Pavlov Institute of Physiology
Russian Academy of Sciences
Neuroiconics Assistive Co. Ltd.

NEUROTECHNOLOGIES

Russia
2021

UDC 004, 159.91, 159.93, 612.8

Authors: Abramov A. D., Akhmerov A. H., Alekseenko S. V., Babenko V. V., Barabanschikov V. A., Basul I. A., Bondarko V. M., Chikhman V. N., Chuev A. A., Danilova M. V., Dyachuk N. M., Galiavin S. I., Gaugel A. O., Gerasimov A. P., Gorbenkova O. A., Gorbunov I. A., Harauzov A. K., Ishchenko I. O., Ivanova L. E., Ivanova N. E., Kalandarova N. N., Khlynov R. D., Khrennikov A. Yu., Kishinskaja E. I., Komarova D. A., Korolkova O. A., Korotaev V. V., Kosyakova A., Kovaleva V. A., Lebedev V. S., Lobodinskaya E. A., Lobzhanidze T. B.-A., Mamaev A. N., Marinova M. M., Merkulyeva N. S., Moiseenko G., A., Morozov S. A., Murav'eva S. V., Naumova D., Okhrimenko A. S., Oskolkova E. S., Pavlov A. V., Plavelsky I. V., Podvigina D. N., Pronin S. V., Rodikova L. S., Ryzhova V. A., Safonov P. A., Samusenkova D. V., Savelyev S. V., Serov E. L., Shakel A. A., Shelepin E. Yu., Shelepin Yu. E., Shut G. A., Skuratova K. A., Solnushkin S. D., Solovyev N. A., Tanifuji M., Timofeev A. N., Ushanov V. V., Vasilev A. S., Vasiliev P. P., Voronaya V. D., Yavna D. V., Zavrazhnova A. A., Zhegallo A. V., Zhukova O. V., Zhumatov M. M.

Neurotechnologies, SPb, Publish by VVM., 2021, 349 p.

Editors: Yuri Shelepin, **Svetlana Alekseenko** (Pavlov Institute of Physiology), **Narisa Nan Chu** (IEEE Brain Data Bank Challenge, CWLab International)

Neurotechnology is the most important trend of intelligent technologies related to the study of brain activity such as: the perception and recognition, the decision-making mechanisms, the analysis of "maps" of neural network activity during planning and organizing targeted actions in real and virtual environment. Neurotechnologies are a part of the creation of the artificial intelligence systems, the convergence of man and autonomous intelligent artificial devices with targeted activities. This book reflects a wide range of investigations from biological and artificial neural networks to rehabilitation neurotechnologies for neurological patients (or patients with sensory and cognitive dysfunctions). Modern investigations and intellectual technologies open a new way of human cognition, purposeful behavior and «free will» (free choice) concept. This book is result of discussed by the authors at the International Conference «Video and Audio Signal Processing in the Context of Neurotechnologies», 2021 at I.P. Pavlov's Institute of physiology, St. Petersburg, Russia.

Contents

Chapter 1.
Integrative transformations in direct neuronal connections between cortical areas of different hierarchical levels

Alekseenko S. V.

Pavlov Institute of Physiology Russ Acad Sci, St. Petersburg, Russia

The integration of visual information collected by millions of photoreceptors starts in the retinae and continues across the entire hierarchy of cortical areas of the brain. However, the basic features of images, their shape, texture, colour, dynamic parameters are analysed in the visual cortex, in which more than 10–14 structural and functional levels are identified, sequentially and in parallel transmitting visual signals [Felleman, Van Essen, 1991; Scanel et al. 1995; Rockland, 1997]. The selective properties of neurons in the visual areas largely depend on their connections with neurons of higher and lower structural levels, as well as on intra-areal connections. However, direct contacts between neurons, not mediated by interneurons, form the skeleton of these brain networks.

Cortical areas belonging to the lower hierarchical level (areas V1 in primates, areas 17, 18 in the cat) receive direct afferentation from the subcortical structure — lateral geniculate nucleus (LGN) [Felleman, Van Essen, 1991; Dreher 1986; Rosenquist 1986; Wandell et al., 2003]. These areas, in addition to retinotopic ordering, also have a stereotyped columnar organization [Kaas, 1997; Hubel, Wiesel, 2005, review]. Neurons with a similar position of their receptive field centers on the retina and some similar functional properties are located in vertical columns passing through all six cortical layers. Information about vertical connections within these structural and functional cortical units was obtained in numerous morphological and electrophysiological studies [Gilbert, Wiesel, 1985, review]. However, the connections of the cortical columns in the areas of different hierarchical levels have been insufficiently studied, and the changes in neuronal connections during the transition between areas of different cortical levels are also unclear.

Information about the connections of neuronal columns innervated from one or the other eye, which alternate in the primary visual cortex, is important for understanding binocular vision. This is because the visual path-

ways from both retinas converge for the first time in this cortical area and binocular neurons are formed, which are tuned to loci in the surrounding space. The diameter of the columns receiving inputs from the left or right eye, called ocular dominance columns (ODCs), is ~500µm [Hubel, Wiesel, 1959, 1962].

Neuronal connectivity of such ODCs was investigated earlier in the areas 17 and 18 located on the lateral gyrus of the brain in a cat. For this, neurons sending axons to the single ODC were identified in areas 17, 18 of both hemispheres and in the LGN [Алексеенко и др., 1999, 2002]. Retrogradelly transported marker of neuronal connections — horseradish peroxidase (HRP) — was microiontophoretically delivered into the whole depth of the single cortical ODC. After identifying and analyzing the positions of labeled neurons bodies (somatas), the brain slices containing these neurons and the injected ODC were magnified 20 times, and the distance from each neuron to the injection site or its projection on the slice was measured. Three-dimensional reconstruction of the region of neuronal inputs to the ODC was performed using the data of a series of sequential slices. The innervation of the injected columns from the left or right eye was determined by analysis of retrograde-labeled neurons in the eye-specific A and A1 layers of LGN.

Intrinsic neuronal connections in areas 17 and 18.

The ODCs have been shown to receive inputs from neurons located in elongated regions in the cortical areas 17, 18. The maximum length of such intrahemispheric intra-areal neuronal connections is 5.0 ± 0.8 mm for ODCs in area 17 and 4.2 ± 0.7 mm for ODCs in area 18 [Алексеенко и др., 1999; Toporova et al., 2001; Alekseenko et al., 2005]. However, the orientation of the identified regions of intrinsic horizontal inputs to the ODCs in areas 17 and 18 is different, these regions are oriented along the projections of the main meridians of the visual field: along the horizontal meridian in area 17 and along the vertical meridian in area 18. It is due to the anisotropy of cortical magnification factor in these areas, the values of which differ along the projection of the horizontal and vertical meridians [Albus, Beckmann, 1980; Tusa et al., 1981]. Experimental evidence of different orientations of the input neuron zones for individual ODCs (crosses) of areas 17 and 18 are shown on the unfolded surface of the cortex in Fig. 1a-d. The neurons that provide intrinsic horizontal long-range inputs to the ODCs of areas 17, 18 are located in the superficial and deep cortical layers (filled and open symbols on the fig. 1) showing that spatial regions with labeled neurons in these layers lie in the same register.
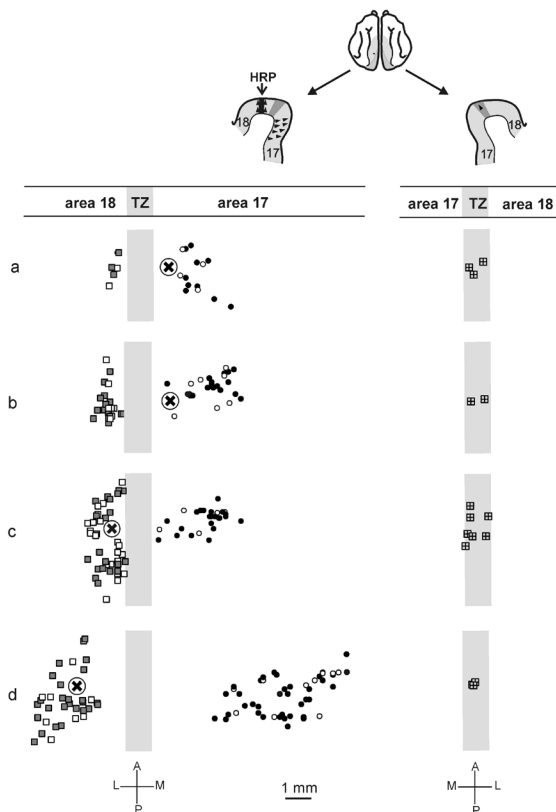
Fig.1. Distribution of labeled neurons in areas 17, 18 of the cat's visual cortex after HRP injection into the ocular dominant columns of area 17 (a, b), area 18 (c, d). Top view of the unfolded cortical surface onto which the labeled neurons from superficial layers (filled simbols) and deep layers (open simbols) were projected. TZ— the transition zone 17/18. Large crosses– injected cortical columns. In upper part of the figure, the locations of injected neuronal column in area 18 and labeled neurons in both hemispheres are shown on the frontal brain section. A—anterior, P—posterior, L—lateral, M—medial.

The ODCs, located in the region of interhemispheric connections, also receive inputs from a small number of neurons in the opposite hemisphere. Callosal inputs come from neurons of superficial cortical layers only. Moreover, they are located in non-mirror-symmetric positions in relation to the

injected ODC (fig. 1). The asymmetry of interhemispheric direct neuronal connections was first discovered by Olavarria [Olavarria, 1996]. To understand the reasons for the revealed asymmetry of connections consider the visual fields projections in cortical areas 17 and 18. These areas are located in both hemispheres. In each hemisphere there is a projection of the contralateral (opposite) half of visual fields of both eyes and also the projection of a portion of the ipsilateral hemifield which is localized in the transition zone between these areas (TZ 17/18) [Payne, 1990]. Callosal neurons sending axons to the injected ODC are located in the TZ 17/18. The source of inputs to asymmetrically located and connected via corpus callosum parts of areas 17, 18 is the zone of naso-temporal overlap on the retina [Stone, 1966]. This retinal zone has mixed projections to both hemispheres. The injected ODCs of areas 17, 18 and their callosal input-neurons are innervated from the naso-temporal overlap zone of the same retina (fig. 2).
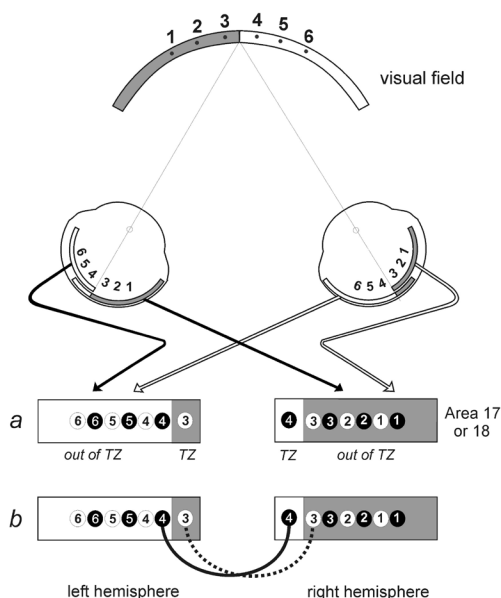
Fig. 2. Simplified scheme of the projections of objects 1–6 to the retinae and their representation in the cortical areas 17, 18 shown on the diagram of flattened cortical surface, where horizontal meridian of visual fields is represented (a). Projections from the left eye are shown by black circles, from the right eye by white circles. Direct callosal neuronal connections are shown in b. TZ—the transition zone 17/18.

The eye-specificity of direct interhemispheric connections in the cortical areas of the lower hierarchical level was shown by different methods in the cat [Olavarria, 2001, 2002; Алексеенко и др., 2002] and in other animals [Laing et al., 2015]. Such interhemispheric connections combine the projections of the left and right halves of the visual field of the eye, presented in different hemispheres.

Connections between areas 17 and 18

Intrahemispheric neuronal connections between areas 17 and 18 have many similarities with intrinsic connections of these areas. Input-neurons from area 17 to the ODCs of area 18 are located in an elongated zones oriented along the projection of the horizontal meridian of the visual field (fig. 1c, d). Analogically, input-neurons from area 18 to the area 17 ODCs are located in an elongated zones oriented along the projection of the vertical meridian (fig. 1a, b). Moreover, the neurons that provide intrahemispheric connections between areas 17 and 18 are located in the superficial and deep cortical layers (fig. 1).

The ODCs of both areas 17, 18 receive inputs from the TZ 17/18 in the opposite hemisphere like the intrinsic connections of these areas (fig.1). Moreover, the callosal neurons that provide these inter-areal connections are also located in the superficial cortical layers, as are the callosal neurons that provide intrinsic connections to the ODCs of these areas.

Thus the obtained information about a similar asymmetry in the organization of intra- and inter-areal connections of ODCs in areas 17 and 18 and the same laminar structure of their intrahemispheric as well as interhemispheric connections is an additional evidence that these areas in a cat belong to the same primary level in the hierarchy of cortical regions.

Connections of areas 17, 18 with areas 19 and 21a

It is of interest how the outputs from areas 17 18 to the columns of areas of higher hierarchical levels are organized. Cortical areas 19 and 21a, located on the cat's suprasylvian gyrus, do not receive inputs from LGN, but have connections with areas 17, 18 [Dreher, 1986]. According to the data of electrophysiological and behavioural studies, it is believed that areas 19 and 21a refer to the information channel "what?" for analyzing the shape and texture of images [Dreher, 1986; Dreher et al., 1993; Khayat et al., 2000; Harutiunian-Kozak et al., 2008; Villeneuve et al., 2009; Kim et al., 2019]. We have microiontophoretically injected HRP into the single cortical columns of areas 19, 21a and analysed the locations of labeled neurons in areas 17, 18.

The belonging of the labeled neurons to these cortical areas was determined using known maps of the cat's brain [Tusa et al., 1978, 1979], in which the boundaries between these areas, as well as the projections of the meridians and parallels of the visual field, are indicated on successive frontal brain slices. Individual cortical columns do not receive afferents from neurons of the whole cortical area. Therefore, it was not a problem with the assignment of the neuron to a particular cortical area, since there were empty regions between the labeled neurons of different areas. The spatial coordinates of the injected columns (azimuth and elevation of their receptive fields) were also determined from the maps of these areas [Tusa et al., 1979; Tusa, Palmer, 1980].

It was obtained that each column of areas 19, 21a receives inputs from four subzones in areas 17 and 18. The labeled neurons were located in the TZ 17/18 of hemisphere ipsilateral to injected column, in the TZ 17/18 of contralateral hemisphere, and in regions located outside the TZ 17/18 in areas 17, 18. Examples of the distribution of labeled neurons on the unfolded cortical surface are shown in Fig. 3a for a column of area 19 and in Fig. 3b for a column of area 21a.

Labeled neurons of all subzones of inputs were located in the superficial layers. It confirms that areas 19, 21a belong to higher levels in the hierarchy of cortical areas than areas 17, 18 [Scanel et al., 1995].

The distance between the subzones outside the limits of TZ 17/18 was different and depended on the distance (azimuth) of the injected column from the projection of visual field CVM. In areas 19, 21a columns closer to the projection of CVM (fig. 3a), the distance between these subzones in area 17 (open diamonds) and area 18 (gray squares) was relatively less than the distance between the subzones of the same pairs of areas obtained in more distant columns (fig. 3b). These differences are due to the mirror arrangement of the projection maps in areas 17, 18 relative to each other and confirm the retinotopic correspondence of intrahemispheric neuronal connections existing between these areas [Salin et al., 1992].

The zone of afferent inputs, identified in area 18 outside the TZ 17/18 (grey squares in fig. 3) is oriented parallel to the projection of CVM like the zone of intrinsic area 18 connections (fig. 1c, d). However, the zone of inputs from area 17 (open diamonds), as compared to the zone of this area intrinsic connections (fig. 1a, b), as well as the zone of afferentation of area 18 columns from area 17 (fig. 1c, d) (Алексеенко и др., 1999; Alekseenko et al., 2001), is widened mainly along the rostro-caudal (A-P) axis of the brain. It indicates the summation of inputs to areas 19, 21a along the projection of the visual field vertical meridian.
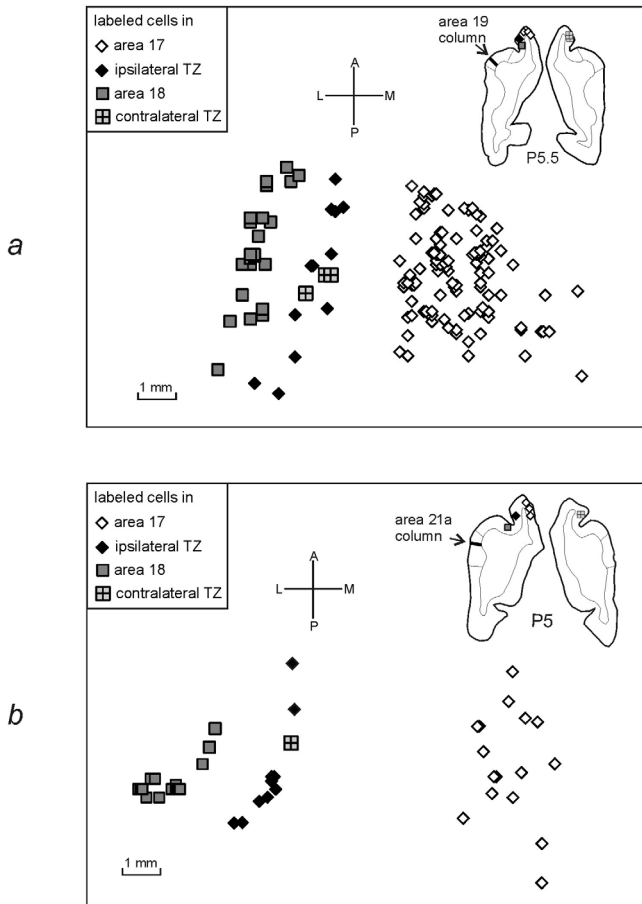
Fig. 3. Distribution on unfolded cortical surface of labeled neurons in cat's areas 17, 18 after HRP injection into a column of neurons in area 19 (a), area 21a (b). Visual coordinates of the column (azimuth/elevation) are in *a* 3/+5 deg, in *b* 7/+ 5 deg. TZ—the transition zone 17/18. The upper right inserts show the frontal brain sections with injected neuronal columns and labeled neurons in areas 17, 18. P— Horsley-Clark coordinates of these brain sections. A, P, L, M as in fig. 1. Note that the location of the callosal inputs to the injected column is mirror-symmetrically su- perimposed on the distribution of the intrahemispheric inputs.

Unlike the inter-areal connections between areas 17 and 18, whose ODCs receive input from one intrahemispheric subzone (fig. 1), the columns of areas

19, 21a do receive additional inputs from the ipsilateral hemisphere. These additional inputs are located in the TZ 17/18 (fig. 3, black diamonds), where a portion of the ipsilateral visual hemifield is represented [Payne, 1990]. For today, the intrinsic structure of TZ 17/18 is known; it consists of two ret- inotopically ordered subzones belonging to area 17 and area 18. The central vertical meridian (CVM) is presented on both borders of TZ 17/18 (fig. 4). Such organization of TZ 17/18 was shown in the studies of interhemispheric connections in areas 17, 18 using local injections of neuronal markers [Ola- varria, 2001; Алексеенко и др., 2002], as well as assessing the sensitivity of TZ 17/18 neurons to spatio-temporal frequencies [Rochefort et al., 2007].
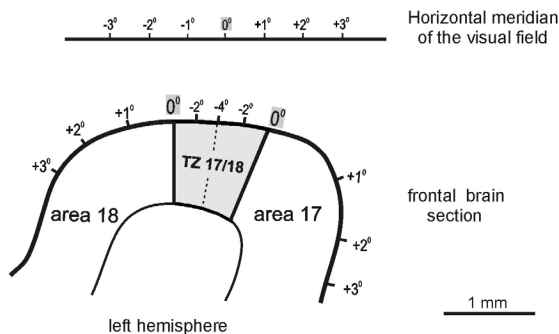
Fig. 4. Simplified scheme of horizontal meridian of the visual field representation in the areas 17 and 18, shown on frontal brain section of the left hemisphere. TZ — the transition zone 17/18, which is about one mm wide [Payne 1990].

The identified inputs from the ipsilateral TZ 17/18 to the columns of areas 19, 21a are arranged in rows oriented vertically in the visual field projection (black diamonds in fig. 3 and fig. 5). Note that in fig. 5 the inputs from only TZ 17/18 are shown for the five columns of areas 19, 21a. The location of these input-neurons in one or two rows depended on the distance of the column of areas 19, 21a from the projection of CVM. Columns closer to the CVM pro- jection receive input from two rows (fig. 5a-c). Taking into account the known design of TZ 17/18 (fig. 4), one of the rows belongs to area 17, and the other to area 18. However, more distant columns of areas 19, 21a receive inputs from a single row only (fig. 5d, e); this row is located on the border between areas 17 and 18. According to Payne (1990), the border between areas 17 and 18 runs along the edge of the representation of a portion of the ipsilateral visual hemifield, the width of which is 3.6 deg on the "null" horizontal meridian

and increases to 20–25 upward and downward. The retinotopic ordering in the TZ 17/18 and the spatial orientation of inputs from this zone can provide neurons of area 19, 21a columns with sensitivity to the orientation component of images [Dreher, 1986; Wimborne, Henry, 1992; Huang et al., 2006; Harutiunian-Kozak et al., 2008], in particular, to the vertical orientation of images or their fragments presented in the ipsilateral eye because the majority of TZ 17/18 neurons are monocular and driven by ipsilateral eye [Berman et al., 1982].
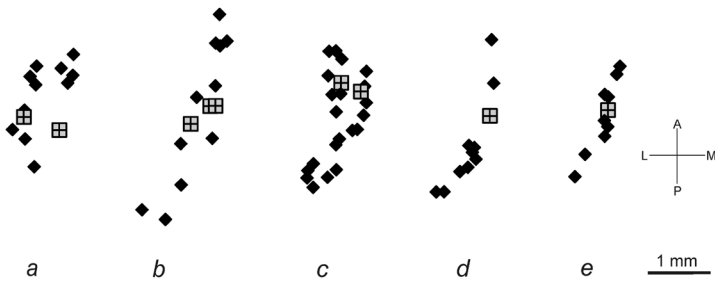
Fig. 5. Location of labeled neurons in the transition zones 17/18 of both brain hemispheres after HRP injection into the columns of areas 19, 21a. Black diamonds — neurons of the hemisphere, ipsilateral to the injected column; squares with crosses — callosal neurons, their positions were mirror-symmetrically superimposed on distribution of input-neurons in the ipsilateral hemisphere. A, P, L, M as in fig. 1. Visual coordinates (azimuth/elevation) of injected columns: 2/–1 deg in a; 3/+5 deg in b; 3/–1 deg in c; 7/+5 deg in d, 10/–3 deg in e.

The fourth subzone of inputs to areas 19, 21a columns is located in the TZ 17/18 of the opposite brain hemisphere. In the figures 1 and 5, these callosal neurons (squares with crosses) are plotted on the diagrams of labeled neurons in the hemisphere ipsilateral to the injected column. These callosal inputs, like callosal inputs to ODCs in areas 17, 18, were not numerous (only 1 to 4 neurons). However, their position in the TZ 17/18 was mirror-symmetric with respect to the position of the input neurons from the rows obtained in the TZ 17/18 of the opposite hemisphere. It was defined by mirrored overlapping of the brain slices of different hemispheres. It is obvious that the distance to the projection of CVM (azimuth) for input-neurons from both TZs 17/18 was the same. Such an arrangement of inputs from the TZs 17/18, in which

different visual hemifields are presented, and, moreover, from different eyes might have some functional meaning for neurons of the column in areas 19, 21a. Convergence of equidistant inputs from both TZs 17/18 (callosal inputs and the inputs from some neurons in the rows of the ipsilateral TZ 17/18, both located in the projection of the same horizontal meridian) may generate binocular neurons in area 19, 21a columns and provide tuning to the loci in the sagittal plane of surrounding space. In the cat, these loci lie in this plane in front of the eyes fixation point. This is because the retinal naso-temporal overlap zones (the origin of projections to the TZ 17/18) are located in the temporal parts of the retinae where objects from the nasal visual hemifields are projected. Information from such binocular neurons can be transmitted to the centers of vergent eye movements and used in the psychophysical process of fusion of two images of an object. Moreover, information about the position of the object in the sagittal plane can serve as a reference point for a comparative assessment of the position of other objects in space.

It is known that the majority of neurons in areas 17, 18 with the exception of layer IV and TZ 17/18 neurons are binocular [Berman et al., 1982; Hubel, Wiesel, 2005]. Our studies of the neuronal connections of the columns of areas 19, 21a showed that the extent of input subzones located in area 17 outside the TZ 17/18 roughly corresponds to the size of intrinsic neuronal connections in this area. This extent is also comparable to the size of the functional modules (3–4 mm) for encoding the range of disparities identified in the isotropic modular structure of area V1 in macaques [Parker et al., 2016]. The convergence of outputs from binocular neurons of such a module and neurons of both TZ 17/18 can provide the neurons of areas 19, 21a with the encoding of more complex stereo-features of objects that are described in primates: the position and inclination of stereo planes, boundaries between them, surface configuration [Bakin et al., 2000; Li et al., 2017; Lu et al., 2018; Li, Shigemasu, 2019; Pasupathy et al., 2020].

Conclusion. The revealed structure of the direct connection of the neuronal columns of the areas located at the first and higher levels of the cortical hierarchy indicates that the integrative processes of information transfer are caused by the appearance of additional neuronal connections in the efferent pathways.

Summary

The aim of this article is to summarize experimental data on the intrinsic and inter-areal neuronal connections of the ocular-dominant columns of areas

17, 18 in cats, as well as on efferents from these areas into the columns of areas 19, 21a of higher hierarchical levels. Comparing the patterns of connectivity between areas 17 and 18 with their intrinsic connections, additional evidence was obtained showing that both these areas belong to the same primary level in the hierarchy of the cortical areas. The columns of areas 19, 21a, in addition to the inputs from the subzones in areas 17 and 18, which provide connections between these areas, receive afferentation from the ipsilateral transition zone 17/18. This study indicates the importance of studying the microstructure of direct neuronal connections between different cortical areas for understanding the integrative processes of visual information transmission.

References

1. *Алексеенко С. В.*, *Топорова С. Н.*, *Макаров Ф. Н.* Микротопография корковых полей 17 и 18 у кошки / Сенсорные системы. 1999. 13(4): 278–283.

2. *Алексеенко С. В.*, *Топорова С. Н.*, *Макаров Ф. Н.* Нейрональные связи, обеспечивающие объединение полуполей зрения / Сенсорные системы. 2002. 16 (2): 83–88.

3. *Alekseenko S. V.*, *Toporova S. N.*, *Makarov F. N.* Spatial distribution of neuronal connections between areas 17 and 18 in the cat's cortex / Georgian J. Neurosci. 2001. 1(1): 59–64.

4. *Alekseenko S. V.*, *Toporova S. N.*, *Makarov F. N.* Neuronal connections of the cortex and reconstruction of the visual space / Neurosci. Behav. Physiol. 2005. 35(4): 435–442.

5. *Albus K., Beckmann R.* Second and third visual areas of the cat: interindividual variability in retionotopic arrangement and cortical location / J. Physiol. 1980. 299: 247–276.

6. *Bakin J. S.*, Nakayama K., *Gilbert C. D.* Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations / J. Neurosci. 2000. 20(21): 8188–8198.

7. *Berman N.*, *Payne B. R.*, *Labar D. R.*, *Murhy E. H.* Functional organization of neurons in cat striate cortex: variations in ocular dominance and receptive-field type with cortical laminae and location in visual field / J. Neurophysiol. 1982. 48: 1362–1377.

8. *Dreher, B.* Thalamocortical and corticocortical interconnections in the cat visual system: relation to the mechanisms of information processing / In: *Pettigrew J. D.*, *Sanderson K. J.*, *Levick W. R.*, editors. Visual Neurosci-

ence. 1986. London, England: Cambridge University Press. pp. 290–314.

9. *Dreher B., Michalski A., Ho R. H., Lee C. W.,* Burke W. Processing of form and motion in area 21a of cat visual cortex / Visual Neurosci. 1993. 10: 93–115.

10. *Felleman D. J.*, Van *Essen D. C.* Distributed hierarchical processing in the primate cerebral cortex / Cereb Cortex. 1991. 1(1):1–47.

11. *Harutiunian-KozakB.A.,Grigorian G. G.,Kozak J. A.,Sharanbekian A. B.*, *Sarkisyan G. S., Khachvankian D. K.* Orientation sensitive properties of visually driven neurons in extrastriate area 21a of cat cortex / Arch. Ital. Biol. 2008. 146(2): 119–130.

12. *Hubel D. H.*, *Wiesel T. N.* Receptive fields of single neurones in the cat's striate cortex / J. Physiol. (Lond). 48(3):574–591.

13. *Hubel D. H.*, *Wiesel T. N.* Receptive fields, binocular interaction and functional architecture in the cat's visual cortex / J. Physiol. (Lond). 1962. 160(1):106–154.

14. *Hubel D. H.*, *Wiesel T. N.* Brain and visual perception. Oxford University Press. N. Y. Oxford. 2005. 744p.

15. *Kaas J. H.* Topographic maps are fundamental to sensory processing / Brain Res Bull. 1997; 44(2):107–112.

16. *Khayat P. S.*, *Saint-Amour D., Molotchnikoff S., Lepore F., Guillemot J.-P.* Cellular response to texture and form defined by motion in area 19 of the cat / Eur J Neurosci. 2000. 12(5):1727–1738.

17. *Kim T., Bair W., Pasupathy A*. Neural Coding for Shape and Texture in Macaque Area V4 / J. Neurosci. 2019. 39(24): 4760–4774.

18. *Laing R. J.*, *Turecek J., Takahata, Olavarria J. F.* Identification of eye-specific domains and their relation to callosal connections in primary visual cortex of Long Evans Rats /Cereb Cortex. 2015; 25(10): 3314–3329.

19. *Li Z., Shigemasu H*. Generalized representation of stereoscopic surface shape and orientation in the human visual cortex / Front. Hum. Neurosci. 2019. 13:283.

20. *Li Y., Zhang C., Hou C., Yao L., Zhang J., Long Z.* Stereoscopic processing of crossed and uncrossed disparities in the human visual cortex / BMC Neurosci. 2017. 18(1):80.

21. *Lu Y., Yin J., Chen Z., Gong H., Liu Y., Qian L., Li X., Liu R., Andolina I., Wang W.* Revealing detail along the visual hierarchy: neural clustering preserves acuity from V1 to V4 / Neuron. 2018. 98(2): 417–428.

22. *Olavarria J. F.* Non-mirror symmetric patterns of callosal linkages in areas 17 and 18 in cat visual cortex / J. Comp.Neurol. 1996. 366(4):643–655.

23. *Olavarria J. F.* Callosal connections correlate preferentially with ipsilateral cortical domains in cat areas 17 and 18, and with contralateral do-

mains in the 17/18 transition zone / J. Comp. Neurol. 2001. 433(4):441–457.

24. *Parker A. J.*, *Smith J. E.* T., *Krug K.* Neural architectures for stereo vision / Phil.Trans. R. Soc. B. 2016. 371(1697):20150261.

25. *Pasupathy A., Popovkina D. V., Kim T:* Visual functions of primate area V4 / Annu. Rev. Vis. Sci. 2020. 6:363–385.

26. *Payne B. R.* (1990) Representation of the ipsilateral visual field in the transition zone between areas 17 and 18 of the cat's cerebral cortex. Visual Neurosci. 1990. 4(5): 445–474.

27. *Rockland K.* Elements of cortical architecture. Hierarchy revised / In: Cerebral cortex, v. 12. eds.: Rockland et al. Plenum Press, NY. 1997. p. 243–294.

28. *Rochefort N. L.*, *Buzas P., Kisvarday Z. F., Eysel U. T., Milleret C.* Layout of transcallosal activity in cat visual cortex revealed by optical imaging / Neuroimage. 2007. 36: 804–821.

29. *Rosenquist A. C.* Connections of visual cortical areas in the cat. In: *Jones E. G.*, Peters A. eds., Cerebral cortex.V.3. Visual cortex., N.Y., Plenum Press. 1986, 81–117.

30. *Salin P. A.*, *Girard P., Kennedy H., Bullier J.* Visuotopic organization of corticocortical connections in the visual system of the cat / J. Comp.Neurol. 1992. 320(4): 415–434.

31. *Scannell J. W., Blakemore C., Young M. P.* Analysis of connectivity in the cat cerebral cortex. J. Neurosci. 1995.15(2): 1463–1483.

32. Stone J. The naso-temporal division of the cat's retina / J. Comp.Neurol. 1966. 126(4): 585–599.

33. *Toporova S. N.*, *Alekseenko S. V.*, *Makarov F. N.* The spatial distribution of horizontal connections in area 18 of the cortex in cats / Neurosci. Behav. Physiol. 2001. 31(4): 345–348.

34. *Tusa R. J.*, *Palmer L. A.* Retinotopic Organization of Areas 20 and 21 in the Cat / J. Comp. Neurol. 1980. 193(1): 147–164.

35. *Tusa R. J.*, *Palmer L. A.*, *Rosenquist A. C.* Multiple cortical visual areas: Visual field topography in the cat / In: Cortical Sensory Organization, v.2. Multiple Visual Areas. ed.: *Woolsey C. N.* Humana Press, Clifton, NY, USA. 1981. 1–31.

36. *Tusa R. J.*, *Palmer L. A.*, *Rosenquist A. C.* The retinotopic organization of area 17 (striate cortex) in the cat / J. Comp. Neurol. 1978. 177(2): 213–236.

37. *Tusa R. J.*, *Rosenquist A. C.*, *Palmer L. A.* Retinotopic organization of areas 18 and 19 in the cat / J. Comp. Neurol. 1979. 185(4): 657–678.

38. *Huang L., Shou T., Chen X., Yu H., Sun C., Liang Z.* Slab-like functional architecture of higher order cortical area 21a showing oblique effect of

orientation preference in the cat / NeuroImage. 2006. 32: 1365–1374.

39. *Villeneuve M. Y.*, *Vanni M. P.*, Casanova C. Modular organization in area 21a of the cat revealed by optical imaging: comparison with the primary visual cortex / Neuroscience. 2009. 164: 1320–1333.

40. *Wandell B. A.*, *Dumoulin S. O.*, *Brewer A. A.* Visual field maps in human cortex / Neuron. 2007. 56(2):366–83.

41. *Wimborne B. M.*, *Henry G. H.* Response characteristics of the neurons in the cortical area 21a of the cat with special reference to orientation specificity / J. Physiol. 1992. 449: 457–478.

Chapter 2.
Evaluation of video images of the chimeric face in statics and dynamics[1]

Barabanschikov V. A., Marinova M. M.

Institute of Experimental Psychology,
Moscow State University of Psychology and Education, Russia

In the psychology of interpersonal perception, the mechanisms of generating the observer's perceptions of the communicator's personality by his or her facial expressions are of particular interest. With the help of what means or internal technologies this process is carried out? What information basis are impressions about the state or character traits of the communicant built on? How does the general idea of personality correspond to the diagnostic attributes of emotional expressions? What functional load is carried by individual elements and parts of the face and how do they fit into the perception of the face as a whole? The experimental solution of these and similar questions is often carried out on the material of static images of the person: photos, portraits, drawings, allowing to vary easily the spatial relations of the face and to change the content of its components. The method of evaluation of the "impossible face"—artificial assembly of the image on the basis of elements or parts of the face, which it consists of or can consist, but which occupy "not their places", are differently oriented in space or belong to different people—has gained popularity. Stimulus models of a "mirror face" assembled only from the left side, "composite" and "chimerical" faces, combining the upper and lower or right and left halves of faces of different people in one image, a "tethered face" containing inverted eyes and mouth, and others are effectively used (Barabanschikov, 2012; de Heering, Wallis, Maurer, 2012; Demuthova, Démuth, 2018; Dole, Pascalis, 2017; McKone et al., 2013; Stephanie et al., 2020; Weibert et al., 2017; Williams et al., 2016). The value of artificial perceptual objects is that they contain contradictions not present in everyday objects and have unusual properties. During the exposition of the "impossible face" a problem situation is formed, which requires the coordination of unfamiliar spatial properties and relations with the observers' experience. The ways of its resolution—the emerging perceptions of the depicted person—give the researcher useful information about the relationship between the parts and the whole in interpersonal perception, the

---

role of the egocentric orientation of the face, the regularities of its personification, the mechanisms of generating the image of the communicant.

In recent years, researchers of interpersonal communication have increasingly turned to the perception of a moving or "living" face, its changes in real time. In contrast to statics, visible changes in facial expressions, eye movements, head shaking, lip and jaw movements during articulation, bring still pictures to life, endowing them with animistic qualities. The subject of perception becomes not so much the spatial organization of facial slices as the structure of their relations: the emergence, development, transitions of facial states into another modality, as well as the role of information received from other sources. The expressions of the moving face characterize the activity of the person as a whole, which constitutes the interpersonal situation and regulates the flows of subject-subject interactions (Barabanshchikov, Korolkova, 2020).

In the context of dynamics, appeals to the analysis of static images of the "impossible face" become ineffective. The methods of transformation of the "live" face inscribed in the processes of interpersonal communication are required (Barabanschikov, 2017). In the future, this will expand the potential of the popular methodology and make the research results more suitable for solving practical problems. In this work we propose one of such tools, using the possibility to control the internal content and spatial relations of the moving face in a real communicative situation. The basic purpose of the executed research: description and analysis of features of perception of stimulus models of a new type, on an example of the movable chimerical face, which right and left sides belong to different people. During the development of a technique, we were interested in ways of manifestation of key features of a chimerical face, earlier registered in conditions of statics:

(1) the integrity of perception, bounded by the pregnancy of form (symmetry, non-inverted orientation in the visual field, perceptual complementation of missing elements and relations, unstructured contrasting background) and the attribution of the face to a certain personality (its personification);

(2) distraction effect—impressions and actions caused by the spatial displacement of the facial halves relative to each other, or by a change in the nature of their boundaries;

(3) the inversion effect—a decrease in the sensitivity of the perceptual system to the introduced transformations of the face when it is flipped by 180°.

Particular attention was paid to (4) the correlation between perceptions of the static and dynamic of the chimeric face, as well as (5) the influence on its

evaluation of sound speech and (6) perceptions of the personality and states of the speaker.

Research method

The research method is based on the information technology of image synthesis using Deepfake artificial intelligence. Synthesis is performed by superimposing and combining selected images on the original photos or video clip. As a result of training (generative-adversarial neural networks GAN are applied) elements of the original image are changed into the desired ones, maintaining the impression of integrity and naturalness of the object both in statics and dynamics. Deepfake is designed to solve applied computer graphics problems, including the creation of clips in which some people's faces are changed to others (Karamazova, 2015) or desired facial expression patterns are superimposed on someone else's face and implemented in real time (Thies, 2018). The sound channel of video images is preserved and open to possible transformations.

Stimulus model. Relying on the DeepFaceLab software, we constructed a model of a virtual Sitter, reproducing a moving chimerical face in full-face, whose right and left sides belong to different people. DeepFaceLab performed a partial merging of videos featuring two woman interviewed in front of the camera at different times (Barabanschikov, Marinova, 2020; 2021; Barabanschikov, Marinova, Abramov, 2021). TThe original images were brought to the same format, and the corresponding part of the face surface of the beginning model (20 years old) [Sitter 1] was superimposed on the left half of the face of the another woman (35 years old) [Sitter 2]. In the final model, the upper half of the combined the upper half of the combined video image had no pronounced border; the lower half included a slight mismatch of surfaces - a bend playing the role of an external distractor (Fig. 1)

To create a complex video image using DeepFaceLab 2.0 software tools, both videos were storyboarded into 2000 images, from which 300 frames tak-en from each video were manually selected and correlated (600 images in total). This allowed the algorithm training to be adjusted in two steps: 1) su-perimposing one face image on another, taking into account the unique facial expressions, movements, and different positions of the head, eyes, mouth, and nose; and 2) creating a kinked section at the bottom of the moving face. The first stage of training lasted 12 hours without interruption (120 thousand iterations), in the second stage the training lasted 120 hours without interrup-

tion (600 thousand iterations). In total, more than 130 hours of GAN training and more than 720 thousand iterations were performed to create a two-minute video clip. With the end of machine learning, all the settings were transferred to the video editor, where the parameters of the superimposed mask were adjusted: size, edge smoothing, resolution value, new superposition algorithm, color and sharpness. After rendering, the video was entered into the professional video editor Adobe After Effects 2020 for further processing: selecting a video segment and reducing it to 15 seconds, creating a video with no sound, as well as retouching the image so that only the head and neck were visible.

A   B   C
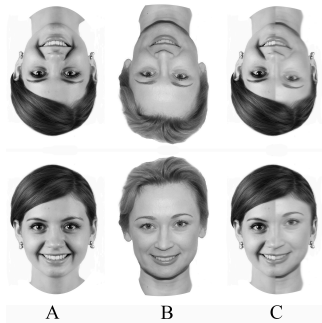
Fig. 1. Static stimulus material in forward (A) and inverted (B) position: original still frame of the Sitter 1, original still frame of the Sitter 2, still frame of the synthesized (chimeric) face

Since the study required not only dynamic, but also static exposure, the same frame with the smiling Naturist 1 was taken from the original as well as from the created video with two faces and subjected to editing. The phases of the video images of the chimeric face at 1-second increments are shown in Figure 2.

0   1   2   3   4
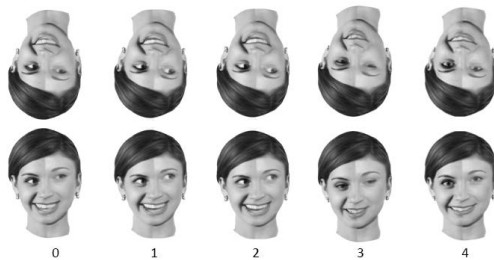
Fig. 2. Freeze frames in one-second increments of dynamic expressions of the chimeric face.

Hardware and software. The following equipment was used to create the stimulus material: Dell U2715H 27" monitor with a resolution of 2560×1440 and frame rate 86GHz, video card NVIDIA GeForce Asus TUF-RTX3090 24Gb, 32 Gb RAM, Core i7–4790K CPU @ 4.00GHz, software "DeepFaceLab" version 2.0, professional video editor Adobe After Effetcs version 2020, and professional photo editor Adobe Photoshop version CC2020.

Subjects. The study involved 42 participants: 30 females and 12 males, aged 18 to 44 years (M = 26; SD = 4.8), students of Moscow universities with normal or corrected vision. Level of English language proficiency: in the scope of secondary school or non-language high school.

Experimental procedure. The research was conducted face-to-face and online. The stimulus material used was 6 video and 4 face exposures, combined into 5 Episodes (Table 1): (I) inverted [Ex1] and straight [Ex2] video exposures of the chimeric face, (II) inverted [Ex3] and straight [Ex4] still image of the chimeric face, (III) inverted [Ex5] and straight [Ex6] original video exposures of Naturist 2; (IV) inverted [Ex7] and straight [Ex8] still image of the original video image of Sitter 1, (V) inverted [Ex9] and straight video image [Ex10] of the chimeric face with the voice of Sitter 1. The order of exposures remained unchanged for all subjects: the inverted face images were shown first in each episode.

Table 1. Sequence of presentation of facial images: Ex1 — Ex10-sequence number of exposure; below — content of the Episode.

Episode I		Episode II		Episode III		Episode IV		Episode V	
inverted video Ex1	non-inverted video Ex2	inverted photo Ex3	non-inverted photo Ex4	inverted video Ex5	non-inverted video Ex6	inverted photo Ex7	non-inverted photo Ex8	inverted video Ex9	non-inverted video Ex10
Video model (Sitter #1+#2)		Freeze Frame Model (Sitter #1+#2)		Original video (Sitter #2)		Freeze frame (original video, Sitter #1)		Video model (Sitter #1+#2 +speech sound)	

The subjects were warned that in some cases the visual exposures might be accompanied by English speech.

After each presentation of images (duration 15 s.), the following require-ments and questions were sequentially displayed on the PC monitor screen:

1) Rate the facial qualities of the character seen on a five-point scale (5–4–3–2–1, where 5 is the highest and 1 is the lowest quality value. Each scale requires only one number: (a)"Harmonious—Disharmonious," (b)"Attrac-tive—Unattractive," (c)"Normal—Bizarre," (d)"Natural—Artificial;

2) What state is he/she in? Choose one of seven emotions: neutral, happy, surprise, fear, disgust, anger, sadness;

The answers to the questions were read aloud or typed by the test taker on the PC keyboard in a free or preset form; there was no time limit for answer-ing. Changing questions and/or starting a new exposure were regulated by pressing the "next" key. At the end of the main experiment, participants were asked to answer additional questions relevant to their motivation and the way they performed the tasks.

The questionnaire was created through the Internet platform Google Forms for the study in the laboratory (face-to-face) and on the Internet (on-line). In the face-to-face format, the subject was positioned in front of a wide-screen monitor with a resolution of 1280×720 px at a distance of 60 cm from the center; the face image size in the center of the screen was $13° \times 8.5° \pm 1.5°$; the displacement of the halves in the lower part of the face was up to 30′ (angular minutes). The online questionnaire was administered only when the subjects had technical capabilities for the experiment: a 17-inch monitor with a resolution of 1280×720 px at a distance of 60 cm, as well as a microphone and audio output devices (headphones or speakers). Display settings of the Google-form were made via Skype under the guidance of the experimenter.

The experiment itself was preceded by an acquaintance with the instruc-tions and with the type of face image exposures (using examples of other Sitters' images).

Data processing. When analyzing the perceived quality of images, for each of the ten facial exposures, we calculated the average frequency values of the ratings (f1, f2, f3, f4, f5) received on the scales of harmony, attractiveness, commonness, and naturalness for all subjects and the group as a whole. The asymmetry coefficient (Cas), which recorded the ratios of conditionally posi-tive (f1, f2) and negative (f4, f5) categories for each of the scales, was used as

the main parameter for comparing assessments: $\text{Cas} = \dfrac{(f_1 + f_2) - (f_4 + f_5)}{f_1 + f_2 + f_4 + f_5}$.

An additional parameter was the average frequency of the central position (f3) of each of the scales. If Cass indicates the advantage of one of the polar categories (Harmoniousness vs. Disharmoniousness, Attractiveness vs. Unattractiveness, Ordinary vs. Bizarre, Naturalness vs. Artificiality), then the f3 value indicates the difficulty of the choice made, including the actual refusal to perform the task (Barabanschikov, 2009). The significance of the differences was established using the chi-square criterion. SPSS21.0 package and R environment were used to process the obtained data.

Results

Experimental data are grouped under three headings that describe the phenomena under study from different angles: perceived qualities of images, perceptions of the Sitter's personality, and observer reports and behavior.

Perceived qualities of images. Examples of primary evaluations of chimeric face videos and photographs (Episode I) are shown in Fig. 3. As a rule, the distribution of choices on the scales "Harmonious — Disharmonious" and "Ordinary — Whimsical" coincide, similar scores are obtained for most of the Episodes of the scale "Natural — Artificial", the answers on the scale "Attractive — Unattractive" have a smoothed form.
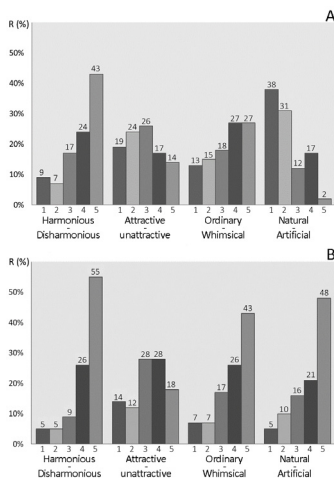
Fig. 3. Fractions of observers' responses (*R*) on the main scales when exposing video images of a chimeric face: *A* — inverted, *B* — non-inverted exposure.

The general dependence of the perceived quality of face images (Cas) on its exposure conditions (Ex1Ex10) is shown in Fig. 4. It is necessary to pay attention to a number of features.
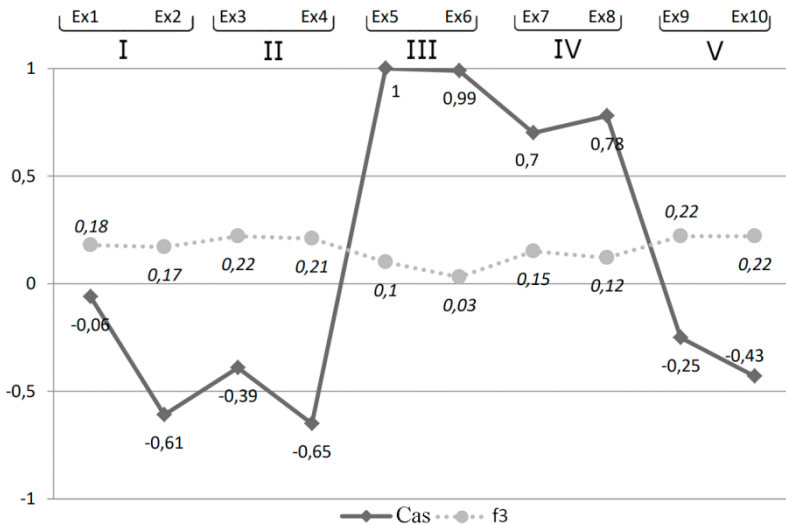
Fig. 4. Averaged dependence of perceived qualities of facial images on exposure conditions. Cas — coefficient of asymmetry, f3 — frequency of central position selection (dashed line), Ex1 — Ex10 — ordinal numbers, I — V — Episodes.

First, the original faces, undistorted by distractor and binary gluing, whether they are inverted or not (Episodes III, IV) are evaluated positively on all four scales and have high values. However, while Sitter 2 is perceived at the limit of the upper level of values, Sitter 1 is a notch lower (0.7, 0.78). The difference between them lies not only in the presence or absence of movement, but also in which side, right or left, they represent in the component face (Episodes I, II, V).

Second, chimerical faces (Episodes I, II, V) are mostly evaluated negatively, as disharmonious, unattractive, bizarre, artificial, although their deviations from the original (zero) level are not so significant. The Episodes of this type are characterized by the fact that the values of the assessments of the inverted images of a moving face are located closer to the zero-mark compared to the assessments of the straight-faced videos and photos. This means that observers of the inverted chimeric face do not perceive the objective distor-

tions as sharp (–0.06,—0.39) as they appear in the non-inverted exposure (–0.61,—0.65), and the apparent deformations of the moving chimeric face are smaller than in the photo-image exposure. The inversion effect occurs in both cases, but it is significantly higher for video images.

Third, in terms of visual content, Ex1 and Ex9 and Ex2 and Ex10 are identical. They are distinguished by their place in a series of exposures, as well as the absence or presence of sound accompaniment. In the latter case (Ex9), the evaluations occupy an intermediate position, when the inversion effect of the chimerical face dynamics weakens, and the non-inverted video evaluations (Ex10) become more independent of the deformations.

Fourth, fluctuations in the frequency of selecting the central position of the scales (f_3) are sensitive to changes in the content of the episodes. When Cas decreases, there is a tendency for f_3 to increase, indicating an increase in the difficulty of the task being performed.

The general dependence of perceived facial image quality on exposure conditions is concretized differently on the scales used. Local peculiarities of perception may include the following:

– Most of the scores on the Attractive—Unattractive scale are in the positive zone of Cass. The exceptions are non-inverted video and chimeric face photos (Ex2 and Ex4). At the same time, the estimates of the non-inverted exposure of the moving chimeric face accompanied by English speech (Ex10) are higher than those of the inverted one (Ex9).

  – The photograph of the inverted chimeric face (Ex3) is perceived at the borderline of normal and bizarre. The non-inverted and inverted chimeric face videos with sound (Ex9, Ex10) are recognized in the negative zone as equivalent.

  – The inverse (Ex1) and non-inverted (Ex2) videos, which have diametrically opposite values on the "Natural—Artificial" scale, appear to be the most contrasting. The positive values of the frequency of selection of the inverted original photo of the Naturist 1 (Ex7) on the same scale are significantly reduced.

Let us note that the influence of the dynamic distractor has a variable character: the distortions of the same moving face image on still frames 0, 2, 3 (Fig. 2) look smoothed in relation to still frames 1, 4, 5.

Assessments of the Sitter's emotional state depend on the specific Episode. Two basic emotions—happiness and serenity—and a number of addi-

tional emotions — surprise, sadness, fear and disgust — are differentiated in the answers of the subjects. The basic emotions in various combinations are preserved under all conditions, while the content and number of additional emotions vary; in eight out of ten exposures, states of sadness are noticed. Thus, when shown an inverted video image of a chimerical face (Ex1), 52 % of responses are for the emotion of happiness, 22 % for a neutral state, 9 % for surprise, and 6 % for sadness. During non-inverted exposure (Ex2) the main emotions in general retain their values, but the volume of additional emotions increases: surprise (12 %), sadness (7 %), fear (3 %) (Fig. 5.).
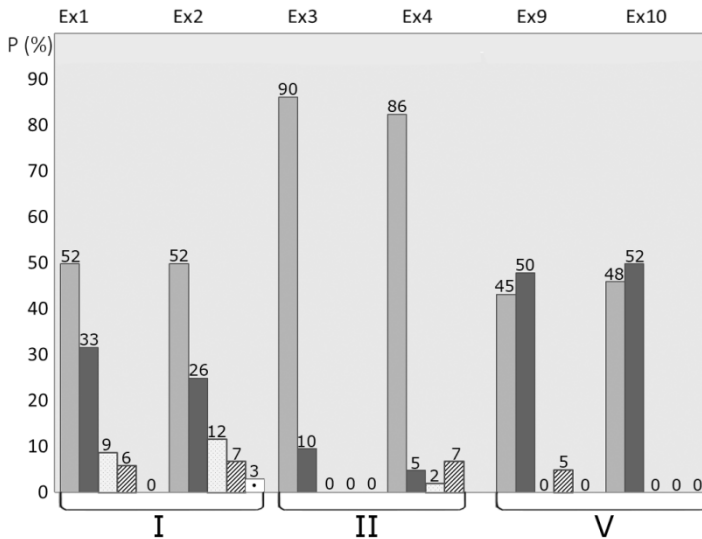
Fig. 5. Averaged estimates of the states of mind of the Sitters depending on the conditions of exposure. Answer options: ☐happiness, ■neutral, ☐surprise, ▨sadness, ⊡fear.

With the demonstration of an inverted still chimeric face (Ex3) the pattern of responses changes: only happiness (90 %) and neutral (10 %). The non-inverted exposure of the photo-image (Ex4) narrows the amount of tranquility to 5 % at the expense of sadness (7 %) and surprise (2 %).

The state of Sitter 2 in the inverted video (Ex5) is evaluated in a more balanced way: happiness (81 %), neutral (14 %), sadness (2 %). The non-inverted video of the same Sitter (Ex6) reinforces only additional emotions: happiness (79 %), neutral (10 %), sadness (7 %), disgust and surprise (2 % each). Similar

results are obtained with the inverted exposure of the original photo of Naturist 1 (Ex7): happiness (76 %), neutral (19 %), disgust (3 %), sadness (2 %). Her non-inverted exposure scores (Ex8): happiness (69 %), neutral (26 %), sadness (2 %).

The appearance of sounding speech again changes the ratio of emotions: in moving face inversion (Ex9): happiness (45 %), neutral (50 %), sadness (5 %); in non-inverted exposition (Ex10): happiness (48 %), neutral (52 %).

