# Feature Binding in the Visual Modality Depends on Attention: Analysis of Mismatch Negativity

K. E. Sayfulina,<sup>1,2,</sup> I. E. Lazarev,<sup>1,3</sup> E. G. Chernysheva,<sup>1</sup> and B. V. Chernyshev<sup>1,2</sup>

## UDC 612.821.2+159.91+159.931+159.952

Translated from Zhurnal Vysshei Nervnoi Deyatel'nosti imeni I. P. Pavlova, Vol. 69, No. 1, pp. 37–50, January–February, 2019. Original article submitted October 23, 2017. Revised version received January 15, 2018. Accepted March 5, 2018.

This study address the question of whether attention is required for binding of features in the visual modality. Subjects performed a task based on discrimination of visual stimuli – Gabor grids – characterizing two features: the spatial frequency and the tilt angle. Deviant stimuli could be detected on the background of standard stimuli only using a combination of features, not single features. Event-related potentials were analyzed in four experimental conditions: selective attention to the target stimulus; selective neglect of the nontarget stimulus; distributed attention to all visual stimuli; intermodal distraction of attention from the visual modality to the auditory modality. Mismatch negativity was significantly present only in the situation of attention to visual stimuli – both selective and distributed. These results showed that binding of features occurred only in the situation of attention to visual stimuli.

Keywords: event-related potentials, feature binding, mismatch negativity, oddball, visual perception.

**Introduction.** Sensory perception of objects in the surrounding world and initialization of adaptive behavioral responses requires not only analysis of a multitude of the individual features of the object, but also their integration, supporting formation of integral representations of each object in the brain [Kahneman et al., 1992; Hommel, 1998, 2004; Schneegans, 2017]. The basic physiological mechanisms encoding the features of sensory stimuli are now generally well known [Bartels and Zeki, 1998; Livingstone and Hubel, 1988], though the search for the mechanisms of integration, or feature binding, continues to remain relevant. It is significant that the extent of the involvement of top-down processes in the mechanism remains unclear: primarily, whether attention is required [Treisman, 1996] or whether binding occurs exclusively as a result of low-level sensory processes.

According to Treisman's feature integration theory, the perception of objects occurs in two stages: at the first step, the features of the object are processed by the brain separately, automatically, in parallel, and independently, and this does not require attention; the second step, whose success requires attention, combines these features, resulting in the formation of an overall representation of the object [Treisman and Gelade, 1980]. The theory is based mainly on the results of behavioral experiments: thus, visual seeking for a specified object among others takes longer when it differs from the others in terms of a combination of features than when it differs in terms of a single feature; false conjunctions have been shown to be made in conditions of insufficient attention, i.e., erroneous binding of features belonging to different objects [Treisman and Gelade, 1980].

The opposite point of view holds that feature binding is an automatic process occurring in the brain at the early stages of processing of sensory signals. This view is based on the results of analysis of mismatch negativity (MMN), which is regarded as a neurophysiological correlate of early pre-attention processes [Näätänen, 1990; Kimura et al.,

<sup>&</sup>lt;sup>1</sup>National Research University Higher School of Economics (HSE), Moscow, Russia; e-mail: kseniasayfulina@gmail.com.

<sup>&</sup>lt;sup>2</sup> Lomonosov Moscow State University, Moscow, Russia.

<sup>&</sup>lt;sup>3</sup> Institute of Medical Biological Problems, Russian Academy of Sciences, Moscow, Russia.

## Sayfulina, Lazarev, Chernysheva, and Chernyshev



Fig. 1. Visual stimuli and procedural scheme of experiment. *a*) Visual stimuli (Gabor grids): left rare (LR), right rare (RR), left frequent (LF), right frequent (RF), horizontal rare (HR), vertical rare (VR). *b*) Procedural scheme of experiment. Sequences of visual and auditory stimuli are presented simultaneously and independently of each other.

2011; Michie et al., 2016]. Winkler et al. [2005] demonstrated the occurrence of mismatch negativity in response to deviant stimuli characterized by a combination of features in conditions of intermodal distraction of attention. This led to the conclusion that attention is not required for feature binding. It should be noted that the distracting task in these experiments may have been insufficiently difficult, such that the subject's attention to the deviant stimulus could not in fact be excluded. In addition, this study did not analyze the effects of the distribution of attention between the target and nontarget stimuli within a single visual stream of stimuli.

The aim of the present work was to study the influence of attention on feature binding in visual stimuli by analysis of MMN. The study was carried out in experimental conditions with deeper manipulations of targeted attention: 1) attention was directed to a particular combination (conjunction) of features in the visual modality (selective attention); 2) attention was directed to another combination of features in the same modality, while this combination of stimuli was subjected to neglect (intramodal distraction of attention); 3) attention was distributed for processing of all stimuli in the ongoing modality (distributed attention); 4) attention was distracted to solving a task in another modality (intermodal distraction of attention). We used a sequence of visual stimuli in which the deviant (common) stimuli differed from the standard (rare) stimuli only in terms of a combination of features and not by individual features. Focusing on the fact that the MMN phenomenon is highly reliable and stable [Kimura et al., 2011] and is clearly apparent even when there are complex contextual differences between a specific stimulus from the series of preceding stimuli [Sussman et al., 2007], and also on the fact that the

# Feature Binding in the Visual Modality

deviant stimuli used in our sequences could be identified exclusively from a combination of features and not from features taken separately, the present study considered MMN as an indicator of feature binding.

