrus in the left hemisphere. The fusiform gyrus is a classical structure of the ventral stream. A novel result is also a locus of activation within the lobule VI of the right cerebellum. One has to note that the right cerebellar lobule VI is involved in the tasks which have a cognitive component and simultaneously activate regions of the left hemisphere [29]. Overall, cross-lateral connectivity seems to be a frequent case in building functional systems of neocortical areas and that of the cerebellum [30].

The current reconciliation of competing views on the brain mechanisms behind ambient and focal fixations poses at least three further research questions. The first one is about the way these mechanisms actually work: it is necessary to clarify cause-and-effect connections among the described structures, e.g. by applying the dynamic causal modelling [10, 31, 32]. The second question is about general biological and computational reasons behind this and other forms of hemispheric asymmetries [33]. Lastly, an unexplored issue remains the role of brain structures located upstream of the dorsal and ventral visual streams. In a speculative way, one can suppose that ventromedial vs. ventrolateral locations of activity clusters that were attributed to ambient and focal modi of eye movements in one of the studies [9] could be related to a projection of the dorsal and ventral visual pathways on the deep cortical layers around hippocampal formation [34], where they also become to play a role in different forms of memory [35]. A low temporal resolution of previous studies may be the reason that only these late components of brain activity involved with free visual behavior were revealed.

## Conclusion

In the present study, we realized the use of fixations in free viewing as events to model brain activation in an experiment using ultrafast multi-band scanning

protocol and FIBER paradigm of data analysis. This allowed us for the first time to demonstrate contrasting lateralization of brain mechanisms belonging to dorsal and ventral streams in charge of visual perception and eye movements in free processing of multiobject images. Therefore the new multimodal imaging methodology could make a definitive contribution to the exploring brain mechanisms behind ambient and focal modes of active vision. A next step in this line of research would be a modelling of effective connections among neural networks involved with active vision. The conclusion also poses a number of further questions about a possible relation between two modes of active vision and other structural and functional asymmetries found at different levels of the human brain functional organization.

**Acknowledgment.** This work was partially supported by the NRC “Kurchatov Institute” (grant No.1359). The development of algorithms for classification of eye movements in fMRI studies was funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) as part of Germany’s Excellence Strategy — EXC 2050/1 — Project ID 390696704 — Cluster of Excellence “Centre for Tactile Internet with Human-in-the-Loop” (CeTI) of Technische Universität Dresden.

We thank Dr. V. Verkhutov for sharing with us his knowledge on the travelling-wave processes in the human brain.