Discussions of the results

According to the findings, observer responses are closely related to the type of Episode and the egocentric orientation of face images. The inversion effect—decrease of sensitivity to details and spatial relations of the transformed face when it is turned by 180°—is a cross-cutting pattern, which becomes a precondition for formation of the most plausible (pregnant) image of the Sitter from the observer's point of view. The effect was registered both in static conditions (Episodes II, IV) and in dynamic conditions (Episodes I, V). In the latter case its values on all scales are higher. The greatest influence of inversion was detected in the exposure of video images of a chimerical face (Ex1). Compared to photographs, the "glued" halves of the moving face and the presence of distortions are noticed less frequently, although there is a feeling that the face is "as if not real" or has a vague defect. In Episode II the fracture of the lower part and then the difference of the sides are quickly detected; the frozen chimerical face seems artificial, inanimate, often causing irritation, sometimes fear. The inversion effect of the moving face persists in subsequent expositions (Episode V), but in a less vivid form. Its influence is clearly detectable in the inverted and non-inverted demonstration of a still frame of the original face of Sitter 1, devoid of any transformation (Episode IV). Despite the recognition of the Sitter and the positive feedback, her evaluations are weighted and always suggest alternatives (especially in the inversion). It is not difficult to assume that the phenomenon is based on the inclusion of the Sitter's face in Episodes I and II, generating an expectation or attitude of face distortion. However, the original face of the Sitter 2 in Episode III is not subject to such expectations, although it was also partially present in the preceding Episodes. Moreover, the demonstration of Sitter 2's moving original face in both forward and reverse exposures receive the highest values on the scales of facial harmony, attractiveness, ordinariness, and naturalness. The decision-making difficulties exhibited in Episodes I and II are reduced, especially in non-inverted exposure. Finally, the right side of the moving face is also associated with the localization of the audible speech in

Episode V. All of this points to the unequalness of the sides referring to the different Sitters and the selectivity of their perception. The combined video differentiates between the main (right side of Sitter 1) and the supplementary (left side of Sitter 2) halves. One assumes the function of the whole and plays an active role in the process of interpersonal perception, the other realizes service functions, supporting this process as a "person of the second plan". In the course of perception, the functions of the sides can change, but the leading role of the right half is preserved.

The perceived states of the Sitter in the video images of the chimeric face (Episode I) are divided into three groups: happiness (52 %), neutral (26–33 %), a combination of emotions including sadness, surprise, and fear (15–26 %). The fullest palette of additional expressions is recognized in the conditions of inversion. The chimeric face photos (Episode II) are completely dominated by the expression of happiness (86–90 %). When the original faces of Sitters 2 and 1 are exposed, these scores decrease: 79–81 % for video (Episode III) and 69–70 % for photographs (Episode IV). With the inclusion of the sound channel (Episode V), parity is established between the assessments of happiness and neutral, and the probability of additional emotions tends to zero. Essentially, the structure of assessments of the Sitter's state reflects the most important characteristics of the exposed situation — statics and dynamics, non-inverted and inverted images, presence and absence of speech accompaniment, original faces and their transformations. Another method of experimental research into the perception of video images of a chimerical face opens up.

Undoubtedly, the distractor plays a key role in the perception of the constructed model. The fracture of the left lower part of the face enters directly into the content of his image, formalizing manifestations of emotional states and perceptions of the Sitter's personality. The evaluated qualities of the image tell us how the face is perceived as a surface. The relationship of content and expression of the same subject of perception is established. The distractor introduced has a number of properties that determine its methodological value.

Firstly, it is *dynamic*, and its changes are synchronized with changes of the face as a whole. This is what makes it possible to use dynamic distractors as a tool for studying the perception of a "living face", and not only a chimerical one.

Secondly, the synchronization is not rigid; at different moments of time the distractor is differently included in the process of image formation, having

different influence on it. One may believe that this feature is the basis for the increased effect of the inversion of video images as compared to photographs.

Thirdly, the dynamic distractor is a kind of *navigator*, leading the observer to an understanding of the duality of the image of the chimerical face and its medium. The first thing the subjects notice at the beginning of the experimental session is a vague peculiarity of the speaking female face, some strangeness, which is hidden or masked in the course of the inversion. Non-inverted exposure allows to clarify the localization and categorize the distractor as a discrepancy of a part of the face and its whole, and later during the exposure of a still image — as a distortion of the face. Expansion of the area of attention leads to the discovery of asymmetry of features and individual details — skin folds, position of the pupils, eyebrow shape, etc., which are then grouped into an independent whole, opposing the right side of the image. Perceptions of the contradictory nature of the character, which unites two different people, are cemented in the course of perception and comparison of the original faces of each of the Sitters, and resolved through the hierarchization of the relations between the right and left sides. The integrity of the perception of the faces on the video is confirmed by the observer's conviction of the real existence of the person being exhibited, who adequately behaves in an understandable situation: habitually moving, saying something, expressing emotions, looking at the observer as if addressing him, etc. The compromise found is realized in assessments of multimodal expositions of the chimerical face. From this point of view, the described technique can be referred to the technologies of studying the generation, formation and functioning of the image of the impossible biological object, in which the facial surface can take a variety of forms and be included in the communicative process in different ways.

Finally, fourth, the constructive possibilities of the dynamic distractor are limited. In particular, it has practically no influence on establishing the sex of the Sitter and the current content of his activity.

Concluding the discussion, let us note the motivating potential of the described technique. Voluntary participants, both in the laboratory and in the remote format of work, successfully coped with the proposed tasks, showed interest and readiness for other "tests" of an unusual type. Test takers were attracted and retained by the form of organization of the study, which included elements of a psychophysical experiment and a case study, united by the task of identifying a person with a biologically impossible facial device. The change of Episodes with similar Sitters and unpredictability of new conditions of facial exposures dramatized the content of the experiment. An in-

trigue emerged, a plot—a course of events requiring the active involvement of the subject. This gave rise, even if small, unexpected personal discoveries, the advancement and verification of hypotheses, and positive and negative emotions of the observers themselves, spontaneously implementing the research scenario.

Conclusion

The current trend in the psychology of interpersonal cognition is related to the study of the patterns of perception of a moving, "living" person included in natural communicative situations. To date, this trend does not have sufficient experimental support, and traditional analysis of perception of static stimulus patterns proves ineffective. In our work, we propose a new method for studying the mechanisms of perception of human psychological features in ecologically and socially valid conditions. It is based on a dynamic model of a "chimerical face" constructed with the help of Deepfake video image synthesis information technology. The right and left sides of the generated face belong to movie actors of different age and popularity; the upper half has no pronounced border, while the lower half contains a slight side shift, acting as a distractor. The observers were required to evaluate the perceived qualities of the stimulus model and the images of the original prototype faces, as well as to determine the state, activity content and personality characteristics of real and virtual Sitters under conditions of dynamics and statics, non-inverted and reverse exposure, monovisual and multimodal representation of the stimuli. Cross-analysis of the obtained data leads to the following conclusions:

- In contrast to statics, the perception of a "living" chimerical face relies on a broader informational basis, conditioned not only by the temporal organization of expressions, but also by the entire visually unfolding communicative situation. Phenomena of perception of models of this type—integrity, distraction and inversion effects are filled with new content and characterized by regularities unknown before.
- The integrity of the perception of a moving chimeric face is strengthened by (1) including in the exposition an ecologically and socially significant context representing additional (multimodal) sources of information, and (2) completeness of adequate perceptions of the personality of the virtual Sitter as a medium-formed one with typical properties.
- The dynamic distractor participates in the formation of the general

facial expression, attracts the attention of observers, and plays the role of a navigator, leading to an understanding of the duality of the character. Although changes in the distractor are synchronized with changes in the main elements of the chimeric face, its influence changes during perception.

- A decrease in the sensitivity of the visual system to transformations of the face as it rotates 180° — the inversion effect — takes place in both static and dynamic conditions; in the latter case, its values are higher. In multimodal perception, the magnitude of the effect decreases, and the adequacy of non-inverted exposure estimates increases.
- Original faces in motion and statics, regardless of egocentric orientation on all scales are evaluated positively at the level of high values. For almost all exposures, chimerical faces are generally evaluated negatively: as disharmonious, unattractive, bizarre, and artificial.
- During the exposition of Episodes I–II, a distortion of the perceived face is formed, which is selective in nature and is associated with the image of Sitter 1. The leading role in the perception of the moving chimerical face is played by its right half, while the left half supports this process by playing the role of the background. The hierarchization of the perceived relations of the parties removes the contradictory nature of the artificial character.

References

1. *Barabanshchikov V.A.* Vospriyatie vyrazhenij lica: monografiya. Moscow: Institut psihologii RAN, 2009. (In Russ.).
2. *Barabanshchikov V.A.* Ekspressii lica i ih vospriyatie. Moscow: Institut psihologii RAN, 2012. (In Russ.).
3. *Barabanshchikov V.A.* (otv. red.) Kognitivnye mekhanizmy neverbal'noj kommunikacii. Moscow: Kogito-Centr, 2017. (In Russ.).
4. *Barabanschikov V.A.*, *Korolkova O.A.* Perception of "Live" Facial Expressions. // Eksperimental'naya psikhologiya = Experimental Psychology (Russia), 2020. Vol. 13, no. 3, pp. 55–73. doi: 10.17759/exppsy.2020130305 (In Russ.).
5. *Barabanschikov V.A.*, *Marinova M.M.* Perception of video images of the chimeric face // Poznanie i perezhivanie [Cognition and Experience]. 2020. V. 1. № 1. P. 112–134. doi: 10.51217/cogexp_2020_01_01_07 (in Russ.).
6. *Barabanschikov V.A.*, *Marinova M.M.* Deepfake in Face Perception Re-

search. Eksperimental'naâ psihologiâ = Experimental Psychology (Russia), 2021. Vol. 14, no. 1, pp. 4–19. doi:10.17759/exppsy.2021000001. (In Russ., abstr. in Engl.)

7. *Barabanschikov V. A.*, *Marinova M. M.*, *Abramov A. D.* Virtual Personality of a Moving Thatcherized Face. Psikhologicheskaya nauka i obrazovanie = Psychological Science and Education, 2021. Vol. 26, no. 1, pp. 5–18. doi:10.17759/pse.2021000001. (In Russ., abstr. in Engl.)

8. *Karamazova Z. H.* Dipfejki v iskusstve: kak Sal'vador Dali voskres iz mertvyh // https://screenlifer.com URL: https://screenlifer.com/trends/dipfejki-v-iskusstve-kak-salvador-dali-voskres-iz-mertvyh/ (data obrashcheniya: 01.04.2020).

9. *de Heering A, Wallis J, Maurer D.* The composite-face effect survives asymmetric face distortions. Perception. 2012; vol. 41, pp. 707–716. doi: 10.1068/p7212.

10. *Demuthova S., Démuth A.* Handedness and the Preference of the Visual Field in Face Perception. European Scientific Journal. 2018. doi: 10.19044/esj.2018.c3p8.

11. *Dole M, Méary D., Pascalis O.* Modifications of Visual Field Asymmetries for Face Categorization in Early Deaf Adults: A Study With Chimeric Faces. Front. Psychol. 2017. vol. 8, no. 30. doi: 10.3389/fpsyg.2017.00030.

12. *McKone E, Davies AA, Darke H, et al.* Importance of the inverted control in measuring holistic face processing with the composite effect and part-whole effect. Front. Psychol. 2013, vol. 4, no. 33, pp.1–21. doi:10.3389/fpsyg.2013.00033.

13. *Stephanie S. A.* H. Blom, Henk Aarts & Gün R. Semin. Lateralization of facial emotion processing and facial mimicry, Laterality, 2020, vol. 25 no. 3, pp. 259–274, doi: 10.1080/1357650X.2019.1657127.

14. *Thies J., Zollhöfer M., Stamminger M., Theobalt C., Nießner M.* Face-2Face: Real- time Face Capture and Reenactment of RGB Videos // https://arxiv.org/ URL: https://arxiv.org/abs/2007.14808 (circulation date: 01.04.2020).

15. *Weibert K., Müller V., Sänger J.* Familiarity abolishes right-hemispheric bias in face perception. Journal of Vision. 2017. vol. 17. doi: 10.1167/17.10.998.

16. *Williams L. R.*, *Grealy M. A.*, *Kelly S. W.*, *Henderson I.*, *Butler S. H.* Perceptual bias, more than age, impacts on eye movements during face processing. Acta Psychol (Amst), 2016, vol. 164, pp. 127–135. doi: 10.1016/j.actpsy.2015.12.012.

Chapter 3.

The virtual personality of the Thatcherized face in statics and dynamics[1]

Barabanschikov V.A., Marinova M.M., Abramov A.D.

Institute of Experimental Psychology,
Moscow State University of Psychology and Education, Russia

Among psychological problems of human behavior in virtual reality a special place belongs to the concept of virtual personality and forms of its activity in the artificially created environment (Selivanov, 2016; 2019). Experimentally this problem can be solved both by immersing the subject in the VR environment, and in a more general form by studying the properties of the induced personality possessing an "impossible face".

The subject of perception of the "impossible face" has long been developed in psychological science on the material of collaged images of a person, his photographs, portraits, drawings, schemes. Effectively applied are stimulus models of "mirror faces" composed of the same sides (right or left), "composite" and "chimerical" faces, combining images of upper and lower or right and left halves of faces of different people, "unbalanced faces", which components are not located in their places, etc. (Barabanschikov, 2012; de Heering, Wallis, 2012; Stephanie, 2019; Williams et al., 2015). The use of artificial objects having unusual properties allows to reveal latent laws of interpersonal perception, interrelation of parts of a face and its whole, ways of personification of the sitter, their dependence on various conditions, etc.

In the last decade the interests of researchers are increasingly shifting towards the perception of a "living" face, changing in real time (Barabanschikov, Korol'kova, 2020). There is a need for methods to collage video images of a person in the very process of communication. Solving this problem, we have proposed a technique based on the information technology of image synthesis using artificial intelligence Deepfake designed for solving applied tasks

of computer graphics including the creation of clips in which faces of some people change into other or the required facial expression patterns are superimposed on another person's face and implemented in real time (Chawla, 2019; Dolhansky, et al., 2019; Maras, Alexandrou, 2019; Perov et al., 2020).

Using DeepFaceLab software the authors of this article performed a partial merging of videos of two woman being interviewed (Barabanschikov, Marinova, 2020, 2021; Barabanschikov, Marinova, Abramov, 2020, 2021). The left half of the face of a Model 1 (20 years old) was overlaid with the corresponding part of the face surface of a Model 2 (35 years old). In the final model, the upper half of the combined video image had no pronounced border and the lower half included a small mismatch of surfaces — a kink playing the role of a distractor (Fig. 1.)

A   B   C
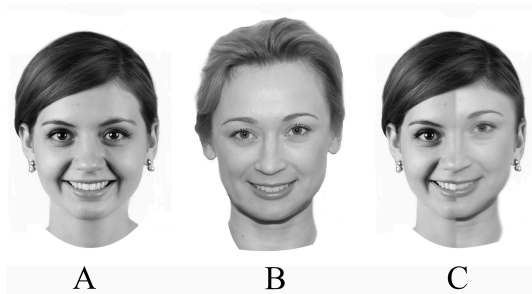
Fig. 1. Stimulus material: A—original photo-image of sitter 1, B—original photo-image of sitter 2, C—combined (chimeric) face;

Observers evaluated: perceived qualities of the image, personality features and states of real and virtual sitters under conditions of statics and dynamics, non-inverted and reverse exposure, monovisual and multimodal representation of stimuli. We have shown that the phenomena of chimeric face perception registered earlier under conditions of statics are preserved but acquire a new content when exposed to a dynamic model. The original faces both in statics and in motion, regardless of egocentric orientation, are evaluated positively at the level of high values. Chimeric faces under all varying conditions are presented as disharmonious, unattractive, bizarre and artificial. The inversion effect is stronger in dynamics than in statics. In multimodal exposure, its magnitude decreases, and the adequacy of non-inverted exposure estimates increases. The gender and behavioral content of the virtual sitter is determined adequately and the perceived age is overestimated. As exposure conditions change, the estimates of emotional expressions of the same chimerical face change.

The present work is devoted to verification of the discovered regularities on the material of Margaret Thatcher illusion. It consists in the difference in perceptions of images of the same cheerful face whose eyes and mouth are flipped 180°. Whereas in the straight egocentric position such a face appears grotesque, containing unconcealed defects and expressing anger, when the face is reversed the grotesque and obvious distortions seem to disappear, and the observer is presented with the pleasant face of a smiling woman (Thompson, 1980). The illusion shows that the configuration uniting the key elements of facial expressions (eyes, mouth) carries different content and appears in different quality depending on the egocentric orientation (Barabanschikov, 2012). Two questions needed to be answered.

(1) Is Thatcher's illusion limited to the exposure of collaged photographs and, if not, how does it manifest in the exposure of video images of the face during the virtual sitter's communication with other people?

(2) How do the manifestations of the chimerical and the Thatcherized face relate, how do they coincide and how do they differ from each other?

Research method

The stimulus model of a Thatcherized face was constructed based on two methods of video image synthesis: IT technology DeepFaceLab 2.0 (DFL), professional video editor "Adobe After Effects" 2020 (Barabanschikov. Marinova, 2020, 2021; Barabanschikov, Marinova, Abramov, 2020, 2021).

At the first stage of the stimulus model creation the neural network application DeepFaceLab was used, which applies the method of generative-adversarial neural networks of GAN type[2] (Perov et al., 2020). Using the built-in functions of the program, the original video with the face of a model (20 years old) with a duration of 15 s was split into 500 frames in JPG format. In automatic mode, the program extracted all the images that contained the face from the obtained frames and also aligned its edges. The next step was to pre-train the face model using XSeg, the DFL's built-in tool for manually masking a whole face or its parts. Using this tool, key points were marked on 20 images, which included the area around the eyes and mouth (Fig. 2).

Next, machine learning was performed on these points in order to automatically mark the selected facial fragments in the remaining frames and their

---

[2] URL: https://neurohive.io/ru/frameworki/deepfacelab-instrument-dlya-generacii-dipfejkov/

correlation with the position and rotation of the head. The training lasted for several hours, during which more than 17 thousand iterations were performed.
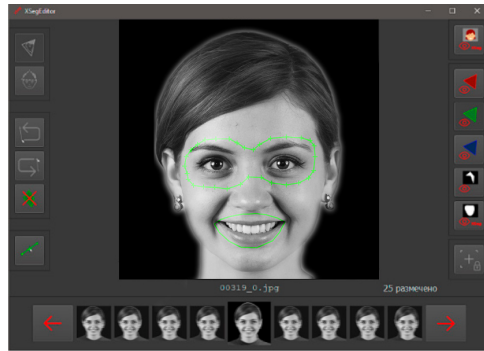
Fig. 2. A screenshot from the XSeg interface of the original sitter's face; the highlighted lines indicate the fragments that are converted to a mask.

After performing the above steps, the function of merging the trained model with the original video sequences was started. When the function finishes, the file containing the trained mask is converted without affecting the rest of the video sequence. This completes the first step.

In the second step all the adjustments were made entirely manually using the video editing program Adobe After Effects, in which the file with the mask obtained from DFL was imported. Each frame which contained parts of the face (eyes, mouth) was inverted one by one, which allowed to achieve a full fusion with the original face of the sitter and with her facial expressions (Fig. 3). One and the other technology provided an opportunity to use the original soundtrack.

A   B
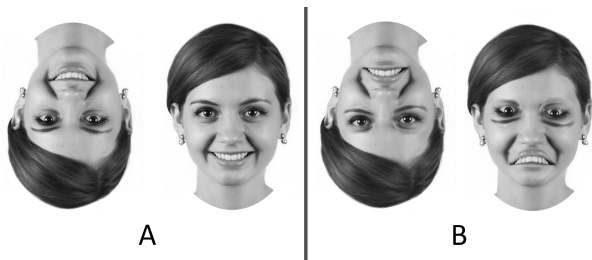
Fig. 3. Freeze frames of the stimulus models used in the experiment inverted and upright positions, A—original image of the sitter's face, B—Thatcherized face;

The study involved 42 university students in Moscow (17 men and 25 women, 18 to 40 years old, M = 23.4; SD = 8.2); with normal or corrected vision. The level of knowledge of English in the volume of a secondary school or non-language high school.

During the main part of the experiment, subjects were sequentially presented with 6 videos and 4 photographic images (Fig. 4): (I) inverted [Ex1] and non-inverted video [Ex2] of Thatcherized face; (II) inverted [Ex3] and non-inverted [Ex4] of the original video without face transformations; (III) inverted [Ex5] and non-inverted [Ex6] photoimage (still image) of Thatcherized face; (IV) inverted [Ex7] and straight [Ex8] original photograph of the sitter without facial transformations; (V) inverted [Ex9] and straight [Ex10] video of the Thatcherized face with English speech included.

Episode I		Episode II		Episode III		Episode IV		Episode V	
inverted video Ex1	non-inverted video Ex2	inverted video Ex3	non-inverted video Ex4	inverted photo Ex5	non-inverted photo Ex6	inverted photo Ex7	non-inverted photo Ex8	inverted video Ex9	non-inverted video Ex10
Video model (Thatcherized Face)		Original video		Freeze frame of the Thatcherized Face		Original freeze frame		Video model (Thatcherized Face + speech sound)	

Fig. 4. Sequence of presentation of facial images: Ex1—Ex10-exposure number; at the bottom—the content of the episode.

After each exposure (15 s) we were required to determine the gender, age, profession of the sitter, as well as to evaluate the attractiveness/unattractiveness of the face on a five-point scale and determine one of the seven basic emotional states it expresses (neutral, happiness, surprise, fear, disgust, anger, sadness).

The experiment was conducted using remote technologies; the questionnaire was created on the basis of the Internet platform Google Form. In order to participate in the experiment, the subjects were required to have the appropriate technical capabilities: a 15" display with a resolution of 1280×720 at a distance of 60 cm from the observer, as well as the presence of audio output devices (microphone, headphones, or speakers). Settings of the Google-form

display were made using video conferencing applications under the guidance of the experimenter.

When analyzing facial attractiveness, the average frequency of each of the ten exposures was calculated. The asymmetry coefficient (Kas) was calculated, establishing the ratio of positive and negative evaluations. Characteristics of the perceived personality — gender, age, profession, and the emotional states of the sitter — were identified by analyzing the subjects' verbal and/or written verbalizations. The shares of key concepts were counted. The significance of the differences was established using Pearson's Chi-square test. R environment and Microsoft Excel 2016 were used to process the obtained data.

Results

The dynamics of the perceived attractiveness of the face (Cas) depending on exposure conditions (Ex1 — Ex10) is shown in Fig. 5. The original face of the sitter regardless of the egocentric orientation both in statics and in dynamics is evaluated as attractive at the level of high values (Cas = 0,8–0,8). The inversion effect is manifested in the decrease of the share of positive answers in comparison with the perception of the straight-centered face: in the conditions of statics Cas = 0.7 in the conditions of dynamics Cas = 0.9. The Thatcherized face is evaluated differently. The sitter is seen as attractive in the first and fifth episodes (Ex1, Ex9). In other cases, regardless of egocentric orientation, the Thatcherized face is perceived as unattractive (Cas = –1 — –0.3). The inversion effect manifests itself in an increase in the number of positive responses and, accordingly, in the asymmetry coefficient: Cas = –0.3 in the static conditions, and Cas = 0.2 in the dynamic conditions.

Thus, depending on the ratio of the egocentric orientations of the face as a whole and its internal structure (eye and mouth areas), the inversion effect — the central characteristic of the Thatcher illusion — appears in a diametrically opposite way, reducing (a regular face) or increasing (a Thatcherized face) the estimate of the attractiveness of the image. A dynamic rectangular Thatcherized face is rated as more attractive (Cas = –0.3) compared to a static face (Cas = –1). The inversion effect has higher values when exposing a static image of a virtual sitter than when exposing a video image. Moreover, it is sensitive to the sounding speech and its intonations: the shares of positive evaluations increase (Cas = 0.3).

The majority of participants in the experiment (94 %) correctly perceive the faces of the sitters as female. The exceptions are a few cases at the be-

ginning (Ex1—Ex3) in the middle (Ex5) and at the end of the series (Ex10), where observers reported a male (1–12 %) or a hybrid (2–21 %) face, which combines male and female features.
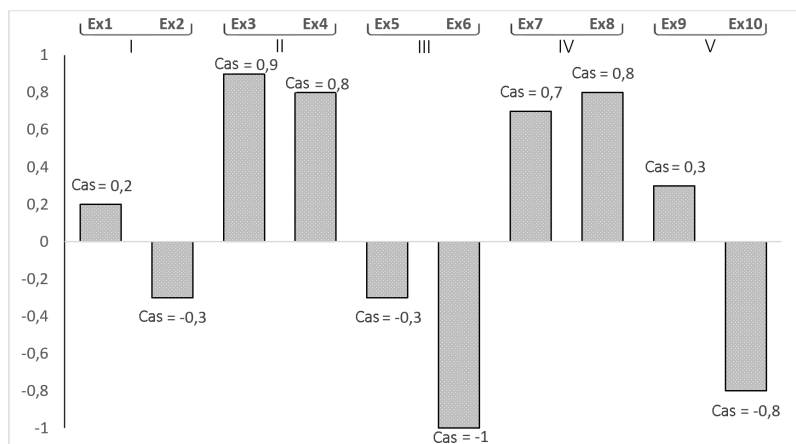
Fig. 5. Dynamics of face attractiveness depending on the type of exposure of the stimulus model. Cas—coefficient of asymmetry, Ex1—Thatcherized face in dynamics, Ex2—non-inverted Thatcherized face in dynamics, Ex3—inverted original face of the sitter in dynamics; Ex4—straight original face of the sitter in dynamics, Ex5—inverted still frame of the Thatcherized face, Ex6—straight still frame of the Thatcherized face, Ex7—inverted still frame of the original face, Ex8—straight still frame of the original face, Ex9—inverted Thatcherized face in dynamics + English speech, Ex10—straight still frame of the Thatcherized face in dynamics + English speech, I—V—episodes.

Assessments of the sitter's emotional state also depend on the specific episode. Two basic emotions—happy and neutral, and a number of additional emotions—surprise, fear, anger, disgust—are identified throughout the experiment. The main emotions in different combinations are preserved under all conditions, the content and number of additional emotions vary; in six out of ten exposures, states of surprise are noticed. Thus, when an inverted video image of a Thatcherized face is shown (Ex1), 67 % of responses are for the emotion of happy, 16 % for a neutral state, 10 % for surprise, and 7 % for fear. At non-inverted exposure (Ex2) the main emotions change their values, tranquility increases threefold (50 %), happiness decreases in almost the same proportion (29 %), mentions of fear increase slightly (11 %), and a new additional state also appears: anger (10 %) (Fig. 6).
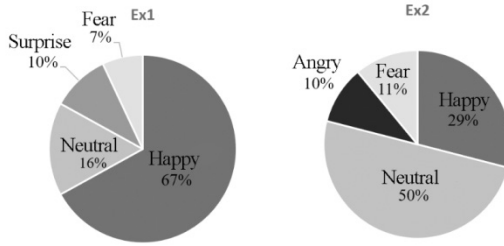
Fig. 6. Median estimates of the states of the sitters during the exposure of inverted [Ex1] and non-inverted [Ex2] video images of Thatcherized face.

When showing an inverted still original face (Ex3) the pattern of responses changes: only happiness (64 %) and tranquility (36 %) remain. Non-inverted exposure of the photo-image (Ex4) narrows the volume of happiness to 5 % at the expense of neutral (41 %).

The state of the sitter in the inverted freeze frame of the Thatcherized face (Ex5) has the following estimates: happiness (60 %), neutral (29 %), surprise (11 %). The straight still image of the same still image (Ex6) brings a greater variety of additional emotions: disgust (33 %), neutral (29 %), happiness (14 %), fear (12 %), anger (10 %), surprise (2 %) (Fig. 7).
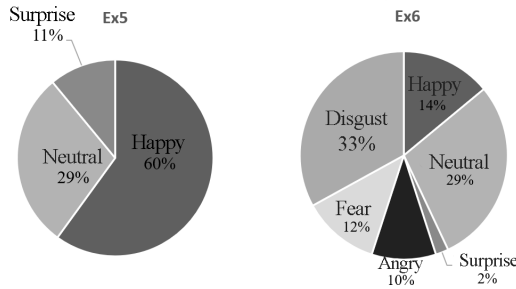
Fig. 7. Median estimates of the states of the sitters during exposure of the inverted [Ex5] and non-inverted [Ex6] freeze frame of the Thatcherized face.

The demonstration of the inverted original still frame of the sitter (Ex7) leaves only two main states: happiness (88 %) and neutral (12 %). Similar scores for the straight-oriented original still frame: happiness (92 %), neutral (5 %), additionally including surprise (3 %).

The appearance of sounding speech again changes the ratio of emotions: in the inversion of the moving tethered face (Ex9): happiness (40 %), neutral (43 %), surprise (17 %); in the non-inverted exposure (Ex10): happiness (33 %), neutral (33 %), surprise (16 %), disgust (11 %), fear (7 %) (Fig. 8).
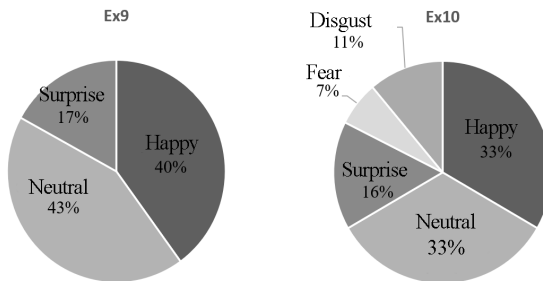
Fig. 8. Median estimates of sitter states during exposure to inverted [Ex5] and non-inverted [Ex6] video images of a Thatcherized face with speech turned on.

Discussion of results

The results of this study show that the Margaret Thatcher illusion described in static conditions is regularly reproduced in video images of the face and has a number of essential features.

The inversion effect — a decrease in the perceived deformations of the face when it is rotated by 180° — is more pronounced in dynamics than in statics. In multimodal exposure, the proportion of positive evaluations increases. A dynamic rectilinear thetcherized face is evaluated as more attractive than a static face. The original face, both statically and dynamically, appears most attractive regardless of egocentric orientation. Evaluations of the attractiveness of the face and the emotional state of the sitter are closely related to the type of episode and its egocentric orientation. Two main emotions, happiness and neutral, as well as a number of additional emotions (surprise, anger, fear, disgust) are highlighted in the subjects' responses. Changes in exposure conditions change impressions of the virtual sitter and perceived emotions. For the theatrical face exposition, the predominant emotion is surprise, while for the chimerical face, sadness is the predominant emotion;

A comparative analysis of the results of the experiments performed with the results of studies of chimerical face perception (Barabanschikov, Marinova, 2020, 2021) suggests that the main regularities found on the material of

the tethered face reproduce the tendencies previously described in studies of chimerical images: (1) despite the face transformations and their attribution to different characters, the artificial (collaged) face is perceived as a whole in all the episodes tested. No reports of a split personality of the virtual sitter were received. (2) assessments of perceived video quality, personality characteristics, and the emotional state of the sitter are related to the same episode and the egocentric orientation of the face. Changes in exposure conditions alter impressions of the chimerical face and the virtual sitter. (3) natural (original) faces both statically and dynamically, regardless of egocentric orientation, are evaluated positively, at the level of high values. Collaged faces, with all variables varied, are perceived as disharmonious and unattractive. (4) The invariance effect of the perception of collaged faces is evident in both statics and dynamics, but is most pronounced in video images. Multimodal exposures reduce the significance of the inversion effect of the collaged face, but increase the adequacy of the evaluations under non-inverted orientation conditions.

This means that, within varying conditions, the perception of the collaged face does not depend on the content of the internal transformation. It is permissible to transfer them to the perception of "mirror", "composite", "unbalanced" and other types of "impossible face". The influence of concrete forms of collage affects the peculiarities of perception of the dynamics of states or in the differences of perceptions of real and virtual sitters, in the nature of associations with their professional activity, age, manifestations of emotional states. Perceived qualities of a Thatcherized face are evaluated more adequately than those of a chimerical face.

Conclusion

The study of the regularities of "impossible face" perception in statics and dynamics is one of the topical directions of modern research in general and social psychology. The work presents a new method of studying the mechanisms of interpersonal perception in ecologically and socially valid conditions. It is based on a dynamic model of "impossible face" created with the help of IT-technology of image synthesis — Deepfake. The constructed model of a virtual sitter reproduces a moving Thatcherized face, the central part of which (eye-mouth triangle) is flipped by 180°. Perceived qualities of facial images, individual psychological characteristics of virtual personality and emotional states of real and virtual sitters under conditions of static and dynamic, non-inverted and reverse exposures, monovisual and multimodal representation of the stimuli were evaluated.

The analysis of the obtained data allows to conclude that the Thatcher illusion is not limited to the exposure of collaged photos, and is reproduced in the exposure of video images of the face in the processes of communication. The main regularities of the appearance of the movable chimerical and Thatcherized face coincide.

The conducted research allows to refer the described technique to the system technologies of studying the generation, formation and functioning of the image of an impossible biological object, in which the facial surface can take a variety of forms and be included in the communicative process in different ways. The technique allows to reveal a plurality of the relations developing in processes of identification of the person and emotional states of the person in real life conditions. New perspectives open up for the development of the nature and mechanisms of interpersonal perception of a person.

References

1. *Barabanshchikov V.A.* Ekspressii litsa i ikh vospriyatie [Facial expressions and their perception].Moscow: IP RAN Publ., 2012.(In Russ.)
2. *Barabanshchikov V.A.*, *Korol'kova O.A.* Vospriyatie ekspressii «zhivogo» litsa [Perception of "live" facial expressions].Eksperimental'naya psikhologiya = Experimental Psychology, 2020.Vol.8, no.3, pp.55–73. DOI:10.17759/exppsy.2020130305 (In Russ.)
3. *Barabanshchikov V.A.*, *Marinova M.M.* Vosprinimaemye kachestva i otsenka ekspressii podvizhnogo khimericheskogo litsa [Perceived qualities and evaluation of expressions of the mobile chimeric face].Litso cheloveka v kontekstakh prirody, tekhnologii i kul'tury [The Face of Man in Contexts of Nature, Technology, and Culture].K.I.Anan'eva, V.A. Barabanshchikov, A.A. Demidov (eds.).Moscow: Moskovskii institut psikhoanaliza, Kogito-Tsentr, 2021. pp. 247–265. (In Russ.).
4. *Barabanshchikov V.A.*, *Marinova M.M.* Vospriyatie videoizobrazhenii khimericheskogo litsa [Perception of video images of a chimerical face]. Poznanie i perezhivanie = Cognition and Experience, 2020.Vol.1, no.1, pp.85–105.(In Russ).
5. *Barabanshchikov V.A.*, *Marinova M.M.* Deepfake v issledovaniyakh vospriyatiya litsa [Deepfake in Facial Perception Research]. Eksperimental'naya psikhologiya = Experimental Psychology, 2021.Vol.14, no.1, pp.4–18 (In Russ., abstr. in Engl.)
6. *Barabanschikov V.A.*, *Marinova M.M.*, *Abramov A.D.* Virtual Personality of a Moving Thatcherized Face. Psikhologicheskaya nauka i obrazovanie = Psychological Science and Education, 2021. Vol. 26, no. 1, pp. 5–18. doi:10.17759/pse.2021000001. (In Russ., abstr. in Engl.)

7. *Selivanov V. V.* Sub"ekt i virtual'naya real'nost': psikhicheskoe razvitie, obuchenie [The subject and virtual reality: mental development, learning].Smolensk: «SmolGU», 2016.430 p.(In Russ.).

8. *Selivanov V. V.* Vzaimodeistvie lichnosti i virtual'noi real'nosti: psikhicheskoe razvitie i lichnostnaya determinatsiya [Interaction of Personality and Virtual Reality: Mental Development and Personal Determination].V.A.Barabanshchikov, V. V. Selivanov (eds).Moscow: Universum, 2019.452 p.(In Russ.).

9. DeepFaceLab: instrument dlya generatsii dipfeikov [Electronic resource] [DeepFaceLab: deepfake generation tool]. Neurohive.URL: https://neurohive.io/ru/frameworki/deepfacelab-instrument-dlya-generacii-dipfejkov/ (Accessed: 30.11.2021).

10. *Chawla R*. Deepfakes: How a pervert shook the world.International Journal of Advance Research and Development, 2019.Vol.4, no.6, pp.4–8.

11. *de Heering A., Wallis J., Maurer D*. The composite-face effect survives asymmetric face distortions.Perception, 2012.Vol.41, no.6, pp.707–716. DOI:10.1068/p7212.

12. *Dolhansky B., Bitton J., Pflaum B., Lu J., Howes R., Wang M., Ferrer C. C.* The DeepFake Detection Challenge (DFDC) Dataset [Electronic resource].ArXiv: Computer Vision and Pattern Recognition.2020.URL: https://arxiv.org/abs/2006.07397 (Accessed: 13.12.2021).

13. *Maras M., Alexandrou A*. Determining Authenticity of Video Evidence in the Age of Artificial Intelligence and in the Wake of Deepfake Videos.International Journal of Evidence and Proof, Vol.23, pp.255–262. DOI:10.1177/1365712718807226.

14. *Perov I., Gao D., Chervoniy N., Liu K., Marangonda S., Ume C., Mr.Dpfks., Facenheim C. F., RP L., Jiang J., Zhang S., Wu P., Bo zhou, Zhang W*. DeepFaceLab: A simple, flexible and extensible face swapping framework [Electronic resource].ArXiv: Computer Vision and Pattern Recognition.2020. URL: https://arxiv.org/pdf/2005.05535.pdf (Accessed: 13.10.2020).

15. *Sharma D., Gulati R., Misra I*. Exploring Consistency in Right Hemispheric Hypothesis and Valence Hypothesis for Perception of Emotions in Brain.Psychological Studies, Vol.65, pp.318–326.DOI:10.1007/ s12646–020–00562-y

16. *Stephanie S. A.H.Blom,* Henk Aarts & Gün R. Semin.Lateralization of facial emotion processing and facial mimicry.Laterality, Vol.25, no.3, pp.259–274.DOI:10.1080/1357650X.2019.1657127.

17. *Thompson P*. Margaret Thatcher: A new illusion.Perception, 1980.Vol.9, pp.483–484.DOI:10.1068/ p090483.

18. *Weibert K., Müller V., Sänger J*. Familiarity abolishes right-hemispheric bias in face perception.Journal of Vision, Vol.17.DOI:10.1167/17.10.998.

19. *Williams L.R.*, *Grealy M.A.*, *Kelly S.W.*, *Henderson I.*, *Butler S.H.* Perceptual bias, more than age, impacts on eye movements during face processing.Acta Psychologica, 2016.Vol.164, pp.127–135.DOI:10.1016/j. actpsy.2015.12.012

Chapter 4.
Image Segmentation, Size Estimation and the Model of Modules

Bondarko V. M., Danilova M. V., Solnushkin S. D., Chikhman V. N.

Pavlov Institute of Physiology of the Russian Academy of Sciences,
St. Petersburg, Russia

In this chapter we consider the problem of the relationship between size estimation and image segmentation and discuss a possibility of describing these processes by the model of modules.

Physiological studies of receptive fields in the striate cortex led V. D. Glezer to propose a model of modules (1985), that was developed further by Kaliteevsky et al. (1994). The model resembles a pyramid model with a discrete and a finite set of spatial frequency filters and provides description of images invariant to their size. In neurophysiological studies V. D. Gleser et al. (1989) obtained discrete distributions of the receptive field (RF) sizes and optimal spatial frequencies of neurons in the striate cortex. This was the basis for the proposed model which considers two types of basic functions (1985, 1995): Gabor elements or trigonometric functions (cosines and sines). The two types are equally possible because these functions are similar and it is difficult to distinguish them in neurophysiological experiments. We will consider both types of these basis functions.

Estimation of the object size in most cases depends on the surrounding stimuli and context, the fact that was already known by ancient Greeks. Many optical visual illusions give striking examples of context dependence of size perception. In this chapter we will focus on one of them, the Oppel-Kundt illusion. A parallel to this illusion can be made by considering classical architectural columns. A plain column looks narrower than a column with cannelures, the observation mentioned by Pére Goùye in Annales of the French Academy of Sciences (Goùye, 1700).

Visual scene segmentation is one of the most important tasks in visual perception and a challenge to artificial intellect and neural networks developers. In the first quarter of XX century gestalt psychologists formulated several empirical laws of segmentation and grouping, however, no clear modelling

recipe exists. The most studied is texture segmentation: if parts of the texture differ in lightness, color, element size, or the slope of the lines, the process of segmentation is quick and effortless (Beck, 1983). Several reviews give a detailed account of this visual task (in Russian: Shelepin et al., 1992; in English: Reed and Dubuf, 1993; Yeshurun, Montagna, Carrasco, 2008). In many cases texture segmentation is well described by a procedure that takes into account differences in local amplitude spectra (Julesz, 1975, 1981). It is important to select an appropriate window when calculating local spectral characteristics, since these characteristics significantly depend on the size of the window. In the chapter we make an attempt to estimate the sizes of such windows for simple geometrical figures and show that the optimal window for local description depends on the object's shape and size. Mechanisms of segmentation and size estimation are simulated and compared using the model of modules.

Segmentation of nested images: Experimental data and modelling

To study the problem of selecting the minimum "window" for image segmentation, we used oblique and upright crosses surrounded by a frame as stimuli. The width of the lines in the frame was always 1/10 of the size of the frame. The lines of the crosses were of the same width, and their sizes varied from being adjacent to the frame to small and well separated. Fig.1 shows the examples of stimuli and the response card for the upright crosses.
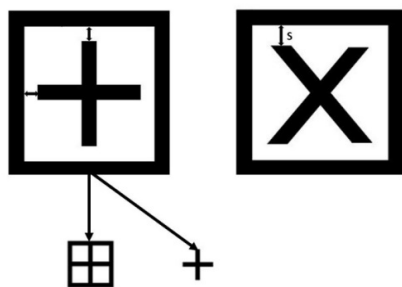
Fig. 1. Examples of the stimuli and the response card.

Top row: upright (left) and oblique (right) crosses surrounded by a frame. Separation *s* is measured between the edges of the frames and of the crosses. The size of the frame was fixed in each series, and *s* varied in each presentation.

Bottom row: the example of the response card for upright crosses.

In each experimental session the size of the frame was fixed and the size of the crosses varied. The stimuli were presented in the center of the monitor screen. The observer was given a response card where four images were drawn (bottom row in Fig.1 shows part of the response card for the upright crosses, for the oblique crosses, the images in the response card were similar). The observer's task was to say which image in the card the stimulus resembles more — the isolated cross (response "cross") or a cross adjacent to the frame (response "frame"). The sizes of the frames were 12–120 arcmin in different experimental series. Eleven observers with normal vision took part in the experiments.

The probit analysis was used to determine the size at which the response changes from "cross" (Fig.1, bottom row, right) to "frame" (Fig.1, bottom row, left). For each observer and for each size of the frame the psychometric functions were based on the total number of trials. They were fit with normal distribution functions using the least square method. The means of the normal distributions corresponded to the size of the crosses when observers separated crosses from the frames. These sizes of the crosses were defined from the fitting and then the distances s between the edges of the frames and the edges of the crosses were calculated. It should be noted that for the small sizes of the frames these distances corresponded to the size of the critical area in the crowding effect (deterioration of stimulus recognition in the presence of the surround) at the resolution limit of the visual system (Danilova, Bondarko, 2007). Then the ratios of sizes of the frames to the sizes of the crosses were calculated. Significant differences in segmentation of oblique and upright crosses were found. For the oblique crosses segregation occurred at smaller sizes, than for the upright crosses. Fig. 2a shows averaged ratios of the frame sizes to the sizes of the oblique and the upright crosses as a function of the frame size at the moment when separation took place. The ratios reached plateau of about 1.5 for the frame sizes larger than 20–30 arcmin (size of the central part of the foveola) and they are larger for the oblique crosses. The data obtained indicate that the segmentation of stimuli is possible at any of their sizes, but only when the stimulus size is larger than the size of the central part of the foveola, the segmentation occurs at the constant ratio between the sizes of crosses and of the frames.

To model the results we applied the model of modules. The images were decomposed into two-dimensional finite Fourier series using sine and cosine functions on square domains. Three, four or five harmonics in modules were

used. The squared areas varied from being larger or equal to the sizes of the images. For each image, the size of the module was selected in such a way as to preserve the largest amount of energy calculated as the sum of the squares of the coefficients of the finite Fourier series. Such module was called optimal, since it ensures the least loss of information in the image.
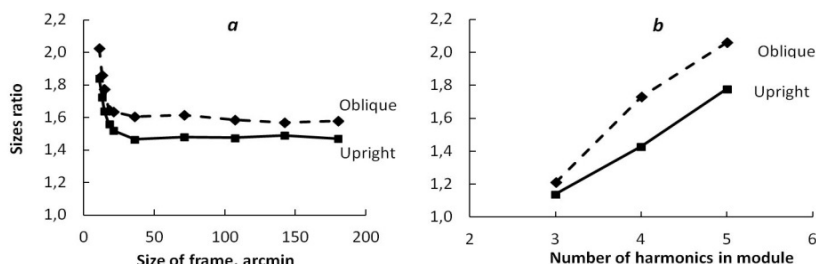
Fig. 2. *a*—ratios of the size of the frame to the size of oblique (dashed line) or upright (solid line) crosses as a function of the frame size at the moment when segmentation occurs. *b*—ratios of the size of the module to the oblique and the upright crosses' sizes as a function of the number of harmonics in the module.

Fig. 2b shows results of modelling: ratios of the sizes of the modules to the sizes of the images are presented as a function of the number of harmonics in the module. The model utilizing four harmonics fits best the experimental data. The modules sizes exceeded the images sizes 1.4–1.5 times which is similar to experimental ratios.

Thus, segmentation of nested black-and-white objects (crosses surrounded by a frame) showed dependence on shape and size. Separation into two figures was size invariant when frames exceeded the size of the central part of the foveola (20–30 arcmin). Ratios of the sizes correspond to the size of the module that optimally describes the crosses, i.e. preserves most of the energy in the image. Segmentation takes place when the frame exceeds the size of this module. For the module of the same size, the oblique cross must be smaller than the upright one.

Relation between segmentation and size estimation results

We compared the differences in the sizes of the oblique and upright crosses sizes in the segmentation experiment and in the experiment that studied size estimation of the same stimuli. The sizes of upright crosses

varied but sizes of oblique crosses were fixed in each series of experiment and equal to 5.4, 8.1, 10.8, 13.5, 16.2, 18.9, 37.8, 56.7 and 75.6 arcmin. Observers' task was to say whether the upright crosses are larger or smaller compared to the oblique crosses. The sizes of crosses were measured as lengths of the sides of the squares in which the crosses were inscribed. The averaged results of these experiments are shown in the Fig. 3a. In Fig. 3b example of the stimuli of identical size (taken as the size of the square) are demonstrated.
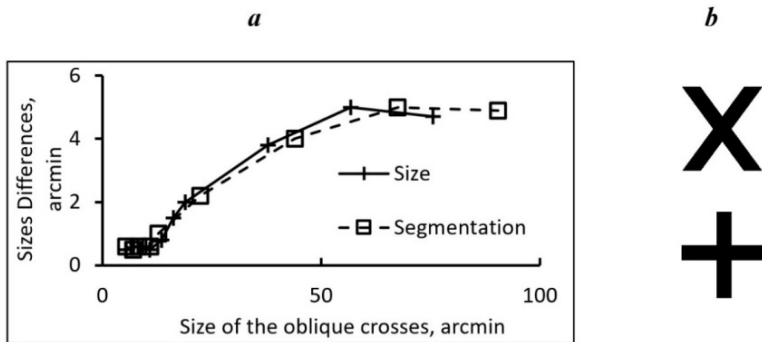
Fig.3. *a* — the difference between the sizes of upright and oblique crosses in the size estimation experiment (crosses) and the segmentation experiment (squares). *b* — both crosses are inscribed in squares of the same size but the oblique cross seems larger. It should be noted, that if the crosses are inscribed in the equal circles, the difference in the perception of size is still present.

It was found that the differences between the sizes of two crosses of different shape are similar in both experiments (Fig. 3a). The data confirm the assumption that for the module of the same size the oblique cross should be smaller than the upright one. Thus, we showed that processes of segmentation and of size estimation are related and we conclude that for images to be perceived of the same size, their optimal modules should have the same size. We tested this assumption in studies of Oppel-Kundt illusion and run psychophysical experiments and modeling of the results.

The Oppel-Kundt illusion: Experimental data

Fig.4 shows examples of the stimuli used to study the Oppel-Kundt illusion: When the widths of two spatial interval delineated by two vertical lines are equal, the interval with additional lines seems larger.
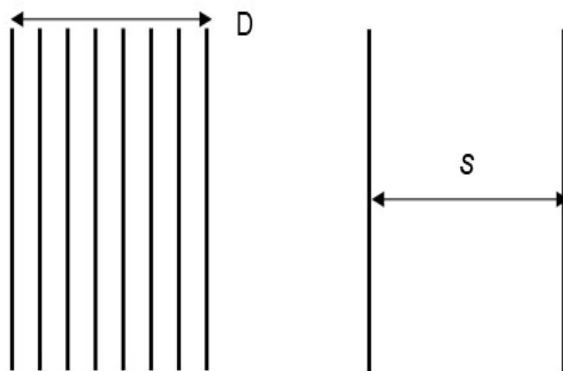
Fig. 4. Example of stimuli in the studies of the Oppel-Kundt illusion. The physical widths of the right and the left (delineated by two vertical outer lines) spatial intervals are equal, but the left interval with the additional lines inside seems wider than the right one.

In the experiment, one spatial interval consisted of two lines, another interval had inside different numbers of lines. The widths of the reference intervals were 1.56, 1.95, 2.34, 2.73 or 3.12 deg. Additional 0, 1, 2, 5, 9 or 14 lines were equally spaced within the interval, i.e., the reference stimulus consisted of 2, 3, 4, 7, 11, or 16 lines. The observer's task was to estimate the separation between the two outer lines by comparing the widths of intervals with additional lines (D) and without them (S). The separation between the two test lines (S) varied during each experimental session to find the width at which two spatial intervals look identical. Four observers took part in these experiments. For each observer the psychometric functions were fitted with normal distribution functions. The means of the normal distributions (points of subjective equality, PSE) correspond to the sizes when the widths of the intervals were perceived as equal. These values were used for estimating the magnitude of the illusion as differences between the PSE from the fitting and physical widths of the intervals. The relative magnitudes were calculated as ratios between the PSE and the physical size of the interval. The discrimination thresholds were measured as standard deviations (84 % responses of the same category at psychometric functions) and then converted to Weber fractions (thresholds divided by interval widths). Fig. 5 shows experimental data: Weber fractions are shown in Fig.5a and the relative magnitude of illusions are given in Fig.5b.
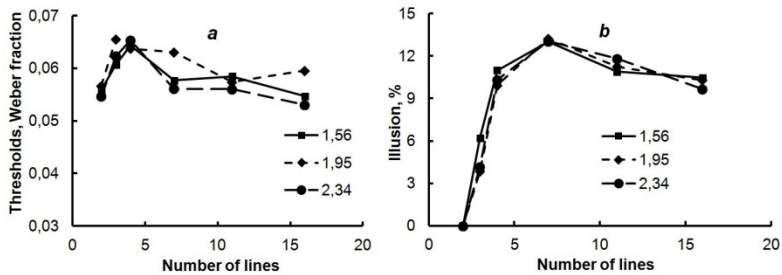
Fig. 5. Examples of the experimental data for the Oppel-Kundt illusion for intervals of 1.56, 1.95 and 2.34 degrees width only. Abscissa: the total number of lines in the interval; two lines mean that there are no additional lines inside the interval. Legends give the widths of the referent spatial intervals in angular degrees. *a* — Weber fractions. *b* — relative magnitudes of the illusion.

Neither Weber fractions, nor the relative magnitudes of the illusion depend on the width of the spatial interval as the three curves nearly coincide, which evidence in favor that the Oppel-Kundt illusion is invariant to the width of the spatial interval. The relative magnitude of the illusion is maximal when two to five internal lines are added to the spatial interval, i. e., the reference stimulus consists of four to seven lines. With the increasing number of lines the relative magnitude reaches plateau. The similar trend between illusion magnitude and number of lines was shown by Bulatov and Bertulis (1997, 1999).
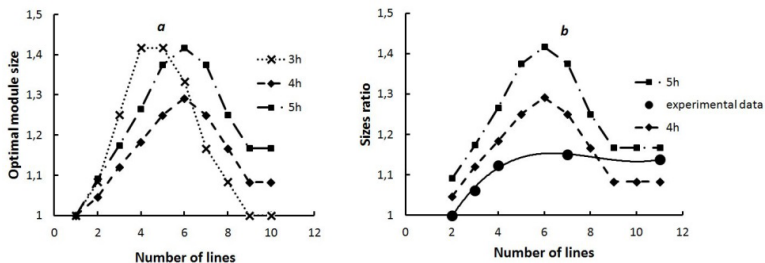
Fig. 6. Comparison of the magnitude of the Oppel-Kundt illusion in experiments and modeling. Abscissa: the total number of lines in the interval; two lines mean that there are no additional lines inside the interval. *a* — modelling results: sizes of the modules divided by images sizes; the number of harmonics is given in the legend. *b* — averaged over the widths of all used intervals experimental ratios of the perceived sizes to the physical sizes (solid line), modelling sizes of the modules divided by images sizes for four (dashed line) and for five (dashed-dotted line) harmonics.

The Oppel-Kundt illusion: Modelling.

The invariance of size in the Oppel-Kundt illusion was modelled by the model of modules with different number of harmonics in the module. Filters of modules were Gabor elements. These modules can be formed by simple and complex receptive fields of neurons in the striate cortex (Glezer, 1985). The size of the module which gave the maximal response to the spatial interval with additional internal lines was taken as the magnitude of the illusion. It should be noted, that Gabor elements are defined on the infinite plane. The size of the module was taken as area at the level of 95 % of the Gaussian envelope.

Fig.6a shows the modelling results for three, four or five harmonics in the module. In Fig. 6b experimental and modelling results are compared. When there are more than five lines within the spatial interval (in the fig.6 it corresponds to seven lines on the abscissa), the experimental ratio reaches plateau and also the size of the optimal module with either four or five harmonics shows flat part of the function. We conclude that under this experimental condition, for images to be perceived as equal in size, they should be described by the optimal module of the same size.

In our earlier studies of spatial intervals (Bondarko, Danilova, 1999) we asked observers to discriminate the width of the spatial interval surrounded by two additional lines by comparing it to the spatial interval of the same width but without the additional lines. We found that both the precision of size discrimination and overestimation of the width (points of subjective equality) depend on the distance to the additional lines. The maximal overestimation was observed at separations equal to quarter of the interval width. This separation corresponds to the distance between the two additional outer lines which is 1.5 times larger than the spatial interval itself and agrees well with the average data in the current segmentation experiments.

Thus, we showed that the magnitude of the Oppel-Kundt illusion depends on the number of lines within the interval and is invariant to the width of the spatial interval. The model of modules was used to fit the experimental data in the Oppel-Kundt illusion. This model involves image processing with spatial-frequency filters tuned to different frequencies and defined on the same area of vision field. The perceived size of the interval corresponds to the size of the optimal module. A similar interpretation of the illusion is given by Bulatov and Bertulis (1997, 1999). These researchers first processed the images with spatial frequency filters. Then the sizes of the areas of the visual cortex

are compared, which change their activity upon presentation of the corresponding stimuli.

Size contrast, segmentation, modelling and art.