We suggested that mismatch negativity would be clearly apparent only in the situation of attention to the corresponding deviant stimuli (both selective and, to a lesser extent, distributed). In the situation of a lack of attention to the deviant stimulus, mismatch negativity will be absent because the deviant nature of the stimuli cannot be detected without the necessary integration of features.

**Methods.** A total of 28 volunteer aged 18–28 (mean  $\pm$  standard deviation 20.7  $\pm$  2.8) years, 14 female, with no neurological or mental disorders, and with normal vision and hearing took part in the study. All subjects were instructed to have a good sleep before the experiment and to avoid consuming stimulants such as strong coffee or tea. All subjects reported good wellbeing before the experiment started after provision of signed informed consent to take part.

Visual stimuli were monochrome Gabor grids of size  $4.9^{\circ}$  with spatial frequencies of  $0.6^{\circ}$  and  $0.3^{\circ}$  and tilt angles of  $45^{\circ}$ ,  $0^{\circ}$ ,  $-45^{\circ}$ , and  $90^{\circ}$  (Fig. 1, *a*). Stimuli were presented at the center of the screen on a gray background; stimulus presentation time was 200 msec; a central fixation stimulus (a small black cross) was displayed at the center of the screen during the intervals between stimulus presentations; subjects were instructed to watch the center of the screen. Auditory stimuli were seven Russian vowels [a], [o], [u], [i], [yu], [ye], and [e] presented using a text-to-speech conversion program. The duration of each sound stimulus was 350 msec.

Stimuli were presented and subjects' responses were recorded using the Presentation program (Neurobehavioral Systems Inc., USA).

The main part of the experiment included four types of blocks in which a completely identical procedure was used for presentation of visual and auditory stimuli (sequences of visual and sequences of auditory stimuli were presented in all blocks in parallel and independently throughout the block (see Fig. 1, b) and blocks differed exclusively in terms of the instruction given to the subject.

Visual stimuli were four types of Gabor grid, differing in terms of the combination (conjunction) of two features (spatial frequency and tilt angle): rare grids with a left tilt (LR, with spatial frequency  $0.6^{\circ}$  and tilt angle  $-45^{\circ}$ ), rare grids with right tilt (RR,  $0.6^{\circ}$ ,  $+45^{\circ}$ ), frequent grids with left tilt (LF,  $0.3^{\circ}$ ,  $-45^{\circ}$ ), and frequent grids with right tilt (RF,  $0.3^{\circ}$ ,  $+45^{\circ}$ ) (Fig. 1, *a*). Two stimuli – LR and RF – were standard (frequent), each with a presentation frequency of 45% in visual stimulus sequences, while the other two – LF and RR – were deviants (rare), each presented at a frequency of 5%. Stimuli were presented in pseudorandom order. The time interval between the moments at which visual stimulus presentations started varied pseudorandomly over the range  $800 \pm 100$  msec (uniformly distributed). In all the main blocks of the experiment, auditory stimulation was carried out simultaneously with visual stimulation – presentation of seven Russian vowels in pseudorandom order with intervals of 2200 msec.

The duration of each block was 8 min; each block included presentation of 565 visual stimuli and 205 auditory stimuli.

*Visual block (VB).* The subject was instructed to press a button in response to presentation on the screen of only one of the two deviant (rare) stimuli, as specified by the instruction (Fig. 1, *a*); during reading of the instruction, the image of one target stimulus and the other three stimuli were shown to the subject on the screen. Each subject performed two different VB, in which the target and nontarget deviant stimuli were swapped (in one block the RR deviant was the target stimulus and the LF stimulus the nontarget stimulus, and the stimuli were the other way round in the other block – LF was the target stimulus and RR was the nontarget stimulus); the order of the two variants of this block was balanced between groups of subjects. The subject was given the instruction to ignore auditory stimuli.

*Motor block (MB)*. The subject was given the instruction to press the button in response to all visual stimuli (without carrying out the stimulus discrimination task) and ignore auditory stimuli. MB was always presented between two VB with the aim of minimizing learning effects from the preceding blocks after swapping the target and nontarget stimuli.

Auditory block (AB). The subject was given the instruction to carry out the "2-back" auditory task – to press the button at the moment at which he or she heard the sound, which was presented before this one position back (for example, with the sequence of presentation [a], [o], [a], the subject had to respond to the second [a] by pressing the button). A pseudorandom sequence was generated such that all such combinations were used at a frequency of 1:3. These subject received the instruction to ignore the visual stimuli, but to watch the screen in the same way as in the other blocks.