**Conflicts of interest.** The authors have no conflicts of interest to disclose.

## References

- Ito J., Yamane Y., Suzuki M., Maldonado P., Fujita I., Tamura H., Grün S. Switch from ambient to focal processing mode explains the dynamics of free viewing eye movements. *Sci Rep* 2017; 7(1): 1082, <https://doi.org/10.1038/s41598-017-01076-w>.
- Pannasch S., Velichkovsky B.M. Distractor effect and saccade amplitudes: further evidence on different modes of processing in free exploration of visual images. *Visual Cognition* 2009; 17(6–7): 1109–1131, <https://doi.org/10.1080/13506280902764422>.
- Unema P., Pannasch S., Joos M., Velichkovsky B.M. Time-course of information processing during scene perception: the relationship between saccade amplitude and fixation duration. *Visual Cognition* 2005; 12(3): 473–494, <https://doi.org/10.1080/13506280444000409>.
- Velichkovsky B.M., Joos M., Helmert J.R., Pannasch S. Two visual systems and their eye movements: evidence from static and dynamic scene perception. In: Bara B.G., Barsalou L., Bucciarelli M. (editors). *Proceedings of the XXVII annual conference of the Cognitive Science Society*. Mahwah: Lawrence Erlbaum 2005; p. 2283–2288.
- Velichkovsky B.M., Rother A., Kopf M., Dornhöfer S.M., Joos M. Towards an express-diagnostics for level of processing and hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour* 2002; 5(2): 145–156, [https://doi.org/10.1016/s1369-8478\(02\)00013-x](https://doi.org/10.1016/s1369-8478(02)00013-x).
- Mills M., Alwatban M., Hage B., Barney E., Truemper E.J., Bashford G.R., Dodd M.D. Cerebral hemodynamics during scene viewing: hemispheric lateralization predicts temporal gaze behavior associated with distinct modes of visual processing. *J Exp Psychol Hum Percept Perform* 2017; 43(7): 1291–1302, <https://doi.org/10.1037/xhp0000357>.
- Velichkovsky B.M., Cornelissen F., Geusebroek J.-M., Graupner S.-Th., Hari R., Marsman J.B., Shevchik S.A. Measurement-related issues in investigation of active vision. In: Berglund B., Rossi G.B., Townsend J., Pendrill L. (editors). *Measurement with persons: theory and methods*. London-New York: Taylor and Francis 2012; p. 281–300.
- Marsman J.B., Renken R., Velichkovsky B.M., Hooymans J.M.M., Cornelissen F.W. Fixation-based event-related fMRI analysis: using eye fixations as events in functional magnetic resonance imaging to reveal cortical processing during the free exploration of visual images. *Hum Brain Mapp* 2012; 33(2): 307–318, <https://doi.org/10.1002/hbm.21211>.
- Marsman J.C., Renken R., Haak K.V., Cornelissen F.W. Linking cortical visual processing to viewing behavior using fMRI. *Front Syst Neurosci* 2013; 7: 109, <https://doi.org/10.3389/fnsys.2013.00109>.
- Parr T., Mirza M.B., Cagnan H., Friston K.J. Dynamic causal modelling of active vision. *J Neurosci* 2019; 39(32): 6265–6275, <https://doi.org/10.1523/jneurosci.2459-18.2019>.
- Golay X., Pruessmann K.P., Weiger M., Crelier G.R., Folkers P.J., Kollias S.S., Boesiger P. PRESTO-SENSE: an ultrafast whole-brain fMRI technique. *Magn Reson Med* 2000; 43(6): 779–786, [https://doi.org/10.1002/1522-2594\(200006\)43:6<779::aid-mrm1>3.0.co;2-4](https://doi.org/10.1002/1522-2594(200006)43:6<779::aid-mrm1>3.0.co;2-4).
- Lin F.-H., Tsai K.W.K., Chu Y.-H., Witzel T., Nummenmaa A., Raji T., Ahveninen J., Kuo W.-J., Belliveau J.W. Ultrafast inverse imaging techniques for fMRI. *Neuroimage* 2012; 62(2): 699–705, <https://doi.org/10.1016/j.neuroimage.2012.01.072>.
- Korosteleva A., Ushakov V., Malakhov D., Velichkovsky B.M. Event-related fMRI analysis based on the eye tracking and the use of ultrafast sequences. *Advances in Intelligent Systems and Computing* 2017; 636: 107–112, [https://doi.org/10.1007/978-3-319-63940-6\\_15](https://doi.org/10.1007/978-3-319-63940-6_15).
- Epstein R., Kanwisher N. A cortical representation of the local visual environment. *Nature* 1998; 392(6676): 598–601.
- Grill-Spector K., Knouf N., Kanwisher N. The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci* 2004; 7(5): 555–562, <https://doi.org/10.1038/nn1224>.
- Kanwisher N., McDermott J., Chun M.M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 1997; 17(11): 4302–4311, <https://doi.org/10.1523/jneurosci.17-11-04302.1997>.
- Friston K.J., Josephs O., Rees G., Turner R. Nonlinear event-related responses in fMRI. *Magn Reson Med* 1998; 39(1): 41–52, <https://doi.org/10.1002/mrm.1910390109>.
- Pannasch S., Helmert J.R., Roth K., Herbold A.K., Walter H. Visual fixation durations and saccade amplitudes: shifting relationship in a variety of conditions. *Journal of Eye Movement Research* 2008; 2(2): 4.
- Brett M., Anton J.L., Valabregue R., Poline J.B. Region of interest analysis using an SPM toolbox. In: *Presented at the 8th International Conference on Functional Mapping of the Human Brain*. Sendai, Japan; 2002.