In psychophysical experiments we showed that segmentation of simple geometrical figures is related to size estimation. Segmentation of two images took place when objects were surrounded by blank areas of a certain size. In neurophysiological studies V. D. Gleser et al. (1989) obtained a discrete distribution of the sizes of receptive fields of neurons in the striate cortex. On the basis of the experimentally obtained distribution, the sizes of the modules were chosen with a difference of half of the octave with the relative values 1.0, 1.41, 2.0, 2.83, etc. Thus, we can assume that at the moment when segmentation occurs or the geometrical illusion disappears, the size of the optimal module corresponds to the size of the next module in this sequence. The choice of the module larger than the internal object demonstrates the fact that a surrounding uniform field is necessary to describe the object by a finite set of filters with the least loss of information. The presence of additional objects within the area of the module leads to deterioration in the description of the test image.

We showed that for successful segmentation the object should be surrounded by a uniform field. This circumstance is widely used in painting. The simplest example would be the painting "Black Square" by Kazimir Malevich. Its schematic image is shown in fig. 7.

There are several versions of this painting. In Fig.7 we show a version where the outer size of the frame is 1.4 times larger than the size of the black square. Paintings with such ratios of sizes are in the Tretyakov Gallery. There are also paintings with other ratios at the Russian Museum and the Hermitage in St. Petersburg. In different versions the author varied not only black paints and textures, but also the ratios between the frame and the black square itself. We can only guess what governed him in such modifications of the relative sizes, but the variations are within the range that we found is optimal for segmentation of nested objects. Everyone can search for specific personal meaning in this example of art, but the author intuitively had chosen the ratio such that it enhances segmentation. We also showed that according to the model of modules, such a ratio provides optimal description of images and corresponds to the next size of the module in the sequence, suggested by Glezer et al. (1989) and by Kaliteevsky et al. (1994).
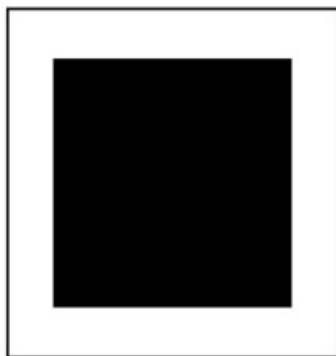
Fig. 7. Illustration to our study.
Schematic image of "Black Square" by K. Malevich.

Relationship between segmentation of small stimuli and crowding

An analogy can be drawn between segmentation of the nested images and the data obtained in studying the crowding effect at the resolution limit of the visual system (Danilova, Bondarko, 2007). We showed that for the smallest images the separation between frames and crosses is equal to the size of critical area in crowding effect. Size of critical area is equal to the maximal separation between stimulus and surroundings when surroundings impair stimulus perception. Outside this critical spacing the crowding effect disappears and recognition of the test stimulus is not deteriorated. The size of this critical area equals the size of the inhibition areas of the smallest RFs. At the resolution limit, perception is degraded when additional images fall into the inhibitory areas of the RFs of the neurons that process the stimuli. Thus, segmentation occurs when the additional objects fall beyond the RFs that describe the image.

The role of the central part of the foveola in perception

Segmentation of nested black-and-white objects (upright and oblique crosses surrounded by a frame) showed dependence on their shape and size. Segregation into two figures was size invariant when frames exceeded the size of the central part of the foveola (20–30 arcmin). Our data confirmed the hypothesis about the specific role of the central part of the foveola and retinotopically associated RFs in different areas in the processes of image recognition, segmentation and integration (Campbell, Shelepin, 1990). According to

the hypothesis, this structure was formed in evolution as a minimum window for the identification of small images and individual details. It can play a role of the pyramid layer of the minimum size but with maximal resolution, and one of the subsequent layers in the pyramid is the fovea. For the central part of the foveola the limitation of four harmonics is natural due to the optics of the eye (Bondarko et al., 2021). Moreover, perception is size invariant only when images are larger than the size of central part of the foveola (Chikhman et al., 2011).

Summary.

In the chapter we show that the physiologically based model of modules allows to select the optimal module which preserves most of the energy and provides optimal description of objects. Using experimental data on image segmentation and size estimation we show correspondence between psychophysical data and the modelling results using a physiologically based model. It is shown that the optimal description of images with the least loss of information by a finite set of filters, implemented in the model, is associated with the segmentation of objects and the estimation of their sizes: the sizes of the images correspond to the sizes of the modules that describe the crosses in the optimal way. The size estimation and segmentation can be described by the model of modules: segmentation occurs when the frame exceeds the size of such module. "Black Square" by K. Malevich is a good example how painters use frames to enhance perception of the painting.

References

1. *Beck J.* Textural segmentation, second-order statistics, and textural elements // Biol. Cybern. 1983. V. 48. P. 125–130.
2. *Bondarko V. M.*, *Danilova M. V.* The estimation of the size of visible objects of various shapes depends on the distance between the objects and the conditions under which they are presented // Sensory Systems. 1997. 11(3). P. 300–311.
3. *Bondarko V. M.*, *Danilova M. V.* Spatial interval discrimination in the presence of flanking lines // Spatial Vision. 1999. 12(2). P. 239–253.
4. *Bondarko V. M.*, *Danilova M. V.*, *Chikhman V. N.* Segmentation of visual images: experimental data and modelling //J. Opt. Technol. 2021. V. 88(12). P. 756–766.

5. *Bulatov A., Bertulis A.* Distortions of length perception // Biol. Cybern. 1999. V. 80. P. 185–193.

6. *Bulatov A., Bertulis A., Mickiene L.* Geometrical illusions: study and modelling // Biol. Cybern. 1997. Vol. 77. P. 395–406.

7. *Campbell F. W., Shelepin Y. E.* The mechanics of the foveola and its role in defining an object // Perception. 1989. Vol. 18. № 4. P. 532.

8. *Campbell F. W., Shelepin Y. E.* Possibilities of the foveola in various objects // Sensory Systems. 1990. V. 4(2). P. 181–185.

9. *Campbell F. W., Shelepin Y. E., Tegeder R. W., Pavlov N. N.* Psychophysical measurement of the intercone separation and object recognition in the human foveola // Ophthal. Physiol. Opt. 1992. Vol. 12. N1. P. 101–102.

10. *Chikhman V. N., Shelepin Y. E., Pronin S. V.* Experimental study of invariant perception of wavelet images // J. Opt. Technol. 2011. V. 78(12). P. 803–807.

11. *Danilova M. V., Bondarko V. M.* Foveal contour interactions and crowding effects at the resolution limit of the visual system // J. Vision. 2007. 7(2). P. 1–18.

12. *Glezer V. D.* Vision and mind. 1985 (in Russian, 1995 in English.) L.: Nauka. 283 p.

13. *Glezer V. D., Yakovlev V. V., Gauselman V. E.* Harmonic basis function for spatial coding in the cat striate cortex // Visual Neurosci. 1989. V. 3. P. 351–383.

14. *Goùye P. T.* Diverses observations de physique générale. 1. Histoire de l'Académie royale des sciences. Avec les mémoires de mathématique et de physique. Tirez des registres de cette Académie, Année MDCC, 1700. V. 2. P. 11–13.

15. *Julesz B.* Experiment in the visual perception of texture. // Sci. Am. 1975. Vol. 232. P. 34–43.

16. *Julesz B.* A theory of preattentive texture discrimination based on first-order statistics of textons. // Biol. Cybern. 1981. Vol. 41. P. 131–138.

17. *Julesz B.* Textons, the elements of texture perception, and their interactions. // Nature (London). 1981. Vol. 290. P. 91–97.

18. *Kaliteevsky N. A., Semenov V. E., Glezer V. D., Gauselman V. E.* Algorithm of invariant image description by the use of a modified Gabor transform // Applied optics. 1994. V. 33. N23. P. 5256–5261.

19. *Reed T. R., Dubuf J. M.H.* A Review of Recent Texture Segmentation and Feature Extraction Techniques // CVGIP: Image Understanding. 1993. V. 57(3). P. 359–372.

20. *Shelepin Y. E., Glezer V. D., Bondarko V. M., Pavlovskaya M. B., Vol I. A., Danilov Y. P.* Spatial vision // Physiology of vision / Ed. A. L. Byzov. Moscow: Nauka, 1992. P. 528–586.

21. *Yeshurun Y., Montagna B., Carrasco M.* On the flexibility of sustained attention and its effects on a texture segmentation task // Vision research. 2008. P. 80–95.

Chapter 5.
Semantic analysis of the computer environment for the formation of adolescent writing

Chuev A. A., Savelyev S. V.

A. P. Avtsyn Research Institute of Human Morphology,
Moscow, Russia

Comparison of the oral speech of Soviet and Russian schoolchildren showed fundamental differences. The new method by which the analysis was carried out makes it possible to assess both quantitative and qualitative differences between similar linguistic products.

Young people, especially schoolchildren, are interesting because they are a relatively unified social group with similar social experiences. At the same time, in our country, such a historical phenomenon has developed, when, with a difference of thirty years, teenagers who speak the same language, and, in the same location, ended up, in fact, in completely different countries. With different living conditions, value systems, educational programs and communication methods [1]. For this reason, it is extremely interesting to determine the patterns of rapidly occurring evolutionary changes in oral speech from Soviet to Russian young people.

Many have seen videos on the Internet with similar names "as Soviet teenagers used to say" [2]. Indeed, some fragments of the TV programs "Up to 16 and Older" [3] sound rather contrasted and more pleasant for listeners in comparison with the reasoning of modern young hominids. Answering the questions: why and what is it connected with? The first thing that comes to mind is a well-structured speech and a rich vocabulary. However, an analysis of the corpus of decrypted texts showed that the volume of the active vocabulary of a dozen speakers of the "Up to 16 and Older" program and modern Russian video bloggers aged 13 to 23 is approximately the same. About 36 percent of the words are unique. The figure may fluctuate, but not a lot.

Comparing the letters and essays of Soviet schoolchildren with modern essays and posts on social networks, it turned out that no significant discrepancies were found in the number of unique words in the texts either — depending on the increase in the volume of the text, the figure decreases from

40 to 30 percent for all groups. Moreover, A. P. Chekhov has the same results. His texts were used in the experiment as a well-known additive method. That is, in percentage terms, Russian-speaking people use approximately the same number of unique words in their written speech.

There is an interesting study that confirms this [4]. It compares the richness of active vocabulary among writers of different eras. And the most prolific of all time was the writer Vladimir Sorokin. He uses just 36 percent of the original words in his texts. For comparison, Pushkin uses about 25 percent of the original words. This does not mean that the written language of students is richer. In the study of the works of writers, rather large corpora of texts were used initially. 25 thousand words each. In our study, we did not take more than 5 thousand words. And the more words there are in the corpus, the lower the percentage of original ones. By the way, after Vladimir Sorokin, the second number is Viktor Pelevin. That is, contemporaries who have access to the Internet and more varied personal experience than, say, the writers of the 19th century, also compose texts that are richer in original words. We leave aside their artistic value.

However, purely intuitively, Internet users still see the difference between Soviet and modern speech. In addition, there are publications on the Internet of the essays and letters of Soviet teenagers [5], who also collect positive comments. That is, the effect, most likely, has nothing to do with the so-called tube sound of the speakers of Soviet television programs. It is important.

What then? In our opinion, the answer to the question can be given by another approach. An analysis of the text corpora of three specific authors showed a significant difference in the energy they spent on their production. We took one of the participants in the previous study, he prepared a letter to the newspaper in 1981. The second one was a modern schoolchild who takes part in mathematical olympiads. Conditionally exemplary. He actively talks about his life on social media. The third was an absolutely marginal young man. At the time of this writing, he is in custody, awaiting a preventive measure for his actions. All three spared no effort in writing texts. Their volumes are relatively large—about 1–2 thousand words. Young people wrote without coercion or necessity. These are not comments, not answers to questions, but finished literary products. For public reading. We did not compare the texts entirely, but selected random fragments of 500 words. And they began to analyze what they consist of. Much like a newsroom employee does. For this, a special algorithm was developed that can be used by both humans and computers.

Most of all the energy was spent by the Soviet schoolchild. About a quarter of its text is seriously considered written speech. Let me emphasize what

he thought about before writing. In some cases, he obviously chose from several options, and somewhere he tried to construct a new image on his own. Accordingly, everything else is a kind of reflexive speech, the reproduction of which does not cause any difficulties. In the text of a Russian participant in mathematical Olympiads, a similar model of behavior can be traced only in a tenth of the text. That is, two and a half times less than that of a Soviet schoolchild. But the most interesting result is that of a marginal young man. The originality of his work was less than a percent. About 10 times less than his more successful contemporary and 25 times less than a Soviet youth.

In addition, all authors made mistakes in the construction of their written language. The Soviet schoolchild did not succeed in correctly connecting words in that part of the text that is less than a percent. The Russian mathematician did not succeed in that part of the text, which is slightly more than a percent. And unsuccessful attempts to take advantage of the charms of the Russian language occupied 7 percent of the text of the modern marginal. About 7 times more than that of other study participants.

In terms of energy savings, the third participant also surpassed the other two in quality. Its text is one and a half times more filled with dubious linguistic constructions that cover up the lack of content.

We assume that the characteristics of a particular speech or text sample directly depend on the amount of energy spent by the author. It is on this indicator that one should be guided when comparing the oral and written speech of Soviet and Russian young people, and not explore the original words, meanings of statements or their subjective beauty.

We again took transcripts of interviews with Soviet young people between the ages of 13 and 23, these are the heroes of the "Up to 16 and Older" program. And also transcripts of interviews with popular Russian video bloggers of the same age range. Corpuses of texts in each group made up several thousand words, from which random fragments of about 500 words were taken. The comparison showed that a Soviet teenager spent 16 times more on his speech than a young video blogger. At the same time, energy savings are perceptible to the ear, that is, what the listener must already spend in order to understand what the speaker is talking about at all—this category among Soviet youth is about 7 percent, and among the mentioned video bloggers—more than half of the speech content or 51 percent.Normal speech is the part that all speakers sound about the same as most native speakers. For the Soviet group, this category was 77 percent, and for the Russian—49 percent.

The conclusion suggests itself. The speech of Soviet youth is more pleasant for the listener, because they seriously expended energy on its production, which saved the listener's energy. The principle has been known in linguistics for a long time [6]. This was confirmed once again, but quantitatively and qualitatively. Accordingly, Russian video bloggers cannot connect two words in half of the attempts, which greatly complicates the perception, not to mention the aesthetic side. We are not interested in the content. But most importantly, the Soviet youth, unlike video bloggers, took care that the performance was not only utilitarian, but also, in a sense, artistic. Video bloggers did not even make such attempts.

The question of the nature of the observed changes is quite legitimate. Naturally, some understanding of the neurological mechanisms of speech formation and an objective quantitative assessment of the energy expenditure of the brain will be needed. Let us consider the neurological substrate of written speech, which is formed in children regardless of the social environment of the Soviet or modern periods. Learning to write at the age of 6–7 occurs with a very immature bark, which accumulates social instincts from the outside world in its connections. These vast areas of the brain continue to actively form the entire complex of both near and distant intercortical connections. In addition to the very fact of differentiation, under a meaningful load, complex memorization of sensorimotor, speech and auditory speech algorithms occurs. By pronouncing the words written out with a pen, the child creates an unprecedented situation in the development of the brain. To write a coherent sentence, he must deal with the semantic planning of the arrangement of words that should be depicted according to the complex laws of written speech. All this must happen simultaneously and very quickly. The brain of a growing person is very rarely under such cognitive pressure.

This has several major consequences. On the one hand, connections are intensively formed between the most "human" centers of the brain, where interneuronal synaptic contacts create reliable storage systems for algorithms of complex movements. The development of writing skills is impossible without the involvement of memory, auditory and speech centers of the cortex. The result is the acquisition of precious social benefits that are highly valued in the hominid community. In this respect, writing makes the child use and strain his memory. This property of the brain is very important in biological processes.

Thus, in the study of conditioned reflex and figurative memory, it was shown that in the series of rays—sharks—bony fish—turtles—snakes—rats—monkeys, the retention of event traces increases from several days to 120 days. This valuable quality of our brain depends a little on our desires.

After the start of teaching written language, the ability to memorize approximately doubles, which was found by comparing illiterate and trained children.

On the other hand, written language creates the basis for individual cognitive success in mastering the world around him. The logic of constructing phrases, the meaningfulness of the text presented make the child think over and model the result of the work. When writing texts, we are no longer faced with biological, but with human laws that affect the maturation of the nervous system. The meaningfulness of the written text is a non-genomic way of inheriting all forms of social instincts developed by humanity. For this reason, mastering written language was the first and most successful attempt to break the unreliable circle of demonstration-speech methods of transferring hominid knowledge outside the genome.

Speaking about written speech, it should be noted that the ability to read and write is not given to everyone. Children regularly appear who are practically incapable of mastering this method of communication. As a result of titanic pedagogical efforts, it turns out to teach them the most primitive writing, which causes an almost physiological rejection in them. This is due to the fact that the so-called writing area for some people can be very small. If this is the case, then it is impossible to teach writing to some children due to the individual characteristics of the organization of the cerebral cortex. Fortunately, such features of the structure of the cortex are extremely rare, and the inability to learn is usually a consequence of the traditionally poor upbringing and social savagery of parents. However, most children learn to write. This is facilitated by competition and herd development of general skills in primary school.

The basis of children's behavior is not yet a conscious and rational choice, but a primitive imitation, which is burdened with following innate and social instincts. The saddest thing is that hormonal regulation and the desire to improve one's biological state through systems of social domination act as incentives. After learning to write, children quickly move on to the more primitive finger pointing at a tablet or phone. As a result, a situation arises of replacing the forced load on the developing brain with computer prompts, artificial syntax and punctuation. Ultimately, the exclusion of handwritten texts from child development reduces the formation of the structural foundations of self-awareness and cognitive abilities. At the same time, our community is focused on ensuring that creative deviations never appear. Only primitive demonstrative dominance in populations of extremely limited social imitators and conformists guarantees the illusion of personal prosperity. For this reason, developed written language is gradually being replaced by indistinct and illiterate phrases,

pictures and emotional emoticons. This situation begins to be realized among those who are capable of non-biological thinking. An alternative to simplifying and biologizing the development of the human brain are artificial methods of controlling the differentiation and maturation of the nervous system.

We settled on written speech as the main source of cognitive impressions and a way of interfering with the formation of the rational part of the brain. Let's try to consider options for counteracting the degradation of individual education. In this respect, handwriting and calligraphy, known to all for a long time, play a huge role. It is about the neat and beautiful spelling of letters and words that add up to sentences and texts. If a person is engaged in calligraphy, then his ability to control his own brain increases many times over. This is due to the fact that, in addition to the ability to write and the semantic content of the text, there are artistic criteria for assessing the fruits of child labor. The aesthetics of handwriting is not only a sign of belonging to a hominid, but also a way of self-affirmation. Showcasing well-written text is considered a great method of motivating teenagers to calligraphic feats. The emerging semantic and aesthetic discipline leads to the development of unique motor skills, the ability to concentrate and elementary accuracy. These patterns have been known for a long time and cannot be doubted.

Nowadays, many simple thoughts began to reach the prophets of computer religions. For this reason, two popular stylus writing programs for active computer screens have been released (abcPocketPhonics and WritePad). In privileged schools in America and Europe, calligraphy takes at least an hour a week, while in Japan and China calligraphy is cultivated as gymnastics and a way of organizing thoughts. It is widely known that the owners of calligraphic handwriting in most cases perform twice as better than sloppiness with poor handwriting. It is interesting to compare the ability of children to write on a computer keyboard or pen. Writing with a pen doubles the ability to master new educational material and the logical presentation of one's own thoughts [7].

Mobile services, and messengers in particular, have provided a huge number of users with an environment for a new communication format. This is written speech that does not require the skills of a pen or even a stylus. Moreover, such speech by its nature is more consistent with the spoken style, it is a written form of oral speech, in contrast to the book style of writing [8]. As a result, the user uses the most primitive tools in the form of buttons on the screen or a microphone, and practically unlimited possibilities for saving the volume of writing. Abbreviations of words and whole expressions, the so-called emoticons, gifs and literal expressions of onomatopoeia, are used.

Added to this are voice typing, simple video and audio recording.

Oral speech is undergoing a transformation. It is associated with the access of any individual to a mass audience. As a result, information products created not by professional actors, broadcasters, screenwriters and translators, but by ordinary people themselves, become models for linguistic imitation. Often, even less socialized speakers use the language worse than the public itself. The general decline in the level of spoken language in popular Internet shows removes responsibility from speakers in real life. If before the appearance of social networks, tongue-tiedness was followed by a lowering of social status and public censure, now even speaking in public, a speaker who, in half of his attempts, fails to combine words into a sentence, retains his high social status. On the contrary, in some cases, some kind of disregard for the listener even adds solidity to the speaker. Simply put, those who have a poor command of the language no longer need to hide it, and everyone else can not strain to demonstrate their linguistic capabilities. Speech is reproduced according to the principle of sufficiency, and not on the basis of striving for a standard.

We found a relationship between simplification of speech and simplification of social behavior in general. And vice versa. The approach allows you to quantitatively and qualitatively assess the degree of marginalization of an individual and a social group as a whole.

Bibliography

1. *Kara-Murza S. G.* Soviet civilization. M.: Rodina. 2019 .— 1280 p.
2. Now I know. Children of the USSR. Just listen to their literate speech! YouTube. 2020 (https://youtu.be/K6xV1rt2jJI)
3. Soviet television. GOSTELERADIOFOND. Up to 16 and older. Fartsovschiki, cooperative (1988). YouTube. 2018 (https://youtu.be/TCokgKj4FHU)
4. VOS. Who has a bigger dictionary? W-O-S.ru. 2014 (https://w-o-s.ru/article/9037)
5. News of Togliatti. Letter from 1981: Soviet schoolchildren's predictions about Togliatti, Russia and the world. Tltgorod.ru. 2017 (https://www.tltgorod.ru/news/?news=72735)
6. *Alpatov V. M.* Linguistics: From Aristotle to Computational Linguistics. M.: Alpina non-fiction. 2018 .— 253 p.
7. *Saveliev S.V.* Morphology of consciousness: In 2 volumes. M.: VEDI, 2021.— T. 1.— 3rd ed., Erased.— 224 p.
8. *Misonzhnikov B. Ya.* Phenomenology of the text (the ratio of the content and formal structures of the printed edition). SPb.: Publishing house of St. Petersburg. University, 2001 .— 490 p.

Chapter 6.
Developmental and epileptic encephalopathies as an experiment set by nature

**Gerasimov A. P., Ushanov V. V., Galiavin S. I.,
Oskolkova E. S., Shakel A. A.,
Zhumatov M. M., Lobzhanidze T. B.‑A., Ivanova N. E.**

Almazov National Medical Research Centre,
Pavlov First St.Petersburg State Medical University,
Mechnikov North-Western State Medical University,
St. Petersburg State Pediatric Medical University, St. Petersburg, Russia

According WHO position epilepsy is one of the most common neurological diseases, affecting around 50 million people of all ages around the world. It has a bimodal distribution according to age with peaks in the youngest individuals and in those over 60 years of age (World Health Organization, 2019).

Classification of epilepsy, presented by International League Against Epilepsy at 2017 distributed forms according epilepsy types (focal, generalized, combined generalized & focal, unknown) und etiology (structural, genetic, infectious, metabolic, immune, unknown) (Scheffer et al., 2017). According this classification developmental and epileptic encephalopathies (DEE) are generalized genetic epilepsy.

Pavlov considered disease as an experiment set by nature. In context of visual perception we have many such "nature experiments". Genetic epileptic disorders are usually associated with cortical part of analyzer.

Cortical blindness may be result of many diseases: MELAS syndrome; Alpers syndrome; Mohr-Tranebjaerg syndrome; seizures, cortical blindness, and microcephaly syndrome; Leber congenital amaurosis; developmental and epileptic encephalopathies (DEE) etc. In OMIM request "cortical blindness" leads to 68 nosology positions with many DEE.

We observed the patient with genetically verified DEE started from cortical blindness as the primary symptom.

Developmental and epileptic encephalopathies (DEEs) represent a clinically and genetically heterogeneous group of age-dependent neurologic

disorders characterized by onset of refractory seizures in infancy or early childhood. Affected individuals have delayed psychomotor development or developmental regression, particularly after the onset of seizures. DEE incorporates the previous grouping of 'early infantile epileptic encephalopathies (EIEE).

Advanced genetic methods have enabled more specific delineation of these clinically defined disorders at the molecular level, which has led to their classification as types of DEEs based on the genetic defect. However, the historically delineated clinical entities are still often used to describe the phenotype, regardless of the gene involved (Auvin et al., 2016; Shbarou, Mikati, 2016; Zhou et al., 2018; Steel et al., 2017).

For the future discussions several definitions are necessary.

System genomics is part of system biology. It studies genome as a system (not a random group of genes) using methods of the system analysis.

Now there is big group of OMICs (genomics, transcriptomics, proteomics, metabolomics, etc.). We would like to accent two members of this family.

Interactomics is scientific direction studying interactome — complex of interactions between genes and, secondary, coded by these genes proteins.

Pathwayomics is scientific direction studying pathwayome — complex of physiological and biochemical regulatory consequences (pathways).

We tried to search DEE from positions of system genomics. As the basic instrument of search resource OMIM was used. Standard requests were: developmental and epileptic encephalopathy (DEE). As the additionally requests used: sodium channels; potassium channels; calcium channels; chloride channels; GABA (receptor); glycine (receptor); NMDA (receptor); ion pomp (Na/K ATPase). Directions of search were distributed between members of scientific team.

Information was transformed into database (including phenotype — DEE number, OMIM number of disease, inheritance type, gene name, OMIM number of gene, cytogenetic localization, physiology functions). Analysis includes physiology aspects and interaction between genes and groups of genes (PheneGene "green" button in OMIM). Cytogenetic location of genes was analyzed as an important aspect of genetic interaction.

OMIM data is current for December 2021.

Finally 101 genetic variant was analyzed.

Results

DEE is big group of clinically similar, but physiologically different situations. Accumulation of information is very fast: between presentation and preparing this text for publishing (near 1 month) number of analyzed genetic variants increased from 99 to 101 — DEE98 and DEE99 were published and officially included in OMIM database (Vetro et al., 2021).

Types of inheritance are very different: autosomal dominant, autosomal recessive (most often), X-linked dominant, X-linked recessive. There are different clinical phenotypes (like Ohtahara, West, Dravet syndrome).

Methodical difficulties may be illustrated by short description of DEE1.

Developmental and epileptic encephalopathy 1; DEE1 (OMIM: 308350). This disease has alternative titles: early infantile epileptic encephalopathy 1; EIEE1; infantile spasm syndrome, X-linked 1; ISSX1; West syndrome, X-linked; Ohtahara syndrome, X-linked; infantile epileptic-dyskinetic encephalopathy; XMESID. Clinical description of the syndrome was published by Ohtahara, S. et al. in 1976 (in Japanese!). Phenotype of West syndrome was published by W. J. West, a 19th century neurologist who described the syndrome in his own son (Foldvary-Schaefer, Wyllie, 2003).

DEE1 is part of a phenotypic spectrum of disorders caused by mutation in the ARX gene comprising a nearly continuous series of developmental disorders ranging from lissencephaly (LISX2; 300215) to Proud syndrome (300004) to infantile spasms without brain malformations (DEE) to syndromic (309510) and nonsyndromic (300419) mental retardation. Gene ARX (aristaless-related homeobox, X-linked) is situated at Xp21.3.

For details see DEE1 and ARX description in OMIM (https://www.omim.org/entry/308350, https://www.omim.org/entry/300382).

Functional different groups of genes were selected: ion channels, neurotransmitters and receptors, intracellular signaling, sprouting and synaptogenesis, control of expression, (neurospecific) enzymes, mitochondria-associated genes.

Ion channels. Genes of ion channels were described in epileptology context in previous chapter. 5 sodium channel genes have association with DEE (with modifying effect from SCN9A on SCN1A) (tab.1) (Harkin et al., 2007; Wolff et al., 2017; Ramadan et al., 2017).

DEE and sodium channels

Gene	OMIM number	Locus	Diseases
SCN1A	182389	2q24.3	DEE6 (Dravet and not-Dravet types) with modifying effect from SCN9A at the same locus, GEFSP2, FEB3A, FHM3
SCN2A	182390	2q24.3	DEE11, BFIS3
SCN3A	182391	2q24.3	FFEVF4, DEE62
SCN8A	600702	12q13.13	DEE13, MYOCL2, CIAT, BFIS5
SCN1B	600235	19q13.11	DEE52, GEFSP1, BRGDA5, ATFB13

6 genes of potassium channels have association with DEE — KCN2A, KCNT2, HCN1, KCNT2, KCNB1, KCNQ1 (Gururaj et al., 2017).

3 genes of calcium channels have association with DEE — CACNA1E, PACS2, CACNA1A (Epi4K Consortium, 2016; Olson et al., 2018).

Different ion pumps genes are associated with DEE — ATP6V1A, PIGB, ATP1A2, ATP1A3 (Vetro et al., 2021; Murakami et al., 2019).

Neurotransmitters and receptors. Only 3 neurotransmitter systems have significant association with DEE.

2 genes encoding NMDA receptors — GRIN2B and GRIN2D — are associated with DEE. But additionally several genes are functionally linked with NMDA system (amino acid transporters, ion channels etc.).

GABA system as the leading inhibitory mechanism in CNS is important part of anticonvulsant system. Mutations in 8 receptor genes (GABRA2, GABRB1, GABRB2, GABRA1, GABRG2, GABBR2, GABRB3, GABRA5) are associated with DEE.

Glycine system is linked with DEE indirect way — through interaction with GABA and NMDA systems.

We discovered no DEE association with acetylcholine, norepinephrine, dopamine, serotonin neurotransmitter systems.

Intracellular pathways. Big group of DEE is result of intracellular signaling errors. But these systems are working in complex and linked with extracellular receptors. This way our distribution between groups has not absolutely character.

The epileptic pathology of GTP/GDP/GMP system is associated with genes DNM1, DENND5A, TBC1D24, ARHGEF9.

DEE may be associated with UTP/UDP system and incorrect phosphorylation and transport of glucose and galactose. Genes UGP2, UGDH, SLC35A2 are important.

Kinase/phosphrylase genes in DEE group are well-presented. Genes DOCK7, UGP2, CDK19, NTRK2, PNKP, SIK1, CDKL5, SYNJ1 code enzymes with different localisation.

Calcium-calmodulin signaling pathway in DEE context is directly linked with PPP3CA, secondary — with many genes.

MTOR pathway has association with 2 DEE genes — SZT2 and DOCK7.

With the ubiquitin are associated 3 DEE genes: RHOBTB2, UBA5, NEUROD2.

Inositol 1,4,5-trisphosphate (IP3) and diacylglycerol (DAG) — are directly linked with DEE by mutations in PLCB1, ITPA and SYNJ1.

Glycosylphosphatidylinositol (GPI) is a membrane anchor for cell surface proteins. PIG proteins are subunits of the GPI complex that take part in reception and transduce signal into cell. This very big family of proteins and genes is described at (https://www.omim.org/entry/311770). Mutations in PIGB, PIGQ, PIGS, PIGP, PIGA lead to DEE.

Sprouting and synaptogenesis, synaptic function. Genes associated with DEE may have influence the actin dynamics, actin remodeling, axon elongation, dendritic spine morphogenesis, and synaptic plasticity, cortical migration, dendrite branching. This effect is more significant in prenatal period. Associated genes are: DOCK7, CYFIP2, PHACTR1, ACTL6B, SPTAN1, DENND5A, CUX2, PCDH19.

Disorders of synaptic functions (vesicle transport, neurosecretion, endocytosis), more actual in postnatal period, associated with genes CPLX1, ADAM22, DNM1, NECAP1, DMXL2, AP3B2, NSF, SYNJ1 are linked with DEE too.

Transcription, translation, reparation, proliferation. DEE may be associ-

ated with disorders of control of DNA state and transcription (genes CDK19, YWHAG, ACTL6B, CHD2, PNKP). It is important fact, because understanding of genome instability as mechanism of CNS diseases appeared only several years ago (Iourov et al., 2021).

RNA-assotiated proteins (including tRNA-binding proteins) may be damaged with mutations in the genes PARS2, HNRNPU, DALRD3, CELF2, AARS1, EEF1A2.

Enzymes and mitochondria. Many enzymes may be changed in the situation of DEE. Kinases and phosphrylases were discussed above in signaling section. Classical enzymes (and genes) with possibility of enzymatic block are:

- beta-galactoside-alpha-2,3-sialyltransferase-III — ST3GAL3;
- pyrimidine biosynthesis, carbamoyl phosphate synthetase — CAD;
- malate dehydrogenase — MDH1 (cytoplasmatic) and MDH2 (mitochondrial);
- glutamate decarboxylase — GAD1;
- mitochondrial glutamate oxaloacetate transaminase, also called aspartate aminotransferase, a pyridoxal 5-prime phosphate-dependent enzyme — GOT2.

Damage of mitochondria-associated genes (PARS2, MDH2, SLC25A12, SLC25A22, GOT2, GUF1) may lead to DEE.

Discussion

We demonstrated distribution of DEE genes between several physiological groups. The same time many genes have presented in two or more groups and interaction between groups is physiologically valid. For example, gene ACTL6B encodes a component of the neuron-specific BAF complex. This complex has a role in chromatin remodeling and histone acetylation, which regulates gene expression during development, particularly in the process of dendritic outgrowth (Bell et al., 2019).

Sometimes interaction is unexpected. So, small cytosolic proteins coded FGF12 and FGF13 genes from fibroblast growth factor homologous factor (FHF) family interact with the cytoplasmic tails of voltage-gated sodium channels and elevate the voltage dependence of neuronal sodium channel fast inactivation (Siekierska et al., 2016).

Sometimes interpretation is not clear. We have a patient with West-phe-notype of DEE. He has verified mutations in genes PNKP and SYNJ1. The PNKP gene, associated with microcephaly, seizures, and developmental delay or DEE10, codes a polynucleotide kinase 3-prime phosphatase, which cata-lyzes the 5-prime phosphorylation of nucleic acids and also has an associat-ed 3-prime phosphatase activity, predictive of an important function in DNA repair following ionizing radiation or oxidative damage (Jilani et al., 1999). The SYNJ1 gene, associated with DEE53, codes synaptojanin-1, a polyphos-phoinositide phosphatase that has a role in clathrin-coated pit and synaptic vesicle dynamics (Hardies et al., 2016). Both DEE have autosomal recessive inheritance, but both mutations are in heterozygous state. Compound hetero-zygous mutations are described for the both diseases! Members of family were examined: father has no mutations and is healthy; mother has both mu-tations in PNKP and SYNJ1 and is healthy. This way it is difficult to select causative mutation.

Additionally problem is somatic mosaicism (Iourov et al., 2019). Study-ing of blood is incorrect in this situation, but getting the nervous tissue is near impossible.

Cytogenetic aspects of DEE are very interest. We selected several "hot" regions:

1p31.3–1p34.2–4 genes

2q24.3–4 genes (SCN)

2q31.1–2q32.2–3 genes

4p12–3 genes (2GABR)

5q33–5q34–4 genes (3GABR)

7q11.23–2 genes

9q34.11–3 genes

16q21–16q23.2 (including fragile site 16q23.2) — 3 genes

19q13.11–19q13.33–4 genes

20q12.13–20q12.33–4 genes

21q22.11–22q22.3–3 genes

The same time only 1 gene is situated at chromosomes 8, 10, 14, and there are no DEE genes at chromosomes 13, 18, 22.

Our opinion is this distribution is not random. Indirect argument is discovering gene ATP1A3 of DEE99 in "hot" region 19q13.2.

Physiological processes are provided by proteins (receptors, signal proteins, ion channels, ion pumps, enzymes etc.). Dysfunction of genetic apparatus leads to not adequate quality or quantity of proteins with loss of function. There are many weaknesses at primary level, but then turn on mechanisms of channeling, and results are similar.

DEE and visual disorders are associated with pathology of ion channels, neurotransmitters (NMDA, GABA), synaptic transduction, intracellular signalling, expression and reparation, gistogenesis, metabolic reaction. There is overlapping with mitochondrial pathology. One gene may be associated with 2–3 processes and one pathway may be provided by several genes. Function of genes may depend from pathogenic mutations, formally neutral polymorphisms, genetic interactions.

There are several cytogenetic loci of high presence of DEE genes.

Method of systemic analysis may be useful in neurogenetics.

Acknowledgments

References

1. World Health Organization. (2019). Epilepsy: a public health imperative. World Health Organization. https://apps.who.int/iris/handle/10665/325293.
2. *Scheffer I. E, Berkovic S, Capovilla G, Connolly M. B., French J., Guilhoto L., Hirsch E, Jain S., Mathern G. W., Moshé S. L., Nordli D..R., Perucca E., Tomson T., Wiebe S, Zhang Y..H, Zuberi S..M.* ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. Epilepsia. 2017 Apr;58(4):512–521. doi: 10.1111/epi.13709.

3. *Auvin S, Cilio M. R., Vezzani A*. Current understanding and neurobiology of epileptic encephalopathies. Neurobiol Dis. 2016 Aug;92 (Pt A):72–89. doi: 10.1016/j.nbd.2016.03.007.

4. *Shbarou R, Mikati M. A.* The Expanding Clinical Spectrum of Genetic Pediatric Epileptic Encephalopathies. Semin Pediatr Neurol. 2016 May;23(2):134–42. doi: 10.1016/j.spen.2016.06.002.

5. *Zhou P, He N, Zhang JW, Lin ZJ, Wang J, Yan LM, Meng H, Tang B, Li BM, Liu XR, Shi YW, Zhai QX, Yi YH, Liao WP*. Novel mutations and phenotypes of epilepsy-associated genes in epileptic encephalopathies. Genes Brain Behav. 2018 Nov;17(8): e12456. doi: 10.1111/gbb.12456.

6. *Steel D, Symonds JD, Zuberi SM, Brunklaus A*. Dravet syndrome and its mimics: Beyond SCN1A. Epilepsia. 2017 Nov;58(11):1807–1816. doi: 10.1111/epi.13889.

7. *Vetro A, Nielsen H.N., Holm R., Hevner R.F., Parrini E., Powis Z., Møller R.S., B.ellan C., Simonati A., Lesca G, Helbig K.L., Palmer E.E., Mei D., Ballardini E, Van Haeringen A, Syrbe S, Leuzzi V, Cioni G, Curry CJ, Costain G, Santucci M, Chong K, Mancini GMS, Clayton-Smith J, Bigoni S, Scheffer I.E., Dobyns W.B., Vilsen B., Guerrini R.*; ATP1A2/A3-collaborators. ATP1A2- and ATP1A3-associated early profound epileptic encephalopathy and polymicrogyria. Brain. 2021 Jun 22;144(5):1435–1450. doi: 10.1093/brain/awab052.

8. *Foldvary-Schaefer, N., Wyllie, E.* Epilepsy. In: Goetz, C. G. (ed.): Textbook of Clinical Neurology. (2nd ed.) Philadelphia: Saunders (pub.) 2003. P. 1167.

9. https://www.omim.org/entry/308350.

10. https://www.omim.org/entry/300382.

11. *Louise A. Harkin, Jacinta M. McMahon, Xenia Iona, Leanne Dibbens, James T. Pelekanos, Sameer M. Zuberi, Lynette G. Sadleir, Eva Andermann, Deepak Gill, Kevin Farrell, Mary Connolly, Thorsten Stanley, Michael Harbord, Frederick Andermann, Jing Wang, Sat Dev Batish, Jeffrey G. Jones, William K. Seltzer, Alison Gardner,* The Infantile Epileptic Encephalopathy Referral Consortium,, Grant Sutherland, Samuel F. Berkovic, John C. Mulley, Ingrid E. Scheffer, The spectrum of SCN1A-related infantile epileptic encephalopathies, Brain, Volume 130, Issue 3, March 2007, Pages 843–852, https://doi.org/10.1093/brain/awm002.

12. *Markus Wolff, Katrine M. Johannesen, Ulrike B.S. Hedrich, Silvia Masnada, Guido Rubboli, Elena Gardella, Gaetan Lesca, Dorothée Ville, Mathieu Milh, Laurent Villard, Alexandra Afenjar, Sandra Chantot-Bastaraud, Cyril Mignot, Caroline Lardennois, Caroline Nava, Niklas Schwarz, Marion Gérard, Laurence Perrin, Diane Doummar, Stéphane*

Auvin, Maria J. Miranda, Maja Hempel, Eva Brilstra, Nine Knoers, Nienke Verbeek, Marjan van Kempen, Kees P. Braun, Grazia Mancini, Saskia Biskup, Konstanze Hörtnagel, Miriam Döcker, Thomas Bast, Tobias Loddenkemper, Lily Wong-Kisiel, Friedrich M. Baumeister, Walid Fazeli, Pasquale Striano, Robertino Dilena, Elena Fontana, Federico Zara, Gerhard Kurlemann, Joerg Klepper, Jess G. Thoene, Daniel H. Arndt, Nicolas Deconinck, Thomas Schmitt-Mechelke, Oliver Maier, Hiltrud Muhle, Beverly Wical, Claudio Finetti, Reinhard Brückner, Joachim Pietz, Günther Golla, Dinesh Jillella, Karen M. Linnet, Perrine Charles, Ute Moog, Eve Õiglane-Shlik, John F. Mantovani, Kristen Park, Marie Deprez, Damien Lederer, Sandrine Mary, Emmanuel Scalais, Laila Selim, Rudy Van Coster, Lieven Lagae, Marina Nikanorova, Helle Hjalgrim, G. Christoph Korenke, Marina Trivisano, Nicola Specchio, Berten Ceulemans, Thomas Dorn, Katherine L. Helbig, Katia Hardies, Hannah Stamberger, Peter de Jonghe, Sarah Weckhuysen, Johannes R. Lemke, Ingeborg Krägeloh-Mann, Ingo Helbig, Gerhard Kluger, Holger Lerche, Rikke S Møller, Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders, Brain, Volume 140, Issue 5, May 2017, Pages 1316–1336, https://doi.org/10.1093/brain/awx054.

13. *Ramadan, W, Patel, N, Anazi, S, et al*. Confirming the recessive inheritance of SCN1B mutations in developmental epileptic encephalopathy, Clin Genet, 2017; 92: 327–331. https://doi.org/10.1111/cge.12999.

14. *Gururaj, S., Palmer, E. E., Sheehan, G. D., Kandula, T., Macintosh, R., Ying, K., Morris, P., Tao, J., Dias, K.-R., Zhu, Y., Dinger, M. E., Cowley, M. J., Kirk, E. P., Roscioli, T., Sachdev, R., Duffey, M. E., Bye, A., Bhattacharjee, A*. A de novo mutation in the sodium-activated potassium channel KCNT2 alters ion selectivity and causes epileptic encephalopathy. Cell Rep. 21: 926–933, 2017. doi: 10.1016/j.celrep.2017.09.088.

15. Epi4K Consortium. De Novo Mutations in SLC1A2 and CACNA1A Are Important Causes of Epileptic Encephalopathies. Am J Hum Genet. 2016 Aug 4;99(2):287–98. doi: 10.1016/j.ajhg.2016.06.003.

16. *Olson H.E., Jean-Marçais N., Yang E, Heron D., Tatton-Brown K., van der Zwaag P.A., Bijlsma E.K., Krock B.L., Backer E., Kamsteeg E.J., Sinnema M., Reijnders M.R.F., Bearden D., Begtrup A., Telegrafi A., Lunsing R.J., Burglen L., Lesca G., Cho M.T., Smith L.A., Sheidley B.R., Moufawad E.l. Achkar C., Pearl P.L., Poduri A., Skraban C.M., Tarpinian J., Nesbitt A.I., Fransen van de Putte D.E., Ruivenkamp C.A.L., Rump P., Chatron N., Sabatier I., De Bellescize J., Guibaud L., Sweetser D.A., Waxler J.L., Wierenga K.J.;* DDD Study, *Donadieu J., Narayanan V., Ramsey K.M.;* C4RCD Research Group, *Nava C., Rivière J.B., Vitobello A., Tran Mau-Them F., Philippe C., Bruel A.L., Duffourd Y., Thomas L.,*

Lelieveld S.H., Schuurs-Hoeijmakers J., Brunner H.G., Keren B., Thevenon J., Faivre L., Thomas G., Thauvin-Robinet C. A. Recurrent De Novo PACS2 Heterozygous Missense Variant Causes Neonatal-Onset Developmental Epileptic Encephalopathy, Facial Dysmorphism, and Cerebellar Dysgenesis. Am J Hum Genet. 2018 May 3;102(5):995–1007. doi: 10.1016/j.ajhg.2018.03.005.

17. *Murakami Y., Nguyen T.T.M., Baratang N., Raju P.K., Knaus A., Ellard S., Jones G., Lace B., Rousseau J., Ajeawung N.F., Kamei A., Minase G., Akasaka M., Araya N., Koshimizu E., van den Ende J., Erger F., Altmüller J., Krumina Z., Strautmanis J., Inashkina I., Stavusis J., El-Gharbawy A., Sebastian J., Puri R.D., Kulshrestha S., Verma I.C., Maier E.M., Haack T.B., Israni A., Baptista J., Gunning A., Rosenfeld J.A., Liu P., Joosten M., Rocha M.E., Hashem M.O., Aldhalaan H.M., Alkuraya F.S., Miyatake S., Matsumoto N., Krawitz P.M., Rossignol E., Kinoshita T., Campeau P.M.* Mutations in PIGB Cause an Inherited GPI Biosynthesis Defect with an Axonal Neuropathy and Metabolic Abnormality in Severe Cases. Am J Hum Genet. 2019 Aug 1;105(2):384–394. doi: 10.1016/j.ajhg.2019.05.019.

18. https://www.omim.org/entry/311770.

19. *Iourov, I.Y.; Yurov, Y.B.; Vorsanova, S.G.; Kutsev, S.I.* 2021. Chromosome Instability, Aging and Brain Diseases. Cells 2021, 10, 1256. https://doi.org/10.3390/cells10051256.

20. *Bell S., Rousseau J., Peng H., Aouabed Z., Priam P., Theroux J.F., Jefri M., Tanti A., Wu H., Kolobova I., Silviera H., Manzano-Vargas K., Ehresmann S., Hamdan F.F., Hettige N., Zhang X., Antonyan L., Nassif C., Ghaloul-Gonzalez L, Sebastian J, Vockley J, Begtrup AG, Wentzensen IM, Crunk A, Nicholls R.D., Herman K.C., Deignan J.L., Al-Hertani W., Efthymiou S., Salpietro V., Miyake N., Makita Y., Matsumoto N., Østern R., Houge G., Hafström M., Fassi E., Houlden H., Klein Wassink-Ruiter JS, Nelson D, Goldstein A, Dabir T, van Gils J, Bourgeron T, Delorme R, Cooper G.M., Martinez J.E., Finnila C.R., Carmant L., Lortie A., Oegema R., van Gassen K., Mehta S.G., Huhle D., Abou Jamra R., Martin S., Brunner H.G., Lindhout D., Au M., Graham J.M. Jr., Coubes C., Turecki G., Gravel S., Mechawar N., Rossignol E., Michaud J.L., Lessard J., Ernst C., Campeau P.M.* Mutations in ACTL6B Cause Neurodevelopmental Deficits and Epilepsy and Lead to Loss of Dendrites in Human Neurons. Am J Hum Genet. 2019 May 2;104(5):815–834. doi: 10.1016/j.ajhg.2019.03.022.

21. *Siekierska A., Isrie M., Liu Y., Scheldeman C., Vanthillo N., Lagae L, de Witte P.A., Van Esch H., Goldfarb M., Buyse G.M.* Gain-of-function FHF1 mutation causes early-onset epileptic encephalopathy with cer-

ebellar atrophy. Neurology. 2016 Jun 7;86(23):2162–70. doi: 10.1212/WNL.0000000000002752.

22. *Jilani A, Ramotar D, Slack C, Ong C, Yang XM, Scherer SW, Lasko DD*. Molecular cloning of the human gene, PNKP, encoding a polynucleotide kinase 3'-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage. J Biol Chem. 1999 Aug 20;274(34):24176–86. doi: 10.1074/jbc.274.34.24176.

23. *Hardies K, Cai Y, Jardel C, Jansen AC, Cao M, May P, Djémié T, Hachon Le Camus C, Keymolen K, Deconinck T, Bhambhani V, Long C, Sajan SA, Helbig KL; AR working group of the EuroEPINOMICS RES Consortium, Suls A, Balling R, Helbig I, De Jonghe P, Depienne C, De Camilli P, Weckhuysen S*. Loss of SYNJ1 dual phosphatase activity leads to early onset refractory seizures and progressive neurological decline. Brain. 2016 Sep;139(Pt 9):2420–30. doi: 10.1093/brain/aww180.

24. *Iourov I. Y., Vorsanova S. G., Yurov Y. B., Kutsev S. I.* 2019. Ontogenetic and Pathogenetic Views on Somatic Chromosomal Mosaicism. Genes 2019, 10, 379; doi:10.3390/genes10050379.

Chapter 7.
Genetic control of ion channels in epilepsy from the standpoint of functional genomics and interactomics

Gerasimov A. P., Zavrazhnova A. A., Komarova D. A., Samusenkova D. V., Okhrimenko A. S., Kishinskaja E. I., Kalandarova N. N., Ishchenko I. O., Dyachuk N. M., Ushanov V. V.

Almazov National Medical Research Centre,
Pavlov First St.Petersburg State Medical University,
Mechnikov North-Western State Medical University,
St.Petersburg State Pediatric Medical University, St. Petersburg, Russia

Situation in clinical diagnostic and fundamental search of genetic and multifactorial diseases changed notably in recent years. Cytogenetic, biochemical and molecular genetic diagnostics are available in principle (Mosele et al., 2020). Organization of centralized laboratory services, the outposts (offices), biocourier system and biobanks made molecular diagnostic principally accessable (Linsen et al., 2019; Beskow, 2019; Hewitt, Watson, 2013; Rush et al., 2020). Electronic communications network from local hospital net up to internet and OMIM database help the diagnostics significantly (Dive et al., 2020).

Another problem exists. "Gene search" may be incorrect task. Disease may be result of copy number variation (CNV) (Iourov et al., 2019) and mutations in noncoding sequences (Luchini et al., 2019). Disease (and normal state) may be result of combination of effects of several genes, and there may be helpful gene maps and genomic nets (Iourov et al., 2021). Disease may be result from the summation in a dangerous direction of the effects of formally neutral polymorphisms. Finally, the clinical associations are described not for all genes, the databases like OMIM and ClinVar are constantly under updating.

Our search was performed in the context of molecular epileptology. Classification of epilepsy, presented by International League Against Epilepsy at 2017 distributed forms according epilepsy types (focal, generalized, combined generalized & focal, unknown) und etiology (structural, genetic, infec-

tious, metabolic, immune, unknown). This way not only generalized genetic epilepsy (including developmental and epileptic encephalopathies — DEE) but focal genetic epilepsy is possible! (Scheffer et al., 2017).

For the future discussions several definitions are necessary.

System genomics is part of system biology. It studies genome as a system (not a random group of genes) using methods of the system analysis.

Now there is big group of OMICs (genomics, transcriptomics, proteomics, metabolomics, etc.). We would like to accent two members of this family.

Interactomics is scientific direction studying interactome — complex of interactions between genes and, secondary, coded by these genes proteins.

Pathwayomics is scientific direction studying pathwayome — complex of physiological and biochemical regulatory consequences (pathways).

As the basic instrument of search resource OMIM was used. Because ion channels are voltage-gated and ligand-gated we included in analysis the most associated with epilepsy neurotransmitter systems. Standard requests were: sodium channels; potassium channels; calcium channels; chloride channels; GABA (receptor); glycine (receptor); NMDA (receptor); ion pomp (Na/K AT-Pase). Directions of search were distributed between members of scientific team.

Information was transformed into formalized database (including gene name, OMIM number, cytogenetic localization, associated diseases, physiology functions). Analysis includes physiology aspects and interaction between genes and groups of genes (PheneGene "green" button in OMIM). Cytogenetic location of genes was analyzed as an important aspect of genetic interaction.

Results

Sodium channels. Physiology functions of voltage-gated sodium channel are generation and propagation of action potentials, mainly in nerve and muscle. Voltage-sensitive sodium channels are heteromeric complexes consisting of a large central pore-forming glycosylated alpha subunit and 2 smaller auxiliary beta subunits (Baroni, Moran, 2015).

Physiology functions of nonvoltage-gated sodium channel (the superfamily of DEG/ENaC) are directed transport of Na + ions, through which sodium channels are involved in neurotransmission, pain perception, mechanotransduction, control of the balance of Na + and water in the body. Subtypes of sodium

channels are ENaC (subunits: α, β, γ, δ) and ASIC (subunits: ASIC1a, ASIC1b, ASIC2a, ASIC2b, ASIC3, ASIC4, ASIC5) (Hernandez et al., 2021).

Generally association of sodium channels with different diseases (including epilepsy) is well described (Catterall, 2017). We added actual information from OMIM source.

Association with diseases was found in 16 sodium channels genes, 13 of them are voltage-gated and 3 are nonvoltage-gated.

Diseases of CNS associated with sodium channels are: developmental and epileptic encephalopathy, 6 (Dravet and non-Dravet types), 11, 13, 52, 62 types; generalized epilepsy with febrile seizures plus, types 1, 2, 7; familial focal epilepsy with variable foci 4; familial febrile seizures 3A and 3B; familial hemiplegic migraine 3; benign familial infantile seizures 3, 5 types; cognitive impairment with or without cerebellar ataxia. In the peripheral nervous system association with sodium channels have such variants of pathology as: primary erythermalgia; congenital insensitivity to pain, paroxysmal extreme pain disorder, small fiber neuropathy, familial episodic pain syndrome 2 and 3 types; hereditary sensory and autonomic neuropathy, type VII (Trump et al., 2016). Sodium channel associated myopathies are: hyperkalemic periodic paralysis, type 2; hypokalemic periodic paralysis, type 2; congenital myasthenic syndrome, type 16; acetazolamide-responsive atypical myotonia congenital; paramyotonia congenital; familial myoclonus, type 2. Big group of heart diseases is linked with mutations in genes of this group: Brugada syndrome 1, 5, 7 types; dilated cardiomyopathy 1E; nonprogressive heart block; progressive heart block, type IA; long QT syndrome 3, 10; sick sinus syndrome 1; familial ventricular fibrillation 1; familial atrial fibrillation, 10,13, 14, 16, 17 types; nonspecific cardiac conduction defect.

It is important, that for gene SCN1B are described both epileptic and cardiology disorders.

Association with diseases was found in 16 sodium channels genes, 13 of them are voltage-gated and 3 are nonvoltage-gated.

Preferred loci with a high representation of genes were identified:

2q24.3: SCN1A, SCN2A, SCN3A, SCN7A/SCN6A, SCN9A

3p22.2: SCN5A, SCN10A, SCN11A

11q23.3: SCN2B, SCN3B, SCN4B

For genes of sodium channels and associated diseases modifying genes

are described.

Potassium channels. Physiology functions of potassium channel are: conducting potassium ions down their electrochemical gradient; setting or resetting the resting potential in many cells; shaping the action potential in neurons by the delayed counterflow of potassium ions.

There are several genetic and functional groups, the biggest (with subgroups) is group of KCN genes.