The order of blocks was balanced over groups of subjects. Before experiments started, subjects were assigned one of four sequences of blocks, with the same probability: AB-VB1-MB-VB2,AB-VB2-MB-VB1,VB1-MB-VB2-AB, and VB2-MB-VB1-AB, where VB1 is the visual block with the RR target stimulus, VB2 is the active visual block with the LF target stimulus, MB is the motor block, and AB is the auditory block.

The experimental procedure also included a "passive oddball" control block and a "training block."

*Passive oddball block.* This additional block was the first in the experiment and lasted 230 sec. Visual stimuli were Gabor grids in the vertical and horizontal orientations and low spatial frequency (Fig. 1, a). The ratio of standard and deviant stimulus presentation frequencies was 9:1. The time interval between the starts of stimulus presentations was  $1000 \pm 50$  msec. The duration of stimulus presentation on the screen was 200 msec. Auditory stimulation was not



Fig. 2. Averaged visual ERP for the "passive oddball" block. *a*) Averaged ERP in the occipital region of interest (O1, O2, Oz); *b*) averaged map of difference ERP in the interval 100–180 msec.

presented and subjects were instructed to watch the screen passively without any additional tasks. This control block was used to assess the level, spatial localization, and time of generation of MMN in response to deviant stimuli differing from the standard stimuli in terms of only one feature rather than the combination of two features in the conditions of the present experiments.

The training block before the auditory block was used to train subjects and always preceded the auditory block; its duration was 5 min. During this block, no visual stimuli were presented on the screen, and only sound stimuli were delivered. If the proportion of correct responses by the subject was below 0.5, the block was repeated.

Electroencephalogram traces were made using an acti-CHamp electroencephalograph (Brain Products, Germany) and the program package PyCorder (Brain Products, Germany). Traces were made from 60 active silver chloride electrodes in an actiCap (Brain Products, Germany) positioned in accordance with the international 10–10 system. The reference consisted of combined mastoid electrodes. Electrooculographic electrodes were also positioned for recording the vertical and horizontal oculogram. The resistance of all electrodes was less than 10 k $\Omega$ .

Processing was in Matlab (MathWorks Inc., USA) using EEGLAB [Delorme and Makeig, 2004] and ERPLAB [Lopez-Calderon and Luck, 2014]. Processing used scripts accessing the built-in functions of this software. Primary data processing included removal of various types of artifact not associated with eye movements from traces and interpolating channels with low-quality EEG traces from neighboring channels. The next stage in processing consisted of removal of eye movement artifacts using the independent components analysis (ICA) method. EEG traces were filtered in the range 0.1–30 Hz. The null line for construction of event-related potentials (ERP) was adjusted using the 100-msec prestimulus interval. ERP were averaged separately for each condition.

The experiment design provided for analysis of ERP for four conditions:

1) active selective attention (ERP to target deviant stimuli in active visual blocks, TVB);

2) active neglect with intramodal transfer of attention (nontarget ignored deviant stimuli in active visual blocks, NVB);

3) distributed attention in the visual modality (ERP to deviant stimuli in the motor block – MB);

4) intermodal distraction of attention (ERP to deviant stimuli with "auditory block," AB).

The differential ERP needed for detection of MMN were calculated for each subject by subtracting the mean amplitude of ERP to standard stimuli from the mean ERP amplitude to deviant stimuli (the analysis used only standard stimuli from the epoch immediately preceding the deviant stimuli in accordance with the scheme of each of the four conditions defined above).

The time interval for measurement of mismatch negativity of 100–180 msec was selected on the basis of published data [Czigler and Csibra, 1990; Winkler et al., 2005; Farkas et al., 2015] and in accordance with results from visual analysis of ERP in terms of the presence of a negative deviation on differential ERP (Fig. 2, a).

The occipital area of interest for measurements of mismatch negativity was in electrodes O1, Oz, and O2, selected on the basis of published data [Kimura et al., 2011; Pazo-Alvarez et al., 2003; Winkler et al., 2005] and also on the basis of visual analysis of maps of difference potentials showing that the maximum amplitude of the negative wave was seen in occipital electrodes O1, Oz, and O2 (Fig. 2, *b*).

The N2c component was measured using the time interval 250–300 msec in the area of interest O1, Oz, and O2 on the basis of published data [Luck and Kappenman, 2011] and visual analysis of ERP.

The P300 component was analyzed over the time interval 360–600 msec in the parietal area of interest: electrodes Pz, CPz, P1, P2, and POz, selected on the basis of published data [Polich, 207] and visual analysis of ERP with the observation of a maximum extent of this component to the target stimulus.

Amplitudes in different leads in the zones of interest were averaged separately for each ERP component and each condition, as well as for each time point in the intervals selected.



Fig. 3. Averaged visual ERP for deviants and standards and difference ERP for four conditions: target stimulus in active visual block (TVB); nontarget stimulus in active visual block (NTB), motor block (MB); auditory block (AB). *a*) Averaged ERP in the occipital region of interest (O1, O2, Oz); gray rectangles identify the intervals 100–140, 140–180, and 250–300 msec; \*p < 0.05, \*\*p < 0.01 (significant differences between ERP to deviants and standards in the corresponding intervals, with Bonferroni correction). *b*) Averaged maps of difference ERP for four conditions in the interval 100–180 msec, corresponding to the MMN generation time. *c*) Averaged maps of difference ERP for four conditions, averaged amplitude in the interval 250–300 msec, corresponding to the N2c generation time.