20. Hsu C.-T., Clariana R., Schloss B., Li P. Neurocognitive signatures of naturalistic reading of scientific texts: a fixation-related fMRI study. *Sci Rep* 2019; 9(1): 10678, <https://doi.org/10.1038/s41598-019-47176-7>.
21. Schuster S., Hawelka S., Himmelstoss N.A., Richlan F., Hutzler F. The neural correlates of word position and lexical predictability during sentence reading: evidence from fixation-related fMRI. *Lang Cogn Neurosci* 2019, <https://doi.org/10.1080/23273798.2019.1575970>.
22. Kaltwasser L., Hildebrandt A., Recio G., Wilhelm O., Sommer W. Neurocognitive mechanisms of individual differences in face cognition: a replication and extension. *Cogn Affect Behav Neurosci* 2014; 14(2): 861–878, <https://doi.org/10.3758/s13415-013-0234-y>.
23. Fodor J.A. *The modularity of mind: an essay on faculty psychology*. Cambridge, MA: MIT Press; 1983.
24. Velichkovsky B.M. Modularity of cognitive organization: why it is so appealing and why it is wrong. In: Callebaut W., Rasskin-Gutman D. (editors). *Modularity: understanding the development and evolution of natural complex systems*. Cambridge, MA: MIT Press; 2005; p. 335–356.
25. Aquino K.M., Schira M.M., Robinson P.A., Drysdale P.M., Breakspear M. Hemodynamic traveling waves in human visual cortex. *PLoS Comput Biol* 2012; 8(3): e1002435, <https://doi.org/10.1371/journal.pcbi.1002435>.
26. Genç E., Bergmann J., Singer W., Kohler A. Surface area of early visual cortex predicts individual speed of traveling waves during binocular rivalry. *Cereb Cortex* 2015; 25(6): 1499–1508, <https://doi.org/10.1093/cercor/bht342>.
27. Thiebaut de Schotten M., Urbanski M., Valabregue R., Bayle D.J., Volle E. Subdivision of the occipital lobes: an anatomical and functional MRI connectivity study. *Cortex* 2014; 56: 121–137, <https://doi.org/10.1016/j.cortex.2012.12.007>.
28. Milner A.D., Goodale M.A. Two visual systems reviewed. *Neuropsychologia* 2008; 46(3): 774–785, <https://doi.org/10.1016/j.neuropsychologia.2007.10.005>.
29. Stoodley C.J., Valera E.M., Schmahmann J.D. Functional topography of the cerebellum for motor and cognitive tasks: an fMRI study. *Neuroimage* 2012; 59(2): 1560–1570, <https://doi.org/10.1016/j.neuroimage.2011.08.065>.
30. Van Overwalle F., Van de Steen F., Marien P. Dynamic causal modeling of the effective connectivity between the cerebrum and cerebellum in social mentalizing across five studies. *Cogn Affect Behav Neurosci* 2019; 19(1): 211–223, <https://doi.org/10.3758/s13415-018-00659-y>.
31. Chen J., Snow J.C., Culham J.C., Goodale M.A. What role does “Elongation” play in “Tool-Specific” activation and connectivity in the dorsal and ventral visual streams? *Cereb Cortex* 2018; 28(4): 1117–1131, <https://doi.org/10.1093/cercor/bhx017>.
32. Ushakov V.L., Sharaev M.G., Kartashov S.I., Zavyalova V.V., Verkhlyutov V.M., Velichkovsky B.M. Dynamic causal modeling of hippocampal links within the human default mode network: lateralization and computational stability of effective connections. *Front Hum Neurosci* 2016; 10: 528, <https://doi.org/10.3389/fnhum.2016.00528>.
33. Velichkovsky B.M., Nedoluzhko A.V., Goldberg E., Korosteleva A.N., Efimova O.I., Sharaev M.G., Ushakov V.L. New insights into the human brain’s cognitive organization: views from the top, from the bottom, from the left and, particularly, from the right (submitted).
34. Dickerson B.C., Eichenbaum H. The episodic memory system: neurocircuitry and disorders. *Neuropsychopharmacology* 2010; 35(1): 86–104, <https://doi.org/10.1038/npp.2009.126>.
35. Ranganath C., Ritchey M. Two cortical systems for memory-guided behavior. *Nat Rev Neurosci* 2012; 13(10): 713–726, <https://doi.org/10.1038/nrn3338>.