Many paroxysmal disorders are associated with potassium channels: developmental and epileptic encephalopathy, 7, 14, 24, 57 types; generalized epilepsy with febrile seizures plus, type 10; epilepsy, progressive myoclonic 3, with or without intracellular inclusions; myokymia; benign neonatal seizures, type 1 and 2; cerebellar atrophy, developmental delay, and seizures; Liang-Wang syndrome; paroxysmal nonkinesigenic dyskinesia, 3, with or without generalized epilepsy; epilepsy, intellectual/developmental delay (Inuzuka et al., 2020; Mao et al., 2020; Marini et al., 2018; Nava et al., 2014; Bailey et al., 2019; Liang et al., 2019). Other disorders associated with potassium channels are: spinocerebellar ataxia 13 and 19; enlarged vestibular aqueduct, digenic, SESAME syndrome; autosomal dominant mental retardation, 46; Birk-Barel mental retardation dysmorphism syndrome; microcephaly, cataracts, impaired intellectual development, and dystonia with abnormal striatum; Temple-Baraitser syndrome; Birk-Barel mental retardation dysmorphism syndrome; Bartter syndrome, type 2; susceptibility to thyrotoxic periodic paralysis, type 2. This way here is presented wide range of diseases — from epileptic encephalopathy up to endocrinology.

Cytogenetic distribution of genes was wide, but preferred locus with a high representation of genes was identified — 1p13.2–1p13.3: KCND3, KCNC4, KCNA10.

Calcium channels. Physiology functions of calcium ion channels are mediation Ca^{2+} influx in response to action potentials, initiation contraction in cardiac, smooth and skeletal muscle cells, initiation of synaptic transmission in neurons, initiation of secretion of hormones in endocrine cells, activation of calmodulin-dependent protein kinase II (CaMKII) and protein kinase C (PKC).

Calcium channels may be separated into two groups by activation mechanism: voltage-gated (genes CACN group) and ligand-gated (TRM group).

The voltage-gated calcium channels convert membrane electrical signals to intracellular Ca^{2+}-mediated events.

They are formed as a complex of several different subunits: α1, α2, β, and γ. The α1 subunit forms the ion conducting pore while the associated subunits have several functions including modulation of gating.

Genes of CACN group have association with retinal pathology (Aland Island eye disease; cone-rod dystrophy, X-linked, 3; night blindness, congenital stationary (incomplete), 2A, X-linked; retinal cone dystrophy 4), seizures (developmental and epileptic encephalopathy, 42 and 69 types; primary aldosteronism, seizures, and neurologic abnormalities; susceptibility to childhood absence epilepsy 6; susceptibility to idiopathic generalized epilepsy 6 and 9; susceptibility to juvenile myoclonic epilepsy 6), heart pathology (Brugada syndrome 3 and 4; long QT syndrome 8; Timothy syndrome; sinoatrial node dysfunction and deafness), other diseases (malignant hyperthermia susceptibility 5; susceptibility to thyrotoxic periodic paralysis 1; hypokalemic periodic paralysis, type 1; episodic ataxia, type 2; familial hemiplegic migraine 1; spinocerebellar ataxia 6 and 42; familial hyperaldosteronism, type IV) (Liu et al., 2020; Du et al., 2019; Matthews et al., 2019).

Genes of TRM group are associated with: primary avascular necrosis of femoral head, 2; sodium serum level QTL 1; brachyolmia type 3; familial digital arthropathy-brachydactyly; hereditary motor and sensory neuropathy, type IIc; metatropic dysplasia; distal hereditary motor neuronopathy, type VIII; parastremmatic dwarfism; scapuloperoneal spinal muscular atrophy; spondyloepiphyseal dysplasia, Maroteaux type; spondylometaphyseal dysplasia, Kozlowski type; transient neonatal hyperparathyroidism; susceptibility to amyotrophic lateral sclerosis-parkinsonism/dementia complex (Blum et al., 2019; Dyment et al., 2019).

Cytogenetic distribution of genes was wide, but preferred loci with a high representation of genes were identified:

3p21.31–3p14.3: CACNA1D, CACNA2D2, CACNA2D3.

19q13.42: CACNG6, CACNG7, CACNG8.

<u>Chloride channels.</u> Functions of chloride channels are: Cl transport in the electrochemical gradient direction, making constant potential difference in a resting cell and a cell with changed functional activity, the cell volumoregulation, substance transport through the epithelium, involved in fluid secretion (Pauli et al., 2000).

Group CLCN genes code different variants of chloride channels. The chloride homeostasis of neurons and nonneuronal cells is maintained in part

by chloride conductance through the CLCN2 channel.

CLCN genes are associated with: myotonia congenital; myotonia levior; leukoencephalopathy with ataxia; familial hyperaldosteronism, type II; susceptibility to idiopathic generalized epilepsy, 11; susceptibility to juvenile absence epilepsy, 2; susceptibility to juvenile myoclonic epilepsy, 8; Raynaud syndrome; Dent's disease; hypopigmentation, organomegaly, and delayed myelination and development (Sun et al., 2010; Claverie-Martin et al., 2005). 2 of 7 genes (CLCN3 and CLCN6) have no clinical associations.

CLCN genes have wide cytogenetic distribution.

Na^+/K^+ pump. Physiology functions of Na^+/K^+-ATPase are maintenance of transmembrane electrochemical gradient and cell volume regulation.

Na^+/K^+-ATPase consist of 2 subunits, a large catalytic subunit (alpha) and a smaller glycoprotein subunit (beta). Subunit alpha encoded by ATP1A1, ATP1A1-AS1, ATP1A2, ATP1A3, ATP1A4. Subunit beta encoded by ATP1B1, ATP1B2, ATP1B3. Modulatory subunit gamma, encoded by FXYD2, FXYD4, FXYD6.

Many other genes are associated with Na^+/K^+-ATPase functions, including MECP2 (Rett syndrome) (Li et al., 2020). Totally 46 genes were analyzed.

Genes of ATP1 group have clinical association with: Charcot-Marie-Tooth disease, axonal, type 2DD; hypomagnesemia, seizures, and mental retardation 2, alternating hemiplegia of childhood 1 and 2; familial basilar migraine; familial hemiplegic migraine, 2; Developmental and epileptic encephalopathy 98 and 99; fetal akinesia, respiratory insufficiency, microcephaly, polymicrogyria, and dysmorphic facies; CAPOS syndrome; dystonia type 12 (Vetro et al., 2021; Lassuthova et al., 2018).

From 46 analyzed genes only 27 have clinical phenotype. 15 (55,5 %) genes with phenotype relate to neurological diseases, 13 (48 %) genes-disorders of conduction (neuropathy, deafness, problems with conduction system of the heart, 2-relate to leukoencephalopathy).

Cytogenetic distribution of genes was wide, but preferred loci with a high representation of genes were identified:

1q23.2–1q24.2: ATP1A2, ATP1A4, ATP1B1, KCNJ10, LY9.

11q23.3–11q24.2: FXYD2, FXYD6, HEPACAM

19q13.12–19q13.2: FXYD1, FXYD5, FXYD7.

GABA receptors. GABA neurotransmitter system is one of the most important mechanisms of postsynaptic inhibition in CNS. Group of GABAA ionotropic receptors consist of alpha, beta, gamma (often), delta (rare) subunits. GABAB metabotropic receptors have subunits R1 and R2.

Action of GABA system is associated with big number of proteins and genes, very important are GABA transporters of SLC family.

We analyzed 33 genes, including 12 GABARA receptor and 2 GABARB receptor genes.

GABA receptors are associated with: developmental and epileptic encephalopathy 19, 43, 45, 54, 74, 78, 79 types; infantile or early childhood epileptic encephalopathy, 2; familial febrile seizures 8; generalized epilepsy with febrile seizures plus, type 3; susceptibility to generalized epilepsy with febrile seizures plus, type 5; susceptibility to childhood absence epilepsy 4, 5; susceptibility to juvenile myoclonic epilepsy, 5; neurodevelopmental disorder with poor language and loss of hand skills; susceptibility to autism 1; susceptibility to alcohol and nicotine dependence (Hernandez, Macdonald, 2019; Burgess, et al., 2019).

Cytogenetic distribution of genes was wide, but preferred loci with a high representation of genes were identified:

4p12: GABRA2, GABRA4, GABRB1, GABRG1.

5q34: GABRA1, GABRB2, GABRG2.

Glycine receptors. Glycine neurotransmitter system is essential preventer of overstimulation of the central neural system. Glycinergic synapse inhibits α-motoneurons, that allows to make alternating muscle contractions. Glycine receptors consist of alpha (genes GLRA1, GLRA2, GLRA3, pseudogene GLRA4) and beta (gene GLRB) subunits. For the work of this system are important D-amino acid oxidase (DAO) with activator (DAOA) and glycine transporters of SLCA family.

Genes of glycine system are associated with hyperekplexia 1, 2, 3 types and glycine encephalopathy with normal serum glycine (Mademont-Soler et al., 2021; Wu et al., 2020).

Genes of glycine system have wide cytogenetic distribution.

NMDA receptors. N-methyl-D-aspartate receptor is a specific type of ionotropic glutamate receptor and ion channel. Physiology functions of this system are: synaptic plasticity, learning and memory, control of apoptosys

in complex with other mediator systems. Subunits of NMDA receptors are typically two GluN1 and two GluN2 subunits, but combination of subunits depends from part of brain and age.

Genes of receptor subunits are in GRIN and GRIK groups, but there are several functional associated genes.

21 genes were analyzed, 12 genes have association with diseases, but 9 have not. NMDA receptor genes are associated with developmental and epileptic encephalopathy 27 and 46 types; neurodevelopmental disorder with or without hyperkinetic movements and seizures, autosomal dominant and autosomal recessive types; focal epilepsy with speech disorder and with or without impaired intellectual development; intellectual developmental disorder, autosomal dominant 6, with or without seizures; mental retardation, autosomal recessive, 6 (Chen et al., 2017; Platzer et al., 2017).

Locus of high concentration of glutamate receptors genes is discovered: 9q31.1–9q34.3 (GRIN1, GRIN3A, NSMF)

Discussion

In all of the analyzed groups of genes we see wide cytogenetic distribution, but the same times there are compact loci with big number of analyzed genes (table1).

Table 1

Distribution of subunits

System	Families of genes	Number of genes	Number of loci
Na	SCN	14	7
K	KCN	31	18
Ca	CACN	25	11
Cl	CLCN	7	7
Na/K	ATP	8	5
Gly	GLR	4+1	3
GABA	GABR	12	7
NMDA	GRIN, GRIK	8+5	10
5HT	HTR	17	11

There is big number of interaction between groups. So, hypokalemic periodic paralysis is associated with mutations in genes of sodium and calcium channels. Glycine receptors function is linked with NMDA and GABA systems.

Possible mechanism of interaction is influence of near situated genes. We demonstrated several "hot points". As the final illustration locus 1p36, very important in clinical aspect, is demonstrated (table 2).

Table 2

Analyzed genes of cytogenetic locus 1p36.

System	Gene	OMIM	Locus	Diseases
K	KCNA2B	601142	1p36. 31	
Cl	CLCN6	602726	1p36.22	
Na/K	AGRN	103320	1p36.33	Myasthenic syndrome, congenital, 8, with pre- and postsynaptic defects
GABA	GABRD	137163	1p36.33	{Epilepsy, generalized, with febrile seizures plus, type 5, susceptibility to}, {Epilepsy, idiopathic generalized, 10}, {Epilepsy, juvenile myoclonic, susceptibility to}

Important methodic conclusion is need for NGS-sequencing and array-CGH for complex genetic analysis.

Epilepsy is multifactorial group of diseases. There are many weaknesses at primary level and different injuring factors, but then turn on mechanisms of channeling.

Epilepsy, especially early forms, is associated with intellectual and perceptive disorders. Epilepsy, including DEE, is associated with genes of many ions and neurotransmitter systems. Our search demonstrates possibility of systematization — from chaotic random list of genes to functional and cytogenetic classification. Genes of each of these systems have wide cytogenetic distribution, but with loci of high presentation.

This way method of system genomic and interactomic may be very useful.

The material of this chapter was discussed by the authors at the V IEEE International Conference «Video and Audio Signal Processing in the Context of Neurotechnologies» November 9 to 13, 2020 in St. Petersburg and, partially, at VI IEEE International Conference "Video and Audio Signal Processing in the Context of Neurotechnologies" SPCN — 2021 8–12 of November 2021

St. Petersburg, Russia.

Acknowledgments

References

1. *Mosele F, Remon J, Mateo J, Westphalen CB, Barlesi F, Lolkema MP, Normanno N, Scarpa A, Robson M, Meric-Bernstam F, Wagle N, Stenzinger A, Bonastre J, Bayle A, Michiels S, Bièche I, Rouleau E, Jezdic S, Douillard JY, Reis-Filho JS, Dienstmann R, André F.* Recommendations for the use of next-generation sequencing (NGS) for patients with metastatic cancers: a report from the ESMO Precision Medicine Working Group. 2020. Ann Oncol. 2020 Nov;31(11):1491–1505. doi: 10.1016/j.annonc.2020.07.014.

2. *Linsen L, T'Joen V, Van Der Straeten C, Van Landuyt K, Marbaix E, Bekaert S and Ectors N.* 2019. Biobank Quality Management in the BB-MRI.be Network. Front. Med. 6:141. doi: 10.3389/fmed.2019.00141.

3. *Beskow, A.* 2019. Uppsala Biobank — the development of a biobank organization in a local, regional, and national setting. Upsala Journal of Medical Sciences, 124(1), 6–8. https://doi.org/10.1080/03009734.2018.1547992.

4. *Hewitt R. & Watson P.* 2013. Biopreservation and Biobanking.Oct 2013.309–315. http://doi.org/10.1089/bio.2013.0042.

5. *Rush A., Catchpoole D. R., Ling R., Searles A., Watson P. H., Byrne J. A.* 2020. Improving Academic Biobank Value and Sustainability Through an Outputs Focus. Value Health. 2020 Aug;23(8):1072–1078. doi: 10.1016/j.jval.2020.05.010.

6. *Dive, L., Critchley, C., Otlowski, M. et al.* 2020. Public trust and global biobank networks. BMC Med Ethics 21, 73 (2020). https://doi.org/10.1186/s12910–020–00515–0.

7. *Iourov I. Y., Vorsanova S. G., Yurov Y. B., Kutsev S. I.* 2019. Ontogenetic and Pathogenetic Views on Somatic Chromosomal Mosaicism. Genes 2019, 10, 379; doi:10.3390/genes10050379.

8. *Luchini C, Bibeau F, Ligtenberg MJL, Singh N, Nottegar A, Bosse T, Miller R, Riaz N, Douillard JY, Andre F, Scarpa A.* ESMO recommendations on microsatellite instability testing for immunotherapy in cancer,

and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach. 2019. Ann Oncol. 2019 Aug 1;30(8):1232–1243. doi: 10.1093/annonc/mdz116.

9. *Iourov, I.Y.; Yurov, Y.B.; Vorsanova, S.G.; Kutsev, S.I.* 2021. Chromosome Instability, Aging and Brain Diseases. Cells 2021, 10, 1256. https://doi.org/10.3390/cells10051256.

10. Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guil*hoto L, Hirsch E, Jain S, Mathern GW, Moshé SL, Nordli DR, Perucca E, Tomson T, Wiebe S, Zhang YH, Zuberi SM.* ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. Epilepsia. 2017 Apr;58(4):512–521. doi: 10.1111/epi.13709.

11. *Baroni D and Moran O* (2015) On the multiple roles of the voltage gated sodium channel β1 subunit in genetic diseases. Front. Pharmacol. 6:108. doi: 10.3389/fphar.2015.00108.

12. *Carlos M. Hernandez; John R. Richards* Physiology, Sodium Channels. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. https://www.ncbi.nlm.nih.gov/books/NBK545257/

13. *Catterall, W.A.* Forty Years of Sodium Channels: Structure, Function, Pharmacology, and Epilepsy. Neurochem Res 42, 2495–2504 (2017). https://doi.org/10.1007/s11064–017–2314–9.

14. *Trump N, McTague A, Brittain H, et al* Improving diagnosis and broadening the phenotypes in early-onset seizure and severe developmental delay disorders through gene panel analysis Journal of Medical Genetics 2016;53:310–317. doi:10.1136/jmedgenet-2015–103263.

15. I*nuzuka, L. M., Macedo-Souza, L. I., Della-Ripa, B., Monteiro, F. P., Ramos, L., Kitajima, J. P., Garzon, E., Kok, F.* Additional observation of a de novo pathogenic variant in KCNT2 leading to epileptic encephalopathy with clinical features of frontal lobe epilepsy. Brain Dev. 42: 691–695, 2020. doi: 10.1016/j.braindev.2020.05.003.

16. *Mao, X., Bruneau, N., Gao, Q., Becq, H., Jia, Z., Xi, H., Shu, L., Wang, H., Szepetowski, P., Aniksztejn, L.* The epilepsy of infancy with migrating focal seizures: identification of de novo mutations of the KCNT2 gene that exert inhibitory effects on the corresponding heteromeric KNa1.1/KNa1.2 potassium channel. Front. Cell. Neurosci. 14: 1, 2020. Note: Electronic Article. doi: 10.3389/fncel.2020.00001.

17. *Marini, C., Porro, A., Rastetter, A., Dalle, C., Rivolta, I., Bauer, D., Oegema, R., Nava, C., Parrini, E., Mei, D., Mercer, C., Dhamija, R., and 58 others*. HCN1 mutation spectrum: from neonatal epileptic encephalopathy to benign generalized epilepsy and beyond. Brain 141: 3160–3178, 2018. doi: 10.1093/brain/awy263.

18. *Nava, C., Dalle, C., Rastetter, A., Striano, P., de Kovel, C. G. F., Nabbout,*

R., Cances, C., Ville, D., Brilstra, E. H., Gobbi, G., Raffo, E., Bouteiller, D., and 26 others. De novo mutations in HCN1 cause early infantile epileptic encephalopathy. Nature Genet. 46: 640–645, 2014. doi: 10.1038/ng.2952.

19. *Bailey, C. S., Moldenhauer, H. J., Park, S. M., Keros, S., Meredith, A. L.* KCNMA1-linked channelopathy. J. Gen. Physiol. 151: 1173–1189, 2019. doi:10.1085/jgp.201912457.

20. *Liang, L., Li, X., Moutton, S., Schrier Vergano, S. A., Cogne, B., de Saint-Martin, A., Hurst, A. C. E., Hu, Y., Bodamer, O., Thevenon J., Hung, C. Y., Isidor, B., and 16 others*. De novo loss-of-function KCNMA1 variants are associated with a new multiple malformation syndrome and a broad spectrum of developmental and neurological phenotypes. Hum. Molec. Genet. 28: 2937–2951, 2019. doi:10.1093/hmg/ddz117.

21. *Liu, G., Papa, A., Katchman, A.N. et al*. Mechanism of adrenergic CaV1.2 stimulation revealed by proximity proteomics. Nature 577, 695–700 (2020). https://doi.org/10.1038/s41586–020–1947-z.

22. *Du X, Wei C, Hejazi Pastor DP, Rao ER, Li Y, Grasselli G, Godfrey J, Palmenberg AC, Andrade J, Hansel C, Gomez CM*. α1ACT Is Essential for Survival and Early Cerebellar Programming in a Critical Neonatal Window. Neuron. 2019 May 22;102(4):770–785.e7. doi: 10.1016/j.neuron.2019.02.036.

23. *Matthews, A.M., Blydt-Hansen, I., Al-Jabri, B. et al*. Atypical cerebral palsy: genomics analysis enables precision medicine. Genet Med 21, 1621–1628 (2019). https://doi.org/10.1038/s41436–018–0376-y.

24. Trpc5 deficiency causes hypoprolactinemia and altered function of oscillatory dopamine neurons in the arcuate nucleus, Ana Moreno-Pérez, Martina Pyrski, Bernd Bufe, Anela Arifovic, Petra Weissgerber, Marc Freichel, Frank Zufall, Trese Leinders-Zufall Proceedings of the National Academy of Sciences Jul 2019, 201905705; DOI: 10.1073/pnas.1905705116.

25. *Dyment, D.A., Terhal, P.A., Rustad, C.F. et al*. De novo substitutions of TRPM3 cause intellectual disability and epilepsy. Eur J Hum Genet 27, 1611–1618 (2019). https://doi.org/10.1038/s41431–019–0462-x.

26. *Pauli, B. U., Abdel-Ghany, M., Cheng, H.—C., Gruber, A. D., Archibald, H. A., Elble, R. C*. Molecular characteristics and functional diversity of CLCA family members. Clin. Exp. Pharm. Physiol. 27: 901–905, 2000. DOI: 10.1046/j.1440–1681.2000.03358.x.

27. S*un, C., Tranebjaerg, L., Torbergsen, T., Holmgren, G., Van Ghelue, M*. Spectrum of CLCN1 mutations in patients with myotonia congenita in northern Scandinavia. Europ. J. Hum. Genet. 9: 903–909, 2001. Note: Erratum: Europ. J. Hum. Genet. 18: 264 only, 2010 DOI: 10.1038/sj.ejhg.5200736.

28. *Claverie-Martin, F., Flores, C., Anton-Gamero, M., Gonzalez-Acosta, H., Garcia-Nieto, V*. The Alu insertion in the CLCN5 gene of a patient with Dent's disease leads to exon 11 skipping. J. Hum. Genet. 50: 370–374, 2005. DOI: 10.1007/s10038–005–0265–5.

29. L*i CH, Coffey EL, Dall'Agnese A, Hannett NM, Tang X, Henninger JE, Platt JM, Oksuz O, Zamudio AV, Afeyan LK, Schuijers J, Liu XS, Markoulaki S, Lungjangwa T, LeRoy G, Svoboda DS, Wogram E, Lee TI, Jaenisch R, Young RA*. MeCP2 links heterochromatin condensates and neurodevelopmental disease. Nature. 2020 Oct;586(7829):440–444. doi: 10.1038/s41586–020–2574–4.

30. *Vetro A, Nielsen HN, Holm R, Hevner RF, Parrini E, Powis Z, Møller RS, Bellan C, Simonati A, Lesca G, Helbig KL, Palmer EE, Mei D, Ballardini E, Van Haeringen A, Syrbe S, Leuzzi V, Cioni G, Curry CJ, Costain G, Santucci M, Chong K, Mancini GMS, Clayton-Smith J, Bigoni S, Scheffer IE, Dobyns WB, Vilsen B, Guerrini R*; ATP1A2/A3-collaborators. ATP1A2- and ATP1A3-associated early profound epileptic encephalopathy and polymicrogyria. Brain. 2021 Jun 22;144(5):1435–1450. doi: 10.1093/brain/awab052.

31. *Lassuthova P, Rebelo AP, Ravenscroft G, Lamont PJ, Davis MR, Manganelli F, Feely SM, Bacon C, Brožková DŠ, Haberlova J, Mazanec R, Tao F, Saghira C, Abreu L, Courel S, Powell E, Buglo E, Bis DM, Baxter MF, Ong RW, Marns L, Lee YC, Bai Y, Isom DG, Barro-Soria R, Chung KW, Scherer SS, Larsson HP, Laing NG, Choi BO, Seeman P, Shy ME, Santoro L, Zuchner S*. Mutations in ATP1A1 Cause Dominant Charcot-Marie-Tooth Type 2. Am J Hum Genet. 2018 Mar 1;102(3):505–514. doi: 10.1016/j.ajhg.2018.01.023.

32. *Ciria C. Hernandez, Robert L*. Macdonald A structural look at GABAA receptor mutations linked to epilepsy syndromes Brain Research, Volume 1714, 2019, Pages 234–247, ISSN0006–8993, https://doi.org/10.1016/j.brainres.2019.03.004.

33. *Burgess, R., Wang, S., McTague, A., Boysen, K.E., Yang, X., Zeng, Q., Myers, K.A., Rochtus, A., Trivisano, M., Gill, D.,, Sadleir, L.G., Specchio, N., Guerrini, R., Marini, C., Zhang, Y.-H., Mefford, H.C., Kurian, M.A., Poduri, A.H. and Scheffer, I.E.* (2019), The Genetic Landscape of Epilepsy of Infancy with Migrating Focal Seizures. Ann Neurol, 86: 821–831. https://doi.org/10.1002/ana.25619.

34. *Mademont-Soler I, Casellas-Vidal D, Trujillo A, Espuña-Capote N, Maroto A, García-González MDM, Ruiz MD, Diego-Álvarez D, Queralt X, Perapoch J, Obón M*. GLYT1 encephalopathy: Further delineation of disease phenotype and discussion of pathophysiological mechanisms. Am J Med Genet A. 2021 Feb;185(2):476–485. doi: 10.1002/ajmg.a.61996.

35. *Wu Z, Lape R, Jopp-Saile L, O'Callaghan BJ, Greiner T, Sivilotti LG.* The startle disease mutation α1S270T predicts shortening of glycinergic synaptic currents. J Physiol. 2020 Aug;598(16):3417–3438. doi: 10.1113/JP279803.

36. *Chen W, Shieh C, Swanger SA, Tankovic A, Au M, McGuire M, Tagliati M, Graham JM, Madan-Khetarpal S, Traynelis SF, Yuan H, Pierson TM.* GRIN1 mutation associated with intellectual disability alters NMDA receptor trafficking and function. J Hum Genet. 2017 Jun;62(6):589–597. doi: 10.1038/jhg.2017.19.

37. *Platzer K, Yuan H, Schütz H, Winschel A, Chen W, Hu C, Kusumoto H, Heyne HO, Helbig KL, Tang S, Willing MC, Tinkle BT, Adams DJ, Depienne C, Keren B, Mignot C, Frengen E, Strømme P, Biskup S, Döcker D, Strom TM, Mefford HC, Myers CT, Muir AM, LaCroix A, Sadleir L, Scheffer IE, Brilstra E, van Haelst MM, van der Smagt JJ, Bok LA, Møller RS, Jensen UB, Millichap JJ, Berg AT, Goldberg EM, De Bie I, Fox S, Major P, Jones JR, Zackai EH, Abou Jamra R, Rolfs A, Leventer RJ, Lawson JA, Roscioli T, Jansen FE, Ranza E, Korff CM, Lehesjoki AE, Courage C, Linnankivi T, Smith DR, Stanley C, Mintz M, McKnight D, Decker A, Tan WH, Tarnopolsky MA, Brady LI, Wolff M, Dondit L, Pedro HF, Parisotto SE, Jones KL, Patel AD, Franz DN, Vanzo R, Marco E, Ranells JD, Di Donato N, Dobyns WB, Laube B, Traynelis SF,* Lemke JR. GRIN2B encephalopathy: novel findings on phenotype, variant clustering, functional consequences and treatment aspects. J Med Genet. 2017 Jul;54(7):460–470. doi: 10.1136/jmedgenet-2016–104509.

Chapter 8.
Gamma band brain activity of an awake and sleeping monkey in response to flashes of light

Alexey Harauzov, Liubov Ivanova, Daria Podvigina

Pavlov Institute of Physiology RAS,
Saint-Petersburg, Russia

The modern understanding of the brain functioning is based on a network approach, when areas of the brain that are distant from each other are combined into so-called large-scale neural networks based on their synchronous activity and participation in the performance of certain tasks. Using tomographic methods, such associations in the brain were identified as the "central executive network" associated with decision-making, a neural network responsible for detecting and switching attention to significant signals, networks responsible for memory functions, language abilities, processing sensory information, and some others. A common feature of all large-scale neural networks is an internal functional connection and opponent interactions with the so-called Default Mode Network (Raichle et al., 2001; Raichle, 2006), which, apparently, coordinates their work by maintaining the balance between activated and deactivated areas of the brain.

However, despite the large amount of experimental data, there is still no unified understanding of the neurophysiological mechanisms of interaction of large-scale neural networks. Of particular interest are opponent interactions, when the activation of some areas of the brain is accompanied by a decrease in activity in other areas. Since most of the experimental data were obtained using the fMRI method, which estimates oxygen consumption by brain tissues and has a low temporal resolution, information about the nature and dynamics of opponent interactions is currently contradictory (Fracasso et al., 2021).

Recording neural activity directly in the structures of interest in the brain would help to clarify these issues. However, in humans, for obvious reasons, this is difficult to do. The optimal object for research on this topic is rhesus monkeys, in which large-scale neural networks similar to humans were previ-

ously found (Mantini et al., 2011). The aim of our experiments was to localize the brain regions involved in opponent interactions by fMRI, the subsequent implantation of microelectrodes in these areas, and the registration of local field potentials (LFP) reflecting the multiunit activity of neurons in the waking state and during sleep.

FMRI Localization of Brian Areas Demonstrating Opponent Interactions

Five experiments were carried out on three male rhesus monkeys (the experiment was repeated twice for two monkeys). It is known that for a successful fMRI study of the brain, the subject must be conscious, but at the same time lie still, in particular, with no head moving. In monkeys, this was achieved using specially developed fixing masks made of silicones of different densities, taking into account the individual characteristics of the animal's head shape. We made each mask on the basis of computed tomography (CT) data, pre-printing a head model on a 3D printer. The animal's body was fixed using a specially made suit and placed in a plastic case with non-magnetic holders on the straps.

Since forced immobilization can cause stress in the animal, which can affect the recorded signal, each monkey underwent 4–5 trainings simulating the experiment before fMRI studies. Macaques were injected with short-term anesthesia, packed in a fixing case and put on a mask. When the animal woke up, sounds were fed through speakers, simulating a working MRI scanner, and flashes of light were fed through light guides. Each workout lasted 2–2.5 hours, during which heart rate (HR) was recorded using an ECG as an indicator of stress levels. The training was carried out once a week to give the animal time to recover. After training, we analyzed the level of cortisol in animals' blood—the hormone, which is involved in the development of stress reactions. Taken together, heart rate and cortisol levels compared to controls from the literature suggest that monkeys are not exposed to severe stress during immobilization. However, by training them before the fMRI procedure, we eliminate the effects of novelty or fear that could affect the results of fMRI studies.

Changes in the level of blood oxygenation (BOLD signal) in response to stimulation by light flashes were recorded on a 1.5 Tesla tomograph. For the first 30 seconds, the monkey lay in complete darkness, then for the next 30 seconds it was stimulated with flickering light using light guides attached close to the animal's eyes. The light-dark cycle was repeated 13 times. The

duration of each flash was 10 ms, the time between flashes varied from 100 to 500 ms. Thus, a variable stimulation frequency was achieved.

Statistical maps of the distribution of activity in the monkey brain were assessed in a standard way by comparing the BOLD signal accumulated over all periods of darkness (resting phase) and over all periods with flickering light (activation phase). When averaging the results of each experiment, we used an additional noise reduction method—clustering, in which unidirectional changes in the BOLD signal in neighboring voxels were combined into a single cluster. In our experiments, the size of one voxel (an element of the volumetric image of the brain) was 2 * 2 * 2 mm. We took 80 voxels for the clustering threshold. We did not consider changes in the BOLD signal in smaller brain regions.

Figure 1 shows the average results of five fMRI experiments obtained on three monkeys. Areas of the brain that significantly increase oxygen consumption during light stimulation relative to the resting phase, when the animal was in complete darkness, are marked in red. The areas of the brain in which oxygen consumption, on the contrary, decreased in response to flashes of light are marked in blue (two-sided t-test, P <0.01).
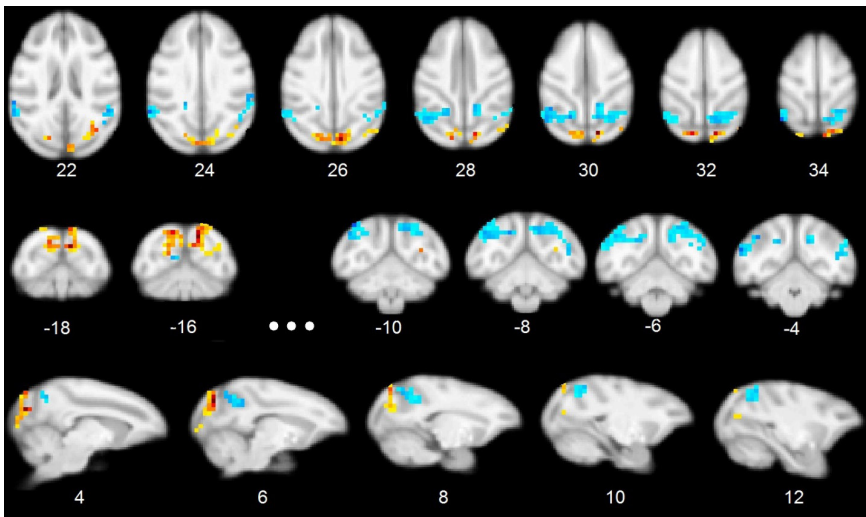
Fig. 1. Statistical maps of the distribution of activity during light stimulation, obtained on the basis of five fMRI experiments on three monkeys (P <0.01).

According to the electronic atlas of the rhesus monkey brain (Bakker et al., 2015) based on data from Paxinos et al. (2000), an increase in the BOLD signal in response to light flashes was observed in the primary visual cortex (area V1) with a maximum in the X, Y, Z coordinates: 2.7,—41.7, 25.9 mm. A decrease in oxygen consumption during light stimulation was observed in various regions of the parietal cortex. The maximum decrease in activity was recorded in the posterior part of the parietal cortex, in the region of the intraparietal sulcus, in the area 7 according to Brodmann. According to the electronic atlas of the brain (Bakker et al., 2015), this is the posterior part of the lateral intraparietal area (LIP) with partial capture of the dorsal parietal area with coordinates X, Y, Z: 5.55,—32.3, 31.5 mm. In the vicinity, deactivation was observed in the posterior parietal area (PPt) with partial capture of the medial temporal area (extrastriate area V5) with coordinates X, Y, Z: 9.5,—30.5, 34.5.

If the increase in oxygen consumption in the primary visual cortex in response to flashes of light was quite expected, then the deactivation in the parietal cortex requires an explanation. According to the literature (Podvigin, Makarov, Shelepin, 1986), neurons in area 7 of the parietal cortex, where we observed deactivation, can respond to visual stimulation, but most often with the active participation of an animal in an experiment, which requires coordination of visual, somatosensory, and motor functions. In particular, the lateral intraparietal region, where we observed the maximum decrease in oxygen consumption during light stimulation, plays a key role in the control of the oculomotor apparatus during visual search and localization of objects in the visual field (Christopoulos et al., 2018). In our case, the monkeys passively viewed the presented flashes, without any task. They could not carry out a targeted visual search, since the light stimulation was diffuse, covering almost the entire visual field, and the eye simply had nothing to fixate on.

Presumably, the multidirectional change in the BOLD signal observed in the occipital and parietal cortex in response to light flashes reflects the principle of maintaining the balance of neural activity in the brain. During light stimulation, the main load fell on the primary visual cortex. At the same time, the areas responsible for the localization of objects in space, the integration of sensory signals, etc., remain less involved. Accordingly, activity in these areas is reduced, which saves energy resources of the brain.

This may explain the decrease in activity in the extrastriate area V5, which is responsible for the perception of movement and complex visual patterns. With diffuse illumination of the entire visual field during stimulation, this

area remains less involved in the processing of the incoming signal. Similar data were obtained in human fMRI studies with stimulation of the entire visual field by light flashes, in which, along with the activation of the primary visual cortex, a decrease in activity in the extrastriate regions of the brain was observed (Maggioni et al., 2013, 2016). The authors did not raise the question of the usefulness of such an opponent relationship; they were more interested in the neurophysiological nature of the BOLD signal decrease, since there is a possibility that the redistribution of blood flow rather than a decrease in neuronal activity may be its reason. The answer to this question will be given in the next section, based on the analysis of the total electrical activity of neurons directly from the regions of the brain involved in opponent interactions.

An alternative explanation for the discovered opponent interactions between the regions of the occipital and parietal cortex can be made on the basis of the existence in the brain of primates of the so-called Default Mode Network (DMN). This functional association of the parietal, frontal, and some other regions of the brain, discovered 20 years ago (Raichle et al., 2001), is in an opponent relationship with all known large-scale neural networks. The activity of DMN is maximal when the subject is not doing anything, and decreases as the cognitive load increases. In monkeys, the main component of this large-scale neural network is located in the parietal region, adjacent to our localized regions in area 7, but closer to the medial surface of the cortex — in the PGM and MIP areas (Mantini et al., 2011). The registration of electrical activity in these structures of the brain in response not only to flashes of light, but also to auditory stimulation would help to answer the question of whether the regions of the parietal cortex we localized belong to the default system. Since the DMN is universal and does not depend on the modality of the incoming signal, visual and auditory stimulation would cause the same response — a decrease in neural activity.

Electrophysiological Studies on Monkeys

Two monkeys, Eduard and Yunt, were implanted with microelectrodes for long-term recording of local field potentials (LFP) in the regions of the occipital and parietal cortex localized by fMRI. Electrodes were implanted in Eduard's right hemisphere, and in the left hemisphere of Yunt. In both animals, the target in the occipital region was the primary visual cortex (V1); in the parietal region, electrodes were implanted into the Lateral Intraparietal area (LIP). A bundle of two or three electrodes glued together was implanted into each area, the active ends of the electrodes were located at a distance of 1.5 mm. Electrical activity was recorded bipolar, as the difference of potentials

between two adjacent active ends of the electrodes. Thus, a local recording of activity in a limited area of the brain was achieved. This is the so-called local field potential, the high-frequency oscillations of which reflect changes in the multiunit activity of neurons surrounding the recording electrode (Leszczyński et al., 2020).

The electrodes were tungsten rods in a Teflon insulating sheath, the ends of which were exposed in such a way as to achieve the required resistance. The thickness of Eduard's tungsten rods was 180 microns, the resistance (measured with a 1 kHz sinusoidal signal) was 3–6 kΩ. Yunt's electrodes were thinner (130 microns), and the size of the exposed area was smaller, respectively, the resistance was about 200 kOhm. The signal was digitized with a sampling rate of 2000 Hz and filtered in the range of 1–500 Hz.

Before the operation of implantation of electrodes, stereotaxic coordinates of penetrations were calculated for each animal, taking into account the thickness of the bones, the shape of the skull and the location of large blood vessels. For this, the coordinates of the targets were transferred from the space of anatomical MRI to the space of computer tomography with angiography data. On the basis of these data, the coordinates of the trepanning holes, the angles and depth of penetration of the electrodes were calculated. In order to save time on operations, all manipulations with stereotaxic equipment were practiced on a copy of the skull of each monkey, printed on a 3D printer.

After all the experiments, before the second operation on electrode removing, a control CT scan of the animal's head was performed to verify the position of the electrodes. The obtained data with the coordinates of the electrodes were transferred into the space of anatomical MRI and compared with the data of functional MRI. Control measurements showed that all the electrodes hit exactly the intended targets.

After the operation to implant the electrodes, the animals were placed in a primatological chair, limiting their movements so that the monkeys could not reach the stitches on the head. A few days after the operation, necessary for the animals to adapt to the new habitat, experiments began to record the electrical activity of the brain in response to flashes of light, similar to those used in the fMRI study. One experiment was carried out per day, lasting one and a half hours. In total for subsequent data averaging, 11 identical experiments were carried out on each monkey during wakefulness and 16 experiments were carried out on one animal (Yunt) during sleep.

During the experiment, the animals stayed in a dark room. Light stimula-

tion was carried out from a computer monitor, synchronized with a computer recording brain biopotentials for subsequent accumulation and averaging of evoked responses. The flicker parameters were similar to those used in the fMRI experiment, however, the duration of light stimulation was shorter — only 3 seconds, and the duration of the rest (darkness) phase varied from 6 to 8 seconds.

Changes in local field potentials caused by flickering were analyzed using wavelet decomposition of the signal in the EEGLAB program (Delorme, Makeig, 2004). This is the so-called time-frequency representation of data, which allows estimating the changes over time in the power of the oscillations at different frequencies caused by flashes of light. Oscillations were analyzed in the frequency range of 3–300 Hz. Figure 2 shows an example of such processing of data obtained on the monkey Yunt in the waking state (left) and in the sleep state (right). An increase in the power of oscillations is coded in red, a decrease in the power of oscillations relative to the background fragment preceding stimulation is coded in blue. Only those changes are shown that exceeded the statistical threshold at $P < 0.01$.

In both monkeys in the waking state, the onset of light stimulation caused a phasic increase in activity in the occipital cortex in a wide frequency range. In Yunt, an increase was observed in the range of 3–150 Hz, in Edward, in the range of 3–100 Hz. Oscillations at higher frequencies (up to 300 Hz), on the contrary, showed predominantly a decrease in power in response to the presentation of flashes (example in Figure 2, top left). Termination of light stimulation caused a phasic change in the oscillation power in a wide frequency range in both animals.

In the parietal region of monkeys in a state of wakefulness, in response to light stimulation, a tonic response developed in the form of an increase in the power of oscillations in the frequency range of 3–20 Hz for Yunt and 3–30 Hz for Edward. In the higher frequency range, up to 150 Hz, a significant decrease in the oscillation power was observed in both monkeys (example in Figure 2, bottom left).

At this stage of the study, we were interested in changes in the oscillation power in the gamma range, since high-frequency LFP oscillations indirectly reflect changes in multicellular neural activity around the electrode (Ray et al., 2008; Leszczyński et al., 2020). On the figure 2, under each raster images, the time variation of the average oscillation power in the frequency range of 50–100 Hz is shown. Both animals in the waking state demonstrated a phasic increase of activity in the V1 area of the visual cortex during the first 100 ms

after the start of stimulation. Then the activity decreased by about two times, but continued to exceed the background level until the end of flickering. In the parietal region, an initial slight increase in activity was also recorded, which coincided in time with that in the V1 zone. However, then the power of gamma oscillations sharply decreased, reaching a minimum at 300 ms after the start of stimulation and remaining significantly below the background level for the rest of the time of presentation of flashes.
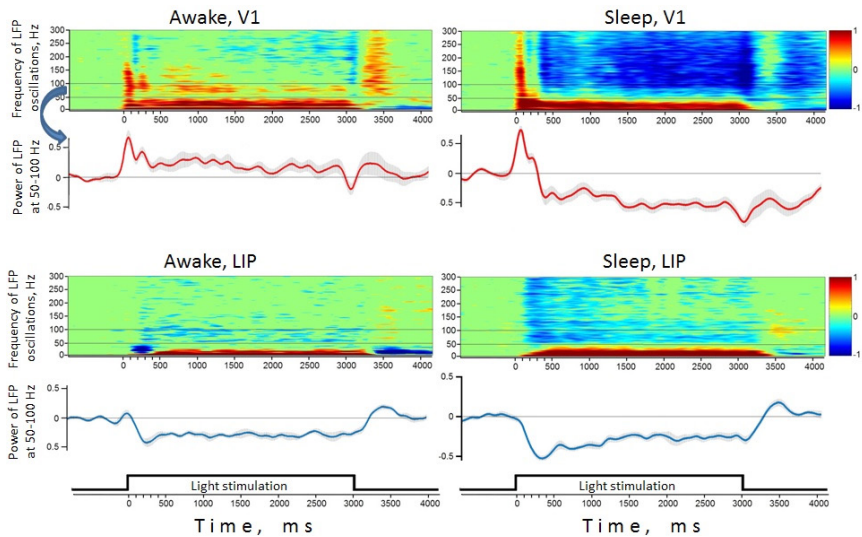
Fig. 2. Changes in the power of LFP oscillations in monkey Yunt in the occipital cortex (area V1) and in the parietal cortex (area LIP) in response to aperiodic flashes of light. Left—the state of wakefulness (11 experiments). Right—sleep state (16 experiments).

In a state of sleep or drowsiness, the reaction to flashes was recorded only in one monkey—in Yunt. The right side of Figure 2 shows the results of averaging of 16 additional experiments on monkey Yunt, in a state of sleep during stimulation. It can be seen that in a sleeping animal, as in the waking state, the flashes first cause a short-term increase in the power of gamma oscillations in the visual cortex. However, then a sharp decrease in power is observed, significantly below the background level preceding stimulation. At the same time, the activity in the parietal region of the brain did not significantly depend on the state of the monkey—both in sleep and in the waking state, a

decrease in the power of gamma oscillations was recorded.

The study of the time dependence of the total power of the oscillations in the range of 50–100 Hz showed that the change in the polarity of the response in the occipital cortex of a sleeping animal occurs approximately 300 ms after the start of stimulation. In the parietal cortex of the sleeping animal, the negative response was more powerful and developed approximately 100 ms earlier than in the waking animal (Figure 2, line graphs).

The level of wakefulness of the animal was assessed using an infrared camera and by the power of the delta rhythm in the ECoG recorded from the surface of the parietal cortex. In monkeys, as in humans, the power of oscillations in the delta range increases during sleep, which was confirmed by our analysis of the video-ECoG (Figure 3).
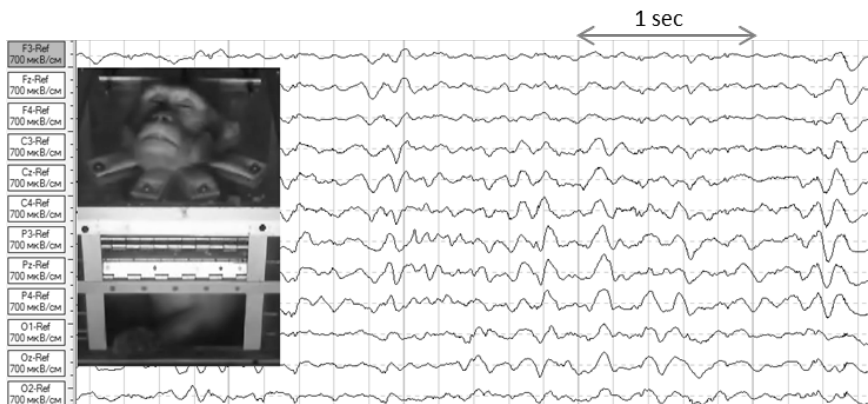
Fig. 3. An example of a video-ECoG of monkey Yunt in a sleepy or drowsy state with pronounced oscillations in the delta range.

Comparison of the power of the delta rhythm in ECoG, recorded from the surface of the parietal cortex, with the total power of LFP gamma oscillations in the visual cortex during stimulation in each of the 27 experiments in monkey Yunt (both in sleep and in the waking state) is shown in Figure 4. The total power of gamma oscillations was calculated in the frequency range 50–100 Hz and in the time interval 160–3000 ms (that is, only the tonic response). Both of these parameters showed a high negative correlation (K = –0.73, P << 0.001, two-tailed Pearson test). That is, the more the animal is immersed in sleep (an increase in delta oscillations in the parietal cortex), the lower the

power of gamma oscillations in the visual cortex.
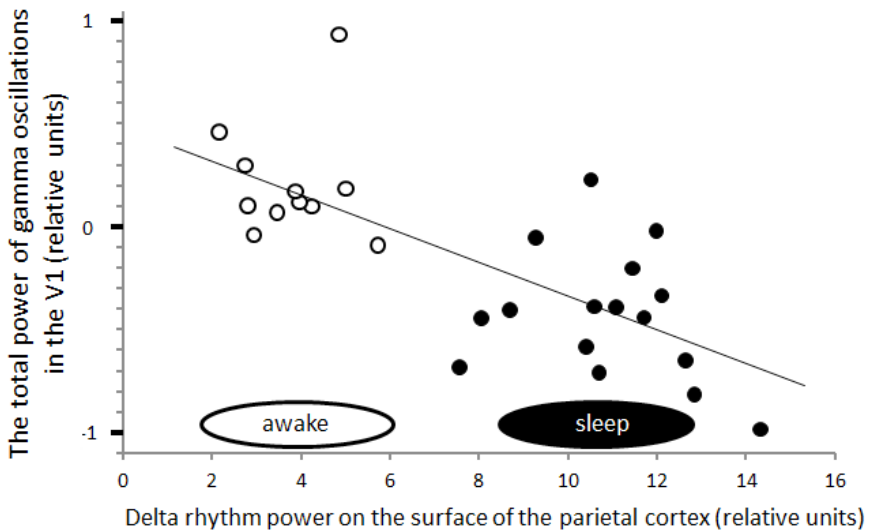
Fig. 4. Comparison of the total power of gamma oscillations in the primary visual
    cortex and the power of the delta rhythm as an indicator of the level of sleep.

On the one hand, this dependence is easy to explain — the longer the animal sleeps or naps during the experiment, the more time its eyes are closed (or partly-closed). Accordingly, the light flux to the retina decreases, which leads to a weakening of the response in the V1 zone. However, in responses of a deeply sleeping animal to light stimulation, we observed a paradoxical decrease in the power of gamma oscillations in the primary visual cortex (Figure 2, right). This is a result of active inhibition of background neural activity during flashes of light. Presumably, this inhibition is a part of the protective mechanism of the brain, which ensures restful sleep in humans and animals, without which we would wake up from any rustle.

Conclusion

As a result of the performed tomographic experiments on monkeys, during aperiodic stimulation with light flashes, stable reciprocal interactions were found between the primary visual cortex and the regions of the parietal cortex responsible for visual search and localization of objects in space. An increase in oxygen consumption in the primary visual cortex during light stimulation

was accompanied by a decrease in oxygen consumption in the parietal regions. Analysis of the local field potentials recorded in the indicated structures of the brain confirmed that a decrease in the level of oxygen consumption is a consequence of a decrease in multiunit neuronal activity.

Thus, we obtained direct electrophysiological confirmation of fMRI data on the opponent interactions between the primary visual cortex and the region of parietal cortex responsible for visual search and localization of objects in the visual field. These interactions are based on excitatory and inhibitory processes, as a result of which we observed changes in the power of LFP oscillations at frequencies of 50–100 Hz. In awake animals, both areas initially respond to flashes with a phasic increase in activity, which may reflect the generalized development of the orientation reflex, which encompasses many brain structures. Then tonic responses develop, coinciding in time with the duration of the stimulus. In the V1 zone, this is a stable increase of activity relative to the background level, associated with the primary processing of the sensory signal. In the parietal region, which is not occupied during the processing of this type of visual stimuli, on the contrary, a tonic decrease in activity is observed.

In the waking state, multidirectional changes in neural activity in the occipital and parietal regions of the brain — activation and deactivation — begin almost simultaneously, about 300 ms after the presentation of flashes, and last throughout the entire period of stimulation. Given the high energy consumption of the brain, it can be assumed that reciprocal interactions observed in the occipital and parietal cortex in response to flashes of light reflect the general principle of maintaining the balance of neural activity. During uniform illumination of the entire visual field, the main load falls on the primary visual cortex. At the same time, areas that carry out more complex processing of visual signals are less in demand, and a decrease in their activity would save the energy resources of the brain.

The visual system of a sleeping animal reacts to flickering in another way — tonic inhibition of activity develops in the area V1. It is noteworthy that during sleep, the primary visual cortex is activated in the first milliseconds of stimulation, and the parietal region of the brain reacts to flashes, as in the waking state, reducing its activity throughout the duration of stimulation. This suggests that in a sleeping animal, information about flickering of light is received and processed by the brain at various levels, however, inhibition of activity in the primary visual cortex leads to ignoring the stimulus, bypassing consciousness. We explain the obtained effect by the work of the protective mechanisms of the brain, which protect sleep from minor stimuli.

References

1. *Christopoulos V. N.*, K*agan I., Andersen R. A.* Lateral intraparietal area (LIP) is largely effector-specific in free-choice decisions // Scientific Reports. 2018. 8.

2. *Maggioni E., Molteni E., Arrigoni F., Zucca C., Reni G., Triulzi F. M.*, *Bianchi A. M.* Coupling of fMRI and NIRS measurements in the study of negative BOLD response to intermittent photic stimulation // 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013, pp. 1378–1381. doi: 10.1109/EMBC.2013.6609766.

3. *Maggioni E., Zucca C., Reni G., Cerutti S., Triulzi F. M., Bianchi A. M., Arrigoni F.* Investigation of the electrophysiological correlates of negative BOLD response during intermittent photic stimulation: An EEG-fMRI study // Human brain mapping. 2016. 37(6): 2247–2262. https://doi.org/10.1002/hbm.23170.

4. *Подвигин Н. Ф. Макаров Ф. Н. Шелепин Ю. Е.* Элементы структурно-функциональной организации зрительно-глазодвигательной системы. Наука, Ленинград. 1986. 252 стр.

5. *Raichle M. E.*, *MacLeod A. M., Snyder A. Z.*, *Powers W. J.*, *Gusnard D. A.*, *Shulman G. L.*) A default mode of brain function // Proc. Natl. Acad. Sci. U.S.A. 2001. 98: 676–682; DOI: 10.1073/pnas.98.2.676.

6. *Raichle M. E.* Neuroscience. The Brain's Dark Energy // Science. 2006. 314: 1249–1250. http://dx.doi.org/10.1126/science. 1134405.

7. *Mantini D., Gerits A., Nelissen K., Durand J. B., Joly O., Simone L., Sawamura H., Wardak C., Orban G. A., Buckner R. L., Vanduffel W.* Default Mode of Brain Function in Monkeys. The Journal of Neuroscience. 2011. 31(36):12954–12962.

8. *Leszczyński M., Barczak A., Kajikawa Y., Ulbert I., Falchier A. Y., Tal I., Haegens S., Melloni L., Knight R. T.*, *Schroeder C. E.* Dissociation of broadband high-frequency activity and neuronal firing in the neocortex // Science advances. 2020. 6(33): eabb0977. https://doi.org/10.1126/sciadv.abb0977.

9. *Delorme A., Makeig S.* EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis // Journal of neuroscience methods. 2004. 134(1): 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

10. *Ray S., Crone N. E., Niebur E., Franaszczuk P. J., Hsiao, S. S.* Neural correlates of high-gamma oscillations (60–200 Hz) in macaque local field potentials and their potential implications in electrocorticography // The Journal of neuroscience. 2008. 28(45): 11526–11536. https://doi.org/10.1523/JNEUROSCI.2848–08.2008.

11. *Bakker R., Tiesinga P., Kötter R.* The Scalable Brain Atlas: instant web-based access to public brain atlases and related content // Neuroinformatics. 2015. 13(3):353–366. doi: 10.1007/s12021–014–9258-x.

12. *The Paxinos et al.* (2000) template contains a subset (cortex, amygdala, thalamus, striatum) of the regions in: G. Paxinos, X. F. Huang, and A. W. Toga "The Rhesus Monkey Brain in Stereotaxic Coordinates" (2000)Academic Press. = Paxinos G., *Huang X. F.*, *Toga A. W.* The Rhesus Monkey Brain in Stereotaxic Coordinates. 2000. Academic Press.

13. *Fracasso A., Gaglianese A., Vansteensel M. J., Aarnoutse E. J., Ramsey N. F., Dumoulin S. O., Petridou N.* FMRI and intra-cranial electrocorticography recordings in the same human subjects reveals negative BOLD signal coupled with silenced neuronal activity // Brain structure & function. 2021, Aug 7. doi: 10.1007/s00429–021–02342–4. Online ahead of print.

Chapter 9.
Four-channel polarization system for hematocrit level monitoring

Ruslan D. Khlynov, Victoria A. Ryzhova,
Valery V. Korotaev

ITMO University
St. Petersburg, Russia

The paper considers the development of an optoelectronic polarization system designed for non-contact measurement and monitoring of hematocrit level. The structural scheme of a four-channel polarization system for hematocrit level monitoring in the medical field is proposed. The principle of operation of the system is based on simultaneous measurement of Stokes vector parameters of polarized radiation scattered at an angle by a biological sample and determination of hematocrit level that depends on blood anisotropy parameters.

The field of medical diagnostics is one of the most promising areas of humanity. In recent years, truly revolutionary discoveries have been made that have changed the world [1]. Cancer of lung, heart, brain, etc., as one of the deadliest diseases of mankind, is now no longer invincible, thanks to the development and implementation of many surgical, X-ray endovascular, therapeutic and other methods based on optical, polarizing and optoelectronic devices and systems, which are now successfully used in various medical departments to detect and treat diseases at an early stage [2]. However, there are still areas of medicine that still use traditional methods and approaches that are difficult to use in the modern XXI century due to the emergence of new diseases that cannot be diagnosed using outdated technologies and devices. One such area is invasive blood testing for the diagnosis of various diseases. The main advantage – highly accurate determination of rheological properties of blood (error 1 % [3]) is no longer able to cover all the shortcomings of this method, such as prolonged acquisition of results, uninformativeness, limited application, the possibility of blood contamination and others. In this connection, there is a need for a new approach in determining various diseases using noninvasive medical diagnosis of blood rheological properties.