The reliability of the presence of different ERP components in each experimental condition was assessed by analysis of variance of ERP amplitude to deviant stimuli and ERP amplitude to standard stimuli. The Bonferroni correction for multiple comparisons was used by multiplying p values by 4 (the number of experimental conditions).

**Results.** The analysis included results obtained in experiments on 21 subjects; seven subjects were excluded from the analysis because of lack of success in performing

the experimental tasks and/or because of poor quality EEG/ ERP traces.

Behavioral results. In the two active visual blocks (VB), subjects gave  $84.1 \pm 12.3\%$  correct responses ( $84.7 \pm 11.9\%$  with the RR target stimulus,  $83.5 \pm 12.7\%$  with the LF target stimulus). The proportion of presses in response to presentation of nontarget stimuli was 0.55% of all erroneous presses (including 0.08% in response to nontarget LF stimuli and 0.47% in response to nontarget RR stimuli). In the



Fig. 4. Averaged visual evoked potentials. *a*) MMN (100–180 msec, occipital region of interest – O1, O2, Oz); *b*) early phase of MMN (100–140 msec, occipital region of interest – O1, O2, Oz); *c*) late phase of MMN (140–180 msec, occipital region of interest – O1, O2, Oz); *c*) late phase of MMN (140–180 msec, occipital region of interest – O1, O2, Oz); *c*) and PO2 (250–300 msec, occipital region of interest – O1, O2, Oz); *e*) P300 (360–600 msec, parietal region of interest (Pz, CPz, P1, P2, and POz). Blocks designated as in Fig. 3.

motor block (MB), subjects responded to presentation of the visual stimulus in  $86.5 \pm 7.0\%$  of cases. In the auditory block (AB), subjects gave correct responses on performance of the "2-back" task in  $88.3 \pm 11.3\%$  of cases.

#### Sayfulina, Lazarev, Chernysheva, and Chernyshev

*Mismatch negativity*. Mismatch negativity in the "passive oddball" block was present in the occipital leads and was marked over the time interval 100–180 msec (Fig. 2).

As shown in Fig. 3, *a*, negative deviations in the difference ERP between deviants and standards were present in at least part of the time interval 100–180 msec for all conditions, though the extent varied strongly. Mismatch negativity was reliably present in the time interval 100–180 msec in the conditions TVB (p < 0.01, Bonferroni correction) and MB (p < 0.01, Bonferroni correction) and had an occipital location (Fig. 3, *a*; Fig. 3, *b*; Fig. 4, *a*). In the other two conditions – NVB and AB – reliable mismatch negativity was not seen ( $p \gg 0.05$  with and without the Bonferroni correction)) (Fig. 3, *a*; Fig. 3, *b*; Fig. 4, *a*).

As shown in Fig. 3, a, the difference event-related potential can be divided into two phases – an early phase (100–140 msec) and a late phase (140–180 msec).

The early phase was significantly expressed in the MB condition (p < 0.01) and was not apparent in the TVB condition, while the late phase was marked for TVB (p < 0.01) and was not seen in MB (p > 0.05) (Fig. 3, a; Fig. 4, b; Fig. 4, c). Thus, in response to target stimuli in VB, the effect was seen in a later time interval than in MB. For the other two conditions – AB and NVB – mismatch negativity was not seen in either of the time intervals ( $p \gg 0.05$ ) (Fig. 3, a; Fig. 4, b; Fig. 4, c).

The N2c component. In the interval 250–300 msec, mismatch negativity of the N2c wave was seen only in the TVB condition (p < 0.05); it was not seen in the other experimental conditions ( $p \gg 0.05$ ) (Fig. 3, *a*; Fig. 4, *d*). This component had an occipital localization (Fig. 3, *c*).

*The P300 component.* The P300 component was marked in response to target stimuli during the time interval 360–600 msec, p < 0.0001 (Fig. 5; Fig. 4, *e*). P300 was not seen in the other experimental conditions ( $p \gg 0.05$ ).

**Discussion.** *Mismatch negativity*. Mismatch negativity in our studies was detected in two experimental conditions based on the conjunction of features: firstly, on selective attention to a specified combination of features in the visual modality (the TVB condition); secondly, on distributed attention in the visual modality (the MB condition). Mismatch negativity was also, as expected, clearly apparent in the "passive oddball" condition, in which the deviant stimulus differed from the standard in terms of only one feature. MMN was suppressed in conditions of distracted attention (NVB and AB).

The time interval for generation of mismatch negativity in our experiments was consistent with published data – in a number of studies (for example, Farkas et al. [2015]), the time interval for mismatch negativity was about 100–200 msec from the moment of stimulus presentation. Winkler et al. [2005] used visual stimuli similar to the stimuli used here and measured mismatch negativity in the interval 108–148 msec, which almost agrees with the time at which mismatch negativity to the deviant stimulus was marked in the MB



Fig. 5. P300 component of averaged visual MMN. *a*) Averaged ERP in the parietal region of interest (Pz, CPz, P1, P2, POz); *b*) averaged map of MMN to target stimulus in the interval 360–600 msec. \*\*\*p < 0.0001 (significant differences between ERP to deviants and standards in the corresponding intervals, with Bonferroni correction.

condition. In the TVB condition, mismatch negativity to the target stimulus was seen somewhat later – in the interval 140-180 msec, which differed from Winkler's study; however, Winkler did not analyze target stimuli and the difference itself could have been caused by the "integrity" of the stimulus. The literature also contains other time intervals for MMN – 100–250 msec [Kimura et al., 2009], 185–205 msec [Astikainen et al., 2008], about 130–190 msec (Czigler and Sulykos, 2010]; 200–250 msec [Nordby et al., 1996]. The spread of values may be associated with the fact that the time and amplitude of MMN generation can be influenced by stimulus parameters – color, spatial frequency, and position [Stefanics, 2014].

Mismatch negativity in all blocks of our experiments had occipital localizations, which is consistent with published data [Winkler et al., 2005; Kimura et al., 2011; Pazo-Alvarez et al., 2003; Nordby et al., 2013; Wang et al., 2014]; there is MEG evidence that the location of the neuronal generators of visual MMN is in the occipital cortex [Susac et al., 2014].

Mismatch negativity and attention. The study reported by Winkler et al. [2005] identified the production of mismatch negativity in response to a conjunction of features independently of attention, though we obtained reliable evidence of MMN generation only in the condition of attention. Presumptively, the attention distraction task in these studies was insufficiently difficult and did not lead to complete switching of attention from the visual modality; the task used in our experiments was in fact complex [Krichner, 1958] and subjectively difficult for the subjects - this was confirmed by results of questionnaires completed by the subjects after completing the experiment. Another feature of our study was that we analyzed ERP to target stimuli, which was not the case in the studies reported by Winkler et al. because these authors tried to avoid analysis of time intervals in which mismatch negativity might overlap with other effects associated with "targetness." As shown by our results (Fig. 3, *a*), mismatch negativity within the interval

analyzed in our studies, 100–180 msec, did not overlap with the N2c component, so we could analyze mismatch negativity in response to the target stimuli just as in other conditions. Both the time interval and the location of the negative wave in response to the combination of features were analogous to those for MMN in the "passive oddball" block, where the nature of the negative wave was not in any doubt.

Thus, our results show that MMN reflects binding of the features of the deviant stimuli and is present only in the situation in which attention to visual stimuli is present – be it selective or distributed. In the more difficult conditions of selective identification of one of the deviant stimuli, MMN arose later than in conditions of nonselective detection of stimuli.

*Mismatch negativity and the mechanisms of analysis of the conjunction of features in the brain.* The possibility that MMN is modulated by attention contradicts a significant volume of literature, as MMN is often named as a neurophysiological correlate of the early, preattention, stage of sensory signal processing occurring at relatively low levels of the sensory areas of the cortex [Kimura et al., 2011; Garrido et al., 2009]. This view is supported by the occurrence of visual MMN in the absence of attention to the stimuli [Winkler, 2005; Czigler and Pató, 2009] – the location and generation time of the difference MMN in conditions of distracted attention were analogous to those in the condition with attention [Harter and Guido, 1980; Czigler and Csibra, 1990; Kenemans et al., 1993; Torriente et al., 1999].

In the studies cited above, which provided evidence supporting the view that MMN is independent of attention, deviant stimuli differed from standards in terms of only one feature. In this case, when the difference in the deviant stimulus from the standard does not require the operation of conjoining features, detection of deviance would appear to be carried out without involving attention, so mismatch negativity in response to it must in fact reflect the automatic, preattention process. This is supported by data from experiments with visual searching demonstrating the "pop-out"

# Sayfulina, Lazarev, Chernysheva, and Chernyshev

phenomenon – if the target object differs from others in terms of one feature, it is detected instantaneously, i.e., in terms of directed attention [McElree and Carrasco, 1999].

What might distinguish the operation of conjunction of features from the processing of single features? One point of view is that the early stages of processing sensory signals from complex stimuli involve recursive interactions between low-level and high-level sensory signal processing loci in the cerebral cortex [Bullier, 2001]; this led to the suggestion that such mechanisms could explain the binding phenomenon [Di Lollo, 2012]. If analysis of complex sensory signals needs the involvement of high-level processes at the early stages of sensory processing (before or during MMN generation), the influence of attention on this process and the resulting modulation of MMN amplitude is consistent. The literature now also contains the view that recursive interaction processes between different levels are always involved in MMN generation (for any stimuli, including simple stimuli not requiring conjunction), opening up the pathway to studies of the mechanisms of MMN modulation by attention [Auksztulewicz and Friston, 2015].