According to the World Health Organization (WHO) statistics for 2019-2020, pathological changes in blood rheological properties remain the most dangerous for mankind, the mortality from which in the world is about 50 % of all deaths, of which about 85 % account for heart attacks and strokes. The whole world medical community decides to reduce the mortality rate from cardiovascular diseases. This requires the development of new modern technologies and methods that will allow determining the rheological properties of blood in a noninvasive way and will be the basis of optoelectronic devices.

Status of the issue

One of the significant indicators of blood rheological properties is hematocrit level. It gives an idea of the ratio between erythrocytes and plasma in the total blood volume. Hematocrit level monitoring allows not only to assess the functional state of the patient's blood and hematopoietic systems, but also to detect infectious processes, as well as to detect timely somatic abnormalities indicating the presence of cardiovascular diseases in a patient associated with arterial and venous pathologies [4]. A decreased hematocrit level (15 % – 25 %) in an individual indicates the presence of acute bleeding, anemia, hemodilution, hyperhydration, etc., high hematocrit level (50 % – 55 %) in human indicates the presence of erythraemia, erythrocytosis, dehydration, etc. [5]. Determination of hematocrit level at an early stage will allow timely correction of blood hyperviscosity to prevent the development of severe disease complications and detect thrombotic neoplasms leading to fatal outcomes (heart attacks and strokes).

The rapid pace of development of optoelectronic devices and systems opens a new promising direction in the research and development of automated video information devices and systems for non-contact measurement and monitoring of hematocrit level, particularly active polarimetry methods based on registration of changes in anisotropic blood state. They allow to carry out the analysis of several parameters of blood state, characterizing its form elements, as well as significantly expand the possibilities of noninvasive glucometry [6].

Implementation of a four-channel polarization system

The structural scheme of a noninvasive system for hematocrit level monitoring is based on the principle of visual Stokes polarimeter with separation of the intensity of radiation falling on the device input by amplitude without rotating elements [7,8] and is presented in the figure 1.
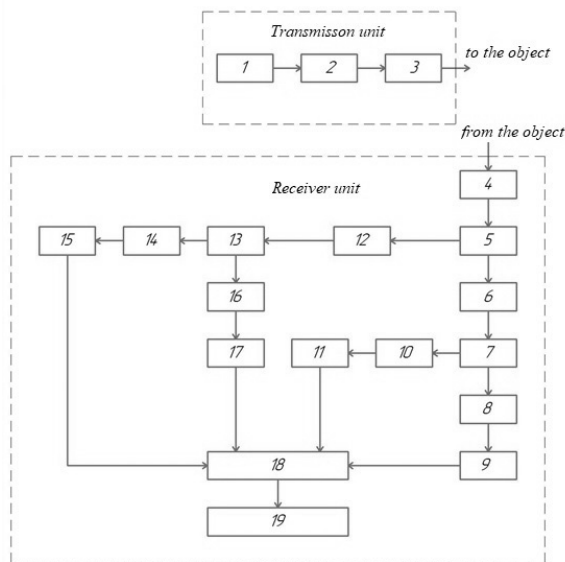
Fig.1. Structural scheme of the Four-Channel System.

The transmitting unit contains a source of optical radiation 1 and optical elements (polarizer 2; phase plate 3) for formation of a given state of polarization of radiation probing the biotissue. Receiving unit contains optical elements (lens 4; beam splitting cubes 5, 7, 13; phase plate 6, 12; quartz plate 8, 10, 14, 16) to obtain four intensity distributions for each matrix receiver of optical radiation (9, 11, 15, 17) for their further pixel-by-pixel processing in the electronic signal processing unit 18. For visualization and receipt of numerical results, information display device 19 is used.

Polarizing blood probing through the little finger nail is implemented in the polarizing system that allows to identify the states of the body with different hematocrit level based on the algorithm of combining the visualized images processing results and comparative analysis of polarization properties of biotissues.

Thus, the hypothesis of the possibility of detecting significant differences between the states of a biological sample with different hematocrit levels using the polarization method was put forward. In order to confirm it, physical modeling of the system was performed and preliminary experimental research was carried out.

Experimental research

The aim of the research was to measure changes in the polarization state parameters of radiation scattered at an angle to the biobject under research, compared to the parameters of the original, probing object radiation, based on the analysis of spatial distributions of Stokes vector parameters. It was necessary to identify numerical differences in these distributions in the physical simulation of the states of biobjects with different hematocrit levels. In the course of the experiment, the files $,I_{0^\circ}, I_{90^\circ}, I_{45^\circ}, I_{-45^\circ}, I_r, I_l$ respectively, each of which contains information about the distribution of the radiation intensity in the image plane, which is used in the processing of measurements to calculate the Stokes parameters. The calculation of the four Stokes vector parameters for angularly scattered radiation was performed according to the formula [9,10]:

$$S = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} I_{0^\circ} + I_{90^\circ} \\ I_{0^\circ} - I_{90^\circ} \\ I_{45^\circ} - I_{-45^\circ} \\ I_r - I_l \end{bmatrix}, \tag{1}$$

where $I_{0^\circ}, I_{90^\circ}, I_{45^\circ}, I_{-45^\circ}, I_r, I_l$ – are the intensities of waves passed through polarizers transmitting linearly-polarized components of radiation with azimuths, respectively, 0°, 90°, +45°, –45°, and right- and left- hand circularly-polarized components of radiation.

The obtained values of Stokes vector parameters made it possible to perform a complex polarization analysis of the ray beam exiting the system, as well as the residual degree of polarization of the input radiation. It was determined by the formula [9,10]:

$$P = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \tag{2}$$

The azimuth and ellipticity of the output radiation were calculated according to the formulas [9,10], respectively:

$$\alpha = \frac{1}{2} arctg\left(\frac{S_2}{S_1}\right); \tag{3}$$

$$e = tg\left(\frac{1}{2} arcsin\left[\frac{S_3}{\sqrt{S_1^2 + S_2^2 + S_3^2}}\right]\right). \tag{4}$$

As a result of the experiment, using MATLAB software, spatial distributions of Stokes vector parameters were calculated, as well as the degree of polarization of the output radiation, which characterizes the ratio between polarized and unpolarized components of radiation for three states of the body, guaranteed different hematocrit level: initial, decreased and stabilized. Distributions of polarization state parameters (azimuth and ellipticity) were obtained for the polarized component of the output radiation. The obtained numerical values are presented in the table 1.

Table 1

Numerical values of the experimental results

Parameter / Hematocrit	S_0	S_1	S_2	S_3	P	α, *rad*	e
Initial	276	−12	35	−5	0,13	−0,62	−0,06
Decreased	230	−40	−19	−3	0,19	0,22	−0,04
Stabilized	272	59	−28	−4	0,24	−0,22	−0,03

Table 1 gives information about changes in parameters of the total Stokes vector and polarization state parameters (degree of polarization, azimuth and ellipticity of the polarization ellipse) during the experimental research. As can be seen, during artificial decrease of hematocrit level and during its stabilization the parameters change their values dramatically, especially the degree of polarization and azimuth of elliptically polarized laser radiation scattered from a biological sample. By determining the four parameters of the Stokes vector using a four-channel polarization system, we can calculate the degree of polarization, azimuth and ellipticity, by which the hematocrit level is determined by comparing the results with empirical calibration characteristics.

Thus, significant differences in the polarization characteristics of the radiation scattered by the object for different states of the body were revealed, which confirms the hypothesis about the possibility of using the polarization method to monitor the hematocrit level.

Conclusion

A four-channel polarization Stokes polarimeter system for real-time hematocrit level monitoring has been developed. The system contains non-moving elements and allows generating simultaneously four linearly independent Stokes vector parameters of the scattered light. The advantages of this scheme are the following: the optical path lengths through the light splitting blocks to matrix photodetectors are identical; each of the four images has the same magnification, distortion and frequency-contrast response throughout the matrix field of view, which enables accurate image registration without distortion. A physical model of polarization sensing of blood through the nail structure of the little finger has been implemented, allowing the identification of body states with different hematocrit levels based on visualization and comparative analysis of polarization properties of biotissues.

Acknowledgements

References

1. *Fernandez-Perez A., Gutierrez-Saiz O., Fernandez-Luna J.L., Moreno F., Saiz J. M.* Polarimetric Detection of Chemotherapy-Induced Cancer Cell Death // Applied Sciences. 2019, 9(14): 2886–2902. https://doi.org/10.3390/app9142886.
2. *Zavyalova K., Gorst A., Mironchev A.* Non-Invasive Determination of Blood Glucose Concentration Using a Near-Field Sensor // Proceedings. 2020, 60(1): 8. https://doi.org/10.3390/ IECB2020–07022.
3. *Wennecke G.* Hematocrit — a review of different analytical methods // Radiometer Medical ApS, 2700 Bronshoj, Denmark, 2004.
4. *Jalal U. M., Kim S. C., Shim J. S.* Histogram analysis for smartphone-based rapid hematocrit determination // Biomedical Optics Express. 2017, 8(7): 3317–3328. https://doi.org/10.1364/BOE.8.003317
5. *Secomski W., Nowicki A., Guidi F., Tortoli P., Lewin P. A.* Non-invasive measurement of blood hematocrit in artery // Bull. Poc. Ac. Tech. 2005, 53(3): 245–250. https://hdl.handle.net/2158/338990.
6. *Yamakoshi Y., Matsumura K., Yamakoshi T., Lee J., Rolfe P., Kato Y., Shimizu K., Yamakoshi K.* Side-scattered finger-photoplethysmography: experimental investigation toward practical noninvasive measurement of blood glucose // J. Biomedical Optics. 2017. 22(6). https://doi.org/10.1117/1.JBO.22.6.067001.

7. *Howard S., Larry J. P., Chenault D., Roche M., Reinhardt J., Joseph P. P.* Four camera complete Stokes imaging polarimeter // Proc. of SPIE — The international society for optical engineering, 2008, vol. 6972, 69720J. https://doi.org/10.1117/12.784797.

8. *Howard S., Larry J. P., David Chenault D., Roche M., Reinhardt J.* Wave slope measurement using imaging polarimetry // Proc. of SPIE — The international society for optical engineering, 2009, vol. 7317, 73170B. https://doi.org/10.1117/12.819031.

9. *Azzam R. M.A., De A.* Optimal beam splitters for the division-of-amplitude photopolarimeter // Journal of the Optical Society of America. 2003, 20(5): 955–958. https://doi.org/10.1364/JOSAA.20.000955.

10. *Gandorfer A. M., Steiner P., Povel H. P., Aebersold F., Egger U., Feller A., Gisler D., Hagenbuch S., Stenflo J. O.* Solar polarimetry in the near UV with the Zurich Imaging Polarimeter ZIMPOL II // Astronomy and Astrophysics.2004, 422 (2): 703–708. https://doi.org/10.1051/0004–6361:20040254

Chapter 10.
Emotional colouring of perceptions: towards experimental verification of quantum-like model

Andrei Khrennikov

Linnaeus University, International Center for Mathematical Modeling in
Physics and Cognitive Sciences
Växjö, SE-351 95, Sweden

In the quantum-like modelling (see, e.g., review [1]) we do not consider quantum physical processes in the brain. The quantum formalism is used as the most general formalism describing measurement processes. Quantum-like models explore this formalism outside of physics, in cognition psychology, decision making, social and political sciences [2-11].

Quantum measurement theory: observer-system separation

Quantum mechanics has many interpretations. By the Copenhagen interpretation, it describes and predicts outcomes of measurements performed by observers on systems. Separation "system-observer" places the crucial role in quantum methodology. The big foundational problem is establishing the boundary between a system and an observer.

By applying quantum formalism for modelling cognition, we also have to consider system-observer separation in information processing by the brain. How can this be done? Straightforwardly it seems that the brain performs *self-observations*. My suggestion is to couple system-observer separation with unconscious vs. conscious information processing in the brain.

Unconscious and conscious information processing in the brain

As is well known, an essential part of information processing in the brain is performed unconsciously. The information system for such processing (call it unconsciousness) is denoted by the symbol **UC**. The space of its states (unconscious mental states) is denoted by the symbol $\mathbf{H} \equiv \mathbf{H}_{UC}$. In the quantum-like model, this is a complex Hilbert space, states are represented by vectors normalized by 1: $<\psi|\psi>=1$.

The idea about unconscious information processing in the brain and more generally the nervous system was presented in works of leading psychologists, James, Freud, Jung, ... However, we do not try to go deeply into psychological issues. We proceed operationally: **UC** denotes a special information processor of the brain. It performs *pre-observational processing of the mental state*.

In our model [2] of brain's functioning consciousness plays the role of observer, performing observation on **UC** -functioning. The brain contains the information processing system (separated from **UC)** and generating conscious experiences; denote this system by the symbol **C**.

Once again it should be noted that we proceed operationally without going deeply into neurophysiological, psychological, and philosophic issues related to the complex notion of consciousness. We want "simply" to model its functioning (in interrelation with unconsciousness). This is a part of the mathematical modelling paradigm which was advertised in numerous works of the author: formal mathematical modelling, instead of "understanding" of phenomena. Functioning of **C** is represented as measurements on the system **UC**.

Introduction of two information processing systems **UC** and **C** matches the quantum measurement scheme, **UC** is the analogue of a physical system exposed to measurements and **C** is the analogue of a complex of measurement apparatuses.

Basic theories of consciousness

Now we recall two basic theories of consciousness [12-15]:

- First Order Theory of Consciousness.
- Higher Order Theory of Consciousness.

"First-order theorists ... argue that processing related to a stimulus is all that is needed for there to be phenomenal consciousness of that stimulus." (see [12]).

"In contrast, ... higher-order theorists argue that a first-order state resulting from stimulus-processing alone is not enough to make possible the conscious experience of a stimulus. ... consciousness exists by virtue of the relation between the first- and higher-order states." (see [12])

The Higher Order Theory distinguishes between unconscious and conscious processing of mental information; what makes cognition conscious is a higher-order observation of the first-order processing.

This is the good place citation of Bohr [16]:

"This crucial point ... implies the impossibility of any sharp separation between the behaviour of atomic objects and the interaction with the measuring instruments which serve to define the conditions under which the phenomena appear."

This viewpoint matches with the Higher Order Theory of Consciousness. A conscious experience is not simply introspection of the **UC** -state, but the result of complex interaction between two information processors **UC** and **C**.

Perceptions and emotions

We follow to von Helmholtz theory of sensation-perception [17]. Perceptions are not simply the copies of sensations, not "impressions like the imprint of a key on wax", but the results of complex signal processing including unconscious cognitive processing and conscious observation of unconscious states [2] (see also [10] and references therein).

Theory of emotions connects them with contexts representation [16]:

"Emotion schema are learned in childhood and used to categorize situations as one goes through life. As one becomes more emotionally experienced, the states become more differentiated: fright comes to be distinguished from startle, panic, dread, and anxiety."

In our terminology, each emotion-generation scheme is crystallized on of the basic life-contexts. Context-labelling is the basic function of emotions.

Contextualization of surrounding environment was one of the first cognitive tasks of biosystems and this ability was developed in parallel with establishing of sensation-perception system [18]:

"Emotions represent adaptive reactions to environmental challenges; they are a result of human evolution; they provided optimal (from the viewpoint of computational resources) solutions to ancient and recurring problems that faced our ancestors."

We point out that *perceptions and emotions are commonly treated as conscious entities*. (See [12-17]).

Unconscious and conscious counterparts of the processes of generation of perceptions and emotions

We shall be concentrated on joint functioning of two information processors transforming [2,10]:

- sensations → perceptions,
- contexts → emotions.

Both processors have conscious outputs. Their functioning is strongly correlated; in the formalism quantum theory correlations are represented by entangled states.

We denote unconscious counterparts of these processors by the symbols.

UC_{per} and UC_{em}, respectively. In modelling of the emotional colouring of perception (its contextualization), we shall consider the compound information system $(\text{UC}_{per}, \text{UC}_{em})$ which modelled as a compound quantum(-like) system.

Compatible and incompatible observables

Observables C_1 and C_2 are called compatible if they can be jointly measurable and the joint probability distribution (JPD) $p_\psi(C_1 = x_1, C_2 = x_2)$ for each state ψ is well defined; observables which cannot be jointly measured and, hence, their JPD cannot be defined are called incompatible.

In the mathematical formalism, compatibility and incompatibility are formalized through commutativity and noncommutativity, respectively. If observables are described as Hermitian operators C_1, C_2, compatibility is encoded as $[C_1, C_2] = 0$. Incompatibility is encoded as $[C_1, C_2] \neq 0$. For compatible observables, JPD is given by the Born's rule:

$$p_\psi(C_1 = x_1, C_2 = x_2) = |\langle E_{x_1}^{C_1} E_{x_2}^{C_2} \psi, \psi \rangle|^2 = |\langle E_{x_2}^{C_2} E_{x_1}^{C_1} \psi, \psi \rangle|^2.$$

$$(1)$$

We stress that the space of perception-observables O_{per} can contain incompatible perceptions as well as O_{em} can contain incompatible emotions.

In the mental framework, incompatibility can be interpreted very naturally: there exist say emotions which can be experienced simultaneously; say happiness and sadness, pride, and shame. There exists incompatible, i.e., jointly unobservable perceptions and other conscious experiences.

The necessity to operate with various incompatible entities is the main root of the successful use of the quantum(-like) information representation.

In the absence of incompatibility, i.e., if, for the same mental state, the brain was able to construct the consistent probabilistic representation (in the form of JPD) of all possible combinations of say emotions, the quantum state formalism would be unnecessary.

Contextuality

In quantum measurement theory, selection of observables co-measurable with *A* is considered as specification of measurement context of *A*-measurements; the *A*-value in the *B*-context can differ from the *A*-value in the *C*-context, for the same pre-measurement state ψ. This is the essence of contextuality playing so important role in quantum information theory [19]:

Definition 1. (Contextuality) *If A,B,C are three quantum observables, such that A is compatible with B and C, a measurement of A might give different result depending upon whether A is measured with B or with C.*

We note that contextual behaviour corresponds to the case of incompatible quantum observables *B* and *C*, i.e., [B, C] ≠ 0. If all observables pairwise commute, it is possible to construct the non-contextual model of measurement based on the joint probability distribution for triple outcomes

$$p_{ABC}(x_k, b_m, c_n|\psi) = \|E_k^A E_m^B E_n^C \psi\|^2 = ... = \|E_k^A E_n^C E_m^B E_n^C \psi\|^2,$$
(2)

where E with corresponding indexes denote spectral families of operators A, B, C. If *B* and *C* are incompatible, such a model is impossible. This is the contextuality scenario.

Bell-type inequalities and experimental testing of emotional colouring of perceptions

In quantum physics, experimental testing of the Bell type inequalities is the hot topic. In psychology and decision making, they have been tested by a few authors. This paper can stimulate such experimenting in consciousness studies with joint measurements of the pairs (*A,B*) = (perception, emotion) or (decision making, emotion). As in physics and the previous psychological experiments, it is natural to test the CHSH inequality.

Consider two incompatible perceptions A1 and A2 and two incompatible emotions B1 and 2', and form cyclically their correlations. The CHSH correlation function is given by the following combination of correlations:

$$C_{\text{CHSH}} = <A1B1> - <A1B2> +<A2\ B1>+ <A2B2>$$
(3)

and, for dichotomous observables yielding values ±1 and in the absence of contextuality the following inequality holds:

$$|C_{\text{CHSH}}| \le 2.$$
(4)

Violation of this inequality gives the measure of contextuality. In quantum physics, the maximal value of C_{CHSH} is 2 2, *Tsirelson bound*.

In psychology and decision making some experiments have already been done [20-23], but emotional contextualization of perceptions or decisions has not yet been studied. So, this note proposes a new series of exciting experiments which can involve psychologists and neurophysiologists.

References:

1. *Basieva, I., Khrennikov, A., & Ozawa,* M. (2021). Quantum-like modeling in biology with open quantum systems and instruments. Biosystems, 201, 104328.
2. *Khrennikov, A.* (2004). Information dynamics in cognitive, psychological, social, and anomalous phenomena, Ser.: Fundamental Theories of Physics, Kluwer, Dordreht.
3. *Khrennikov, A.* (2010). Ubiquitous quantum structure: from psychology to finances}. Berlin-Heidelberg-New York: Springer.
4. *Busemeyer, J. and Bruza*, P. (2012). Quantum models of cognition and decision. Cambridge: Cambridge Univ. Press.
5. *Haven, E. and Khrennikov, A.* (2013). Quantum social science; Cambridge University Press.
6. *Asano, M., Khrennikov, A., Ohya, M., Tanaka, Y., Yamato, I.* (2015). Quantum adaptivity in biology: from genetics to cognition; Springer, Heidelberg-Berlin-New York.
7. *Haven, E., Khrennikov, A. and Robinson, T. R.* (2017). Quantum Methods in Social Science: A First Course; WSP: Singapore.
8. *Bagarello, F.* (2019). Quantum concepts in the social, ecological and biological sciences. Cambridge: Cambridge Univ. Press.
9. *Khrennikov, A.* (2020). Social laser. Jenny Stanford Publ., Singapore.
10. *Khrennikov A.* Quantum-like model for unconscious-conscious interaction and emotional coloring of perceptions and other conscious experiences. Biosystems. 2021; 208:104471; doi: 10.1016/j.biosystems.2021.104471.
11. *Ozawa, M., & Khrennikov, A.* (2021). Modeling combination of question order effect, response replicability effect, and QQ-equality with quantum instruments. J. Math. Psychology, 100, 102491.
12. *LeDoux, J. E. and Brown, R.* (2017). A higher-order theory of emotional consciousness. PNAS, 114 (10) E2016-E2025.
13. *LeDoux, J. E.* (2008). Emotional colouration of consciousness: how feelings come about. In: Frontiers of Consciousness. Chichele Lectures. L. Weiskrantz and M. Davies (eds). Oxford Univ. Press, Oxford.
14. *LeDoux, J.E.* (1984). Cognition and emotion: processing functions and brain systems. In Gazzaniga, M.S. (ed.) Handbook of Cognitive Neuroscience, pp. 357–368. New York: Plenum.

15. *LeDoux, J.E.* (1987). Emotion. In Plum, F. (ed.) Handbook of Physiology. 1: The Nervous System. Vol. V, Higher Functions of the Brain, pp. 419–460. Bethesda, MD: American Physiological Society.

16. *Bohr, N.* (1987). The Philosophical Writings of Niels Bohr. Ox Bow Press: Woodbridge, UK.

17. *von Helmholtz, H.* (1866). Treatise on Physiological Optics. Transl. by Optical Society of America in English. New York, NY: Optical Society of America.

18. *Ekman P.* (1992). An argument for basic emotions. Cognition and Emotion. 6 (3): 169–200.

19. *Khrennikov, A.* (2021). Can there be given any meaning to contextuality without incompatibility? *Int. J. Theor. Phys.* 60. 106–114.

20. *Conte, E., Khrennikov, A., Todarello, O. and Federici, A.* (2008). A preliminary experimental verification on the possibility of Bell inequality violation in mental states, Neuroquantology, 6(3), 214–221.

21. *Asano, M., Khrennikov, A., Ohya, M., Tanaka, Y. and Yamato, I.* (2014). Violation of contextual generalization of the Leggett-Garg inequality for recognition of ambiguous figures, Physica Scripta, T163, 014006.

22. *Basieva, I., Cervantes, V.H., Dzhafarov, E.N., Khrennikov, A.* (2019). True contextuality beats direct influences in human decision making. J. Exp. Psych.: General 148, 1925–1937.

23. *Cervantes, V.H., and Dzhafarov, E.N.* (2020). Contextuality analysis of impossible figures. Entropy 22, 981.

Chapter 11.
Development and validation of BEVEL dataset of natural dynamic facial expressions

Olga A. Korolkova, Elena A. Lobodinskaya

Moscow State University of Psychology & Education,
Moscow, Russia; e-mail: olga.kurakova@gmail.com
Moscow State University of Psychology & Education; Moscow Institute of Psychoanalysis,
Moscow, Russia; e-mail: elena.lobodinskaya@gmail.com

The number of studies that explore the perception of dynamic facial expressions of emotions has been increasing over the last years, and there is a general trend towards the use of more naturalistic facial stimuli, which allow to bridge the gap between controlled laboratory experiment and real-life human behaviour (Dobs et al., 2018). Such facial stimuli are more ecologically valid and can represent a wider range of emotional states compared to the standardised sets of static images of basic emotions that had been extensively used in the past.

Indeed, early studies of the perception of emotions mostly utilised greyscale or full-colour photos of posed intense facial expressions, displayed by professional or amateur actors according to the rules unveiled by Paul Ekman (Ekman, Friesen, 1978). These sets of images generally allow high control of the actors' pose, light conditions and intensity of expressions, but they lack essential information about the flow and development of the expression in time (Krumhuber et al., 2013). One way to add dynamics to such images is to animate them using linear morphing techniques, although the outcome of such transformation is not necessarily an accurate representation of a real facial movement (Korolkova, 2018). To overcome these limitations, more recent studies used artificial dynamic avatars based on motion captured from the faces of real actors, or the video recordings of the actors themselves, but the expressions on their faces in most cases were still posed according to Ekman's rules (Krumhuber et al., 2016).

To further explore the naturalness and genuineness of the emotions expressed on a face, researchers started using video recordings of induced emotions that naturally emerge as responses to some kind of emotional stimuli or situation. For example, participants might be asked to watch video clips, listen to music, or complete various tasks, e.g., to touch a spider to induce fear (Siedlecka, Denson, 2019). Among these methods, watching video clips has

been proved the most effective way of inducing emotional reactions (Fernández-Aguilar et al., 2019).

In the present study we aimed to develop a new dataset of video-induced dynamic facial expressions displayed by Russian participants and validated on Russian sample. According to a large-scale review of more than 350 datasets of emotional stimuli (Diconne et al., 2021), no such sets has been previously created and validated on Russian-speaking participants. Here we report the full workflow of the dataset creation. Preliminary results of each stage have been presented at local conferences (Korolkova, Lobodinskaya, 2020; Lobodinskaya, Korolkova, 2019, 2020).

The work on the database comprised three main stages. At Stage 1, we selected video clips to induce emotional responses in participants. At Stage 2, we presented the emotion-inducing videos to a separate group of participants and video-recorded their expressions. At Stage 3, we extracted short fragments from the recordings of facial expressions and presented them to observers who evaluated the emotions on these recordings, to select those fragments that were consistently perceived as particular emotions. All procedures comply with the Declaration of Helsinki. Further we provide the details of each stage.

Stage 1

Method. The aim of the first stage of the study was to select videos that can induce basic emotions in Russian participants.

Stimuli. A preliminary set of forty videos was comprised through web search. We intended to select short extracts (1–3 min) from documentary videos and cinema that would induce emotions of happiness, surprise, interest, fear, anger, disgust, contempt and sadness. We based our search on the extensive list of previously used film stimuli (Gilman et al., 2017), however, we tried to use extracts from documentaries or less popular cinema that were not familiar to our participants. The short descriptions of the video clips selected for emotion induction are shown in Table 1. Full list of the evaluated videos and their evaluation data are available as online supplementary materials at https://osf.io/qs3cu/.

Participants. Twenty-five participants (most of them university students from Moscow and Moscow region, 14 women and 11 men, age range 18–42 years, median age 33 years) took part at this stage.

Apparatus. The videos were shown either on a wide wall-mounted screen, or on individual monitors with resolution of at least 1024×768 pixels.

Procedure. Participants watched the videos in a pre-defined order. After watching each video, they selected one or several emotions that this video evoked in them, from the following list: happiness, surprise, interest, fear, anger, disgust, contempt and sadness. They could also write down other emotions that were not listed.

Data analysis. The data were analysed using R (R Core Team, 2020). We aimed to identify the videos that were evaluated as evoking one of the emotions significantly more often than each of the other emotions. For this, we pairwise compared the frequencies of selecting each emotion using Fisher's exact test with FDR adjustment.

Results. Among the forty videos that were shown to the participants, six videos were evaluated as evoking happiness more often than any other emotion (adjusted $ps<0.006$); one video evoked interest (adjusted $ps<0.008$); one video evoked disgust (adjusted $ps<0.0003$), and one video evoked anger (adjusted $ps<0.035$). In addition, seven other videos evoked two emotions, namely surprise and interest (adjusted $ps<0.038$); three videos evoked happiness and interest (adjusted $ps<0.035$); two videos, fear and interest (adjusted $ps<0.038$); two videos, fear and surprise (adjusted $ps<0.037$). The other videos either evoked more than two emotions, or there were no emotions evoked more often than the other ones. The results of the evaluation of video clips selected for further presentation at Stage 2 are shown on Figure 1 and in Table 1.

Discussion. At Stage 1 of the study we obtained self-reported emotions that can be induced by watching short video clips. The results show that most of the clips induced mixed emotions, with only several videos consistently perceived as evoking a single basic emotion. Happiness is most often evoked as the only emotion, while several videos can evoke interest accompanied with another basic emotion (surprise, happiness, or fear). Among the videos that we used, no one could evoke contempt. Sadness was evoked by several videos, although it was accompanied by other emotions (e. g., fear) and it has never been the only emotion that was evoked significantly more often than any other emotions by any particular video.

When selecting the video clips for further use at Stage 2, we based our choice on the results of the Stage 1 evaluation and tried to include videos inducing a single emotion (e. g., video no. 39 that shows a man eating larvae, inducing disgust; video no. 18 that shows geese chasing people, inducing happiness), as well as videos that can induce several emotions (e. g., video no. 21 that shows a tiger attacking people in a forest, inducing fear and sadness;

video no. 25 that shows an illusionist doing magic tricks, inducing surprise and interest).

<div align="right">*Table 1*</div>

Description of the videos selected to induce emotions

No.	Description	Induced emotions (Stage 1)	Duration (min:sec)
1	Werewolf chasing people	happiness; interest; fear; disgust	03:44
2	Swinging bridge contest	happiness; interest	02:35
3	Stand-up comics	happiness	03:52
4	Sleeping dogs	happiness	03:10
5	Singing to the cows	happiness	01:43
6	Speeded oil painting	happiness	03:03
7	Tornado	fear	05:49
8	Geese chasing people	happiness	03:09
9	Singing teacher	happiness	01:39
10	Snakes fighting	fear; disgust	03:54
11	Tiger attacking people in a forest	fear; sadness	03:05
12	Crabs and moray eels	interest	03:03
13	Illusionist showing magic tricks	surprise; interest	03:44
14	Throwing children across the street	interest; fear	01:48
15	Praying mantises	interest; disgust	03:00
16	Eating larvae	disgust	03:12
17	Description of nazis medical experiments	anger	04:00
18	Cartoon about emotions	not evaluated	02:33

Stage 2

Method. The main aim of the second stage was to induce genuine emotions in participants and to obtain video recordings of their facial expressions.

Stimuli. Based on the results of the Stage 1, we selected 17 videos to present to the Stage 2 participants, and added one more video that we showed at the end of the study (this video has not been evaluated for emotional content and was used as a closing video to entertain participants at the end of the recording procedure). Their description and the results of the Stage 1 evaluation are provided in Table 1 and on Figure 1.

Participants. Twenty one participants who did not took part in any other stages of the study (15 women and 6 men, median age 22 years) volunteered to participate. They completed the informed consent form and agreed to the use of video recording of their faces for research purposes.
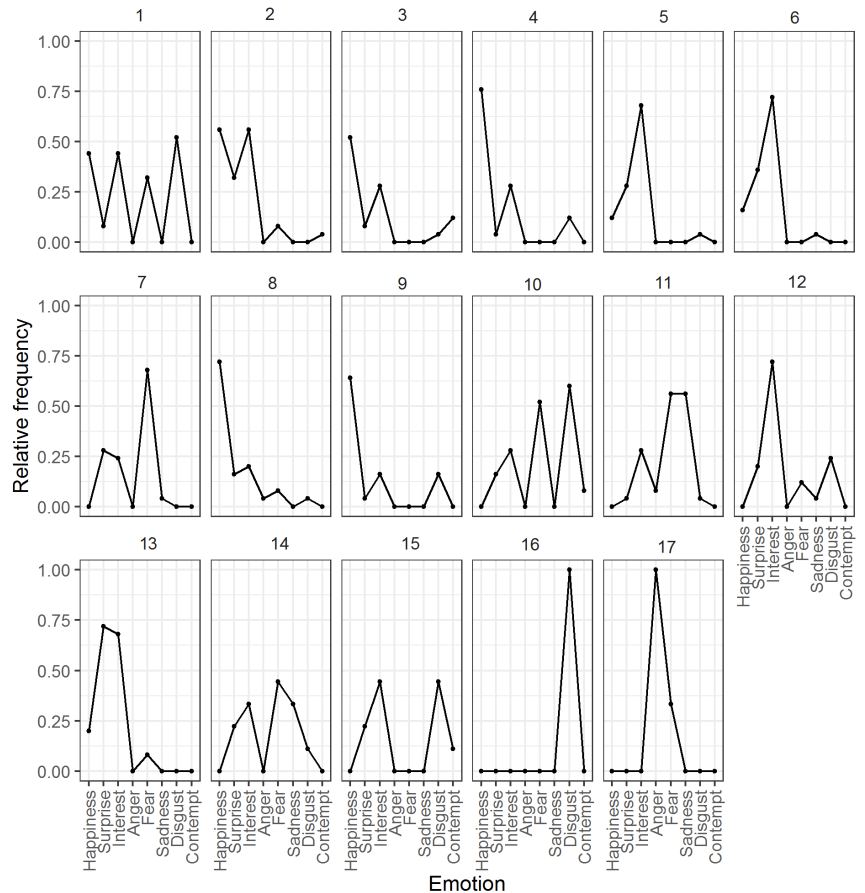
Fig. 1. Relative frequencies of selecting emotions that can be induced by video clips in Series 1. Panels are numbered according to the Table 1.

Apparatus. For the video recording, we used Sony RX10 IV camera in full HD mode in portrait orientation (1920×1080 pixels) at 120 frames per second. The videos were delivered through a display connected to a PC. Sound was

played through professional audio system. Participants sat in front of the display in a separate room with ambient light from 4-point LED light sources. The video camera was mounted just above the display and recorded the participant's face and upper shoulders. The researcher controlled the whole process (playing the videos and recording the face) from outside the recording chamber. The equipment from the point of view of the participant is shown on Figure 2.

Fig. 2. The video recording equipment as seen by the participant.

Procedure. Before the video recording session, participants completed the Toronto Alexithymia Scale (TAS-26) in Russian adaptation (Eresko et al., 1994), and self-evaluated their current emotional state using a modified version of the Differential Emotions Scale (Izard, 1991) translated into Russian. DES-III includes thirty-three Likert type scales that allow to evaluate the intensity of emotions from 1 (the emotion is not present at all) to 5 (the emotion is maximally present). During the recording session the participants were shown the emotion-inducing video clips in fixed order. Their task was to watch the videos and self-evaluate the emotional state after watching each video. Participants were aware of the recording, but were encouraged to behave naturally, as if they were watching these videos at home, and do not try to suppress or emphasise their reaction.

Data analysis. The data were analysed using R, version 3.6.3. For each video fragment, and for self-evaluation data, we fitted a linear model with emotional scale of the DES-III as predictor and the participants' responses

as dependent variable. We used linear contrasts with FDR correction to test whether the mean response on each scale was higher than 2 ("the emotion is slightly present"), and whether it was higher than 3 ("the emotion is moderately present"). We also calculated average TAS-26 score across participants.

Results. The average TAS-26 score of the participants was 57 (range 36–77) which is within the normal range, or slightly higher. Participants evaluated their emotional state before the study as attentive (mean rating 3.6; $p = 0.013$ with FDR correction), with less degree of being calm, serene, relaxed, concentrating, alert, delighted, happy and joyful (mean ratings higher than 2; ps < 0.05 with FDR correction). All other emotions were evaluated as not present at all or slightly present (mean ratings not higher than 2). These results indicate that at the beginning of the study the participants did not experience any intense positive or negative emotions.

The results of the emotional self-evaluation after watching each video clip show that video 1 (Werewolf chasing people) induced states of attention, concentration and alertness (mean rating > 3; ps < 0.002 with FDR correction) with less degree of being calm, serene, relaxed, delighted, happy, joyful and surprised (mean ratings > 2; ps < 0.003 with FDR correction). Video 4 (Sleeping dogs) induced serene and relaxed states (mean ratings > 3; ps < 0.005) and less degree of being calm, attentive, concentrating, alert, delighted, happy, joyful and surprised (mean ratings > 2; ps < 0.001). Video 5 (Singing to the cows) induced calm, serene, relaxed, attentive, concentrating, alert, delighted and happy states (mean ratings > 3; ps < 0.047) as well as less intense joy, surprise, amazement and astonishment (mean ratings > 2; ps < 0.032). Video 6 (Speeded oil painting) induced calm, serene, relaxed, attentive, concentrating, alert, delighted and surprised states (mean ratings > 3; ps < 0.036) and less degree of being happy, joyful, amazed and astonished (mean ratings > 2; ps < 0.001). Video 7 (Tornado) induced attentive, concentrating and alert states (mean ratings > 3; ps < 0.003) as well as less degree of being calm, serene, relaxed, delighted, surprised, amazed and astonished (mean ratings > 2; ps < 0.015). Video 11 (Tiger attacking people in a forest) induced attentive, concentrating and alert states (mean ratings > 3; ps < 0.046) with less degree of being calm and surprised (mean ratings > 2; ps < 0.049). Video 13 (Illusionist showing magic tricks) induced states of being attentive, concentrating, alert, delighted, joyful and surprised (mean ratings > 3; ps < 0.036) as well as calm, serene, relaxed, happy, amazed and astonished (mean ratings > 2; ps < 0.001). Video 14 (Throwing children across the street) induced attentive, concentrating and alert states (mean

ratings > 3; ps < 0.001) and less degree of being surprised (mean rating > 2; p = 0.002). Video 18 (Cartoon about emotions) induced serene, relaxed, attentive, delighted, happy and joyful states (mean ratings > 3; ps < 0.004) as well as being calm, concentrating and alert (mean ratings > 2; ps < 0.001). The other videos that were shown to the participants, did not induce intense emotions with statistical significance, so their influence on the emotional state was less intense.

Discussion. At Stage 2 we obtained high-resolution, high-speed video recordings of induced facial emotional expressions of participants during watching video clips. As we encouraged them to express their emotions naturally, we expected that the intensity and the number of expressive episodes would have high variability between participants. Though, according to the results of the self-evaluation of the emotions evoked by the video-clips, in most cases the participants have indeed experienced the intended emotions, as proved by the group-level statistical significance. Therefore, for the presentation at the next stage of the study, we selected fragments of the video recording based on the experts' opinion and on the participants' self-evaluation.

Stage 3

Method. The aim of this stage was to obtain the normative data regarding the perception of the recorded facial expressions.

Stimuli. Among the individuals whose faces were recorded at Stage 2, we selected four posers who had the highest amount of expressive episodes (please see Table 2). Their recordings were downsampled to 30 frames per second and cropped to show a close-up view of the posers' faces in portrait orientation. Two experts (the authors of the study) independently watched the recordings and selected short (up to 5 seconds) fragments that contained distinct facial expressions. Among these, 33 to 36 fragments of each individual's face were selected jointly by the authors for further evaluation.

Participants. We recruited participants (Russian-speaking students and professionals) who did not take part in the other stages of this study. The data of those participants who had TAS-20 score of 61 or higher, or filled the forms incorrectly (not following the instructions), were excluded from the analysis (please see the description of the final sample in Table 2).

Apparatus. The evaluation of the facial expressions was completed in part in the lab (through paper-and-pencil forms or through web-based forms), in

part online (through web-based forms). In the lab participants viewed the expressions on a wide wall-mounted flat screen and completed the evaluation in small groups. Participants who used web-based version of the study viewed the expressions on their own screens that were at least 1024×768 pixels in resolution and were not mobile devices.

Table 2

Description of the posers and participants of the Stage 3

Series within Stage 3	Posers			Participants (final sample)		
	Gender	Age	TAS-26 score	Sample size (women/men)	Median age (age range)	Average TAS-20 score
Series 1	Female	25	48	37 (32/5)	26 (17–45)	46
Series 2	Male	23	76	35 (26/9)	31 (18–53)	43
Series 3	Female	22	53	38 (31/7)	32 (18–50)	42
Series 4	Male	31	56	47 (44/3)	28 (18–48)	40

Procedure. At the beginning of the study participants completed the Toronto Alexithymia Scale (TAS-20) in Russian adaptation (Starostina et al., 2010), and self-evaluated their current emotional state using DES-III. After that they watched the dynamic expressions presented continuously in fixed order and provided short verbal description of the emotions they perceived on each fragment. After a short break they watched each fragment again and evaluated the expression on the face using DES-III.

Data analysis. The data were analysed using R, version 3.6.3. For each video fragment, and for self-evaluation data, we fitted a linear model with emotional scale of the DES-III as predictor and the participants' responses as dependent variable. We used linear contrasts with FDR correction to test whether the mean response on each scale was significantly higher than 2 ("the emotion is slightly present"), and whether it was higher than 3 ("the emotion is moderately present").

Results. *Data exclusion.* Because of high TAS-20 score, we excluded seven participants in series 1, four participants in series 2, four participants in series 3, and two participants in series 4. In addition, four participants in series 1 completed the DES-III incorrectly, and their DES-III data were also excluded. Several participants did not evaluate all the fragments (five participants

in series 1; four participants in series 2; three participants in series 3; eight participants in series 4), although their (incomplete) data were still included in the analysis.

Self-evaluation using DES-III. In all series, participants evaluated their emotional state before the study as attentive, concentrating and alert (mean ratings higher than 3; $ps < 0.05$ with FDR correction), with less degree of being calm, serene, relaxed, delighted, happy and joyful (mean ratings higher than 2; $ps < 0.05$ with FDR correction). All other emotions were evaluated as not present at all or slightly present (mean ratings not higher than 2). Therefore according to self-evaluation, the participants did not experience any intense positive or negative emotions that could have influenced their evaluation of the facial expressions (Figure 3).
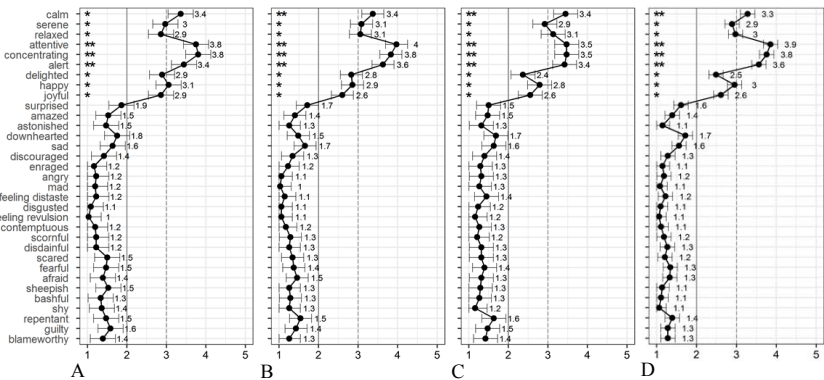
Fig. 3. Self-evaluation of the Stage 3 participants' emotional state before the study in series 1 (A), series 2 (B), series 3 (C) and series 4 (D). On vertical axis, scales of DES-III; on horizontal axis, ratings of experienced emotions. Points represent mean ratings; whiskers, standard error; single asterisks show the emotions that were evaluated significantly higher that 2; double asterisks, the emotions that were evaluated significantly higher than 3.

Evaluation of the dynamic expressions using DES-III. The results of the evaluation of the expressive video fragments at Stage 3 are organised as following. For each series, we present the data regarding those expressions that received average rating significantly higher than 3 ("the emotion is moderately present") on the 5-point Likert scale. Where possible, we group similarly evaluated expressions together and show only one example from each group.

The full evaluation data for each of the dynamic fragments is available as the online supplement at https://osf.io/qs3cu/.

In series 1, one dynamic expression showing a smile with the mouth closed ("happiness") was evaluated as happy and joyful (mean ratings > 3), with moderate degree of being calm, serene, relaxed, attentive, concentrating and delighted (mean ratings > 2). Another expression that showed quick brows flash and eyes widen ("surprise") was evaluated as highly surprised (mean rating > 3) and moderately attentive, concentrating, amazed and astonished (mean ratings > 2). The third type of expression, a slight smile and slightly narrowed eyes ("attention"), was evaluated as attentive and concentrating (mean ratings > 3) while being moderately calm, serene, relaxed, alert, delighted, happy and joyful (mean ratings > 2). The fourth type, a slight smile with the mouth closed and head slightly tilted ("calmness"), was evaluated as calm and relaxed (mean ratings > 3) and moderately serene, attentive, delighted, happy and joyful (mean ratings > 2). The fifth type comprised four dynamic fragments, showing wrinkled nose and mouth and averted eyes ("disgust"). They were evaluated as feeling distaste, and in some cases, being disgusted and feeling revulsion (mean ratings > 3). Moderately intense were the perceived feelings of being contemptuous, scornful, and in one case, also being downhearted and sad (mean ratings > 2). Except the described emotions, no other emotional characteristics have been rated higher than 2 points. The five types of the identified expressions are shown on Figure 4.

In series 2, we were able to identify three groups of expressions that had at least one mean rating significantly higher than 3. The first group ("happiness") included eight expressions that showed smiles or light laughter and were evaluated as delighted, happy and joyful (mean ratings > 3) with moderate intensity of neutral state (calm, serene, relaxed) and interest (attentive, concentrating, and/or alert) and, in some cases, of surprise (mean ratings > 2). Other emotions were not perceived on these video fragments. The second group ("surprise") included seven expressions with briefly rising brows and in some cases with light smile and/or nodding. They were evaluated as surprised and, in some cases, amazed, calm and/or attentive (mean ratings > 3). In some of these expressions, such states as being serene, relaxed, concentrating, alert, delighted, happy, joyful and/or astonished, were evaluated as moderately intense (mean ratings > 2). The third group ("attention") included four expressions of narrowing the eyes and/or lowering the brows that were evaluated as attentive, concentrating and alert (mean ratings > 3) with moderate degree of being calm, serene, relaxed and/or surprised (mean ratings > 2). Apart from

the expressions that comprised these three groups, other video fragments were perceived as less intense emotions. Typical examples from each group are shown on Figure 5.
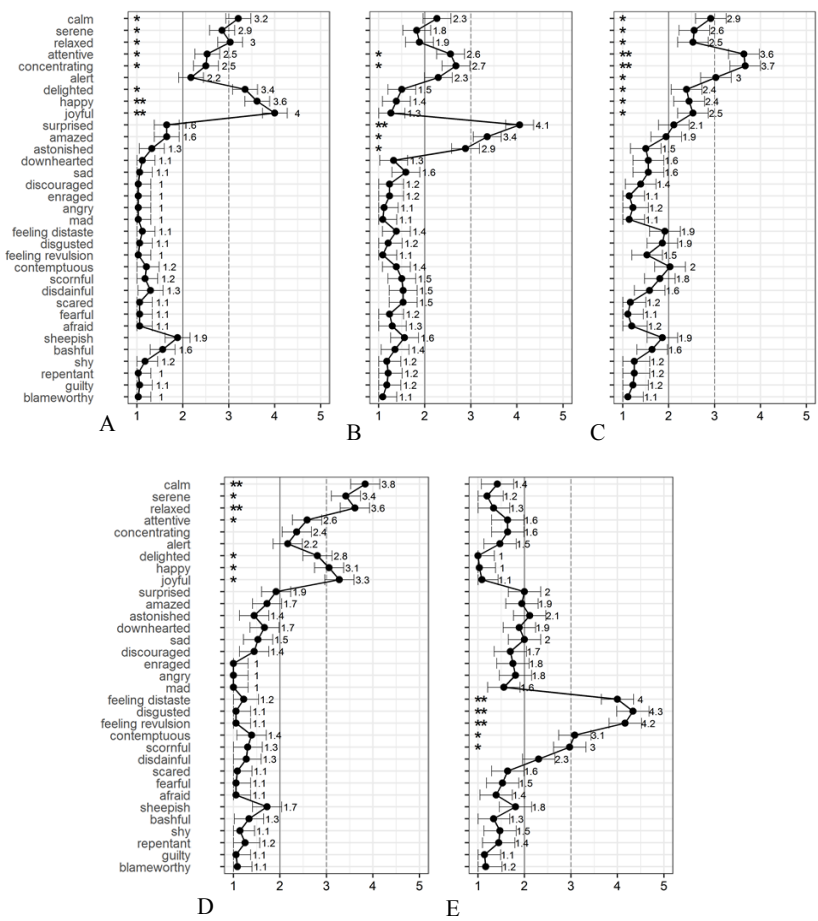
Fig. 4. Examples of the expressions evaluated in Series 1 as "happiness" (A), "surprise" (B), "attention" (C), "calmness" (D), and "disgust" (E). On vertical axis, scales of DES-III; on horizontal axis, ratings of emotions. Points represent mean ratings; whiskers, standard error; single asterisks show the emotions that were evaluated significantly higher that 2; double asterisks, the emotions that were evaluated significantly higher than 3.

In series 3, the analysis revealed five types of perceived expressions. The first one (a burst of laughter) was perceived as delighted, happy and joyful (mean ratings > 3) with moderate intensity of being calm, serene, relaxed, and sheepish (mean ratings > 2). The second type included two expressions perceived as surprised (mean ratings > 3), slightly attentive, concentrating, alert, amazed, astonished, and/or sad (mean ratings > 2). The third type (two expressions) was perceived as intensively sad (mean ratings > 3) and moderately calm, attentive, concentrating, alert, downhearted, and/or repentant (mean ratings > 2). The fourth type (three expressions) was perceived as feeling distaste (mean ratings > 3) as well as disgusted, feeling revulsion, being contemptuous and scornful (mean ratings > 2). Finally, the fifth type (two expressions) was perceived as concentrating (mean ratings > 3), calm, attentive, alert and sad (mean ratings > 2). The examples are shown on Figure 6.
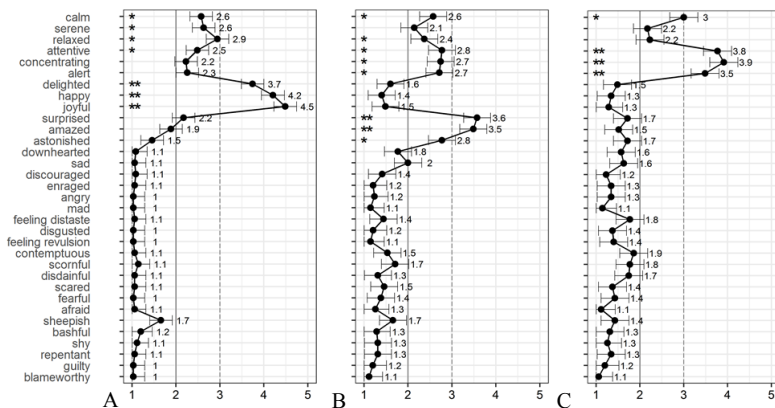
Fig. 5. Examples of the expressions evaluated in Series 2 as "happiness" (A), "surprise" (B), and "attention" (C).

In series 4, we revealed seven different types of expressions (Figure 7). Five fragments showing quick brows rise and/or widening the eyes were evaluated as expression of intense surprise (mean ratings > 3) and moderate state of being alert, amazed, astonished, attentive, calm, concentrating, happy, joyful, relaxed, scared, and/or serene (mean ratings > 2). Three other fragments showing smile or laughter were evaluated as happy and joyful (mean ratings > 3) and calm, serene, relaxed, and delighted (mean ratings > 2). One fragment showing brow rise and eyes widening, and then a smile, was evaluated as both joyful and surprised/amazed (mean ratings > 3) and attentive, delighted,

happy, astonished (mean ratings > 2). Three fragments showing frowning or blinking with the mouth open, were evaluated as attentive, concentrating and alert (mean ratings > 3). One fragment was evaluated as calm (mean ratings > 3) and serene, relaxed, attentive, concentrating, alert, joyful, surprised (mean ratings > 2); one (jaw drop) as astonished, scared and fearful (mean ratings > 3) and attentive, concentrating, surprised, amazed, downhearted, afraid (mean ratings > 2); and one fragment (jaw movements) as angry (mean ratings > 3) and attentive, concentrating, alert, enraged, mad, feeling distaste, disgusted, feeling revulsion, contemptuous, scornful, disdainful (mean ratings > 2).
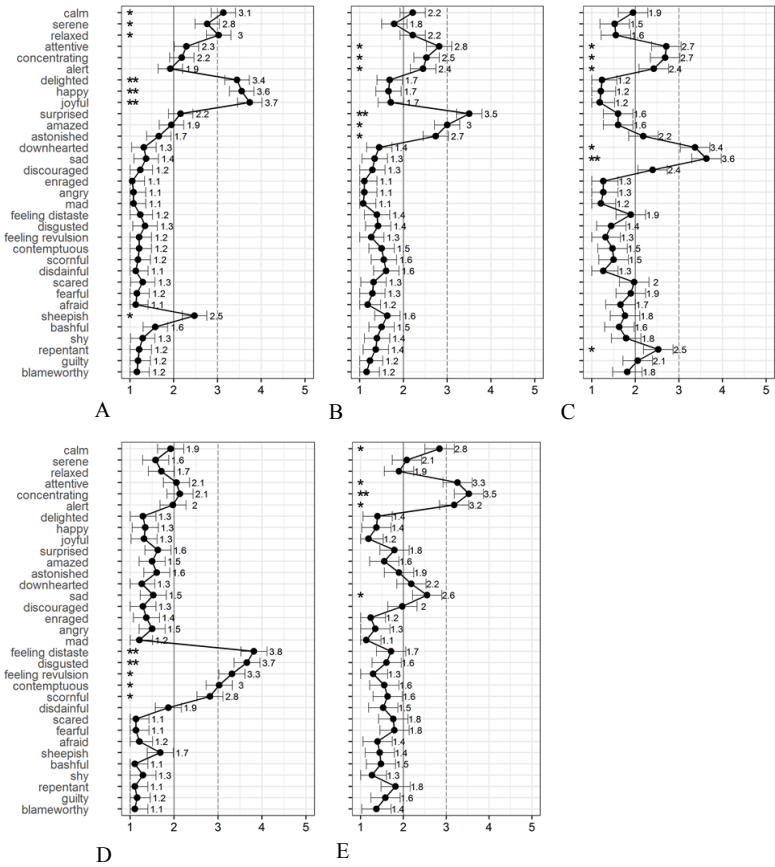
Fig. 6. Examples of the expressions evaluated in Series 3 as "happiness" (A), "surprise" (B), "sadness" (C), "disgust" (D), and "attention" (E).
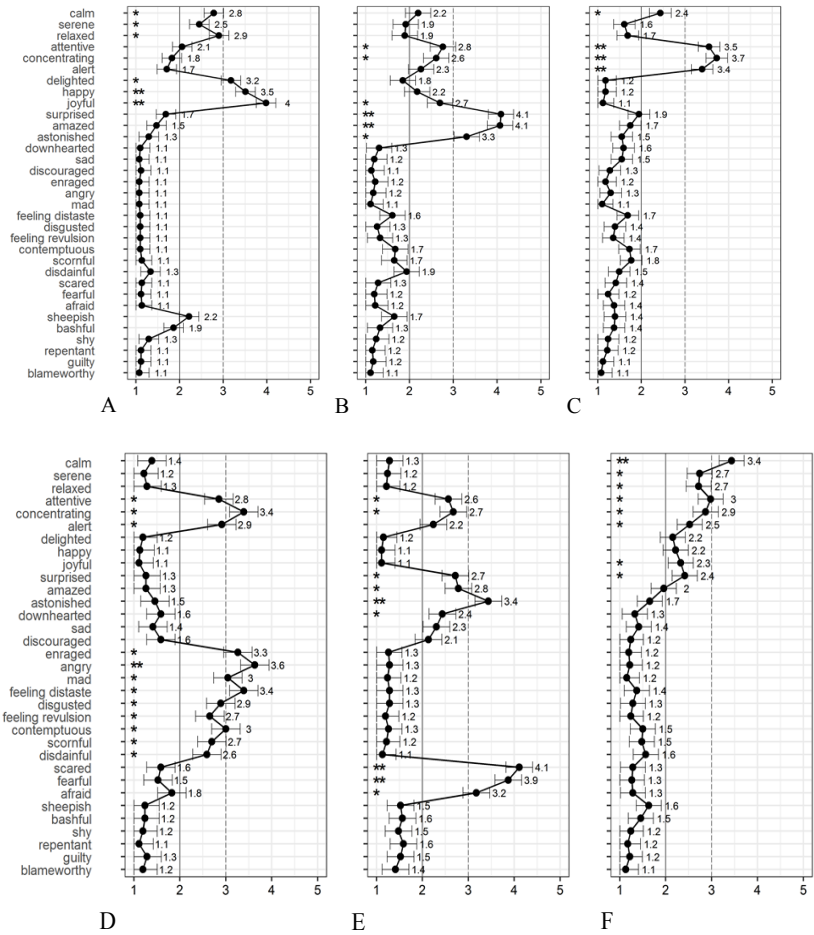
Fig. 7. Examples of the expressions evaluated in Series 4 as "happiness" (A), "surprise" (B), "attention" (C), "anger" (D), "fear" (E), and "calm" (F).