Interpretation of the results obtained in the present work requires consideration from the physiological point of view of whether or not feature binding can occur at the lower levels of sensory processing in the cerebral cortex. The visual cortex is known to contain a population of neurons encoding combinations of features rather than individual features [Matthey et al., 2015]; these are located at higher levels in the visual system as compared with populations encoding individual features. On the other hand, even neurons in the low-level sensory zones of the cortex can undergo simultaneous modulation by multiple physical characteristics of stimuli. In the present study, we used variation of two stimulus features - the spatial frequency of the grids and their tilt angle; these features are known to be among those defining neuron discharges at the lower level of cortical sensory signal processing - in field V1 of the primary visual cortex [De Valois et al., 1982; Hubel and Wiesel, 1959]. Thus, we might expect that all the physiological conditions for unification of the representations of the two visual stimulus features used here are available by level V1 - as selective encoding of each of the four stimuli separately may potentially occur at this level. Thus, the properties of the stimuli used in the present experiments predisposed to detection of combinations of features at level V1 and to generation of MMN independently of attention, analogously to the how this phenomenon has been described for typical experimental conditions of deviance in terms of just one feature. However, according to our results, mismatch negativity to feature combinations does not arise on distraction of attention. Considering the reliability and stability of the mismatch negativity phenomenon [Kimura et al., 2011], this confidently indicates that the feature binding process can in fact be suppressed in the absence of attention. This in turn is consistent with the idea of a distributed recurrent binding

mechanism requiring two-way interactions between different hierarchical levels, i.e., the unavoidable involvement of relatively high levels [Auksztulewicz and Friston, 2015; Bullier, 2001; Di Lollo, 2012], experiencing marked modulation by the attention system.

In our study, visual MMN in the MB condition appeared at an earlier time point than in the TVB condition. MMN generation latency is known to increase with increases in task difficulty [Kimura and Takeda, 2013]. The task in the motor block was simpler than the task in the visual bock: it consisted of undifferentiated responding to all stimuli, which is clearly simpler than responding selectively to one stimulus in a sequence of others. Thus, we can suggest that the more complex processing of the deviant stimulus in the TVB condition takes more time.

The N2c and P300 components. The N2c component was found in ERP to target stimuli (TVB condition), which is consistent with the known properties of this ERP component; published data indicate that it is a member of the family of components arising in response to presentation of the target stimulus [Luck and Kappenman, 2011; Folstein and Van Petten, 2008]. The generation time and location also correspond to published data [Folstein and Van Petten, 2008].

The P300 component was also found only for the target stimulus (TVB condition), which is consistent with published data [Linden, 2005; Polich, 2007]. The time interval and parietal location were also consistent with published data for P3b [Katayama and Polich, 1998; Comerchero and Polich, 1999]. Neither P3b, which should arise in response to the target stimulus, nor P3a, which should arise in response to nontarget deviant stimuli [Polich, 2007] was seen in responses to the nontarget deviant stimuli [Polich, 2007].

The extents of the N2c and P300 components indicate that the target stimulus had not only "deviance," but also "targetness" only in the TVB condition, in accordance with the instructions given to the subjects.

The literature contains indications that feature binding can include two relatively independent levels – an early, sensory level and a late, cognitive level, associated with the high-level processes of attention, awareness, and decision-taking [Takegataetal 2005; Chernyshev et al., 2016]. While MMN allows the early level to be studied, ERP of the N2 and P3 families can, in appropriately designed experiments, be regarded as correlates of binding processes at the late, higher cognitive level [Chernyshev et al., 2016]. The complete absence of these late components in responses to neglected deviant stimuli (the NVB condition) and other conditions of decreased attention to visual stimuli provides further support that binding of the features of this deviant stimulus did not occur either at the early sensory level or at the late cognitive level.

*Can feature binding occur without attention?* If we accept the indication of our results that feature binding actually requires attention, as initially proposed in Treisman's feature integration theory [Treisman and Gelade, 1980], we

# Feature Binding in the Visual Modality

come up against the question of how, despite the narrow bottleneck of attention, cerebral information processing provides a very satisfactory quality of the identification of objects, most of which do not fall into the focus of attention.

It seems very likely that there are two different mechanisms, one of which supports the perception of familiar and important objects, and the other supporting the perception of objects which are new and/or of no practical interest [Colzato et al., 2006; Hommel and Colzato, 2009; Vanrullen, 2009; Velik, 2012]. The first of these may operate on the basis of prepared integral conjunctions of features stored in the brain and evoked in response to sight of a familiar object – this mechanism does not require attention for successful recognition. Rapid recursive processing of sensory information in the brain [Bullier, 2001] presumptively automatically identifies natural categories of objects even from complex combinations of multiple features – without the involvement of attention [Evans and Treisman, 2005].

The second mechanism acts when there is need to obtain a representation of unfamiliar objects in a new environment "on request" – and operates only when attention is being paid to these objects [Vanrullen, 2009]. The procedure and stimulus material in the present experiments appear to have involved the second mechanism of conjunction of features.

**Conclusions.** The results reported here lead to the conclusion that feature binding at the early stages of cortical sensory processing may require attention, which contradicts data obtained in previous studies [Winkler et al., 2005] but supports Treisman's features integration theory [Treisman and Gelade, 1980]. The absence of components of the N2 and P3 families in conditions of reduced attention to visual stimuli indirectly indicates that binding of features outside the focus of attention also did not occur at the later stages of motor processing of sensory stimuli [Chernyshev et al., 2016].