Discussion. At Stage 3 we obtained normative data of the emotion recognition from the dynamic fragments of induced emotional expressions. All of the intended emotions, including happiness, surprise, interest/attention, disgust, sadness, anger, fear and calm state, have been identified among the

evaluated fragments, however, not all emotions have been present or intense enough in all the posers. Moreover, as these fragments present natural emotions, they have greater variability in expression compared to the posed ones. This is reflected in their evaluation data, as some of the fragments are perceived as less intense emotions than the others. Also, most of the evaluated fragments appear to the observers as displaying several emotions at the same time (please see Figures 4–7), which can also reflect inter-subject variability in emotion expression and perception.

Conclusion

In the present study, we describe the development and validation of the first Russian-language set of induced dynamic facial emotional expressions. To do this, at stage 1 we selected the emotion-inducing video clips. These videos were then shown to the participants of the stage 2, whose facial expressions were recorded while they watched the videos. At stage 3, short fragments derived from the face recordings were evaluated by naïve observers, who rated the intensity of thirty-three emotional states on the fragments. In sum, we obtained evaluation data on 138 fragments from recordings of four individuals and were able to identify eight intense emotions displayed on a face: happiness, surprise, interest/attention, disgust, sadness, anger, fear and calm state. We plan to further extend the dataset and include the fragments from recordings of other individuals. The dataset and all the meta-data including validation results can be obtained for use in research purposes at https://osf.io/qs3cu/.

References

1. Diconne, K., Paltoglou, A., Kountouriotis, G., Parker, A., & Hostler, T. (2021). KAPODI – The Searchable Database of 364 Available Emotional Stimuli Sets. https://doi.org/10.31234/osf.io/6j9mx.
2. Dobs, K., Bülthoff, I., & Schultz, J. (2018). Use and Usefulness of Dynamic Face Stimuli for Face Perception Studies — a Review of Behavioral Findings and Methodology. *Frontiers in Psychology*, *9*, 1–7. https://doi.org/10.3389/fpsyg.2018.01355.

3. Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

4. Eresko, D. B., Isurina, G. S., & Kaidanovskaya, E. V. (1994). *Alexithymia and methods for its determination in borderline psychosomatic disorders: Methodological guide*. SPb.: V.M. Bekhterev Psychoneurological Institute.

5. Fernández-Aguilar, L., Navarro-Bravo, B., Ricarte, J., Ros, L., & Latorre, J. M. (2019). How effective are films in inducing positive and negative emotional states? A meta-analysis. *PLOS ONE*, *14*(11), 1–28. https://doi.org/10.1371/journal.pone.0225040.

6. Gilman, T. L., Shaheen, R., Nylocks, K. M., Halachoff, D., Chapman, J., Flynn, J. J., … Coifman, K. G. (2017). A film set for the elicitation of emotion in research: A comprehensive catalog derived from four decades of investigation. *Behavior Research Methods*, *49*(6), 2061–2082. https://doi.org/10.3758/s13428-016-0842-x.

7. Izard, C. E. (1991). *The Psychology of Emotions*. New York: Springer Science & Business Media. Retrieved from https://books.google.com/books?id=RPv-shA_sxMC&pgis=1.

8. Korolkova, O. A. (2018). The role of temporal inversion in the perception of realistic and morphed dynamic transitions between facial expressions. *Vision Research*, *143*, 42–51. https://doi.org/10.1016/j.visres.2017.10.007.

9. Korolkova, O. A., & Lobodinskaya, E. A. (2020). The role of individual characteristics in the perception of emotionally colored video clips. In *IX International Conference on Cognitive Science (INCB-2020)*. Moscow: Russian Academy of Sciences.

10. Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of Dynamic Aspects of Facial Expressions: A Review. *Emotion Review*, *5*(1), 41–46. https://doi.org/10.1177/1754073912451349.

11. Krumhuber, E. G., Skora, L., Küster, D., & Fou, L. (2016). A Review of Dynamic Datasets for Facial Expression Research. *Emotion Review*, 175407391667002. https://doi.org/10.1177/1754073916670022.

12. Lobodinskaya, E. A., & Korolkova, O. A. (2019). Induction of natural emotional facial expressions. In P. N. Ermakov, V. A. Labunskaya, & G. V. Serikov (Eds.), *Appearance in various contexts of interaction: materials of the All-Russian Scientific Conference, October 28-30, 2019* (pp. 41–43). Moscow: CREDO.

13. Lobodinskaya, E. A., & Korolkova, O. A. (2020). Video material for the induction of natural emotional expressions in studies of the mechanisms of perception of a "living" face. In *Psychological and psychoanalytical*

research: Yearbook 2018-2019 of Moscow Institute of Psychoanalysis (pp. 65–71). Moscow: Moscow Institute of Psychoanalysis.

14. R Core Team. (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* Retrieved from http://www.r-project.org/.

15. Siedlecka, E., & Denson, T. F. (2019). Experimental Methods for Inducing Basic Emotions: A Qualitative Review. *Emotion Review*, *11*(1), 87–97. https://doi.org/10.1177/1754073917749016.

16. Starostina, E., Taylor, G., Quilty, L., Bobrov, A., Moshnyaga, E., Puzyreva, N., Bagby, M. (2010). A new 20-item version of the Toronto alexithymia scale: Validation of the Russian language translation in a sample of medical patients. *Social and Clinical Psychiatry*, *20*(4), 31–38.

Chapter 12.
Behavioral markers of the error detection in solving cognitive tasks

V. A. Kovaleva

Saint Petersburg State University, Saint Petersburg, Russia

Errors in solving various tasks can have a significant impact on the effectiveness of further human activity. Numerous studies devoted to the study of the problem of errors indicate the existence of unconscious determinants of erroneous actions in solving various tasks [Bechtereva, Gretchin, 1969; Rabbitt, 1968; Allahverdov, 2000, 2016; Cleeremans, Destrebecqz, 2005; D'Angelo, Humphreys, 2015; Andrianova, 2016; Kuvaldina et al., 2019, etc.].

A person is able to detect his correct and erroneous answers on an unconscious level. This process in cognitive psychology has been called the "phenomenon of error detection". It allowed to make a number of empirically confirmed assumptions that answer the question: why, solving a simple cognitive task, a person still makes mistakes, despite the well-known algorithmic solution of these tasks.

The phenomenon of error detection was first described in the works of Bekhtereva N. P. and Gretchin V. B. in 1968 [Bechtereva, Gretchin, 1969]. They revealed that some neural populations ("error detectors") could not react or react randomly to the task, and some could react only to the correct solving or to the correct and erroneous one. They also obtained data on the existence of a separate neural population that reacts only to the erroneous solving tasks. In 1968, Rabbit also hypothesized the existence of a certain mechanism responsible for the detection, correction and compensation of errors [Rabbitt, 1968].

R. Naatanen, studying the phenomenon of "Mismatch negativity" discovered by him in 1978, noted that the functional principle of error detection is disagreement with a possible program of action [Naata-nen, 2003]. According to the data obtained by recording the electrical activity of the brain, it was revealed that when the subjects make mistakes, a certain neural mechanism is activated. This process has been called "error-related negativity" (ERN). The results of the study of this phenomenon by Goering V., Goss B., Coles, Mey-

er D. and Donchin I. [Gehring, Goss, Coles, Meyer, Donchin, 1993] showed the relationship between ERN and various mechanisms for correcting and compensating errors. However, later Bernstein P., Scheffers M. and Coles M. [Bernstein, Scheffers, Coles, 1995] hypothesized that the correction and compensation process occurs after an error signal has been generated, and ERN is more related to the processes involved in generating error signals than to correcting activities.

The phenomenon of unconscious error detection in the works of Allakhverdov V. M. is interpreted within the theory of unconscious negative choice. It is assumed that a person can unconsciously distinguish between correct and erroneous answers, as well as repeat erroneous and correct answers [Allahverdov, 2000].

To date, there are many studies of markers of unconscious error detection. One of them is the "false > correct-phenomenon", which manifests itself in slowing down the reaction with an erroneous response compared to the correct one [Beckmann, 2000; Rammsayer, Brandler, 2003]. It does not depend on the complexity of the task, nor on the individual level of intelligence [Cleeremans, Destrebecqz, 2005].

Allakhverdov et al. note that confidence in the answer, as one of the markers of error detection, reduces the reaction time when solving tasks [Allakhverdov et al., 2019].

Another marker of the unconscious detection of the accuracy of the answer may be the tendency to repeat previously correct and erroneous answers [Allahverdov, 2000; D'Angelo, Humphreys, 2015; Andrianova, 2016; Kuvaldina et al., 2019, etc.].

In a number of studies, it has been noted that after an erroneous response, the reaction time in the subsequent probe slows down ("Post-error slowing"). There are various theories explaining the cause of the slowdown: the theory of cognitive control [Notebaert et al., 2009], the theory of conflict monitoring [Botvinick, Braver, Barch, Carter, Cohen, 2001]. And it propose that post-error slowing is caused by the relative infrequency of errors which causes attentional capture [Notebaert et al., 2009].

It is noted that a negative priming effect may be involved in the process of the post-error slowing, if the stimulus is repeated in consecutive probes [Hsu, Hsieh, 2021; Keute et al., 2019].

It was also found that under the condition that the stimulus of the masked

prime and the target stimulus are congruent, the number of erroneous responses increases [Keute et al., 2019], however, in the case of congruence of the supraliminal stimulus, the opposite result was obtained. In particular, it was noted that erroneous responses with non-congruent stimulus can also cause conflict and lead to an increase in errors and reaction time, as well as correct responses with non-congruent stimulus [Van der Borght, Braem, Notebaert, 2014].

In the psychology of dysontogenesis, the problem of erroneous actions when solving various kinds of tasks is widely studied due to the fact that this indicator is one of the fundamental markers by which the development of cognitive processes is assessed and the degree of a particular violation is judged. To date, studies have been conducted on the process of erroneous response in respondents with ADHD [Janssen, van Atteveldt, Oosterlaan, 2020; Keute et al., 2019], autism [Sokhadze et al., 2019], Tourette syndrome [Shephard, Jackson, Groom, 2016] and delayed psychomotor development [Schrijvers et al., 2008]. It should be noted that there are practically no studies devoted to the analysis of the process of erroneous response by people with mental retardation. They mainly do not concern the decision-making process about the answer itself, but are aimed at finding the accompanying causes of making mistakes.

Considering this problem from this point of view does not answer the question of the reasons for the increase in the number of errors and reaction time when solving tasks. Ramseyer T. and Brandler S. note that the response delay in simple cognitive tasks does not depend on the level of intelligence, i.e. reaction time is not a measure of the success of tasks, but indicates other individual differences associated with cognitive processes involved in solving various tasks [Rammsayer, Brandler, 2003]. Moreover, the hypothesis that the cause of errors is limited cognitive resources is refuted by various studies of error detection. The phenomenon of unconscious error detection suggests that, despite the erroneous answer, the cognitive system knows the correct answer to the task, but under the influence of various factors decides in favor of the error.

Despite the accumulated empirical data concerning the nature of errors, this problem has not yet been solved. Questions remain open about the mechanisms underlying certain error detection phenomenons, and there are also significant difficulties associated with the inconsistency of empirical data obtained in the study of phenomenons and the heterogeneity of their descriptions.

An additional problem is the process of an erroneous response by people with intellectual disabilities, since insufficient attention is paid to this problem within the psychology of dysontogenesis, and existing studies are limited to explaining the commission of errors by violations in cognitive processes without describing the process of erroneous action itself and the influence of various conditions on this process.

Thus, the current study was conducted in order to replicate the phenomenons identified as possible markers of unconscious error detection, and in order to study various factors that can influence the process of correct and erroneous response, on a sample of adolescents with a normal level of development and adolescents diagnosed with mild mental retardation.

This approach to the study of erroneous answers will allow not only to study the process of unconscious error detection, but also to understand the reasons for making mistakes by people with intellectual disabilities by studying the decision-making process about the answer for further prediction and correction of erroneous answers.

Materials and Methods

Subjects

Two groups of adolescents took part in the study: adolescents with a diagnosis of mild mental retardation (26 people, the age of the subjects was 13–16 years old) and adolescents with a normal level of intellectual development (32 people, the age of the subjects was 12–15 years old).

Research methods

An experiment was conducted using an supraliminal prime acting as a distractor in the previous probe, which had to be ignored.

The subjects were presented with two contour images, one was red, the other was green. The green image had to be ignored, and the red one had to be categorized into one of two groups: eidetic or inanimate, by pressing the keys: to the right is eidetic; to the left is inanimate.

In total, there were 12 different contour images in the experiment: 6 images from the category of eidetic objects and 6 images from the category of inanimate objects. All images were located in the center of the screen on a gray

background. The time from the moment the images appeared on the screen to the moment the subject pressed the key was analyzed.

The experiment was developed using the software "PsychoPy2_PY3 (version v3.0.4)". Incentive materials were presented on a laptop screen with a screen resolution of 1366 x 768.

At the beginning of the experiment, the subject was presented with an instruction on the monitor screen, accompanied by an image of an example task. After that, a series of training tasks followed and, if necessary, the instruction was repeated.

Experiment procedure. At the beginning of the main part of the experiment, the fixation point appeared on the screen for 600 ms, followed by a prime probe (N-1) for 250 ms. and disappeared. The subject was required, ignoring the green image, to classify the red image as eidetic or inanimate by pressing the right or left key. After that, a fixation point for 250 ms immediately appeared, and then a target probe (N) for 250 ms., for which it was also required to ignore the green image, and put the red one into the category by pressing the key.

The main part of the experiment contained 200 pairs of probes (N and N-1), before which there were 6 pairs of training probes. The sequence of pairs of probes was random (randomized). Also, after every 50 pairs of probes, the subjects were offered pauses.

In the experiment, five conditions were used for the presentation of the first probe (prime probe) in relation to the second (target probe):

1. The condition of ignored repetition (IR)—the distractor from the prime probe was repeated as a target stimulus in the target probe.

2. Attended repetition condition (AR)—the target stimulus in the prime probe was repeated as the target stimulus in the target probe.

3. Control conditions (N)—when both the distractor and the target stimulus of the first probe are not repeated as the target stimulus in the second.

4. The condition of categorical congruence of the IR, when the distractor of the prime probe is categorically identical to the target stimulus of the target probe.

5. The condition of categorical congruence of the AR, when the target stimulus of the prime probe is categorically identical to the target stimulus of the target probe.

Research results and their discussion

The percentage of errors in the target probes in the group of adolescents with a normal level of intelligence development was 7.2 %, and in the group of adolescents with mild mental retardation — 19.2 %. The total percentage of errors of prime and target probes in the group of adolescents with a normal level of intelligence development was 7 %, and in the group of adolescents with mild mental retardation — 19 %.

In order to compare the reaction time indicators within the groups of respondents and for the purpose of comparative analysis of the indicators between the two probes, the data were brought to a normal distribution using the Box-Cox transformation method, and then standardized by converting to z-values separately for each respondent. To analyze the transformed data, the Student's t-test for dependent samples, one-factor and two-factor analysis of variance (ANOVA) with subsequent a posteriori comparisons according to the Scheffe criterion were used.

1. Reaction time for correct and erroneous response

In a group of adolescents with a normal level of intelligence development, the average reaction time with a correct answer was statistically significantly higher than with an erroneous answer (t(31)=4.935; p<0.001). In adolescents with mental retardation, the average reaction time with an erroneous response was at the level of a tendency higher than with the correct one (t(25)= −1.854; p=0.064) (Table 1).

Table 1

Reaction time for correct and erroneous answer

	Teenagers with a normal level of intelligence development (M ± SD)	Teenagers with mental retardation (M ± SD)
Correct answer	736±406 ms.	897±694 ms.
Error answer	585±435 ms.	1036±1154 ms.

Discussion of the results

The decrease in reaction time with an erroneous answer in a group of adolescents with a normal level of intelligence development may have been due to the fact that with an erroneous answer, respondents were focused on

the speed of response and, if they were mistaken, then in connection with the strategy to answer as quickly as possible, since the percentage of errors in this group is small. Adolescents with mental retardation were more focused on the accuracy of the answer and erroneous answers were given longer due to the complexity of a certain kind of information. These results correlate with the results of other studies, where the correlation between the speed and accuracy of identification was revealed [Pleskac, Busemeyer, 2010; Rammsayer, Brandler, 2003, etc.]. A similar pattern can be explained by a number of other factors that influenced the decision-making process about the answer: the conditions of presentation of stimulus, categorical congruence of stimulus, priming effect caused by the conditions of tasks and other phenomenons noted in this study as possible markers for the detection of correct and erroneous answers. Further analysis of the results was carried out to verify this assumption.

2. Conditions of presentation of stimuli and priming effect as factors affecting the reaction time for correct and erroneous responses

In adolescents with a normal level of intelligence development, the effect of the type of probe presentation on the reaction time with the correct answer was found only at the tendency level ($F(2.29)=2.898$; $p=0.055$). A posteriori multiple comparisons according to the Scheffe criterion did not reveal the presence of priming effects. Only a tendency to increase the reaction time under the condition of AR was obtained in comparison with the condition of IR ($p=0.057$) (Table 2). With an erroneous answer, a statistically significant effect of the conditions on the reaction time was revealed ($F(2.29)=3.750$; $p=0.024$). There was a tendency to a positive priming effect (116 ms.) under the condition of AR (Scheffe criterion, $p=0.089$). Under the condition of IR, no priming effect was observed ($p>0.1$).

In adolescents with mental retardation, a statistically significant effect of the type of probe presentation on reaction time was found with the correct answer ($F(2.23)=8.605$; $p<0.001$) and with an erroneous answer ($F(2.23)=4.981$; $p=0.007$). There was a statistically significant increase in reaction time under the condition of AR in comparison with the condition of IR for both correct (Scheffe criterion, $p<0.001$) and erroneous answers (Scheffe criterion, $p=0.028$). A negative priming effect was revealed under the condition of AR with correct (–101 ms., Scheffe criterion: $p=0.007$) and erroneous answers (–76 ms., Scheffe criterion: $p=0.013$). A positive priming effect was also found with correct answers, expressed at the tendency level, under the condition of IR (25 ms., Scheffe criterion: $p=0.068$).

Table 2

Reaction time depending on the type of probe presentation with correct
and erroneous response

		Teenagers with a normal level of intelligence development (M ± SD)	Teenagers with mental retardation (M ± SD)
	AR	781±598ms.	981±880ms.
	IR	708±297ms.	855±620ms.
Correct answer	N	731±336ms.	880±625ms.
	AR	498±611ms.	1091±1143ms.
	IR	563±348ms.	1032±1224ms.
Error answer	N	614±404ms.	1015±1122ms.

Discussion of the results

The condition when the target stimulus were identical (AR) may have caused distrust of information in a group of adolescents with a normal level of development, due to which, with correct answers, the reaction took longer, and with erroneous ones — faster, i.e. in a number of trials, a decision was made in favor of an error due to orientation to the speed of response. It was also revealed the presence of a positive priming effect in the case of an erroneous response, provided that AR.

In adolescents with mental retardation, a negative priming effect was found, provided that both the correct answer and the erroneous one were corrected. Under the condition of IR, with the correct answer, there was a tendency to a positive priming effect. This result can also be justified by the presence of distrust of repetitive stimulus, since in the case of repetition of stimulus, on the contrary, a positive priming effect should have been observed. It can also be assumed that the occurrence of a negative priming effect is associated with the perception of the target probe as a new one, since the distractors of the prime and target probes were different, due to which the response could not be given automatically according to a known algorithm for responding to an identical probes. The detected tendency to a positive priming effect under the condition of the IR can be explained by the fact that the contradiction caused by the identity of the stimulus disappeared and the response was given faster. It can also

be assumed that the cognitive system used information about the distractor of the first probe (prime) when predicting subsequent target stimulus, due to which the response was given faster in the conditions of the expected negative priming effect [Norris, Kinoshita, 2008]. Erroneous answers in the conditions of the AR were given longer, most likely due to the presence of distrust.

The results obtained contradict the available data from the study of the relationship between reaction time and congruence of probes [Van der Borght, Braem, Notebaert, 2014]. It was noted that incongruence of stimulus causes an increase in reaction time with both correct and erroneous responses. Since the distractors of the probes in this study could be different when presenting identical target stimulus, the result obtained requires additional verification.

In general, the conditions and priming effects had a stronger effect on reaction time in the group of adolescents with mental retardation.

3. Categorical congruence of stimulus as a factor affecting reaction time in correct and erroneous responses

In adolescents with a normal level of intelligence development, a comparison of categorical conditions for presenting stimulus with the correct answer revealed a statistically significant increase in reaction time under the categorical condition of AR ($t(31)=5,063$; $p<0.001$) (Table 3). With an erroneous answer, no dependence of reaction time on categorical conditions was revealed ($t(31)= -0.108$; $p=0.914$).

Table 3

Reaction time depending on the categorical congruence of the presentation of probe stimulus with correct and erroneous response

		Teenagers with a normal level of intelligence development (M ± SD)	Teenagers with mental retardation (M ± SD)
Correct answer	AR	769±498ms.	942±779ms.
	IR	708±304ms.	856±605ms.
Error answer	AR	614±444ms.	1052±1153ms.
	IR	569±430ms.	1022±1157ms.

In adolescents with mental retardation, a comparison of categorical condi-

tions with correct (t(25)=1,032; p=0.303) and erroneous answers (t(25)=1,571; p=0.118) revealed no differences.

Discussion of the results

The categorical condition of AR with the correct answer increased the reaction time in a group of adolescents with a normal level of intelligence development. Perhaps this pattern was associated with the complexity of analyzing tasks under conditions of identity of target stimulus, due to distrust of repetition. In the group of adolescents with mental retardation, categorical conditions did not affect the reaction time either with correct answers or with erroneous ones.

Thus, it can be assumed that the process of solving tasks in adolescents with a normal level of development was more influenced by categorical conditions than the types of probe presentation, and in adolescents with mental retardation—the type of probe presentation and priming effects rather than categorical congruence of stimulus.

4. Transient responses and markers of unconscious detection of the correctness and inaccuracy of the response

In order to analyze the processes of making an erroneous and correct answer, the processes of transition from one answer to another with successive probes and markers of unconscious error detection when responding to identical stimulus were considered (Fig. 1–2).

4.1. Transient responses, post-error slowing, and post correct answer slowing

The analysis of the differences between the reaction time during the transition from the correct answer to the correct one and the reaction time during the transition from an erroneous answer to an erroneous one in the group of adolescents with a normal level of intelligence development did not reveal significant differences (t(31)= −1,600; p=0.113). In the group of adolescents with mental retardation, differences between these processes were found at the tendency level: when repeating the correct answer, there was a tendency to increase the reaction time (t(25)= −1.730; p=0.085). It was revealed that the change in the erroneous answer was higher than the repetition of the answer only in the group of adolescents with a normal level of intelligence development: the transition from an erroneous to an erroneous answer was shorter in reaction time than the transition from an erroneous to the correct

one (t(31)=3.874; p<0.001). Perhaps this is due to the fact that the repetition of errors appeared in the case of the respondents' strategy to respond faster, ignoring accuracy. In adolescents with mental retardation, there were no differences between the reaction time of repeating an erroneous response and not repeating it (t(25)=1,264; p=0.207). During the transition from the correct answer to the correct one, adolescents with a normal level of intelligence development showed an increase in reaction time relative to the conditions of non-repetition of the correct answer in the next probe (t(31)=4,224; p<0.001). In adolescents with mental retardation, the opposite effect was obtained (t(25)= −3.718; p<0.001), which is associated with a detected slowdown after correct answers, which is confirmed by a comparison of changes in answers: the transition from an erroneous to a correct answer and the transition from a correct to an erroneous answer (t(25)= −3.131; p=0.002). In adolescents with a normal level of intelligence development, an post-error slowing was found: the reaction time during the transition from an erroneous answer to the correct one was higher than during the transition from correct to erroneous (t(31)=8,333; p<0.001), including the transition from erroneous to correct was higher than during the transition from erroneous to erroneous.
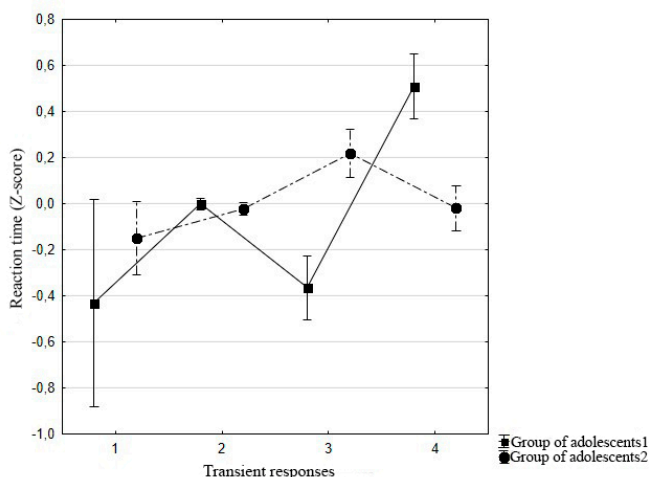
Fig. 1. Reaction time of transitional responses (M±CI) (1—transition from an erroneous answer to an erroneous one, 2—transition from a correct answer to a correct one, 3—transition from a correct answer to an erroneous one, 4—transition from an erroneous answer to a correct one) in a group of adolescents with a normal level of intelligence development (1) and with mental retardation (2).

Discussion of the results

In adolescents with a normal level of intelligence development, an post-error slowing was revealed. The obtained result correlates with the data obtained in a number of other studies on the phenomenon of post-error slowing [Notebaert et al., 2009; Forster, Cho, 2014; Perri et al., 2016, etc.]. Repeated erroneous answers were given faster than non-repeating ones, and the repetition of the correct answer in reaction time was higher than non-repetition, which may be confirmation of the assumption that errors arose as a result of the strategy to respond as quickly as possible.

In the group of adolescents with mental retardation, no post-error slowing was found. This tendency of responses may be due to the fact that adolescents with mental retardation have a higher process of adjustment and adaptation after a mistake. This result can be compared with data obtained in a number of studies on the effect of error awareness on an increase in post-error slowing [Wessel et al., 2018; Kirschner et al. 2020]. There was a slowing in the response after the correct decision: for the transition from the correct answer to the wrong one, adolescents with mental retardation needed more time to adapt than from the error to the correct answer. However, when repeating the correct answer, the reaction time was less than when not repeating, which may be due to the fact that the correct quick answer could be repeated if it was sure of correctness. In case of uncertainty or complexity of the task conditions, the correct answer was replaced by an erroneous one, which required more time.

4.2. Markers of unconscious detection of the correctness and inaccuracy answer

In both groups of adolescents, with a correct answer after an erroneous answer, the reaction time was higher than with an erroneous answer after a correct one: a slowdown in responses after errors was detected ($t(31)= –7.267$; $p<0.001$; $t(25)= –2.049$; $p=0.041$). Repeated correct answers were given longer than non-repeated ones ($t(31)=6,041$; $p<0.001$; $t(25)=2,234$; $p=0.026$).

In adolescents with a normal level of intelligence development, repeated correct answers were given statistically significantly longer than repeated erroneous ones ($t(31)=4,818$; $p<0.001$). There were no differences between these markers in adolescents with mental retardation ($t(25)=1.528$; $p=0.127$). With a correct answer after an erroneous answer, the reaction time was statistically significantly higher than when repeating an erroneous answer in a group of adolescents with a normal level of intelligence development ($t(31)= –5,884$; $p<0.001$). There were no differences between these markers in the

group of adolescents with mental retardation (t(25)= –1,121; p= –0,262).
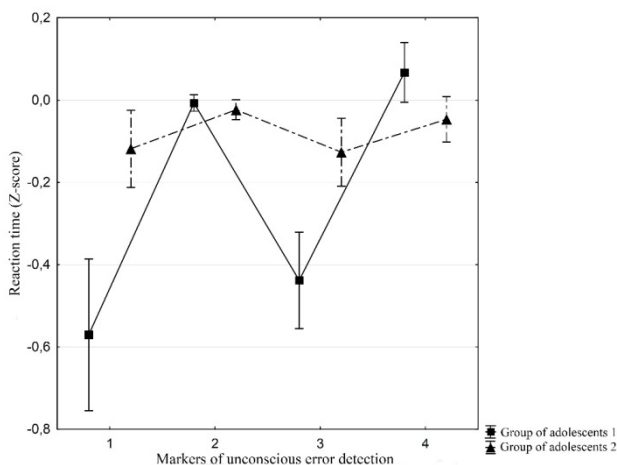
Fig. 2. Reaction time of markers of unconscious error detection to identical stimulus (M±CI) (1—repetition of the wrong answer, 2—repetition of the correct answer, 3—non-repetition of the wrong answer, 4—non-repetition of the correct answer) in a group of adolescents with a normal level of intelligence (1) and with mental retardation (2).

Discussion of the results

The study of markers of unconscious error detection with identical stimulus in adolescents with a normal level of intelligence showed similar results with the analysis of transient responses by respondents: repetition of an erroneous answer and an erroneous answer after correctly were given faster, presumably in connection with the strategy to respond as quickly as possible, and in order to respond to an identical stimulus correctly after an error, more time was required due to the process of correction and adaptation after an error. It is possible that this similarity of the results of the respondents' analysis of transitional responses and identical stimulus is a consequence of the work of the same mechanisms that participate in the decision-making process about the response.

In adolescents with mental retardation, when responding to identical stimulus, as well as with transient responses, there were no differences between the reaction time for repeated erroneous responses and non-repeating ones.

Perhaps, in this group of teenagers, erroneous answers in case of their repetition were given not because of the strategy to answer faster, but the strategy to answer as accurately as possible. The reaction time for repeated correct answers was higher than for non-repeated ones. In this case, erroneous answers were given "by chance". This result is the opposite of the result obtained in the study of transient responses. Also, an increase in reaction time was obtained in case of an erroneous response after the correct one, which may be the result of the adaptation process after the error. This result is also the opposite of the result of the study of transient responses, where there was no slowdown after an erroneous response. Perhaps the differences described above are due to the fact that different decision-making processes were involved in the analysis of transient responses and in the analysis of identical stimulus.

General discussion of the results and conclusions

Analysis of the data obtained in a group of adolescents with a normal level of intelligence development showed that the reaction time with the correct answer was longer than with an erroneous one, which may be due to the orientation of the answers more on speed than accuracy and the priority choice of erroneous "random" answers in case of complexity [Pleskac, Busemeyer, 2010; Rammsayer, Brandler, 2003, etc.].

The analysis of categorically identical target stimulus caused the greatest difficulty in adolescents with a normal level of intelligence development, increasing the reaction time for both erroneous and correct answers. The AR condition also caused complexity, but at the level of a tendency. Perhaps the repetition of target stimulus caused distrust, due to which the reaction time increased with correct answers and decreased with erroneous ones.

With transient responses, adolescents with a normal level of intelligence development experienced post-error slowing [Notebaert et al., 2009; Forster, Cho, 2014; Perri et al., 2016, etc.]. The assumption of error as a result of the strategy to respond as quickly as possible was confirmed by a decrease in reaction time with repeated erroneous responses and a decrease in reaction time in a situation where the correct answer was followed by an erroneous one. The study of the analysis of identical stimulus showed similar results with the process of analyzing transient responses.

The analysis of the data obtained in a group of adolescents with mental retardation showed a tendency to increase the reaction time for erroneous answers in comparison with the correct ones, which may be the reason for the strategy to answer as accurately as possible and the complexity of tasks due to

violations in the decision-making process about the answer.

The condition of repetition of stimulus in the group of adolescents with mental retardation, as well as in the group of adolescents with a normal level of development, caused distrust, due to which there was a negative priming effect. Under the condition of IR, when the contradiction caused by the repetition of stimulus disappeared, a positive priming effect was observed. It can also be assumed that the decrease in reaction time occurred due to the fact that the cognitive system used information about the prime when predicting subsequent target stimulus [Norris, Kinoshita, 2008]. The influence of individual categorical conditions in this group of adolescents was not found.

The analysis of transitional responses in the group of adolescents with mental retardation did not reveal a post-error slowing, which may be due to the fact that this group of adolescents has a higher process of correction and adaptation after an error due to less awareness of the perfect erroneous response [Wessel et al., 2018; Kirschner et al. 2020]. But it was found that the response was slowed down after the correct decision, as a consequence of rechecking the correctness. There was also a decrease in reaction time when repeating the correct answer as an indicator of confidence in the answer. It can be assumed that in some cases, if the respondent was sure of the answer, he gave it quickly and correctly, if not, then he needed to double-check the answer, due to which there was a slowdown after the correct answer.

The study of the analysis of identical stimulus in a group of adolescents with mental retardation and its comparison with the process of analyzing transient responses allowed us to make an assumption about the presence of various mechanisms that participate in the decision-making process about the response in the analysis of transient responses and the analysis of identical stimulus. As with transient responses, approximately the same reaction time was observed with an erroneous response, both after the correct and after the erroneous one. Different results were obtained when comparing repeated correct answers and non-repeating ones. In this case, erroneous answers were given "by chance". Also, an increase in reaction time was obtained with a correct response after an erroneous one, which may be the result of the adaptation process after an error to a previous identical stimulus.

Thus, due to comparative analysis of the processes of correct and erroneous responses, certain patterns of these processes and their dependence on various factors, such as the conditions for the presentation of probes, categorical congruence of stimulus, priming effects, were revealed. Changes were also found in the decision-making processes of the response depending on the

presence or absence of intellectual disabilities in respondents. The process of unconscious error detection was studied and certain patterns of this process were revealed, depending on the presence or absence of developmental disorders. The results obtained in the future will allow us to develop a model for predicting and correcting erroneous responses.

References

1. *Allakhverdov V. M.* Soznanie kak paradox [Consciousness as a paradox]. Saint Petersburg, DNK Publ., 2000. 528 p. (in Russian)
2. *Andriyanova N. V., Karpov A. D.* Rol' slozhnosti zadachi v vozniknovenii povtoryayushchikhsya oshibok [The role of task complexity in the occurrence of repeated errors]. Sed'maya mezhdunarodnaya konferentsiya po kognitivnoi nauke. Tezisy dokladov [The Seventh International Conference on Cognitive Science. Abstracts of reports], 2016, pp. 115–116. (in Russian)
3. *Allakhverdov V., Filippova M., Gershkovich V., Karpinskaia V., Scott T., Vladykina N.* Consciousness, learning, and control: on the path to a theory // Implicit learning: 50 years on. Routledge, 2019. C. 71–107.
4. B*echtereva, N. P., & Gretchin, V. B.* (1969). Physiological Foundations of Mental Activity. International Review of Neurobiology, 329–352.
5. *Beckmann, J. F.* (2000). Differentielle Latenzzeiteffekte bei der Bearbeitung von ReasoningItems. Diagnostica, 46, 124–129.
6. *Bernstein, P. S., Scheffers, M. K., & Coles, M. G. H.* (1995). "Where did I go wrong?" A psychophysiological analysis of error detection. Journal of Experimental Psychology: Human Perception and Performance, 21(6), 1312–1322.
7. *Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D.* (2001). Conflict monitoring and cognitive control. Psychological Review, 108, 624–652.
8. *Cleeremans A., Destrebecqz A.* Real rules are conscious //Behavioral and Brain Sciences.— 2005.—T. 28.— № . 1.—C. 19.
9. *D'Angelo M. C., Humphreys K. R.* Tip-of-the-tongue states reoccur because of implicit learning, but resolving them helps //Cognition.—2015.—T. 142.—C. 166–190.
10. *Forster S. E., Cho R. Y.* Context specificity of post-error and post-conflict cognitive control adjustments //PLoS One.— 2014.—T. 9.— № . 3.—C. e90281.
11. *Gehring, W. J., Coles, M. G. H., Meyer, D. E., & Donchin, E.* (1990). The error-related negativity: An event-related brain potential accompanying errors. Psychophysiology, 27, S34.
12. *Hsu H. M.*, Hsieh S. Age-related post-error slowing and stimulus repeti-

tion effect in motor inhibition during a stop-signal task //Psychological Research.— 2021.— C. 1–14.

13. *Janssen T. W.* P., *van Atteveldt N., Oosterlaan J.* Error and post-error processing in children with attention-deficit/hyperactivity disorder: An electrical neuroimaging study //Clinical Neurophysiology.— 2020.— T. 131.— № . 9.— C. 2236–2249.

14. *Keute M. et al.* Error-related dynamics of reaction time and frontal midline theta activity in attention deficit hyperactivity disorder (ADHD) during a subliminal motor priming task //Frontiers in human neuroscience.— 2019.— T. 13.— C. 381.

15. *Kirschner H. et al.* Neural and behavioral traces of error awareness //Cognitive, Affective, & Behavioral Neuroscience.— 2020.— C. 1–19.

16. *Kuvaldina M., Chetverikov A., Odainic A., Filippova M., Andriyanova N.* Implicit learning from one's mistakes: the negative choice aftereffect // Implicit learning: 50 years on. Routledge, 2019. C. 108–132.

17. Na¨a¨ta¨nen, R., 2003. Mismatch negativity: clinical research and possible applications. Int. J. Psychophysiol. 48, 179–188.

18. *Norris D., Kinoshita S.* Perception as evidence accumulation and Bayesian inference: Insights from masked priming //Journal of Experimental Psychology: General.— 2008.— T. 137.— № . 3.— C. 434.

19. *Notebaert W. et al.* Post-error slowing: an orienting account //Cognition.— 2009.— T. 111.— № . 2.— C. 275–279.

20. *Perri R. L.* et al. How the brain prevents a second error in a perceptual decision-making task //Scientific reports.— 2016.— T. 6.— № . 1.— C. 1–11.

21. *Pleskac T. J.*, *Busemeyer J. R.* Two-stage dynamic signal detection: a theory of choice, decision time, and confidence //Psychological review.— 2010.— T. 117.— № . 3.— C. 864.

22. *Rabbitt, P. M. A.* (1968). Three kinds of error-signaling responses in a serial choice task. Quarterly Journal of Experimental Psychology, 20, 179–188.

23. *Rammsayer, T. & Brandler, S.* (2003). Zum Zeitverhalten beim computergestützten adaptiven Testen: Antwortlatenzen bei richtigen und falschen Lösungen sind intelligenzunabhängig. Zeitschrift für Differentielle und Diagnostische Psychologie, 24, 57–63.

24. *Schrijvers D. et al.* Action monitoring in major depressive disorder with psychomotor retardation //Cortex.— 2008.— T. 44.— № . 5.— C. 569–579.

25. *Shephard E., Jackson G. M., Groom M. J.* The effects of co☐occurring ADHD symptoms on electrophysiological correlates of cognitive control in young people with Tourette syndrome //Journal of neuropsychology.—

2016.— T. 10.— № . 2.— C. 223–238.

26. *Sokhadze E. M.* et al. Comparative event-related potential study of performance in visual oddball task in children with autism spectrum disorder, ADHD, comorbid autism and ADHD, and neurotypical children //NeuroRegulation.— 2019.— T. 6.— № . 3.— C. 134–134.

27. *Van der Borght L., Braem S., Notebaert W.* Disentangling posterror and postconflict reduction of interference //Psychonomic bulletin & review.— 2014.— T. 21.— № . 6.— C. 1530–1536.

28. *Wessel, J. R.* (2018). An adaptive orienting theory of error processing. Psychophysiology, 55(3).

Chapter 13.
The Müller-Lyer illusion in CNN trained for 3D object height estimation

Anton N. Mamaev, Ivan A. Gorbunov

Saint-Petersburg University
Russia

The Müller-Lyer illusion is a classic optical illusion in human perception. However, the causes and underlying principles of the illusions are debated [2]. While some researchers attribute the phenomenon to depth cues, size constancy and spatial perception others explain it with weighted mean and receptive fields activation summation. Furthermore, the research mostly covers the observation of natural phenomena while artificial modelling and simulation of the process are not being conducted. We believe that comprehensive research in this new direction could be valuable for illusions research and propose convolutional neural networks as means to model various phenomena in human perception.

There is at least one published research paper in the scope of optical illusions computational modelling that covers the Müller-Lyer illusion [1]. We trained a feed-forward feature hierarchical model, HMAX, to perform a dual category line length judgment task (short versus long. However, that study also has several problems we attempt to solve in the current research.

Firstly, there was no spatial perception involved. All the stimuli were plain 2D lines and arrows, making it impossible to test if depth cues could be involved in illusion occurrence. Secondly, the neural network architecture used in the study was HMAX, a complicated convolutional neural network with variable filter resolutions and orientations made to represent the visual cortex as closely as possible. While its' strong resemblance to neural structures of the human brain is valuable, usage of a highly-specific and complex architecture makes it hard to both replicate the results and investigate the inner states of the model. Finally, the learning problem in the research was binary classification, enabling only two possible outcomes — the right one and the wrong one. That alone undermines the experimental results as the chance of a lucky guess is a vast 50 %. Moreover, output data of classification problems, in general, lack

the continuity necessary for post-hoc quantitative analysis.

In our study, we revisit convolutional models in the Müller-Lyer illusion research to find evidence whether the illusion is a consequence of spatial perception development as proposed by one of the interpretations. We have to assume that convolutional neural networks share enough similarities with the visual cortex to be a representative model at least for simple visual phenomena. We also have to state that a neural network is a 'blank slate' before any training and the outputs of a trained network depend on the training data and learning problem. Now, as perspective explanation attributes the illusion to the measurement of volumetric bodies, our approach to prove or disprove that was to train a height measurement neural network on a series of 3D imagery and then present it the Müller-Lyer illusion. In case the model estimations are consistently biased in accordance with original studies in humans, we have to further consider the possibility of depth perception development being a major cause of the Müller-Lyer illusion.

Methods

Dataset

While an initial research idea was to fit the model with indoor and outdoor photos of buildings with variable height and position, it was found to be ineffective as it is hard to determine the exact height of a building and the number of photos has to be tremendous. Instead, we came up with a dataset generator — an auxiliary application that renders series of images from a 3D environment built with Godot 3.3.3 open-source game engine. The 3D scene shown in Figure 1 consists of two cuboid meshes and a virtual camera facing the mutual edge of the meshes while being inside one of them. At any given moment one of the meshes is hidden so the picture taken by the camera is either the interior or exterior view.

To prevent overfitting and to ensure that the model performs well regardless of object and camera positions we introduced additional spatial parameters other than height. In total, there were four variables of the scene that are randomized in the given ranges for each picture taken:

Mesh height, also mutual edge length (2,4.2) — target value.

Camera distance (0,5).

Object angle (-10,10)

Camera angle (-25,25)
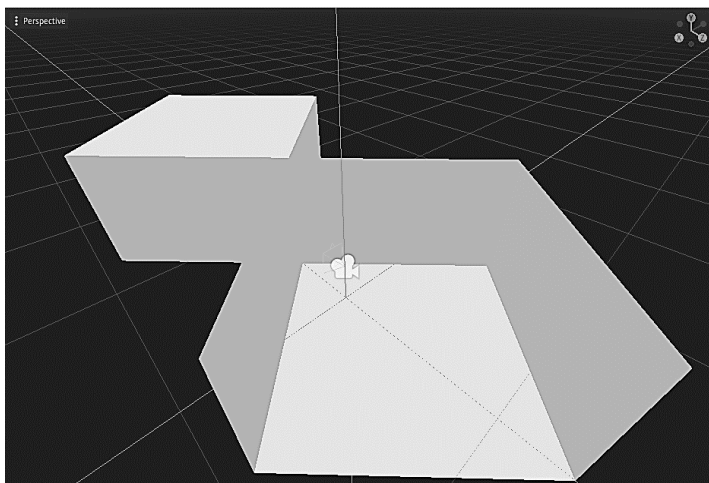
Fig. 1. 3D environment.

The randomization ranges were limited to ensure that the mutual edge is always in the camera's field of view and it is possible to estimate its' height.

All of the output images as shown in Figure 2 were limited to grayscale and 200×200px. resolution to enhance training speed. The filenames included the target value rounded up to sixth number after the decimal point exported directly from 3D scene height variable. We also created four special stimuli similar to the original Müller-Lyer illusion exclusively for testing purposes. They were strictly held out of training dataset to ensure their estimations are based on transfer learning only.
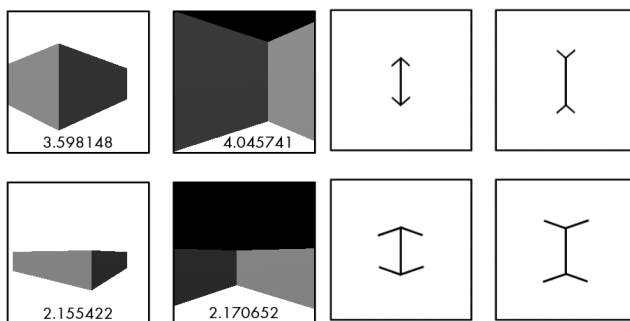
Fig. 2. Input data examples.

Overall, the final training dataset comprised 500 computer-generated images and a batch of 30 other images was held out to test estimation accuracy.

Neural network architecture

We used Keras [3] framework to compile and fit a feed-forward network consisting of ten layers in total:

Convolutional layer 1 ($200 \times 200 \times 1$ input, 32 3×3 filters)

Max Pooling layer 1 (2×2)

Convolutional layer 2 (64 3×3 filters)

Max Pooling layer 2 (2×2)

Convolutional layer 3 (128 3×3 filters)

Max pooling layer 3 (2×2)

Flatten (reshape 3D array to 1D vector)

Fully-connected layer 1 (1024 neurons)

Fully-connected layer 2 (512 neurons)

Fully-connected layer 3 (1 neuron)

A visual overview of the architecture is given in Figure 3. Array dimensions are scaled to fit and should not be referenced as spatially correct.
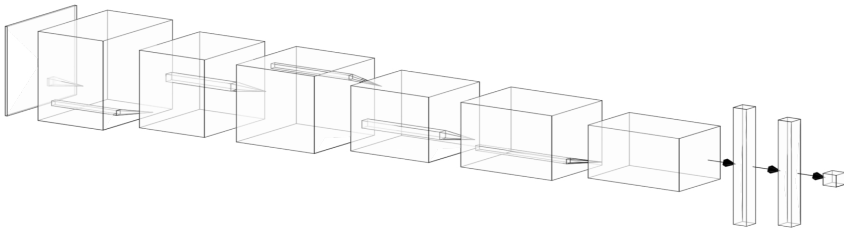
Fig. 3. Neural network overview.

The input layer has the exact shape of a 200×200 pixels grayscale image while the output converges to a single neuron. Most of the layers have ReLU as an activation function with an exception of the output neuron activated lineally. The latter enables the model to provide both positive and negative numeric estimations without being limited by maximum or minimum values.

As the learning problem is regression and the output is a single numeric value, the loss function is Mean Squared Error, as opposed to Cross-Entropy Error for classification problems.

Procedure

The overview of the preparatory and main stages of model training is shown in Figure 4.
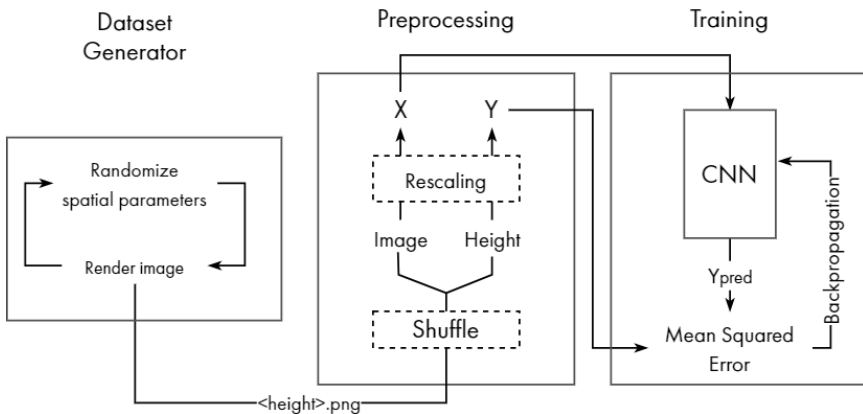
Fig. 4. Data pipeline.

First of all, the dataset generator synthesizes a designated number of images necessary for training, 500 for our final models. The spatial parameters get randomized, the picture gets taken and saved in a folder and the process loops until the set number of images is reached. All of the images are imported by a pre-processing Python script, get shuffled, the target value is extracted from the file's data and the image data is converted to a $200 \times 200 \times 1$ matrix, both get rescaled and appended to the Y and X array respectively. The Y array is composed of standardized height values, while the X array is composed of matrices with pixel values rescaled from (0,255) to (0,1).

The model then gets compiled and fitted for 30 epochs with a learning rate of 0.001. Items of array X get fed through the network and the output value of a single neuron in the last layer is compared with the target value from array Y. The loss is calculated with the mean squared error function and changes are made to the parameters of the model depending on the extent of the discrepancy through backpropagation. Around 30 epochs the loss gets as low as <0.001.

In this study, we had two separate testing phases. The first one tests the accuracy of estimations on a novel dataset unknown to the model while the second one tests whether the model can transfer its' training to a completely different kind of stimuli — the arrows of the Müller-Lyer illusion. The second phase begins only if the model performs with high accuracy at the first phase.

Results

Height estimation accuracy

The fitted model's height estimations for 30 random images not included in the training dataset were compared to the actual height values. The results illustrated in Figure 5 show high estimation accuracy with a mean deviation of 0.053114 and a correlation coefficient of 0.998854. Raw data is presented in Table 1.
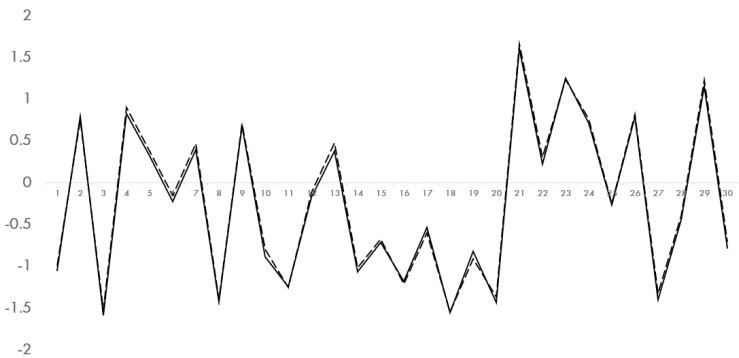
Fig. 5. Estimation accuracy.

Table 1

Accuracy raw data

-0.99725	0.75132	-1.52477	0.896639	0.380402	-0.15586	0.462907	-1.40121	0.704117	-0.79454	-1.25422	-0.12185	0.478146	-1.01683	-0.67781	-1.21133	-0.60048	-1.5506	-0.90883	-1.37262	1.655873	0.305259	1.234963	0.756578	-0.24821	0.819196	-1.3303	-0.3997	1.223734	-0.7062
-1.05337	0.794811	-1.58703	0.821801	0.320473	-0.2264	0.392117	-1.42639	0.675531	-0.8898	-1.24876	-0.17984	0.374568	-1.06528	-0.72024	-1.18067	-0.54084	-1.55434	-0.82615	-1.43319	1.597133	0.224779	1.2505	0.7133	-0.27217	0.784747	-1.39949	-0.44908	1.154376	-0.78613

The Müller-Lyer illusion effect

To verify that the Müller-Lyer illusion consistently affects estimations we fitted a total of 10 models with the same training dataset and recorded the outputs for two pairs of stimuli made to cause the Müller-Lyer optical illusion. As it is shown in Figure 6, all of the models have estimated the lines with inward-facing arrowheads to be slightly longer than with the outward-facing ones. The mean deviations for classic and wide stimuli pairs are 0.130740063 and 0.274046172 respectively. The deviation is higher by 0.143306109 in the second pair of stimuli with wider arrowheads that resemble an inside and outside edge of a 3D object used in training. Raw data is presented in Table 2.
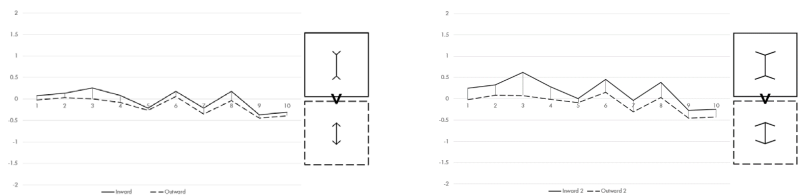
Fig. 6. Illusion stimuli deviation.

Table 2

Illusion stimuli raw data

Classic stimuli									
0.072521	-0.018522	0.072521	-0.018522	0.072521	-0.018522	0.072521	-0.018522	0.072521	-0.018522
0.131267	0.027986	0.131267	0.027986	0.131267	0.027986	0.131267	0.027986	0.131267	0.027986
Wide stimuli									
0.249548	-0.02488795	0.249548	-0.02488795	0.249548	-0.02488795	0.249548	-0.02488795	0.249548	-0.02488795
0.326664	0.08317863	0.326664	0.08317863	0.326664	0.08317863	0.326664	0.08317863	0.326664	0.08317863

Discussion

In the current research, we show that the Müller-Lyer illusion causes a convolutional neural network to provide biased length estimations. The effect is comparable to the effect in human perception. Considering that the model was fitted only to estimate the height of 3D objects, we attribute the occurrence of the Müller-Lyer illusion to it being caused by the particularity of spatial perception itself.

As the Müller-Lyer illusion consistently affects height estimation in convolutional models fitted with spatial imagery, the architecture of a neural network developed for length and height measurement in spatial perception is sufficient to cause the illusion in its classic form. The model fitted with the sole purpose of estimating the height of 3D is essentially a model of a visual subsystem in development, and the fact it is affected by stimuli such as the Müller-Lyer illusion bears evidence of the visual illusions and functions of the vision itself being linked.

While this suggests a strong connection between illusion and spatial perception, the findings do not necessarily conflict with explanations other than the one attributing the illusion to spatial perception. Though spatial vision is viewed as the essential cause of the phenomenon, the weighted mean and neural activation summation could probably be the exact neural processes that provide a foundation for spatial perception itself.