Overall, the results obtained here on the dependence of MMN in responses to combinations of features on attention are in good agreement with current views on distributed cerebral mechanisms of sensory perception based on the interaction of top-down and bottom-up information streams [Auksztulewicz and Friston, 2015; Bullier, 2001].

This study was supported by the Russian Foundation for Basic Research (Grant No. 15-06-10742).

#### REFERENCES

- Astikainen, P., Lillstrang, E., and Ruusuvirta, T., "Visual mismatch negativity for changes in orientation – a sensory memory-dependent response," *Eur. J. Neurosci.*, 28, No. 11, 2319–2324 (2008).
- Auksztulewicz, R. and Friston, K., "Attentional enhancement of auditory mismatch responses: a DCM/MEG study," *Cereb. Cortex*, 25, No. 11, 4273–83 (2015).
- Bartels, A. and Zeki, S., "The theory of multistage integration in the visual brain," *Proc. Roy. Soc. B. Biol. Sci.*, 265, No. 1412, 2327–2332 (1998).
- Bullier, J., "Integrated model of visual processing," *Brain Res. Rev.*, 36, 96–107 (2001).
- Chernyshev, B., Bryzgalov, D., Lazarev, I., and Chernysheva, E., "Distributed feature binding in the auditory modality: experimental evidence

toward reconciliation of opposing views on the basis of mismatch negativity and behavioral measures," *NeuroReport*, **27**, No. 11, 837–842 (2016).

- Colzato, L. S., Raffone, A., and Hommel, B., "What do we learn from binding features? Evidence for multilevel feature integration," *J. Experim. Psychol. Hum. Percept. Perform.*, **32**, No. 3, 705 (2006).
- Czigler, I. and Csibra, G., "Event-related potentials in a visual discrimination task: negative waves related to detection and attention," *Psychophysiology*, 27, No. 6, 669–676 (1990).
- Czigler, I. and Pató, L., "Unnoticed regularity violation elicits change-related brain activity," *Biol. Psychol.*, **80**, No. 3, 339–347 (2009).
- Czigler, I. and Sulykos, I., "Visual mismatch negativity to irrelevant changes is sensitive to task-relevant changes," *Neuropsychologia*, 48, No. 5, 1277–1282 (2010).
- Delorme, A. and Makeig, S., "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," J. Neurosci. Meth., 134, No. 1, 9–21 (2004).
- De Valois, R., Albrecht, D., and Thorell, L., "Spatial frequency selectivity of cells in macaque visual cortex," *Vision Res.*, 22, No. 5, 545–559 (1982).
- Di Lollo, V., "The feature-binding problem is an ill-posed problem," *Trends Cogn. Sci.*, **16**, No. 6, 317–321 (2012).
- Evans, K. and Treisman, A., "Perception of objects in natural scenes: is it really attention free?" J. Experim. Psychol. Hum. Percept. Perform., 31, No. 6, 1476 (2005).
- Farkas, K., Stefanics, G., Marosi, C., and Csukly, G., "Elementary sensory deficits in schizophrenia indexed by impaired visual mismatch negativity," *Schizophr. Res.*, 166, No. 1, 164–170 (2015).
- Folstein, J. and Van Petten, C., "Influence of cognitive control and mismatch on the N2 component of the ERP: a review," *Psychophysiology*, 45, No. 1, 152–170 (2008).
- Garrido, M., Kilner, J., Stephan, K., and Friston, K., "The mismatch negativity: a review of underlying mechanisms," *Clin. Neurophysiol.*, **120**, No. 3, 453–463 (2009).
- Harter, M. and Guido, W., "Attention to pattern orientation: Negative cortical potentials, reaction time, and the selection process," *Electroencephalogr. Clin. Neurophysiol.*, 49, No. 5, 461–475 (1980).
- Hommel, B., "Event files: Evidence for automatic integration of stimulus-response episodes," *Vis. Cogn.*, 5, No. 1–2, 183–216 (1998).
- Hommel, B., "Event files: Feature binding in and across perception and action," *Trends Cogn. Sci.*, 8, No. 11, 494–500 (2004).
- Hommel, B., Colzato, L., and Van Den Wildenberg, W., "How social are task representations?" *Psychol. Sci.*, 20, No. 7, 794–798 (2009).
- Hubel, D. and Wiesel, T., "Receptive fields of single neurones in the cat's striate cortex," J. Physiol., 148, No. 3, 574–591 (1959).
- Kahneman, D., Treisman, A., and Gibbs, B., "The reviewing of object files: Object-specific integration of information," *Cogn. Psychol.*, 24, No. 2, 175–219 (1992).
- Kenemans, J., Kok, A., and Smulders, F., "Event-related potentials to conjunctions of spatial frequency and orientation as a function of stimulus parameters and response requirements," *Electroencephalogr. Clin. Neurophysiol.*, **88**, No. 1, 51–63 (1993).
- Kimura, M., Katayama, J., Ohira, H., and Schröger, E., "Visual mismatch negativity: new evidence from the equiprobable paradigm," *Psychophysiology*, 46, No. 2, 402–409 (2009).
- Kimura, M., Schröger, E., and Czigler, I., "Visual mismatch negativity and its importance in visual cognitive sciences," *Neuroreport*, 22, No. 14, 669–673 (2011).
- Kimura, M. and Takeda, Y., "Task difficulty affects the predictive process indexed by visual mismatch negativity," *Front. Hum. Neurosci.*, 7, 267 (2013).
- Kirchner, W., "Age differences in short-term retention of rapidly changing information," J. Exp. Psychol., 55, No. 4, 352 (1958).
- Lopez-Calderon, J. and Luck S. J., "ERPLAB: an open-source toolbox for the analysis of event-related potentials," *Front. Hum. Neurosci.*, 8, 213 (2014).