Conclusion

Our research has reached two main conclusions. First, it shows that the Müller-Lyer illusion causes a persistent length estimation bias in neural networks largely similar to the effect it causes in humans. As the effect was found in a model fitted from scratch solely for the task of 3D object height estimations, we suspect the illusion to be caused directly by the development of spatial perception.

Second, we have found out convolutional neural networks are capable of accurate estimations for complex spatial scenes and viable to use in modelling and simulation of human sensory and perceptive processes.

Further investigation is required to find out whether the effect exists for other variations of the illusion, compare the model's estimations with human performance and determine the internal parameters and processes of the neural network that contribute to the visual phenomena.

References

1. *Zeman A., Obst O., Brooks K. R., Rich A. N.* The Muller-Lyer illusion in a computation model of biological object recognition // PLoS ONE. 2013 Feb., 8(2): e56126. doi: 10.1371/journal. pone.0056126.
2. *Bermond B., Van Heerden J.* The Muler Lyer illusion explained and its theoretical importance reconsidered // Biol Philos. 1996 Jul., 11(3): 321–338. doi: 10.1007/BF00128785.
3. *Chollet F. and others* // Keras, 2015. http://github.com/fchollet/keras

Chapter 14.
Third parallel visual channel. Short review

Natalia Merkulyeva

Pavlov Institute of Physiology RAS
Saint Petersburg, Russia
mer-natalia@yandex.ru

Visual processing is based upon the integrated activity of various brain networks organized into multiple streams or channels. These channels originate from the retinal ganglion cells and can be traced up to the cerebral cortex. In animals with well-developed vision more studied channels are so-called Y, X, W channels — for Carnivores (Enroth-Cugell, Robson, 1966; Fukada, 1971; Stone, Hoffmann, 1972), or Magnocellular, Parvocellular and Koniocellular channels — for Primates (Livingstone, Hubel, 1987; Merigan, Maunsell, 1993). A homology between Y, X, W and Magnocellular, Parvocellular, and Koniocellular channels was proposed (Perry, Cowey, 1981, 1984). Retinal ganglion cells organizing these channels are different in their soma size, dendritic arborization, axon thickness, and those for Magno(Y) channel have large soma and dendritic arborization, and thick axons, at the same time, those for Parvo(X) channel have smaller soma and smaller dendritic arborization, and thin axons (Polyak, 1936; Boycott, Wässle, 1974; de Monasterio, Gouras, 1975; Leventhal, 1980). In accordance with the morphological peculiarities and response properties, these channels are devoted to the object vision — Parvo(X) channel, and to the spatial vision — Magno(Y) channel. As for third channel (W or Konio), both its structure and function are less systematized, and historically, in most well-known schemes for visual processing, less or no attention has been paid for the third channel.

What can we say about the third visual parallel channel? In Carnivores, retinal ganglion cells organizing the third visual channel comprise as much as 40–60 % of total population of retinal ganglion cells (Fukuda, Stone, 1974; Wilson et al., 1976; Mass, Supin, 1987; Wässle, 2004), in Primates — only 10 % (de Monasterio, Gouras, 1975; Schiller, Malpeli, 1977; Weller, Kaas, 1989) but it well correspondent to the amount of retinal ganglion cells for the Y Magno system (Perry, Cowey, 1984). The most striking feature of the third

channel cells is its heterogeneity: in Carnivores at least 15 types of cells was documented that can be source for the third channel (Kolb et al., 1981), and in Primates — at least 9 types (Polyak, 1936; Dacey, Packer, 2003; Dacey et al., 2003). Examples of these cells are γ-, δ-, ε-, ζ-, η-, θ-, ι-, κ-, λ- types — in Carnivores (Boycott, Wässle, 1974; Leventhal et al., 1980; Berson et al., 1998; Berson et al., 1999; Isayama et al., 1999; Isayama et al., 2000), and small bistratified cells — in Primates (Famiglietti, 1987; Dacey, 1993; Ghosh, Grünert, 1999; Silveira et al., 1999).

Despite the morphological heterogeneity, retinal ganglion cells of third type have several common features like small soma size, large dendritic arborization and very thin axons (Boycott, Wässle, 1974; Famiglietti, 1987; Dacey, 1993; Silveira et al., 1999; Isayama et al., 2000). They are characterized by so-called sluggish response to stimulation (having long latency period and low spike frequency); their axons have very low conduction velocity (Cleland, Levick, 1974; Fukuda, Stone, 1974; Dacey, 1993). One also characteristic feature is an input from the S-cones (Cleland, Levick, 1974; Guenther, Zrenner, 1993; Dacey, Lee, 1994; Rowe, Palmer, 1995; Ghosh, Grünert, 1999).

The projections of the third channel are widely spread across diencephalic and midbrain regions. Two main visual structures within visual thalamus are lateral geniculate nucleus (LGN) and pulvinar complex, and retinal ganglion cells of the third channel organize projections to both. At first, to the ventral most part of the dorsal LGN (LGNd) — in Cparvo layers, in Carnivores (Wilson et al., 1976; Stone et al., 1979; Itoh et al., 1982; Sur, Sherman, 1982), and to the Koniocellular (or S/I) layers — in Primates (Itoh et al., 1982; Lachica, Casagrande, 1993; Feig, Harting, 1994; Casagrande, 1994; Martin et al., 1997; Hendry, Reid, 2000; Szmajda et al., 2008; Baldwin, Krubitzer, 2018). As for pulvinar complex, third channel have projections to the lateral part of the LP-pulvinar complex in Carnivores (Leventhal et al., 1980; Mason, 1981; Itoh et al., 1982), and to inferior pulvinar nucleus — in Primates (Itaya, Van Hoesen, 1983; Nakagawa, Tanaka, 1984; Cowey et al., 1994; O'Brien et al., 2001; Warner et al., 2010). Other main target for the axons of third channel is superior colliculi receiving the 90 % of their fibers (Hoffmann, 1973; Cleland, Levick, 1974; Fukuda, Stone, 1974; Wässle, 1982; Berson, 1987; Chen et al., 1996). Other structures receiving input from third channel are ventral lateral geniculate nucleus (Hughes, Ater, 1977; Mize, Horner, 1984; Leventhal et al., 1985; Chen et al., 1996), pretectal area (Ballas et al., 1981; Perry, Cowey, 1984;

Koontz et al., 1985; Leventhal et al., 1985), nuclei optical acessorii (Farmer, Rodieck, 1982; Telkes et al., 2000), suprachiasmatic nuclei (Koontz et al., 1985; Chen et al., 1996; Murakami et al., 1989; Pu, 1999) etc. The cortical projections of third channel are spread within the primary visual cortex (in layers I–III) (Fitzpatrick et al., 1983; Weber et al., 1983; Diamond et al., 1985; Casagrande, 1994; Ding, Casagrande, 1997), in areas V3 (Yukie, Iwai, 1981; Beck, Kaas, 1998), V4 (Lysakowski et al., 1988; Sincich et al., 2004), medial temporal area (Rodman et al., 2001; Sincich et al., 2004) and inferior temporal areas (Hernández-González et al., 1994) — in Primates, and within the primary visual cortex (in layers I, III, and V) (LeVay, Gilbert, 1976; Leventhal, 1979; Kawano, 1998; Anderson et al., 2009), in areas 19 (Updyke, 1975; Hollander, Vanegas, 1977; Dreher et al., 1980; Kawano, 1998), 20, 21, and areas at the fundus of suprasylvian sulcus — PMLS, PLLS, VLS (Raczkowski, Rosenquist, 1980; Kawano, 1998; Boyd, Matsubara, 1999; Anderson et al., 2009) — in Carnivores.

And about third channel functions. First, this channel responsible for colour processing, since it has an input from S-cones and since it organizes a direct projection to the so-called cytochrome oxidase blobs responsible in particularly to the colour processing (Horton, 1984; Hubel, 1986; Tootell et al., 1988; Valverde, Salzmann et al., 2012; Rockoff et al., 2014).

Second, to the motion detection. At least 70 % of retinal ganglion cells of third type can detect motion (Rowe, Palmer, 1995), and moreover projections of these cells were obtained within cortical areas being the well-known center of motion detection: medial temporal area — in Primates (Stepniewska et al., 1999; Sincich et al., 2004; Nassi, Callaway, 2006; Mundinano et al., 2019), and area PMLS — in Carnivores (Raczkowski, Rosenquist, 1980; Kawano, 1998).

Third, to the oculomotor control. In particularly, a participation into the saccadic suppression (a loss of visual perception during saccadic eye movements) (Casagrande,1994; Xue et al., 1994; Anderson et al., 2009) — due to the tight communication between the superior colliculi, Cparvo(Konio) layers of the LGNd, and visual cortex. A communication between the superior colliculi, Cparvo(Konio) layers, and parabigeminal nucleus, and projections of the third channel to the nuclei optical acessorii and LGNv that are responsible for the stabilization of the retinal image during head movements (Spear et al., 1977; Farmer, Rodieck, 1982; Livingston, Fedder, 2003; Giolli et al., 2006) allows supposing additional responsibility of the third channel to the oculomotor control.

Fourth, to the circadian rhythmicity. Projections of the third retinal ganglion cells was shown to the suprachiasmatic nuclei and LGNv (Costa, Britto, 1997; Costa et al., 1998). And intrinsically photosensitive retinal ganglion cells synthesized a melanopsin being similar to the ganglion cells of W type (Semo et al., 2005) allows supposing a responsibility of the third channel to the circadian rhythmicity regulation.

Fifth, a participation into the blindsight — an ability to respond to visual stimuli in cortically blind persons (after the primary visual cortex damage (Weiskrantz et al., 1974; Denny-Brown, Chambers, 1976)). The candidate for this phenomenon are W/Konio cells in the LGNd (Rodman et al., 1990, 2001; Payne et al., 1996; Schmid et al., 2009, 2010). The importance and in some words power of the third channel was illustrated in investigations of Sur et colleagues. They destroyed in cats a primary visual cortex that led to the retrograde elimination of the A-layers of the lateral geniculate nuclei, and they also destroyed the superior colliculi and inferior colliculi. After surgery, retinal ganglion cells organized an input to the auditory brain regions: to the medial geniculate nucleus and to the primary auditory cortex (Sur et al., 1988). Moreover, it was revealed an ability of W cells to build an orientational cortical columns within primary auditory cortex after combinatory ablation of primary visual cortex and inferior colliculi (Sharma et al., 2000). It is mean that third channel can build a specific feature of the primary visual cortex within the primary auditory cortex even when the "main" visual channels are damaged.

At the same time, in developmental studies a longer period for small sized cells generation was obtained (Walsh et al., 1983; Reese et al., 1994). This fact would point out to the high plasticity of third channel. But at the same time, these cells are less affected by visual deprivation — see for example a high and mainly unchanged percent of W cells in Cparvo layer of cat's LGNd after monocular deprivation (Spear et al., 1989). This fact is for the more resistance of third channel comparing with Magno/Y and Parvo/X (Horton, Hocking, 1998). And more investigations are needed to elucidate the peculiarities of the third channel.

References

1. *Anderson J. C.*, da *Costa N. M.*, *Martin K. A.* The W cell pathway to cat primary visual cortex. J. Comp. Neurol. 2009. 516: 20–35.

2. *Baldwin M. K.L., Krubitzer L.* Architectonic characteristics of the visual thalamus and superior colliculus in titi monkeys. J. Comp. Neurol. 2018. 526: 1760–1776.

3. *Ballas I., Hoffmann K. P., Wagner H. J.* Retinal projection to the nucleus of the optic tract in the cat as revealed by retrograde transport of horse-radish peroxidase. Neurosci. Lett. 1981. 26: 197–202.

4. *Beck P. D.*, *Kaas J. H.* Thalamic connections of the dorsomedial visual area in primates. J. Comp. Neurol. 1998. 396: 381–398.

5. *Berson D. M.*, I*sayama T., Pu M.* The Eta ganglion cell type of cat retina. J. Comp. Neurol. 1999. 408: 204–219.

6. *Berson D. M., Pu M., Famiglietti E. V.* The zeta cell: a new ganglion cell type in cat retina. J. Comp. Neurol. 1998. 399: 269–288.

7. *Boycott B. B.*, *Wässle H.* The morphological types of ganglion cells of the domestic cat's retina. J. Physiol. 1974. 240: 397–419.

8. *Boyd J. D.*, *Matsubara J. A.* Projections from V1 to lateral suprasylvian cortex: an efferent pathway in the cat's visual cortex that originates preferentially from CO blob columns. Vis. Neurosci. 1999. 16: 849–860.

9. *Casagrande V. A.* A third parallel visual pathway to primate area V1. TINS. 1994. 17: 305–310.Hendry S.H.C., *Reid R. C.* The koniocellular pathway in primate vision. Annu. Rev. Neurosci. 2000. 23: 127–153.

10. *Chen B., Hu X. J.*, *Pourcho R. G.* Morphological diversity in terminals of W-type retinal ganglion cells at projection sites in cat brain. Vis. Neurosci. 1996. 13: 449–460.

11. *Cleland B. G.*, *Levick W. R.* Properties of rarely encountered types of ganglion cells in the cat's retina and an overall classification. J. Physiol. 1974. 240: 457–492.

12. *Costa M. S.*, *Britto L. R.* Calbindin immunoreactivity delineates the circadian visual centers of the brain of the common marmoset (Callithrix jacchus). Brain Res. Bull. 1997. 43: 369–373.

13. *Costa M. S.*M.O., *Moreira L. F.*, A*lones V., Lu J., Santee U. R., Cavalcante J. S., Moraes P. R.A., Britto L. R.G., Menaker M.* Characterization of the circadian system in a brazilian species of monkey (Callithrix jacchus): immunohistochemical analysis and retinal projections. Biol. Rhythm Res. 1998. 29: 510–520.

14. *Cowey A., Stoerig P., Bannister M.* Retinal ganglion cells labelled from the pulvinar nucleus in macaque monkeys. Neuroscience. 1994. 61: 691–705.

15. *Dacey D. M.* Morphology of a small-field bistratified ganglion cell type in

the macaque and human retina. Vis. Neurosci. 1993. 10: 1081–1098.

16. *Dacey D. M.*, *Lee B. B.* The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type. Nature. 1994. 367: 731–735.

17. *Dacey D. M.*, *Packer O. S.* Colour coding in the primate retina: diverse cell types and cone-specific circuitry. Curr. Opin. Neurobiol. 2003. 13: 421–427.

18. *Dacey D. M.*, *Peterson B. B.*, *Robinson F. R.*, *Gamlin P. D.* Fireworks in the primate retina: in vitro photodynamics reveals diverse LGN-projecting ganglion cell types. Neuron. 2003. 37: 15–27.

19. de *Monasterio F. M., Gouras P.* Functional properties of ganglion cells of the rhesus monkey retina. J. Physiol. 1975. 251: 167–195.

20. *Denny-Brown D., Chambers R. A.* Physiological aspects of visual perception. I. Functional aspects of visual cortex Arch Neurol. 1976. 33: 219–227.

21. *Diamond I. T.*, Conley M., *Itoh K., Fitzpatrick D.* Laminar organization of geniculocortical projections in Galago senegalensis and Aotus trivirgatus. J. Comp. Neurol. 1985. 242: 584–610.

22. *Ding Y., Casagrande V. A.* The distribution and morphology of LGN K pathway axons within the layers and CO blobs of owl monkey V1. Vis. Neurosci. 1997. 14: 691–704.

23. *Dreher B., Leventhal A. G., Hale P. T.* Geniculate input to cat visual cortex: a comparison of area 19 with areas 17 and 18. J. Neurophysiol. 1980. 44: 804–826.

24. *Enroth-Cugell C., Robson J. G.* The contrast sensitivityof retinal ganglion cells of the cat. J. Physiol. 1966. 187: 517–552.

25. *Famiglietti E. V.* Starburst amacrine cells in cat retina are associated with bistratified, presumed directionally selective, ganglion cells. Brain Res. 1987. 413: 404–408.

26. *Farmer S. G.*, *Rodieck R. W.* Ganglion cells of the cat accessory optic system: morphology and retinal topography. J. Comp. Neurol. 1982. 205: 190–198.

27. Feig S., *Harting J. K.* Ultrastructural studies of the primate lateral geniculate nucleus: morphology and spatial relationships of axon terminals arising from the retina, visual cortex (area 17), superior colliculus, parabigeminal nucleus, and pretectum of Galago crassicaudatus. J. Comp. Neurol. 1994. 343: 17–34.

28. *Fitzpatrick D., Itoh K., Diamond I. T.* The laminar organization of the lateral geniculate body and the striate cortex in the squirrel monkey (Saimiri sciureus). J. Neurosci. 1983. 3: 673–702.

29. *Fukada Y.* Receptive field organization of cat optic nerve fibers with spe-

cial reference to conduction velocity. Vision Res. 1971. 11: 209–226.

30. *Fukuda Y., Stone J.* Retinal distribution and central projections of Y, X, and W-cells of the cat's retina. J. Neurophysiol. 1974. 37: 749–772.

31. *Ghosh K. K.*, *Grünert U.* Synaptic input to small bistratified (blue-ON) ganglion cells in the retina of a new world monkey, the marmoset Callithrix jacchus. J. Comp. Neurol. 1999. 413: 417–428.

32. *Giolli R. A., Blanks R. H., Lui F.* The accessory optic system: basic organization with an update on connectivity, neurochemistry, and function. Prog. Brain Res. 2006. 151: 407–440.

33. *Guenther E., Zrenner E.* The spectral sensitivity of dark- and light-adapted cat retinal ganglion cells. J. Neurosci. 1993. 13: 1543–1550.

34. *Hernández-González A., Cavada C., Reinoso-Suárez F.* The lateral geniculate nucleus projects to the inferior temporal cortex in the macaque monkey. Neuroreport. 1994. 5: 2693–2696.

35. *Hoffmann K-P.* Conduction velocity in pathways from retina to superior colliculus in the cat: a correlation with receptive field properties. J. Neurophysiol. 1973. 36: 409–424.

36. *Holländer H., Vanegas H.* The projection from the lateral geniculate nucleus onto the visual cortex in the cat. A quantitative study with horseradish-peroxidase J. Comp. Neurol. 1977. 173: 519–536.

37. *Horton J. C.* Cytochrome oxidase patches: a new cytoarchitectonic feature of monkey visual cortex. Philos. Trans. R Soc. Lond. B Biol. Sci. 1984. 304: 199–253.

38. *Horton J. C.*, *Hocking D. R.* Monocular core zones and binocular border strips in primate striate cortex revealed by the contrasting effects of enucleation, eyelid suture, and retinal laser lesions on cytochrome oxidase activity J. Neurosci. 1998. 18: 5433–5455.

39. *Hubel D. H.* Blobs and color vision. Cell. Biophys. 1986. 9: 91–102.

40. *Hughes C. P.*, *Ater S. B.* Receptive field properties in the ventral lateral geniculate nucleus of the cat. Brain Res. 1977. 132: 163–166.

41. *Isayama T., Berson D. M., Pu M.* Theta ganglion cell type of cat retina. J. Comp. Neurol. 2000. 417: 32–48.

42. *Isayama T., O'Brien B.J., Ugalde I., Muller J.F., Frenz A., Aurora V., Tsiaras W., Berson D. M.* Morphology of retinal ganglion cells in the ferret (Mustela putorius furo) J. Comp. Neurol. 2009. 517: 459–480.

43. *Itaya S. K., Van Hoesen G. W.* Retinal projections to the inferior and medial pulvinar nuclei in the Old-World monkey. Brain Res. 1983. 269: 223–230.

44. Itoh K., Conley M., *Diamond I. T.* Retinal ganglion cell projections to individual layers of the lateral geniculate body in Galago crassicaudatus. J. Comp. Neurol. 1982. 205: 282–290.

45. *Kawano J.* Cortical projections of the parvocellular laminae C of the dorsal lateral geniculate nucleus in the cat: an anterograde wheat germ agglutinin conjugated to horseradish peroxidase study. J. Comp. Neurol. 1998. 392: 439–457.
46. *Koontz M. A.*, *Rodieck R. W.*, *Farmer S. G.* The retinal projection to the cat pretectum. J. Comp. Neurol. 1985. 236: 42–59.
47. *Lachica E. A.*, *Casagrande V. A.* The morphology of collicular and retinal axons ending on small relay (W-like) cells of the primate lateral geniculate nucleus. Vis Neurosci. 1993. 10: 403–418.
48. *LeVay S., Gilbert C. D.* Laminar patterns of geniculocortical projection in the cat. Brain Res. 1976. 113: 1–19.
49. *Leventhal A. G.* Evidence that the different classes of relay cells of the cat's lateral geniculate nucleus terminate in different layers of the striate cortex. Exp. Brain Res. 1979. 37: 349–372.
50. *Leventhal A. G.*, Keens J., Törk I. The afferent ganglion cells and cortical projections of the retinal recipient zone (RRZ) of the cat's pulvinar complex. J. Comp. Neurol. 1980. 194: 535–554.
51. *Leventhal A. G.*, *Rodieck R. W.*, Dreher B. Central projections of cat retinal ganglion cells. J. Comp. Neurol. 1985. 237: 216–226.
52. *Livingston C. A.*, *Fedder S. R.* Visual-ocular motor activity in the macaque pregeniculate complex. J. Neurophysiol. 2003. 90: 226–244.
53. *Livingstone M. S.*, *Hubel D. H.* Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. J. Neurosci. 1987. 7: 3416–3466.
54. *Lysakowski A., Standage G. P.*, *Benevento L. A.* An investigation of collateral projections of the dorsal lateral geniculate nucleus and other subcortical structures to cortical areas V1 and V4 in the macaque monkey: a double label retrograde tracer study. Exp. Brain Res. 1988. 69: 651–661.
55. *Martin P. R.*, *White A. J.*, *Goodchild A. K.*, *Wilder H. D.*, *Sefton A. E.* Evidence that blue-on cells are part of the third geniculocortical pathway in primates. Eur. J. Neurosci. 1997. 9: 1536–1541.
56. *Mason R.* Differential responsiveness of cells in the visual zones of the cat's LP-pulvinar complex to visual stimuli. Exp. Brain Res. 1981. 43: 25–33.
57. *Mass A. M.*, *Supin A. Y.* Ganglion cells density and retinal resolution in the sea otter, Enhydra lutris. Brain Behav. Evol. 2000. 55: 111–119.
58. *Merigan W. H.*, *Maunsell J. H.* How parallel are the primate visual pathways? Annu. Rev. Neurosci. 1993. 16: 369–402.
59. *Mize R. R.*, *Horner L. H.* Retinal synapses of the cat medial interlaminar nucleus and ventral lateral geniculate nucleus differ in size and synaptic organization. J. Comp. Neurol. 1984. 224: 579–590.

60. *Mundinano I.-C.*, *Kwan W.C.*, *Bourne J.A.* Retinotopic specializations of cortical and thalamic inputs to area MT. Proc. Natl. Acad. Sci. USA. 2019. 116: 23326–23331.
61. *Murakami D.M.*, *Miller J.D.*, *Fuller C.A.* The retinohypothalamic tract in the cat: retinal ganglion cell morphology and pattern of projection. Brain Res. 1989. 482: 283–296.
62. *Nakagawa S., Tanaka S.* Retinal projections to the pulvinar nucleus of the macaque monkey: a re-investigation using autoradiography. Exp. Brain Res. 1984. 57: 151–157.
63. *Nassi J.J.*, *Callaway E.M.* Multiple circuits relaying rimate parallel visual pathways to the middle temporal area. J. Neurosci. 2006. 26: 12789–12798.
64. *O'Brien B.J., Abel P.L., Olavarria J.F.* The retinal input to calbindin-D28k-defined subdivisions in macaque inferior pulvinar. Neurosci. Lett. 2001. 312: 145–148.
65. *Payne B.R.*, *Lomber S.G.*, *Macneil M.A.*, Cornwell P. Evidence for greater sight in blindsight following damage of primary visual cortex early in life Neuropsychologia. 1996. 34: 741–774.
66. *Perry V.H.*, Cowey A. Retinal ganglion cells that project to the superior colliculus and pretectum in the macaque monkey. Neurosci. 1984. 12: 1125–1137.
67. *Perry V.H.*, *Cowey A.* The morphological correlates of X- and Y-like retinal ganglion cells in the retina of monkeys. Exp. Brain Res. 1981. 43: 226–228.
68. *Polyak S.* Minute structure of the retina in monkeys and in apes. Arch. Ophthalmol. 1936. 15: 477–519.
69. *Pu M.* Dendritic morphology of cat retinal ganglion cells projecting to suprachiasmatic nucleus. J. Comp. Neurol. 1999. 414: 267–274.
70. *Raczkowski D., Rosenquist A.C.* Connections of the parvocellular C laminae of the dorsal lateral geniculate nucleus with the visual cortex in the cat Brain Res. 1980. 199: 447–451.
71. *Reese B.E.*, *Thompson W.F.*, *Peduzzi J.D.* Birthdates of neurons in the retinal ganglion cell layer of the ferret. J. Comp. Neurol. 1994. 341: 464–475.
72. *Rockoff E., Balaram P., Kaas J.* Patchy distributions of myelin and vesicular glutamate transporter 2 align with cytochrome oxidase blobs and interblobs in the superficial layers of the primary visual cortex. Eye Brain. 2014. 6 (Suppl 1): 19–27.
73. *Rodman H.R.*, *Gross C.G.*, *Albright T.D.* Afferent basis of visual response roperties in area MT of the macaque: II. Effects of superior colliculus removal. J. Neurosci. 1990. 10: 1154–1164.
74. *Rodman H.R.*, *Sorenson K.M.*, *Shim A.J.*, *Hexter D.P.* Calbindin immu-

noreactivity in the geniculo-extrastriate system of the macaque: implications for heterogeneity in the koniocellular pathway and recovery from cortical damage J. Comp. Neurol. 2001. 431: 168–181.

75. *Rowe M. H.*, *Palmer L. A.* Spatio-temporal receptivefield structure of phasic W cells in the cat retina. Vis. Neurosci. 1995. 12: 117–139.

76. *Schiller P. H.*, *Malpeli J. G.* Properties and tectal projections of monkey retinal ganglion cells. J. Neurophysiol. 1977. 40: 428–445.

77. *Schmid M. C.*, *Mrowka S. W.*, *Turchi J.*, *Saunders R. C.*, *Wilke* M., *Peters A. J.*, *Ye F. Q.*, *Leopold D. A.* Blindsight depends on the lateral geniculate nucleus. Nature. 2010. 466: 373–377.

78. *Schmid M. C.*, Panagiotaropoulos T., *Augath M. A.*, *Logothetis N. K.*, *Smirnakis S. M.* Visually driven activation in macaque areas V2 and V3 without input from the primary visual cortex. PLoS One. 2009. 4: e5527.

79. *Semo M., Llamosas M. M.*, *Foster R. G.*, Jeffery G. Melanopsin (Opn4) positive cells in the cat retina are randomly distributed across the ganglion cell layer. Vis. Neurosci. 2005. 22: 111–116.

80. *Sharma J., Angelucci A., Sur M.* Induction of visual orientation modules in auditory cortex. Nature. 2000. 404: 841–847.

81. *Sincich L. C.*, *Park K. F.*, *Wohlgemuth M. J.*, *Horton J. C.* Bypassing V1: a direct geniculate input to area MT. Nat. Neurosci. 2004. 7: 1123–1128.

82. *Silveira L. C.*, *Lee B. B.*, *Yamada E. S.*, *Kremers J.*, *Hunt D. M.*, *Martin P. R.*, *Gomes F. L.* Ganglion cells of a short-wavelength-sensitive cone pathway in New World monkeys: morphology and physiology. Vis. Neurosci. 1999. 16: 333–43.

83. *Spear P. D.*, *Smith D. C.*, *Williams L. L.* Visual receptive-field properties of single neurons in cat's ventral lateral geniculate nucleus. J. Neurophysiol. 1977. 40: 390–409.

84. *Stepniewska I., Qi H. X.*, *Kaas J. H.* Projections of the superior colliculus to subdivisions of the inferior pulvinar in New World and Old World monkeys. Vis. Neurosci. 2000. 17: 529–549.

85. *Stone J., Dreher B., Leventhal A.* Hierarchical and parallel mechanisms in the organization of visual cortex. Brain Res. 1979. 180: 345–394.

86. *Stone J., Hoffmann K. P.* Very slow-conducting ganglion cells in the cat's retina: a major, new functional type? Brain Res. 1972. 43: 610–616.

87. *Sur M., Sherman S. M.* Linear and nonlinear W-cells in C-laminae of the cat's lateral geniculate nucleus. J. Neurophysiol. 1982. 47: 869–884.

88. *Sur M., Garraghty P. E.*, *Roe A. W.* Experimentally induced visual projections into auditory thalamus and cortex. Science. 1988. 242: 1437–1441.

89. *Szmajda B. A.*, Grünert U., *Martin P. R.* Retinal ganglion cell inputs to the koniocellular pathway. J. Comp. Neurol. 2008. 510: 251–268.

90. *Telkes I., Distler C., Hoffmann K. P.* Retinal ganglion cells projecting to

the nucleus of the optic tract and the dorsal terminal nucleus of the accessory optic system in macaque monkeys. Eur. J. Neurosci. 2000. 12: 2367–2375.

91. *Tootell R. B.*, *Silverman M. S.*, *Hamilton S. L.*, *De Valois R. L.*, *Switkes E.* Functional anatomy of macaque striate cortex. III. Color. J. Neurosci. 1988. 8: 1569–1593.

92. *Updyke B. V.* The patterns of projection of cortical areas 17, 18, and 19 onto the laminae of the dorsal lateral geniculate nucleus in the cat. J. Comp. Neurol. 1975. 163: 377–395.

93. *Valverde Salzmann M. F., Bartels A., Logothetis N. K., Schüz A.* Color blobs in cortical areas V1 and V2 of the new world monkey Callithrix jacchus, revealed by non-differential optical imaging. J. Neurosci. 2012. 32: 7881–7894.

94. *Walsh C., Polley E. H.*, *Hickey T. L.*, *Guillery R. W.* Generation of cat retinal ganglion cells in relation to central pathways. Nature. 1983. 302: 611–614.

95. *Warner C. E.*, Goldshmit Y., *Bourne J. A.* Retinal afferents synapse with relay cells targeting the middle temporal area in the pulvinar and lateral geniculate nuclei. Front. Neuroanat. 2010. 12: 8.

96. *Wässle H.* Parallel processing in the mammalian retina. Nat. Rev. Neurosci. 2004. 5: 747–757.

97. *Wässle H.* Chapter 4 Morphological types and central projections of ganglion cells in the cat retina. Prog. Retin. Res. 1982. 1: 125–152.

98. *Weber J. T.*, *Huerta M. F.*, *Kaas J. H.*, *Harting J. K.* The projections of the lateral geniculate nucleus of the squirrel monkey: studies of the interlaminar zones and the S layers. J. Comp. Neurol. 1983. 213: 135–145.

99. *Weiskrantz L., Warrington E. K., Sanders M. D., Marshall J.* Visual capacity in the hemianopic field following a restricted occipital ablation Brain. 1974. 97: 709–728.

100. *Weller R. E.*, *Kaas J. H.* Parameters affecting the loss of ganglion cells of the retina following ablations of striate cortex in primates. Vis. Neurosci. 1989. 3: 327–349.

101. *Wilson P. D.*, *Rowe M. H., Stone J.* Properties of relay cells in cat's lateral geniculate nucleus: a comparison of W-cells with X and Y-cells. J. Neurophysiol. 1976. 39: 1193–1209.

102. *Xue J. T.*, *Kim C. B.*, *Moore R. J.*, *Spear P. D.* Influence of the superior colliculus on responses of lateral geniculate neurons in the cat. Vis. Neurosci. 1994. 11: 1059–1076.

103. *Yukie M., Iwai E.* Direct projection from the dorsal lateral geniculate nucleus to the prestriate cortex in macaque monkeys. J. Comp. Neurol. 1981. 201: 81–97.

Chapter 15.
Decision-making under contradiction of beliefs and facts: an influence of individual peculiarities of the decision-maker's neural network

Alexander V. Pavlov, Artur O. Gaugel

ITMO University
St. Petersburg, Russia.
E-mail: avpavlov@itmo.ru

The paper is devoted to the role that individual peculiarities of neural system play in decision-making. We consider the situation of contradiction between the beliefs of the decision-maker and the facts that need to be taken into account. To model the situation, we use the "Prisoner's Dilemma" antagonistic game and depict it mathematically by a logic with exclusion. The logic of decision is implemented by tree-layer NN of resonance architecture, which corresponds to the neuro-physiologically motivated concept of the sensation circle by A. M. Ivanitskii. This model of NN allows the phenomenon of cognitive dissonance, that is an attribute of the decision- making under contradiction of preliminary formed beliefs and new facts to be modeled. The interconnection weights are formed and stored in the Fourier-space and we model the individual peculiarities of a decision-maker by the filtration of inputting stimuli at the matrices of the interconnection weights. We study the probability of decision to betray under the two conditions: the information that the second player decides to silence, and his decision is unknown. We study the probability in dependence on the length of correlation of the reference pattern that was used to learn the NN, and its filtration by the matrices. It's demonstrated analytically and confirmed by numerical modeling the range in which probability of decision to betray under the information on the second player's silence, which contradicts to the decision-maker biases, depends on the relation between general and partial features in the inner representation of the world—the "range of betrayal" is increased by decreasing of the general features weight.

A number of classical decision-making systems have been developed for the mathematical description of the decision-making process and its calculation as a choice of alternatives. The approaches were based on the principles

of rationality, however sociological experiments, first of all, studies of cognitive psychologists A. Tversky and D. Kahneman [1–5], clearly revealed that in fact, a human's choice is far from rational—real decisions, first of all—in uncertain and contradictory situations—are very often quite irrational. In this context the term "rational" is understood in the sense of corresponding to such mathematical classical theories as Boolean logic and Kolmogorovian probability, including Savage's Sure Thing Principle [6].

These inconsistencies between classical theories and experiments have stimulated new research into decision-making in real situations [7–11], and a number of hypotheses and approaches have been proposed for wide discussion, including the rather old idea of a "quantum brain". Last years the latter concept was developed by a number of research groups that have achieved good agreement between the results of numerical simulations and experimental data [12–22]. However, this concept has neither physical nor physiological experimental confirmations [23]; today there are analogies only at the level of mathematical description, and from the point of view of reasoning theory, this concept is an abductive hypothesis.

In [24–26], classical neural network mechanisms were shown that causes "violations" of the classical theories and does not require reference to quantum physics either at the level of mathematical formalism or at the level of fundamental physical phenomena. Modeling of the neural network mechanism for decision-making under such real circumstances as inconsistency and uncertainty [26] coincided with the results of sociological experiments [7–12].

At present, the problem of explainable artificial intelligence (XAI) is in the focus of attention, primarily in those areas of AI applications where the cost of error is unacceptably high [27]. The main question for AI within XAI is "Why was such a decision made?" [28]. In this framework, it is practically important that under the same conditions and within the same logic of decision-making, different decision-makers make different, sometimes diametrically opposite decisions. Clearly, individual mental peculiarities of the decision-maker play a major role here [29]. Thus, the explanation should take into account such mental features of the decision-maker as beliefs, biases and prejudices. A possible way to include the mental features in the model is to correlate them with the individual properties of the neural network, which is the material carrier of the decision-making logic [30].

One of the factors influencing the neural network response is the filtering of input stimuli on sensors, sensory tracts and interconnection weight matri-

ces. The latter is due to the nonlinearity of the sensitivity of synaptic contacts, which has a pronounced individual character. This factor plays a key role in the learning of a neural network, i.e. in the formation of the rule of individual decision-making logic. In this article, we investigate the influence of this factor—filtering on the interconnection weight matrices—on decision-making as a choice of alternatives under contradiction of perceived information to the previously formed beliefs of the decision-maker.

Description of the problem

Traditionally, decision-making is studied by the model situations described by the game theory. In particular, to model decision-making under such real circumstances as inconsistency and uncertainty, non-cooperative games, e.g. the Prisoner's Dilemma, are used [31]. The Prisoner's Dilemma game simulates a situation where two criminals A and B are caught and interrogated by the police separately, each has a choice: C (cooperate)—to remain silent or D (defect)—to betray. C and D make up a complete set of alternatives. The player (decision-maker) cannot influence the decision of the accomplice (second player).

Wins/losses (payoff matrix) depend on the situation:

$C^A C^B$—both are silent and receive short prison term;

$C^A D^B$—the betrayer B gets freedom, and the silent A gets the maximum term;

$D^A C^B$—by analogy with the previous one;

$D^A D^B$—both players betray each other, and both receive the average between CC and CD (DC) prison terms.

The CC variant is Pareto-optimal (maximum of total benefit), but in reality [7–12] as a rule, the Nash equilibrium is reached [31] DD—both betray.

In the experiments [7–12], the probability of the decision of the player B to betray was evaluated under the conditions:

a) the decision of player A was unknown to him;

b) he was informed that player A has made the decision C^A;

c) he was informed that player A has made the decision C^A.

Statistical processing of the results gave the order of probabilities, the same for all experiments:

$$\Pr\left(D^{\mathrm{B}}\middle|D^{\mathrm{A}}\right) > \Pr\left(D^{\mathrm{B}}\middle|C^{\mathrm{A}}\right) > \Pr\left(D^{\mathrm{B}}\right). \tag{1}$$

This order contradicts the classical law of total probability for the disjunction of incompatible events, according to which the order was expected

$$\Pr\left(D^{\mathrm{B}}\middle|D^{\mathrm{A}}\right) > \Pr\left(D^{\mathrm{B}}\right) > \Pr\left(D^{\mathrm{B}}\middle|C^{\mathrm{A}}\right). \tag{2}$$

An approach to the problem

In the articles [7–12], as well as a number of others devoted to the problems of choice modeled on the example of the Prisoner's Dilemma game, the following approach was used: conditional probabilities $\Pr(D^{\mathrm{B}}|C^{\mathrm{A}})$ and $\Pr(D^{\mathrm{B}}|D^{\mathrm{A}})$ were taken as facts, on their basis, according to the classical formula of total probability $\Pr(D^{\mathrm{B}})$, the unconditional probability was calculated, compared with experimentally obtained values and possible mechanisms for the latter's exit from the range of conditional probabilities $[\Pr(D^{\mathrm{B}}|C^{\mathrm{A}}), \Pr(D^{\mathrm{B}}|D^{\mathrm{A}})]$ were searched and investigated. This approach focuses on the uncertainty factor.

We have changed our view of the problem: we accept total $\Pr(D^{\mathrm{B}})$ and conditional probabilities under the condition that does not contradict the decision being made, i.e. $\Pr(D^{\mathrm{B}}|D^{\mathrm{A}})$, as facts and focus on the behavior of conditional probability under the condition that contradicts the decision $\Pr(D^{\mathrm{B}}|C^{\mathrm{A}})$. In other words, unlike the abovementioned approach, we focus our attention on the inconsistency factor. The argue this shift of the point of view by the following:

1) according to A. Kolmogorov, probability is an additive measure on the algebra of events;

2) we consider logic of decision as an algebra;

3) total $\Pr(D^{\mathrm{B}})$ and conditional probabilities under the condition that does not contradict the decision being made, i.e. $\Pr(D^{\mathrm{B}}|D^{\mathrm{A}})$ are the measures on the monotonic logic, in which new information doesn't change the logical conclusion;

4) conditional probability under the condition that contradicts the decision $\Pr(D^{\mathrm{B}}|C^{\mathrm{A}})$ is formed by non-monotonic logic [32, 33].

So, from logical point of view, we focus our attention on the neural mechanisms, that allows to implement non-monotonic logic. We use a kind of non-monotonic logic, named logic with exception. This logic is derived from the monotonous logic described by the "Modus Ponens" inference rule: "To

betray is profitable" (to remain silent is unprofitable) by adding to the latter the exception rule: "If the accomplice is not silent".

A model

To implement the logic with exception, we use three-layer neural network of a ring architecture, schematically presented in Fig.1.
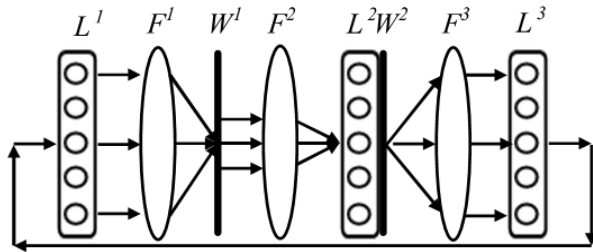
Fig.1. Three-layer neural network of a ring architecture: L^1, L^2, L^3 neural layers: input, estimates and exception, respectively; F — Fourier-transformed elements; W^1, W^2 — matrices of interconnection weights: main rule and exception, correspondingly.

This model of neural network is biologically motivated, as:

1) following by the [34, 35], the interconnection weights are formed and recorded in the Fourier-space;

2) ring architecture of the network corresponds to the concept of the sensation circle by A. M. Ivanitsky [36, 37] and allows a cognitive dissonance, which is an attribute of the decision-making in the Prisoner's Dilemma game [18], to be implemented [26, 33].

To realize logic with exception, it has to implement operation of negation. From an algebraic point of view, negation is a special case of the operation of duality. The Fourier-transform fully satisfies the axiomatic definition of the latter [39], so we use the biologically motivated [34, 35] Fourier-transform as a negation operation — matrices of main rule W^1 and exception W^2 are connected by the Fourier-transform F^2.

The alternatives C and D have both something in common and differences. Let the neural ensembles representing the alternatives $C(x, y)$ and $D(x, y)$ have the same phase spectrum, and different amplitude spectra. So, we can introduce a quantitative assessment of the distinction of alternatives V, defining it as the ratio of the parameters of their amplitude spectra, as which we take

the width of the spectrum at a given α-level ν_α

$$V = \frac{\nu_\alpha^C}{\nu_\alpha^D}. \tag{3}$$

The main rule "To betray is profitable" is stored by the first matrix of the interconnection weights in the "negative form" "to remain silent is unprofitable"

$$W^1(\nu_x, \nu_y) = \Lambda^1\left(F^*\left(C(x,y)\right)F\left(\delta(x_\delta, y_\delta)\right)\right), \tag{4}$$

where ν_x, ν_x are spatial frequencies, Λ is the operator of synaptic contacts W^1 taking into account non-linearity of their sensitivity, F denotes Fourier-transform, asterisk denotes complex conjugation, $C(x, y)$ and $\delta(x_\delta, y_\delta)$ are neural ensembles, representing reference patterns "to be silent" and "unprofitable", respectively.

The exception "If the accomplice is not silent" is stored by the second matrix W^2, which is recorded in accordance with the Heb's learning rule by the two neural ensembles: reference pattern "to be silent", presented in the layer of exclusion L^3, and response of the network on the inputting pattern "to betray", formed by the monotonic logic of the decision $\mathrm{Out}_0^{2(D,C)}(\zeta, \eta)$

$$W^2(\zeta, \eta) = \Lambda^2\left(\left(\mathrm{Out}_0^{2(D,C)}(\zeta, \eta)\right)^* F\left(C(x,y)\right)\right), \tag{5}$$

where (ζ, η) are coordinates in the layer L^2, subscript indicates the number of iteration, zero iteration corresponds to the linear propagation of excitation $L^1 \rightarrow W^1 \rightarrow L^2$, superscript indicates the number of the layer and operands in the brackets. The network's response is as follows

$$\mathrm{Out}_0^{2(In,C)}(\zeta, \eta) = F\left[F\left(In(x,y)\right)\Lambda^1\left(F^*\left(C(x,y)\right)\right)\right] = \left[In(x,y) \otimes C^{\Lambda^1}(x,y)\right], \tag{6}$$

where $\otimes$ denotes correlation.

According to the above described approach, we are interested in the player's evaluation of his own decision to "betray". In accordance with the concept of the "sensation circle" [36, 37], the value is formed on circular iterations that evaluate the individual significance of the input stimulus by comparing it with the contents of the memory.

$$\text{Out}_i^{2(\text{D,C})}(\zeta,\eta) =$$

$$= C^{A^2}(x,y) * F\left(\text{Out}_{i-1}^{2(\text{D,C})}(\zeta,\eta)\right) \otimes F\left(\text{Out}_0^{A^2 2(\text{D,C})}(\zeta,\eta)\right) \otimes C^{A^1}(x,y). \quad (7)$$

In order to correlate measurements in sociological experiments, i.e. probability of the decision, with ones in neural network as a material carrier of the logic, we refer on the Born's rule [38], considering probability as a scalar product of amplitudes. Thus, probability as a quadratic measure is mathematically strictly equivalent to the power (intensity), that is measured in real systems. Since the neural ensembles $C(x, y)$ and $D(x, y)$ have the same phase spectrum, then (6) and (7) are correlation functions and referring to [26], we can use the square of the global maximum of autocorrelation function $s = pr^2$, where r is the radius of correlation, as a quadratic measure, i.e. we will use the notation $(r^D)^2 \equiv \Pr(D)$, $(r^{D|C})^2 \equiv \Pr(D|C)$, etc.. Thus, we can quantify the above-mentioned violation of the classical probability theory by the relation

$$\Omega^{Sq} = \frac{\Pr(D^B)}{\Pr(D^B|C^A)} = \left(\frac{r_0^D}{r_i^{D|C}}\right)^2 \quad (8.1)$$

for neural ensembles described by 2D matrices (functions of two arguments), and

$$\Omega^L = \frac{\Pr(D^B)}{\Pr(D^B|C^A)} = \frac{r_0^D}{r_i^{D|C}}, \quad (8.2)$$

if the neural ensembles are depicted by vectors (functions of an argument).

If $\Omega < 1$, then there is the violation, if $\Omega > 1$, then the experimental result corresponds to classical probability theory. Thus, $\Omega = 1$ is a critical value and we are interested in revealing of the model parameters that affect this value in order to formalize their relationship.

The condition $\Omega = 1$ can be represented as

$$\text{Out}_0^{2(\text{D,C})}(\zeta,\eta) = \text{Out}_i^{2(\text{D,C})}(\zeta,\eta), \quad (9)$$

which can be rewritten as a condition for the stability of the network response

$$\forall k \in \mathbb{N} : \text{Out}_i^{2(\text{D,C})}(\zeta,\eta) = \text{Out}_{i+k}^{2(\text{D,C})}(\zeta,\eta). \quad (10)$$

Taking into account (3), (6), and (7), we can solve (10) with respect to V by denoting its value for stable network response $\Omega = 1$ by V^{Stable}. If $V < V^{Stable}$,

then $\Omega < 1$, and $v.v.$, if $V > V^{Stable}$, then $\Omega > 1$ and the total probability, i.e. unconditional value, describing the decision under uncertainty fits into the interval of conditional values.

For simplicity, but without loss of generality, we adopt Gaussian models for both amplitude spectra and additional low-pass filters that simulate filtration in weight matrices due to the nonlinearity of synaptic sensitivity. The solution for the stable network response is as follows

$$V^{Stable} = \frac{v^C}{v^{Input}} =$$

$$= \sqrt{\sqrt{2\left(2\pi \frac{\zeta}{v}\right)^2 (v^C)^4 + \left[\frac{(R^1)^2 - (R^2)^2}{2} + 1\right]^2} + A - B} \qquad (11)$$

where $A = (R^2)^2 \left[(R^1)^2 + 2\right]$, and $B = \frac{(R^1)^2 - (R^2)^2}{2}$,

where $R^1 = \dfrac{v^C}{v^1}$ and $R^2 = \dfrac{v^C}{v^2}$ are the relative bands of the low-pass filters in the weight matrices W^1 and W^2, v^1 and v^2 are the filters' parameters, respectively. Here in order to avoid cumbersome expressions, we have used functions of only one argument.

From the point of view of information characteristics of images, it is more interesting to describe the reference image not in the frequency (Fourier) space, but in the correlation, i.e. by means of its autocorrelation radius. The latter is linked with the parameter of the Gaussian spectrum v^C by the inverse dependence $r^{CACF} = \dfrac{1}{2\pi v^C}$. Thus, we can rewrite (11) in the following notation

$$V^{Stable} = \frac{v^C}{v^{Input}} =$$

$$= \sqrt{\sqrt{\left(\frac{\zeta}{v}\right)^2 \frac{2}{(2\pi)^2 (r^{CACF})^4} + \left[\frac{(R^1)^2 - (R^2)^2}{2} + 1\right]^2} + A - B} \qquad (12)$$

In (11) and (12) it can be seen that the filtrations on the weight matrices act in the opposite direction, which reflects they Fourier-duality.

Modeling

We simulated decision-making in the Prisoner's Dilemma Game on the neural network shown in Figure 1 in two situations: decision-making under uncertainty, i.e. lack of information about the accomplice's decision, and in a situation where information about the accomplice's decision "to keep silent", which is a condition, contradicts the previously learned decision-making rule "to betray is profitable". Inputting patterns "to keep silent" and "to betray" were represented by stochastic functions of one argument with the same phase spectra and different amplitude Gaussian spectra in accordance with (3), the length of the realizations were 4096 pixels. The parameters of the spectra r^C and r^D were selected based on considerations of the convergence of the process of reducing cognitive dissonance arising from the perception of a condition that contradicts a previously learned rule [33], and the clarity of the results. Low-pass filters are modeled by Gaussian functions too.

Figures 2.a and 2.b illustrate the influence of the reference (training) image's spectra parameter r^C (or, that is equivalent, autocorrelation radius of the training image) on the value of V^{Stable}, which is marked in the figures with vertical dashed lines: Figure 2.a represents the results for $r^C =150$, Figure 2.b — the same for $r^C =75$. The sensitivity of synaptic contacts in the formation of matrices W^1 and W^2 during learning by the main rule and exclusion, correspondingly, was assumed to be linear, i.e. filtering was not modeled here. The value of V^{Stable} according to (8) corresponds to the intersection point of curves describing the dependence of the evaluation of the solution r on the value of V: straight lines represent unconditional estimates r^D at zero iterations, curves 1 and 2 — conditional estimates $r^{D|C}$ at the first and the second circular iterations, correspondingly. To the left of the intersection V^{Stable} point we have $\Omega<1$, that means a violation of classical probability theory, to the right — $\Omega>1$, i.e. the results correspond to the theory.

Figure 3 shows the dependences of V^{Stable} on the autocorrelation radius r^{CACF} of the reference (training) image $C(x)$ for a number of values of the relative band of the low-frequency filter on the matrix of weights W^1, assuming no filtering on the matrix of weights W^2. Standard deviations σ of the numerical results from the theoretical ones (12) are presented in the Table 1.
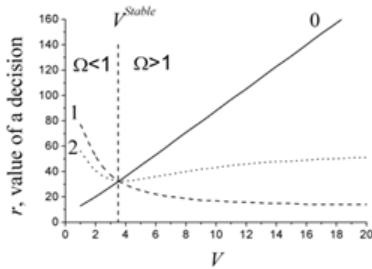
Fig. 2.a. The dependence of the estimates r of the decision to betray on the V—evaluation of the difference of the image to betray from the training image to be silent on the zero (0), first (1) and second (2) circular iterations at r^C=150 periods per length of the image
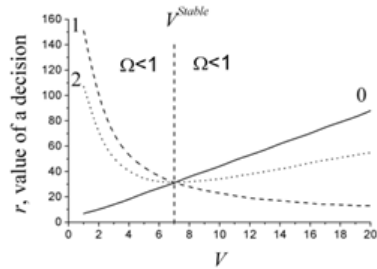
Fig. 2.b. The dependence of the estimates r of the decision to betray on the V—evaluation of the difference of the image to betray from the training image to be silent on the zero (0), first (1) and second (2) circular iterations at r^C=75 periods per length of the image

Figure 4 illustrates the inverse effect of filtering on the matrix of weights of the exclusion rule W^2, described by term $\left(\left(R^1\right)^2 - \left(R^2\right)^2\right)$ in (11) and (12)—it shows the dependence of V^{Stable} on the relative bandwidth of the filter at the value of the relative width of the filter on the matrix W^1 of the main rule of inference R^1 =10 (curve 4 in Figure 3).
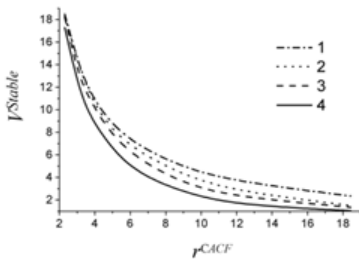
Fig.3. The dependences of the V^{Stable} on the autocorrelation radius of the training image for a set of the ratios of the parameters R^1 presented in the Table 1
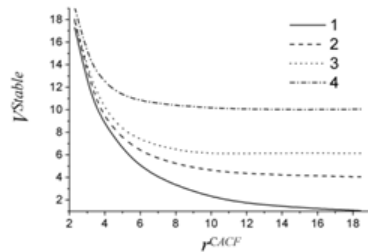
Fig.4. The dependences of the V^{Stable} on the autocorrelation radius of the training image for a set of the ratios of the $R2$ under the filtration on the matrix W^1 by R^1=10

Values of the parameters R^1 and R^2, and standard deviations of
experimental data from the theoretical ones for the results presented in
Figures 3 and 4.

Curve №	Figure 3		Figure 4	
	R^1	Standard deviation, σ	R^2	Standard deviation, σ
1	0	0.035	0	0.198
2	4	0.145	4	0.148
3	6	0.163	6	0.181
4	10	0.198	10	0.112

Discussion

The results presented above show that low-frequency filtering on the matrix of weights of the main decision rule leads to a decrease in the value of V^{Stable} (Fig.3), which corresponds to a shift of the point $\Omega=1$ in Fig.3 to the left. i.e. to the decreasing of the range of values of V, in which there is a violation of classical probability theory. This regularity may be related to the peculiarities of the internal picture of the player's world, represented by the decision-making rule "to betray is profitably". The correlation radius of the "betray" training image reflects its inherent balance of general and particular features: an increase in the correlation radius corresponds to an increase in the specific weight of general features and, conversely, its decrease corresponds to an increase in the specific weight of particular features. Low-frequency filtering of the training image when it is recorded in the memory leads to an increase of the correlation radius of the stored image relative to the reference image received during learning from the outside and, accordingly, to a decrease in the value of V^{Stable}. Thus, a comparison of Figures 2.a and 2.b, and Fig.3 show that with an increase in the proportion of common features in the player's inner picture of the world, i.e., an increase in the generality of his view of the world, the ability to see not only the particular, but also the general, his tendency to betrayal decreases. Conversely, with an increase in the player's attention to particulars to the detriment of the general, i.e., decreasing of the radius of correlation of the remembered image, the tendency to betrayal increases.

Figure 4 shows that low-frequency filtering on the weight matrix, which is an exception to the basic rule, on the contrary, leads to an increase in the value

of V^{Stable} and, thereby, a shift of the value $\Omega=1$ to the right. From the point of view of the model, this is determined by the fact that the weight matrices W^1 and W^2 are linked by the Fourier transform, which sets the duality of the operations defining the algebra of logic, i.e., the negation operation "not", which sets an exception to the basic rule. In mental terms, this can be represented by the following reasoning of the player: "It does not matter that the accomplice is silent, since this is only a special case, only an exception to the general, and it is profitable to betray the general."

Conclusion

Thus, using the non-cooperative game "Prisoner's Dilemma", we have shown that the mental characteristics of the decision-maker influencing the choice of alternatives are determined by both his life experience and the characteristics of the material carrier of his logic — his neural network.

The life experience of the decision-maker in this case is the history of his learning, specifically, the spectral characteristics of the image that was used in learning the basic decision-making rule.

The peculiarities of the material carrier in the studied case are the non-linearity of the sensitivity of synaptic contacts, leading to the filtration of the memorized image relative to the reference one.

Decision-making as a choice of alternatives often has an essential moral aspect, in the Prisoner's Dilemma game this moment is manifested quite clearly. The results show that the features of both learning and the material carrier of the logic of choice are associated with the morality of the decision being made. We do not undertake to assert that they are determinative, since the factor of morality, like freedom of choice, freedom of will as an attribute of a person, in our opinion, is by no means reduced to the properties of the material carrier of the psyche and consciousness, but a certain correlation can be traced.