## Sayfulina, Lazarev, Chernysheva, and Chernyshev

- Linden, D., "The P300: where in the brain is it produced and what does it tell us?" *Neuroscientist*, **11**, No. 6, 563–576 (2005).
- Livingstone, M. and Hubel, D., "Segregation of form, color, movement, and depth: anatomy, physiology, and perception," *Science*, **240**, No. 4853, 740–749 (1988).
- Luck, S. and Kappenman, E., *The Oxford Handbook of Event-Related Potential Components*, Oxford University Press, Oxford (2011).
- Matthey, L., Bays, P., and Dayan, P., "A probabilistic palimpsest model of visual short-term memory," *PLoS Comput. Biol.*, **11**, No. 1, e1004003 (2015).
- McElree, B. and Carrasco, M., "The temporal dynamics of visual search: evidence for parallel processing in feature and conjunction searches," *J. Experim. Psychol. Hum. Percept. Perform.*, 25, No. 6, 1517 (1999).
- Michie, P., Malmierca, M., Harms, L., and Todd, J., "Understanding the neurobiology of MMN and its reduction in schizophrenia," *Biol. Psychol.*, **116**, 1 (2016).
- Näätänen, R., "The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function," *Behav. Brain Sci.*, **13**, No. 2, 201–233 (1990).
- Nordby, H., Brønnick, K. S., and Hugdhal, K., "Processing of deviant visual events reflected by event-related potentials," *Recent Adv. Event-Related Potent. Res.*, 99–104 (1996).
- Pazo-Alvarez, P., Cadaveira, F., and Amenedo, E., "MMN in the visual modality: a review," *Biol. Psychol.*, 63, No. 3, 199–236 (2003).
- Polich, J., "Updating P300: an integrative theory of P3a and P3b," *Clin. Neurophysiol.*, **118**, No. 10, 2128–2148 (2007).
- Schneegans, S. and Bays, P., "Neural architecture for feature binding in visual working memory," J. Neurosci., 37, No. 14, 3913–3925 (2017).
- Stefanics, G., Kremláček, J., and Czigler, I., "Visual mismatch negativity: a predictive coding view," *Front. Hum. Neurosci.*, **8** (2014).

- Susac, A., Heslenfeld, D., Huonker, R., and Supek, S., "Magnetic source localization of early visual mismatch response," *Brain Topogr.*, 27, No. 5, 648–651 (2014).
- Sussman, E., Horváth, J., Winkler, I., and Orr, M., "The role of attention in the formation of auditory streams," *Percept. Psychophys.*, 69, No. 1, 136–152 (2007).
- Takegata, R., Brattico, E., Tervaniemi, M., et al., "Preattentive representation of feature conjunctions for concurrent spatially distributed auditory objects," *Cogn. Brain Res.*, 25, No. 1, 169–179 (2005).
- Torriente, I., Valdes-Sosa, M., Ramirez, D., and Bobes, M., "Visual evoked potentials related to motion-onset are modulated by attention," *Vision Res.*, 39, No. 24, 4122–4139 (1999).
- Treisman, A., "The binding problem," *Curr. Opin. Neurobiol.*, **6**, No. 2, 171–178 (1996).
- Treisman, A. and Gelade, G., "A feature-integration theory of attention," *Cogn. Psychol.*, **12**, No. 1, 97–136 (1980).
- Vanrullen, R., "Binding hardwired versus on-demand feature conjunctions," Vis. Cogn., 17, No. 1–2, 103–119 (2009).
- Velik, R., "From simple receptors to complex multimodal percepts: a first global picture on the mechanisms involved in perceptual binding," *Front. Psychol.*, 3, (2012).
- Wang, W., Miao, D., and Zhao, L., "Automatic detection of orientation changes of faces versus non-face objects: a visual MMN study," *Biol. Psychol.*, **100**, 71–78 (2014).
- Winkler, I., "Interpreting the mismatch negativity," J. Psychophysiol., 21, No. 3–4, 147–163 (2007).
- Winkler, I., Czigler, I., Sussman, E., et al., "Preattentive binding of auditory and visual stimulus features," *J. Cogn. Neurosci.*, **17**, No. 2, 320– 339 (2005).