References

1. *Tversky A, Kahnenman D.* Judgment under uncertainty: heuristics and biases. Science (1974) 185:1124–31. doi: 10.1126/science.185.4157.1124.
2. *Tversky A, Kahneman D*. The framing of decisions and the psychology of choice. Science (1981) 211:453–8. doi: 10.1126/science.7455683.
3. *Tversky A, Kahneman D.* Rational choice and the framing of decisions. J Business (1986) 59:251–78. doi: 10.1086/296365.
4. *Kahneman D., Slovic P., Tversky A.* Judgment Under Uncertainty: Heu-

ristics and Biases. Cambridge, UK: Cambridge University Press (1982).

5. K*ahneman D., Tversky A.* Prospect theory — an analysis of decision under risk. J Econ. (1979) 47:263–92. doi: 10.2307/1914185.

6. *Savage L. J.* The Foundations of Statistics. New York, NY: John Wiley (1954).

7. *Tversky A., Kahneman D*. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review 1983; 90(4): 293–315. DOI: 10.1037/0033–295X.90.4.293.

8. *Tversky A., Shafir E.* The disjunction effect in choice under uncertainty. Psychological Science 1992; 3(5): 305–9. DOI: 10.1111/j.1467–9280.1992.tb00678.x.

9. *Crosson R.* The disjunction effect and reason-based choice in games. Organizational Behavior and Human Decision Processes 1999; 80:118–133. DOI:10.1006/obhd.1999.2846.

10. *Li S., Taplin J.* Examining whether there is a disjunction effect in Prisoner's Dilemma games. China Journal of Psychology 2002; 44: 25–46. http://hdl.handle.net/2440/3240.

11. *Hristova E., Grinberg M.* Disjunction effect in prisoner's dilemma: evidences from an eye-tracking. Proceedings of the 30th Annual Conference of the Cognition Science Society. Washington, DC: 2008; 1225–30.

12. *Busemeyer J., Wang Z., Townsend J.* Quantum dynamics of human decision making. J Math Psychol. (2006) 50:220–41. doi: 10.1016/j.jmp.2006.01.003.

13. *Busemeyer J.* Cognitive science contributions to decision science. Cognition (2015) 135:43–6. doi: 10.1016/j.cognition.2014.11.010.

14. *Busemeyer J., Wang Z.* Quantum cognition: key issues and discussion. Top Cogn Sci. (2014) 6:43–6. doi: 10.1111/tops.12074.

15. *Khrennikov A., Asano M.* A quantum-like model of information processing in the brain. Applied Science 2020; 10(2): 707:1–14. DOI: 10.3390/app10020707.

16. *Adams B., Petruccione F.* Quantum effects in the brain: A review. AVS Quantum Science 2020; 2: 022901–1–15. DOI: 10.1116/1.5135170.

17. *Menskii M. B*. Concept of consciousness in the context of quantum mechanics. Physics — Uspekhi 2005; 48(4): 389–409. DOI: 10.1070/PU-2005v048n04ABEH002075.

18. *Trueblood J. S., Pothos E. M., Busemeyer J. R*. Quantum probability theory as a common framework for reasoning and similarity. Frontiers in Psychology, 11 April 2014; DOI: 10.3389/fpsyg.2014.00322.

19. *Pothos E. M., Busemeyer J.* R. A quantum probability explanation for violations of 'rational' decision theory. Proceedings of the Royal Society Biological Sciences 2009; 279: 2171–8. DOI: 10.1098/rspb.2009.0121.

20. *Busemeyer J. R., Pothos E. M., Franco R., Trueblood J. S.* A quantum theoretical explanation for probability judgment "errors" Psychological Review 2011; 118(2): 193–218. DOI: 10.1037/a0022542.

21. *Broekaert J. B., Busemeyer J. R., Pothos E. M.* The Disjunction Effect in two-stage simulated gambles. An experimental study and comparison of a heuristic logistic, Markov and quantum-like model. Cognitive Psychology 2020; 117: 101262. DOI: 10.1016/j.cogpsych.2019.101262.

22. *Bagarello F., Basieva I., Pothos E. M., Khrennikov A.* Quantum like modeling of decision making: Quantifying uncertainty with the aid of Heisenberg–Robertson inequality. Journal of Mathematical Psychology 2018; 84: 49–56. DOI: 10.1016/j.jmp.2018.03.004.

23. *Zheltikov A. M.* The critique of quantum mind: measurement, consciousness, delayed choice, and lost coherence. Physics - Uspekhi 2018; 61(10): 1016–25. DOI: 10.3367/UFNe.2017.06.038155.

24. *Pavlov A. V., Orlov V. V.* Modelling the mechanisms of quantum logic using the method of superimposed Fourier holograms based on the nonlinearity of the exposure characteristics of holographic recording media. Quantum Electronics 2019; 49(3): 246–252. DOI: 10.1070/QEL16748.

25. *Pavlov A. V.* Simulation of quantum logic in linear recording of superimposed Fourier holograms: Linda phenomenon. Quantum Electronics 2019; 49(8): 777–8. DOI: 10.1070/QEL16939.

26. *Pavlov A.* V. Modeling of quantum-like cognitive phenomena by the Fourier-holography technique under the choice of alternatives. Computer Optics 2021; 45(4): 551–561. DOI: 10.18287/2412–6179-CO-830.

27. *Gunning D, Aha D. W.* DARPA's Explainable Artificial Intelligence Program. AI Magazine 2019; 40(2): 44–58. DOI: 10.1609/aimag. v40i2.2850.

28. *Samek W, Montavon G., Lapuschkin S., Anders C. J., Müller K. — R.* Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications. Proceedings of IEEE2021; 109(3): 247–278. DOI: 10.1109/ JPROC.2021.3060483.

29. *Conati C., Barra O., Putnam V., Rieger L.* Toward personalized XAI: A case study in intelligent tutoring systems. Artificial Intelligence 2021; 298: 103503. DOI: 10.1016/j.artint.2021.103503.

30. *Korteling J. E., Brouwer A. — M, Toet A. A* Neural Network Framework for Cognitive Bias. Front. Psychol. 2018; 9:1561. doi: 10.3389/ fpsyg.2018.0156.

31. *Nash J. F.* Non-Cooperative Games. Annals of Mathematics 1951; 54(2): 286–295. DOI: 10.2307/1969529.

32. *Reiter, R.* A logic for default reasoning. Artificial Intelligence. 1980; 13(1–2): 81–132.

33. *Pavlov A. V.* Holographic memory updated by contradicted information: influence of low frequency attenuation on response stability. Computer Optics 2020; 44(5): 728–736. DOI: 10.18287/2412–6179-CO-668.

34. *Glezer V. D.* The Role of Spatial–Frequency Analysis, Primitives, and Interhemispheric Asymmetry in the Identification of Visual Images. Human Physiology 2000; 26(5): 636–640. DOI: 10.1007/BF02760381.

35. *Glezer V. D.* Matched filtering in the visual system. Journal of Optical Technology 1999; 66(10): 853–856. DOI: 10.1364/JOT.66.000853.

36. *Ivanitskii A. M.* Information synthesis in key parts of the cerebral cortex as the basis of subjective. Neuroscience and Behavioral Physiology 1997; 27: 414–426. DOI: 10.1007/BF02462943.

37. *Ivanitsky A.M, Ivanitsky G.A, Sysoeva O. V.* Brain science: On the way to solving the problem of consciousness. International Journal of Psychophysiology 2009; 73:101–108. DOI: 10.1016/j.ijpsycho.2009.02.004.

38. *Richard F. Feinman, Robert B. Leighton, Matthew Sands.* The Feinman lectures on physics. Vol.3. — Reading, Massachusets, Palo Alto, London: Addison-Wesley Publishing Company, Inc., 1965. URL: https://www.feynmanlectures.caltech.edu/III_toc.html.

39. *Pavlov A. V.* On Algebraic Foundations of Fourier Holography. Optics and Spectroscopy 2001; 90(3): 452–457. DOI: 10.1134/1.1358459.

Chapter 16.
Rhesus monkeys performing cognitive tests in a virtual environment

Daria Podvigina, Liubov Ivanova, Alexey Harauzov

Pavlov Institute of Physiology RAS, Saint Petersburg, Russia

Within the last decades the research of cognitive functions of nonhuman primates has become actual again. This issue is especially important as nonhuman primates are an important tool for the investigation of complex cognitive functions and their impairments in brain disorders of different etiology (Camus et al., 2015; Bondar et al., 2019). Among nonhuman primates, rhesus macaques are the most widely used animal model for studies of neural processes (Freeman and Gosling, 2010). Considering this, the study of cognitive functions in rhesus macaques is of a current interest.

One of the significant tasks in this area is to elaborate a valid and reliable procedure of testing cognitive functions that could be applied to both humans and nonhuman primates. For this purpose, standardized computer-based cognitive tests are often used (Bramlett-Parker & Washburn, 2016; Perdue et al., 2018). In some studies researchers develop tests using one experimental paradigm for monkeys and humans, like, for example, in Gazes et al. (2014); others use test batteries previously designed for humans. For example, a computerized behavioral battery based upon human Cambridge Neuropsychological Test Automated Battery (CANTAB) developed to assess cognitive behaviors of rhesus monkeys (Weed et al., 1999; Ryan et al., 2019). The battery addresses memory, attention, motivation, reaction time and motor coordination—the aspects of cognitive behavior most commonly assessed in nonhuman primate behavioral research.

We also use computer technology in behavioral research on monkeys, which has significantly expanded the possibilities for performing experiments. In particular, we can design complicated research procedures with various visual and sound stimuli of varying complexity and sequence, as well as easily change the temporal characteristics of presentation. This allows us to study the features of both the low-level stages of sensory signal processing and higher cognitive functions.

Experimental set with a touchscreen monitor

Thus transferring animal cognitive tests in a virtual environment is now an important task for researchers. For this purpose, we have designed a hardware-software complex to conduct psychophysical and cognitive studies on monkeys (rhesus macaques, *Macaca mulatta*). The system includes a computer with a touchscreen monitor Iiyama prolite t2250mts with optical technology of touch registration and EventIDE software (OkazoLab Ltd.). It provides visual and auditory stimulation with automatic recording of responses and provision of food reinforcement of correct responses and drink if necessary.

During the experiment, a monkey stays in a special moveable cage made of transparent Plexiglas with an opening for the forelimbs and a built-in reinforcement system (Figure 1). The distance between the cage and the touchscreen monitor is 20 cm. The automatic reinforcement system supplies a food reward to an animal for a correct response. The system includes a round-shaped dispenser propelled by a stepper motor, which is controlled by a programmable microcontroller connected with a computer. Reinforcement system and stimulus presentation system are synchronized using a COM port serial protocol. The experimental system is also described in Ivanova et al., (2016).

Fig.1. A monkey in a movable cage performing a delayed matching-to-sample test on a touchscreen monitor.

The system enables creating complex experimental procedures, adjusting the parameters of stimulus presentation and registering the exact animal response time. Moreover, touchscreen monitor makes it easier to teach monkeys to perform a task, for which various techniques can be additionally used, for example, animating target stimuli, placing bright "hint" images over them, etc.

We have taught rhesus macaques to perform a number of tasks in a virtual environment: a test to assess contrast sensitivity to low spatial frequency images, a task to investigate visuoperceptual and cognitive abilities and a delayed matching-to-sample test for working memory. Below we will describe each task — the stimuli we used, the procedure and the results in short.

Contrast sensitivity test

In this task, the stimuli were grayscale images of Gabor elements of low spatial frequency and different contrast (Fig. 2). The monkeys' task was to touch the screen with their paws in the area of the target stimulus, which was automatically followed by food or drink (juice) reinforcement. After the correct response, the contrast of stimuli gradually decreased. As a result, the threshold contrast at which the monkeys can detect the appearance of low-frequency images was measured by the method of two-alternative forced choice (Ivanova et al., 2016). Besides contrast sensitivity, with this method one can study other characteristics of the low-level processes in visual system, such as color sensitivity, visual acuity, etc.

Test for assessment of visuospatial and cognitive abilities

This test was previously developed for psychophysiological, electrophysiological and tomographic human studies in the tasks of estimation of statistical characteristics of complex images (Shelepin et al., 2009; Harauzov et al., 2016). The task enables to assess both visuoperceptual abilities (contrast sensitivity, spatial grouping) and cognitive abilities such as object classification and decision making. An important advantage of this task is an ability to vary and control its complexity with no change in the physical characteristics of stimuli (size, contrast, color etc.). For the stimuli, we used images of Gabor patches assembled into 10×10 matrices. The images differed in the number of Gabor elements with vertical and horizontal orientations. According to the number of co-oriented (either vertical or horizontal) elements relative to other orientations, matrices could be either "random" with all possible orientations equally represented or "oriented" with the dominance of either vertical, or horizontal, orientations. The matrices were arranged into eight groups accord-

ing to their degree of orderliness: the number of Gabor patches sharing the same orientation was 100 (all 100 Gabor elements are oriented either vertically or horizontally), 64, 56, 48, 40, 32, 24 and 16 (Fig. 3).
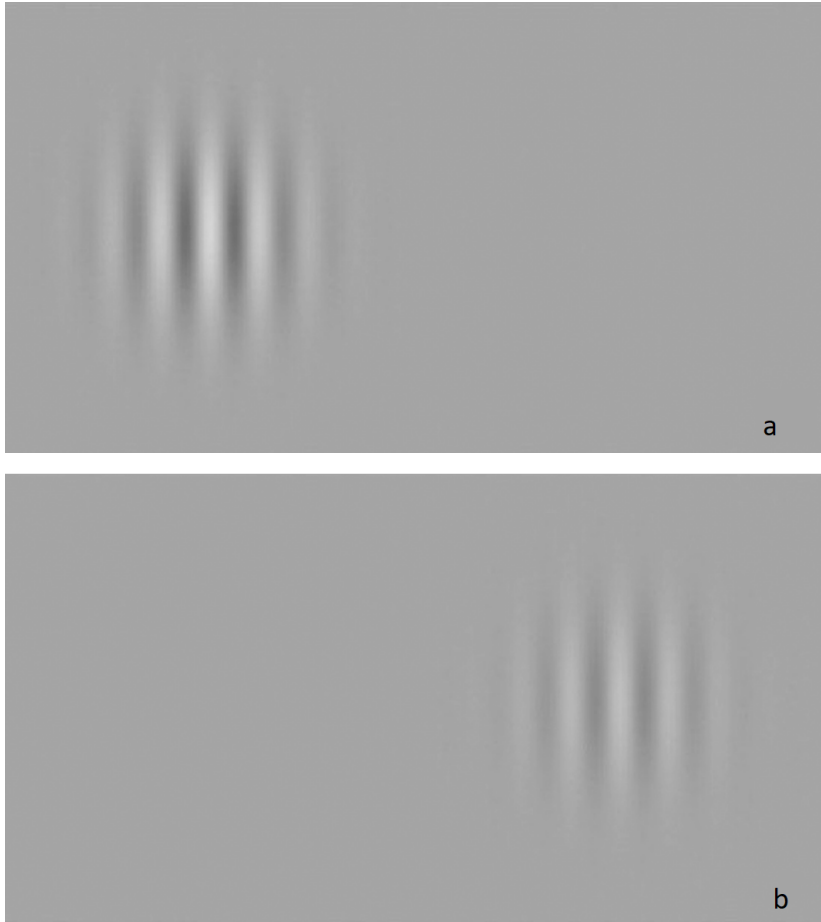
Fig.2. Example of visual stimuli of different contrast in two trials (*a* and *b*).

The observer is to detect a predominant orientation of the elements within a matrix (vertical or horizontal). As one can see from Figure 3, with the decrease of the number of co-oriented elements it becomes harder to detect a predominant orientation.

100 % (vertical)   64 % (horizontal)   56 % (vertical)   48 % (horizontal)

40 % (vertical)   32 % (horizontal)   24 % (vertical)   16 % (horizontal)
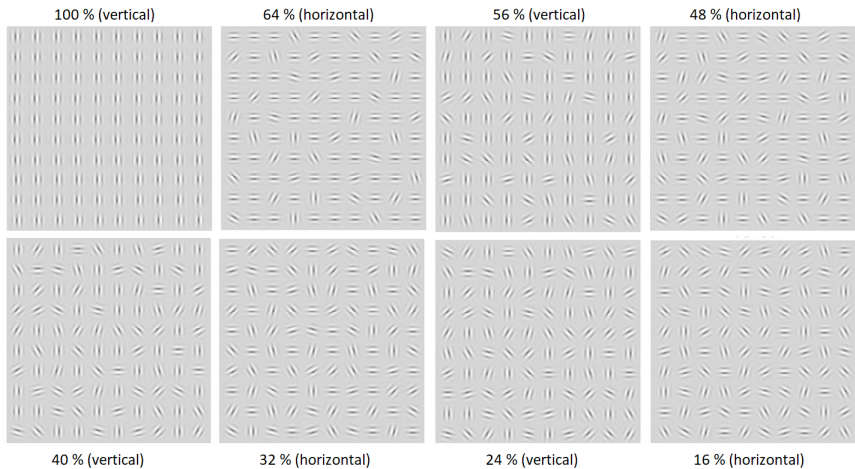
Fig.3. Examples of test matrices of different orderliness level from the stimulus set. Each matrix consists of 100 Gabor elements. Matrices are characterized by the dominance of either vertical or horizontal elements and differ in the number of co-oriented Gabor elements relatively to other orientations. The captions next to the pictures indicate the numbers of co-oriented vertical or horizontal elements within the matrix.

The task of a monkey in this test was to choose a matrix with predominantly vertical orientation between the two matrices presented simultaneously on a touchscreen monitor (two-alternative forced choice; Figure 4).

a   b

Fig.4. The two matrices with 40 co-oriented elements presented in one experimental trial. A—horizontal orientation is predominant, b—vertical one is predominant. A monkey has to touch the vertical one.

To teach monkeys (nine male rhesus macaques) this task we started with black-and-white gratings with vertical and horizontal bars as stimuli for a two-alternative forced choice task (Figure 5). The gratings appeared on a touchscreen for unlimited time, the position of gratings (left or right) was randomized. Five hundred milliseconds before presentation a sound signal was given for a monkey to get ready for visual stimulation. We give monkeys this warning sound before each trial in all experiments with visual stimulation. When a monkey touched the vertical grating, it automatically received food reinforcement.
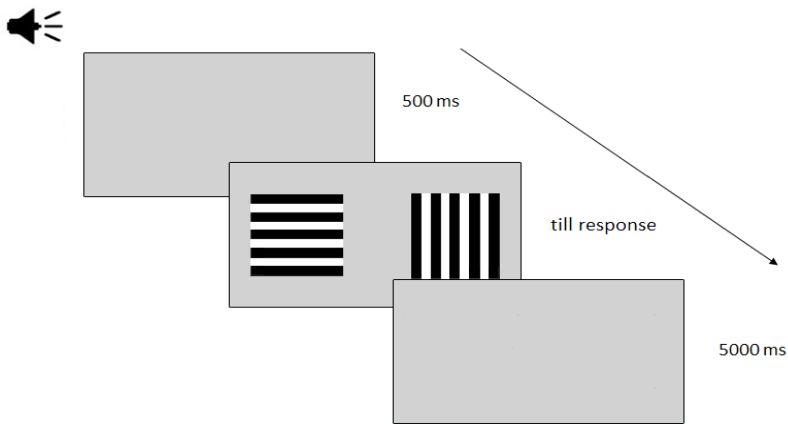
Fig. 5. A trial with black-and-white gratings presentation. At the beginning of each trial there was a sound signal and 500 ms afterwards two gratings were presented until a monkey responded. Another trial began 5000 ms after the response.

After monkeys had learned to choose vertical grating (gave more than 75 % correct responses), they proceeded to the task with 100 % matrices — vertical and horizontal. Then they learned the task with more complex matrices — 64 % and so on. As soon as they got through the whole set of stimuli, we started an experimental session with the four most complex matrices as stimuli (40 %, 36 %, 24 % and 16 %). Figure 6 shows a schematic trial in the experimental session. One trial included two matrices presentation — both with the same percentage of co-oriented elements, one vertical and one horizontal; the position of each matrix on the screen (right or left) in each trial was randomized.
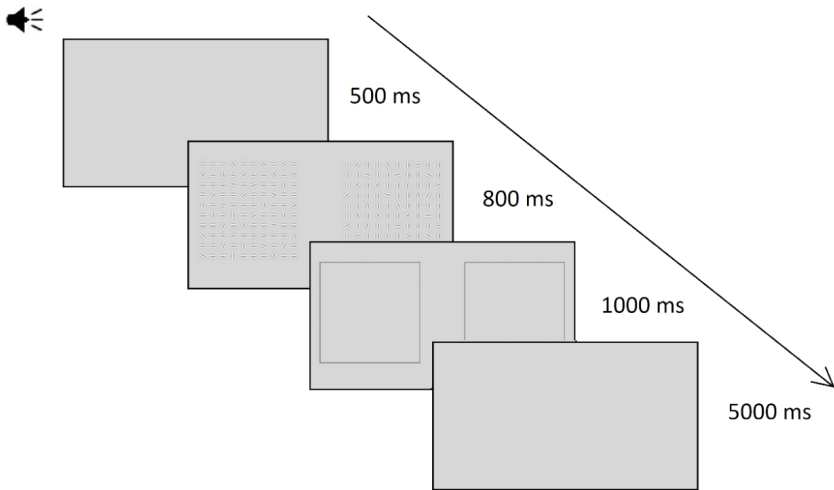
Fig.6. Time course of an experimental trial. At the beginning of each trial there was a sound signal and, after 500 ms, two stimuli were presented for 800 ms. A response area then appeared for 1000 ms and after 5000 ms another trial started.

In the experimental session, we assessed the task performance in terms of reaction time and number of correct responses (accuracy). We found out that reaction time increased with complexity, though not in all monkeys. This increase in reaction time we observed only with the response time limitation (1800 ms in our study). With no such limitation there was no dependence found between reaction time and stimulus complexity (Podvigina et al., 2020).

The number of correct responses depended strongly on the matrix complexity—the number of co-oriented elements (Figure 7, a). We compared the number of correct response monkeys gave in our study and obtained earlier for humans performing the same task (Harauzov et al., 2016; Fig. 7, b). As is seen from the figure, the percentage of correct responses is about 70 for the least ordered matrix (16 % of co-oriented elements) and sharply increases with the degree of "orientational orderliness" both for monkeys and humans. The average number of correct responses to matrices of all types was 85 % for humans (Harauzov et al., 2016) and 86,75 % for macaques. The similarity of the performance curves of humans and monkeys suggests that the same strategies and neural mechanisms are involved while performing this type of task in both species.
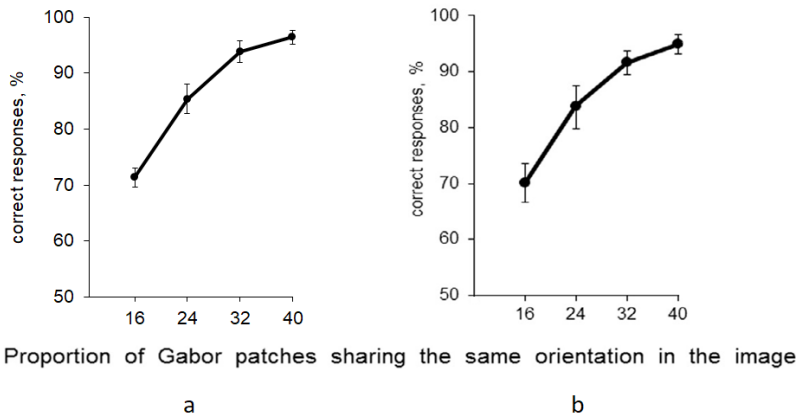
Fig.7. Percentage of correct responses as a function of the number of co-oriented Gabor elements in the matrix for monkeys (a) and humans (b). Trace in (a) is averaged across nine monkeys, (b)—25 participants. Error bars are standard errors of mean. (b) is taken from Harauzov et al., (2016).

Delayed matching-to-sample test

Delayed matching-to-sample task (DMTS) is one of the most commonly used methods for studying mnemonic abilities of both humans and animals. The classical DMTS paradigm is the following: the animal is shown a sample (a three-dimensional object or image), then the sample disappears, and after a delay—from seconds to minutes—two stimuli are presented for choice, one of which is identical to the sample and another one is new. The animal receives reinforcement for choosing a stimulus that is identical to the sample. This task is used to study the characteristics of working memory in animals of different species (Lind et al., 2015), including non-human primates (Rodriguez & Paule, 2009). In the case of primates, various computerized versions of this test are widely used, which use touchscreen monitors to present stimuli and record animal responses (Hoffman et al., 2009; Taubert et al., 2019; Plagenhoef et al., 2021).

We implemented the delayed matching-to-sample task on the experimental setup with a touchscreen monitor described above and trained six rhesus monkeys (males, 6–15 years old) to perform this task with geometric shapes of different colours as stimuli. The stimulus set included 11 geometric shapes of eight colors. Figure 8 shows the scheme of stimulus presentation (one trial).
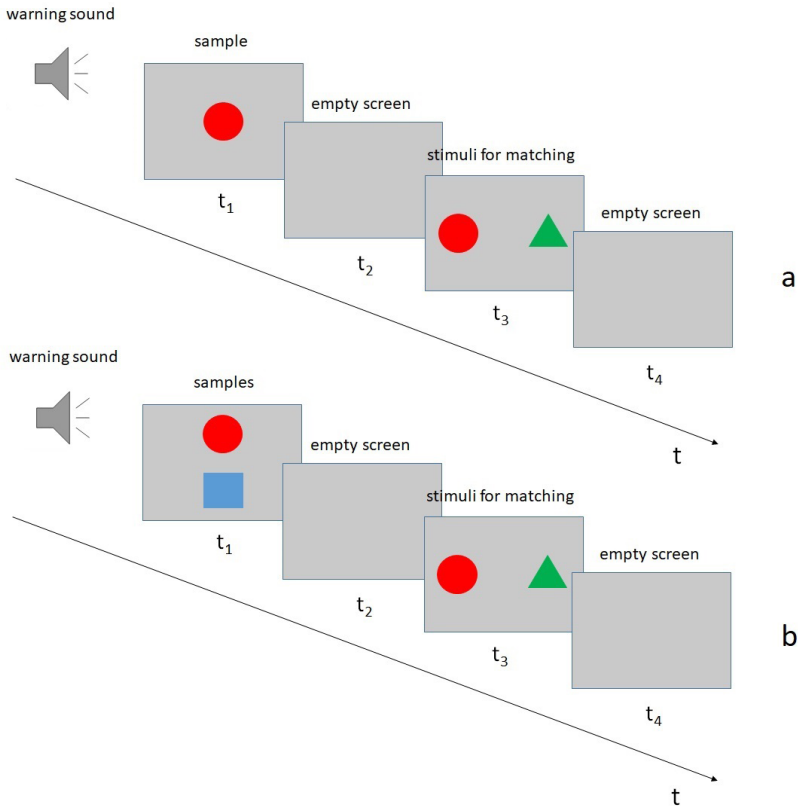
Fig.8. Scheme of a trial in the delayed matching-to-sample test. a—one sample to memorize, b—two samples to memorize. Time intervals t1, t2, t3 and t4 are set by the experimenter at the beginning of the experiment.

At the beginning of the test, the animal is presented with a sample on the screen—a geometric figure. The experimenter sets the time of presentation of the sample before the start of the experiment. Then, after the sample disappears or after a cirtain time (t2), two figures appear on the screen: one is identical to the sample and another one is new. The animal's touching of any of the figures is registered by the program as either a correct or an incorrect response, and after that a blank screen follows for the t_4 time interval (Fig. 8, a). If there is no response, the stimuli also disappear after the time set in the

program. Then the cycle repeats. The position of the correct response — the sample stimulus — on the screen (on the right or on the left) changes in each trial in a random order. For the correct response, an animal automatically receives food reinforcement.

When teaching monkeys to perform this task, we faced the problem of "giving the instruction" to an animal — which stimulus in the sequence should be remembered and which one should be chosen later. To teach the animals the task more quickly, we developed a training version of the task, in which the sample was not simply replaced by stimuli for matching, but was moving towards one of the positions of stimuli for matching on the screen (Fig. 9). That is, at first, the monkey saw a figure in the center of the screen for some time, which then began to move to the left or right side of the screen, and after it stopped in a terminal position, in the opposite part of the monitor an alternative stimulus appeared. Thus, by tracking the movement of the sample — the target stimulus — the animal learned the task. After that, the speed of movement was increased until it completely disappeared — the sample was quickly replaced by stimuli for matching. Then we gradually increased the time between the sample and the stimuli and registered the percentage of correct responses.
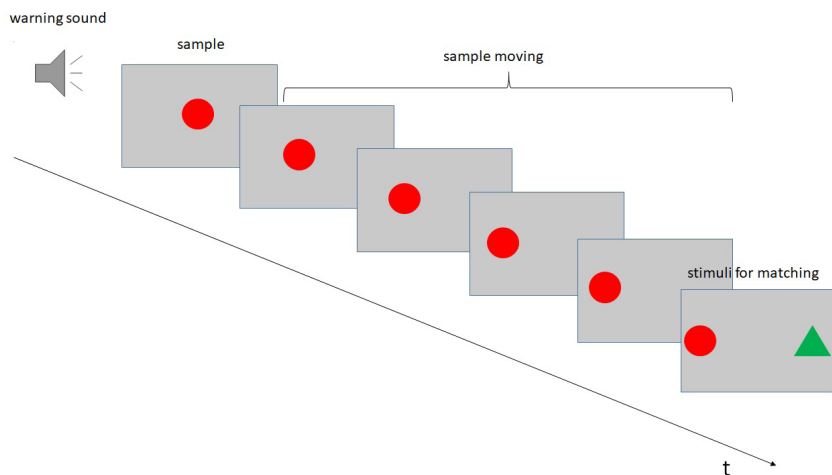

Fig. 9. Schematic representation of a training probe. The sample gradually moves from the center to the edge of the screen. At the end of the movement, an alternative stimulus appears on the opposite side of the screen. The task of the monkey is to touch the sample shape immediately after it stops.

In addition, the program allows us to increase the number of simultaneously presented samples up to five. Those animals that successfully learned the task of matching to one sample went on to a more complex task: they had to memorize two samples, and then choose one of them also between the two stimuli (Fig. 8b). Animals also started this task with a moving sample. Further, the animals went on to the matching to one out of three samples.

Thus, the program allows varying the task difficulty by manipulating two parameters — the delay between the sample and the matching (t2) and the number of samples.

Using the training version of the task with the moving sample, we trained all six monkeys to perform the delayed matching-to-sample task. In the very first experiment with the presentation of a moving sample, the monkeys began to track it and thus gave more than 75 % of correct responses. Then, gradually increasing the speed of the sample movement, we switched to the variant of presentation without sample movement, but with a minimum time interval between the sample and the stimuli for matching — 100 ms. Three of the six monkeys continued to work successfully (give more than 75 % of correct answers) when the presentation parameters were changed. One male named Leo reached 75 % of correct responses in the absence of sample movement after 11 days of training, and the other two, Liman and Filin, needed more training with the sample movement and were able to proceed to the task without movement only after a month of training. Figure 10 shows that in the task with one sample, five animals (all except for Leman) gave on average more than 75 % correct answers. The time of presentation of the sample in this task was 300 ms, the time interval between the sample and the stimuli for selection was 1000 ms.

Then we made the task more complicated by increasing the number of samples. Five animals learned to perform the task with two samples in an average of 6.6 days (SD = 2.7) (here we include only the training sessions without sample movement). Although one of them, a macaque named Edik, overcame the threshold of 75 % of correct responses but did not keep this result in the next series of experiments (Fig. 10). For the sixth animal, named Liman, it took more than two months to learn this task without a prompt of moving sample, and he did not overcome the border of 75 % correct responses and gave about 66 % of correct responses (Fig. 10), while consistently exceeding the level of guessing (50 %). On average, in a task with two samples, the correct response rate of five monkeys was 75 %. The five monkeys proceeded to the three-sample task (without training with sample movement).

They continue to practice, and by now, on average, when performing this task, they give about 66 % of correct answers (Fig. 10).
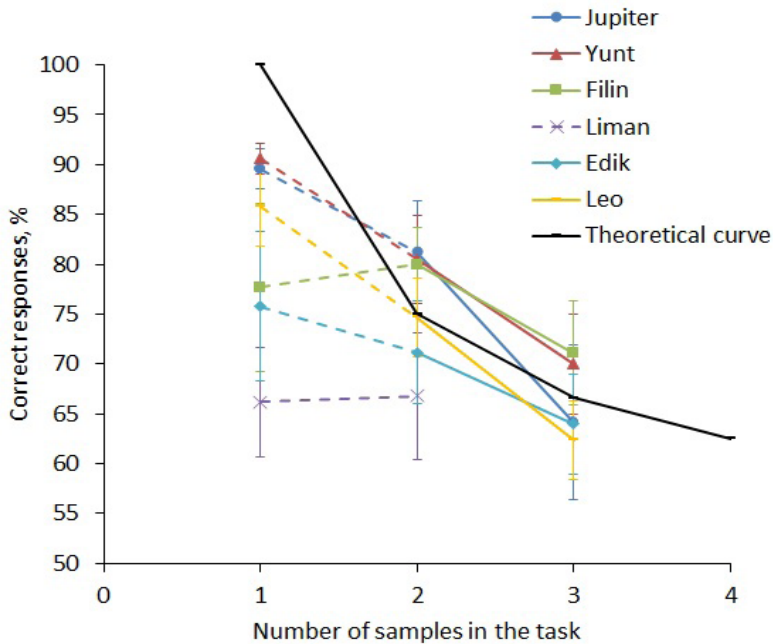
Fig.10. Individual results of animals performing the DMTS task with different numbers of samples. The average percentage of correct responses is shown for each variant of the task. Error bars are the standard deviation. There is a theoretical curve in the plot, which represents the number of correct responses in the case of using the strategy of memorizing only one of the samples (for details see text).

   Observing the animals as they performed the DMTS task, we assumed that as the task became more complicated, they could remember only one sample, ignoring the other ones. Then the level of their correct responses would exceed the level of guessing, but would gradually decrease with an increase in the number of samples. We calculated the percentage of correct responses when using this strategy for each variant of the task (provided that the sample chosen for memorization is correctly identified in 100 % of cases, and the rest of the samples are randomly guessed) and plotted the theoretical curve (Fig. 10). Figure 10 shows that when performing a task with two samples, three monkeys give correct responses with a probability greater than

the calculated one; this suggests that they are not using this strategy and are trying to memorize both figures. In the case of three samples, two out of four monkeys may be using the strategy of memorizing one figure, although they probably simply have not got enough practice yet.

Conclusion

We have designed a hardware-software complex to conduct psychophysical and cognitive studies on monkeys (rhesus macaques). The complex consists of a computer with a touchscreen monitor and an automatic food reward system and allows visual and auditory stimulation, creating complex experimental procedures, adjusting the parameters of stimulus presentation and registering the exact response time. Touchscreen monitor makes it easier to teach monkeys to perform a task, for which various techniques can be additionally used, for example, animating target stimuli, placing bright "hint" images over them, etc. Working in a virtual environment, animals have learned different tasks, which allowed us to assess contrast sensitivity to low spatial frequency images, to investigate visuospatial abilities and to accomplish the delayed matching-to-sample test for working memory, which has traditionally been performed on real three-dimensional objects. Transferring this test in a virtual environment enables us to vary stimulus presentation parameters easily — the number of objects for memorization, their shape and colour, the presentation duration and the retention delay. In all, the use of computerized variants of behavioral tests provides more options for studying the cognitive abilities of monkeys and, in particular, enables to apply one experimental paradigm to both animal and human studies.

References

1. *Bondar, I. V., Vasileva, L. N., Tereshchenko, L. V., Buynevich, A. V. & Latanov, A. V.* (2019). Training of Rhesus Macaques to Complex Cognitive Tasks. *Neuroscience and Behavioral Physiology*, 49(8), 996–1007. https://doi.org/10.1007/s11055–019–00829–6.
2. *Bramlett-Parker, J., & Washburn, D. A.* (2016). Can Rhesus Monkeys Learn Executive Attention? *Behavioral Sciences*, 6(2), 11. https://doi.org/10.3390/bs6020011.
3. *Camus, S., Ko, W. K. D., Pioli, E. & Bezard, E.* (2015). Why bother us-

ing non-human primate models of cognitive disorders in translational research? *Neurobiology of learning and memory*, 124, 123–129. https://doi.org/10.1016/j.nlm.2015.06.012.

4. *Freeman, H. D. & Gosling, S. D.* (2010). Personality in nonhuman primates: a review and evaluation of past research. *American journal of primatology*, 72(8), 653–671. https://doi.org/10.1002/ajp.20833.

5. *Gazes, R. P., Lazareva, O. F., Bergene, C. N., & Hampton, R. R*. (2014). Effects of spatial training on transitive inference performance in humans and rhesus monkeys. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40(4), 477. doi: 10.1037/xan0000038.

6. *Harauzov, A. K., Shelepin, Y. E., Noskov, Y. A., Vasilev, P. P. & Foreman, N. P.* (2016). The time course of pattern discrimination in the human brain. *Vision research*, 125, 55–63. https://doi.org/10.1016/j.visres.2016.05.005.

7. *Hoffman, M. L., Beran, M. J., & Washburn, D. A.* (2009). Memory for "what", "where", and "when" information in rhesus monkeys (Macaca mulatta). Journal of Experimental Psychology: Animal Behavior Processes, vol. 35, no. 2, pp. 143. DOI: 10.1037/a0013295.

8. *Ivanova, L. E., Korjanova, Z. N., Varovin, I. A., Pronin, S. V., Harauzov, A. K. & Shelepin, Y. E.* (2016). Studying the interaction of rhesus monkeys with the tactile monitor during their observation of low-frequency visual test images. *Rossiiskii Fiziologicheskii Zhurnal imeni IM Sechenova*, 102(8), 931–939 [in Russian].

9. *Lind, J., Enquist, M., & Ghirlanda, S.* (2015). Animal memory: A review of delayed matching-to-sample data. Behavioural Processes, vol. 117, pp. 52–58. https://doi.org/10.1016/j.beproc.2014.11.019.

10. *Perdue, B. M., Beran, M. J., & Washburn, D. A.* (2018). A computerized testing system for primates: Cognition, welfare, and the Rumbaughx. *Behavioural processes*, 156, 37–50. https://doi.org/10.1016/j.beproc.2017.12.019.

11. *Plagenhoef, M. R., Callahan, P. M., Beck, W. D., Blake, D. T., & Terry Jr, A. V.* (2021). Aged rhesus monkeys: Cognitive performance categorizations and preclinical drug testing. Neuropharmacology, vol. 187, article no. 108489. https://doi.org/10.1016/j.neuropharm.2021.108489.

12. *Podvigina, D. N., Ivanova, L. E., Varovin, I. A. & Harauzov, A. K.* (2020). Studies of the Cognitive Abilities of Rhesus Macaques. *Neuroscience and Behavioral Physiology*, 50(2), 183–191. https://doi.org/10.1007/s11055–019–00886-x.

13. *Rodriguez J. S., Paule M. G.* (2009) Working Memory Delayed Response Tasks in Monkeys. In: Methods of Behavior Analysis in Neuroscience. 2nd edition. Ed. *Buccafusco J. J.* Boca Raton (FL): CRC Press/Taylor & Francis.

14. *Ryan, A. M., Berman, R. F. & Bauman, M. D.* (2019). Bridging the species gap in translational research for neurodevelopmental disorders. *Neurobiology of learning and memory*, 165, 106950. https://doi.org/10.1016/j.nlm.2018.10.006.

15. *Shelepin, Y. E., Fokin, V. A., Harauzov, A. K., Pronin, S. V. & Chikhman, V. N.* (2009). Location of the decision-making centre during image shape perception. *Doklady Biological Sciences,* 429(1), 511–513. DOI: 10.1134/S001249660906009X

16. *Taubert, J., Flessert, M., Liu, N., & Ungerleider, L. G.* (2019). Intranasal oxytocin selectively modulates the behavior of rhesus monkeys in an expression matching task. Scientific reports, vol. 9, no. 1, pp. 1–13. https://doi.org/10.1038/s41598–019–51422–3.

17. *Weed, M. R., Taffe, M. A., Polis, I., Roberts, A. C., Robbins, T. W., Koob, G. F., & Gold, L. H.* (1999). Performance norms for a rhesus monkey neuropsychological testing battery: acquisition and long-term performance. *Cognitive Brain Researc*h, 8(3), 185–201. https://doi.org/10.1016/S0926–6410(99)00020–8.

Chapter 17.
Symmetrical Patterns in Natural Images

Sergey Pronin

Pavlov Institute of Physiology RAS,
St. Petersburg, Russia

Many studies in image analysis are based on the assumption that images can be described using a limited set of patterns that form a kind of visual alphabet. Elements of these alphabets can be obtained either from a theoretical model describing the basic image structures, or extracted directly from training samples. The second approach is used mainly for narrow classes of images, since in this case it is possible to obtain workable alphabets of small sizes. Most of the alphabets considered in the first approach, on the contrary, are non-specialized. Each element of such alphabets is a function that depends on a number of parameters — orientation, scale, contrast, degree of blur, etc. The types of these functions are most often selected from the same set: these are straight or curved edges and lines, round and elliptical spots, line ends, line intersections and branches, line and edge intersections, corners, areas of constant brightness gradient, sometimes — "saddle-like" function [1,2]. Obviously, all of these patterns are often found in images belonging to a wide variety of classes. But the following question is of interest: is the set of "basic" patterns really exhausted by this set of functions, or are there some frequent patterns that are not included in it? In some works (e. g. [2]) a number of other functions are also used, but they are selected from one or another heuristic consideration. In order to obtain such a set of patterns experimentally, it is necessary to apply the second approach to a sample of images belonging to a fairly wide class. This article describes the results of an experiment where images of natural scenes are used as such a class. This is a fairly broad class of images, and the resulting alphabet of patterns can be expected to be applicable outside of this class. An additional argument in favor of this choice is that the visual systems of animals and humans are adapted to this class of images, and the results obtained can be useful for research in the field of visual physiology.

A typical approach to obtaining an alphabet from a set of images is based on the use of cluster analysis algorithms. The original images are divided into fragments of N*N pixels in size, where N is usually chosen equal to 3–15

pixels. Each such fragment corresponds to a point in the N^2-dimensional image space. It was shown that in the case of images of natural scenes, a significant part of these points are grouped in the region of nonlinear manifolds of relatively low dimensions [3]. Cluster analysis allows these manifolds to be approximated by sets of points that form the visual alphabet.

In the simplest version of this approach, the centers of N*N-pixel windows are located in each pixel of the image, and all the resulting fragments are used for further analysis. This is acceptable in the case of some narrow classes of images, for example, for homogeneous textures. But in the case of images of natural scenes, the absence of preliminary selection of fragments leads to obtaining samples of unacceptably large sizes. The first reason is that the characteristic details of the image, which can be elements of the visual alphabet, when dividing the image into fragments, will have an arbitrary position and orientation relative to the window. As a result, even if a characteristic image detail has a fixed size and shape, it will generate many different fragments that differ only in the relative position of the window and a given image detail. As a result, the size of the resulting alphabet will be much larger than the actual number of types of characteristic parts. Another factor requiring preliminary selection of fragments is the presence of large areas in natural images containing random, noise-like textures. The image space points corresponding to fragments of such textures obviously will not form compact clusters, but they can constitute a significant part of the original sample, increasing its volume and time required for processing.

One of the possible ways of such selection is the use of only fragments with centers at special key points located in the most informative places of the image — corners, intersection points of lines, in the centers of compact spots. As an example, we can cite works [4,5], where the Harris detector was used to search for key points [6]. Although this approach eliminates the variability in the relative position of the window and image details, the existing algorithms for finding key points do not always allow detecting all informative image areas. In addition, the use of key point detectors is undesirable in the case when the obtained elements of the alphabets are supposed to be used as visual stimuli in physiological experiments: there is no guarantee that the visual systems of animals and humans use the same algorithms for selecting key points as these detectors.

In this work, the presence of a sufficiently high degree of symmetry of one type or another is proposed as a criterion for the selection of fragments. Let's say that at some point in the image there is a symmetrical pattern with an axis

and a center of symmetry. A square window with a size of N*N pixels also has a center and axes of symmetry, and the position of the window can be chosen so that its symmetry elements coincide with the corresponding elements of the pattern. This can significantly reduce the variability of the resulting image fragments. In addition, selecting only symmetric patterns allows you to discard most fragments of random textures, which most often do not have symmetry.

Of course, this approach will discard all asymmetric patterns, which can also be elements of visual alphabets. However, there are a number of reasons why symmetrical elements are more important for describing images. Objects of interest usually have at least local symmetry. All living organisms have one or another symmetry, local or general, the same can be said about artificial objects. For example, all symbols (letters, numbers) are composed of line segments and arcs with at least mirror symmetry. As for asymmetric patterns, they are more typical for irregular background textures—ground, tree crowns, clouds, etc. The importance of symmetric patterns is also supported by the presence of a symmetry detection mechanism in the human visual system, which is triggered at the early stages of image analysis. And finally, most of the alphabets described in the literature, based on theoretical models, consist only of elements with one or another symmetry.

Method

The classification of the types of symmetry that can be possessed by compact fragments of images is described in [9]. The final set of symmetry types used in this work is as follows:

1. Translational symmetry, that is, symmetry with respect to displacement by an arbitrary amount. An example of patterns that exhibit this type of symmetry are fragments of straight edges or lines. These fragments can be described by one-dimensional profiles (sections). Three types of these symmetries were used. For the first type, T_{even}, the profile is described by an even function, for the second, T_{odd}—by an odd function, for the third, T_{asym}—by an asymmetric function that is neither even nor odd.

2. Mirror symmetry about one axis. Two types of this symmetry were used. The first type, M_1, is symmetry with respect to reflection without changing the brightness of the pixels. The second type, NM_1, is symmetry with respect to a combination of reflection and contrast negation.

3. Types of symmetry, which are characterized by the presence of a center

of symmetry — a point that remains stationary during transformations, while all other points are displaced. The following types were used:

3a) Symmetry with respect to rotation at an arbitrary angle about the center of symmetry ($R_\ast$). An example of such patterns is round spot or ring.

3b) Symmetry with respect to 120n° rotation, where n is an integer (R_{120}).

3c) Symmetry with respect to 180n° rotation, where n is an integer (R_{180}).

3d) Symmetry with respect to a combination of 180n° rotation and contrast negation (NR_{180}).

Thus, a total of 9 types of symmetry were used.

Since the patterns in images of real natural scenes cannot be absolutely symmetrical, the following criterion for the approximate symmetry of an image fragment was used. If $\mathbf{I}$ is a fragment, F is a transformation that sets the type of symmetry, and $F(\mathbf{I}) = \mathbf{I'}$, then fragment $\mathbf{I}$ is considered symmetric if the Euclidean distance between $\mathbf{I}$ and $\mathbf{I'}$ does not exceed the threshold D_{max}. To calculate this D_{max} value, the same set of images of natural scenes was used, but these images were transformed into random textures with the same basic statistical characteristics as the original images. For this, its spectrum was calculated for each image, the phase component of this spectrum was replaced by a random one, and then the image corresponding to this spectrum was calculated using the inverse Fourier transform. The images obtained in this way have the same mean brightness values, variances and autocorrelation functions as the original ones, but the structure of the objects depicted on them is completely destroyed. Further, for fragments of such "randomized" images, the distribution of the values of the Euclidean distances between $\mathbf{I}$ and $F(\mathbf{I})$ was calculated. As Dmax was chosen the distance which provided the probability of detector triggering equal to 10^{-5} (on "randomized" images). For each type of symmetry, its own Dmax value was calculated.

Fragments of the image were circular and had a diameter of 13 pixels. The brightness values were taken at points located at the grid nodes with a step equal to two pixels; 29 points were placed inside such a circle. The grid could have an arbitrary orientation, the values of the brightness at its points are calculated by interpolation using the algorithm [10]. If the fragment had mirror symmetry (types M_1 and NM_1), then the orientation of the grid was chosen to coincide with the orientation of the axis of symmetry. To describe the profiles of patterns with translational symmetry, we used sections with a length of 13 points, with a distance between the points also of 2 pixels. The choice of the

distance between the points, equal to two pixels, is due to the fact that the images before processing were smoothed by convolution with a Gaussian kernel with a half-width equal to 2.3 pixels. Smoothing is necessary to increase the accuracy of interpolation. All fragments were normalized so that their average brightness was 0 and the standard deviation was 1.

Each image was processed in three stages. At the first stage, areas of the image that did not contain high-contrast details were marked. For this, for each pixel, the contrast value was calculated using the formula $(B_{max} - B_{min}) /$ $(B_{max} + B_{min})$, where Bmax and Bmin are the maximum and minimum brightness in the vicinity of a given pixel with a diameter of 13 pixels. If the contrast was below 0.2, then the pixel was marked as low-contrast and was excluded from further analysis.

At the second stage, we searched for contrasting regions with translational symmetry (straight lines, edges, regions of constant brightness gradient, etc.). In each of these areas, there was a pixel with the maximum value of the gradient, and a section was drawn through it describing the brightness profile of the area. The profile was entered into the corresponding list, then all pixels of the area were marked and excluded from further analysis.

At the third stage, among the remaining unmarked image areas, we searched for fragments with symmetries of types M_1, NM_1, R_{180}, NR_{180}, R_{120}, and $R_¥$. The points of these areas were examined with a step of 0.5 pixels. If a fragment with approximate symmetry was found at some point, the position of its center (with an accuracy of 0.1 pixels) and orientation (with an accuracy of 1°) were determined to ensure the maximum degree of symmetry. If the distance between fragment **I** and F(**I**) did not exceed the D_{max} threshold, it was entered into the corresponding list of fragments.

After processing all images for each of the nine types of symmetry, the nine lists of fragments was obtained. Further, each of these lists was analyzed using the cluster analysis method. In the case of fragments of images of natural scenes, a significant part of points in the image space are grouped in the region of nonlinear manifolds, which can have a complex shape. Therefore, the DB-SCAN algorithm [11] was chosen as an algorithm for cluster analysis, which allows working with clusters of arbitrary shape. This algorithm has two input parameters: the radius **e** and the minimum number of points N_{min}. The choice of the values of the parameters **e** and N_{min} is determined by the specifics of the problem being solved. In this work, we chose the values of the parameters that provide the division of the initial set of symmetric patterns into subsets with

clearly different properties (for example, negative and positive variants of the same class of patterns), and small variations in the values of the parameters should not lead to a significant change in these subsets. Since different clusters can have different point densities, the following technique was used. At first, the value of N_{min} was chosen in the range of 50–200, and the search for clusters with a high density of points was carried out. Then the value of N_{min} was decreased to 10, and the points that were not included in the clusters found at the first stage were analyzed. Small clusters (with less than 50 points) were discarded.

Results

The set of images used in this work consisted of 829 monochrome photographs, which depicted natural objects and textures at various scales — plants, animals, soil, snow, water surface, clouds. Of these, 229 images were taken from the database used in [12].

Fragments of the M_1 symmetry type with the parameters of the DBSCAN algorithm $\mathbf{e} = 0.01$ and $N_{min} = 200$ form two clusters with sizes of 3155 and 21791 points. Sample patterns from the first cluster are shown in Fig. 1 (a). The second cluster contains negative variants of the same patterns. Fig. 1 (a) shows that these patterns can be roughly divided into four subclasses: curved edge, corner, end of line, and a pattern that can be called a ledge. However, there are no sharp boundaries between these subclasses. The first three subclasses are part of the set of commonly used basic patterns described in the introduction. The fourth subclass, "ledge", is not included in this set.

(a)   (b)
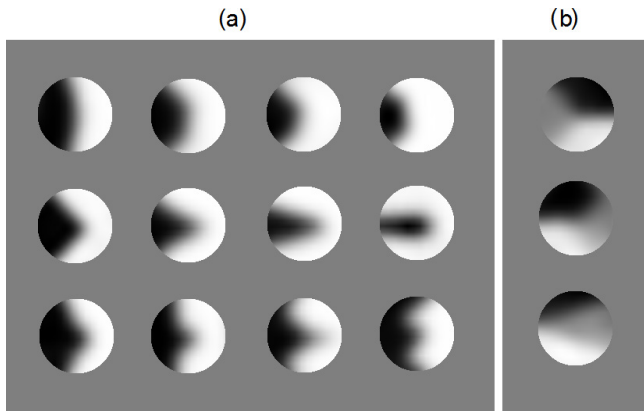
Fig. 1. Examples of patterns with symmetry types M_1 (a) and NM_1 (b).

Fragments of the NM_1 symmetry type at $\mathbf{e} = 0.01$ and $N_{min} = 50$ form two clusters. All elements of these clusters are variations of the patterns shown in Fig. 1 (b). They have two mirror-symmetric variants — "right" and "left", and each cluster contains only one of the variants.

For the R_{180} symmetry type, with $\mathbf{e} = 0.05$ and $Nmin = 100$, the patterns form two clusters, containing the same sets of pattern types, only in "positive" and "negative" variants. By choosing the values of $\mathbf{e}$ and N_{min}, these clusters can be divided into several subsets consisting of four types of patterns shown in Fig. 2 (a): elongated spots (first row), isthmuses (second row), intersections of two lines (third row), and patterns consisting of two closely spaced ends of lines or spots (bottom row). Elongated spots and intersections of lines are often used as elements of visual alphabets. "Isthmuses" and pairs of spots have a distant similarity to the "saddle" functions from [1,2]. But the "saddle" functions are symmetric with respect to the combination of 90-degree rotation and contrast negation, while the patterns in Fig. 2 (a) do not have this property.

(a)   (b)
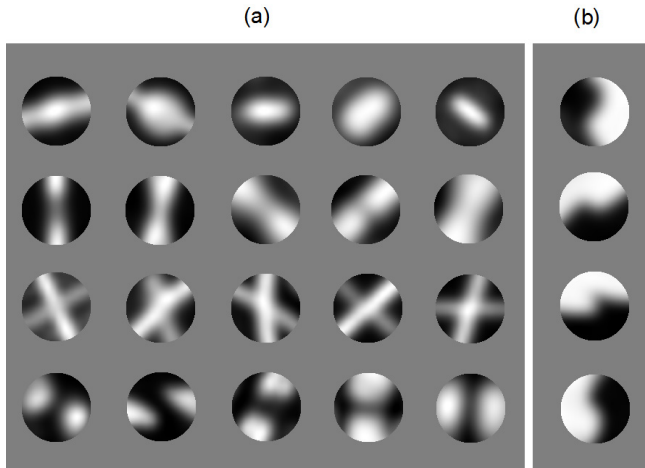
Fig. 2. Examples of patterns with symmetry types R_{180} (a) and NR_{180} (b).

For the NR_{180} symmetry type at $\mathbf{e} = 0.007$ and $N_{min} = 100$, the patterns form two clusters of close sizes (1787 and 1830) containing S-shaped patterns, some of which are shown in Fig. 2 (b). Like the patterns in Fig. 1 (b), they have two mirror-symmetric variants — "right" and "left", each of the clusters contains only one variant. Patterns of this type have not yet been used as elements of visual alphabets.

Fig. 3 (a) shows patterns with the R_{120} symmetry type; Fig. 3 (b) — type of symmetry $R_{\yen}$. In both cases, two clusters were obtained containing similar sets of patterns, only in "positive" and "negative" variants (both variants are shown in Fig. 3). All of these types of patterns are often used as elements of alphabets.

(a)   **(b)**
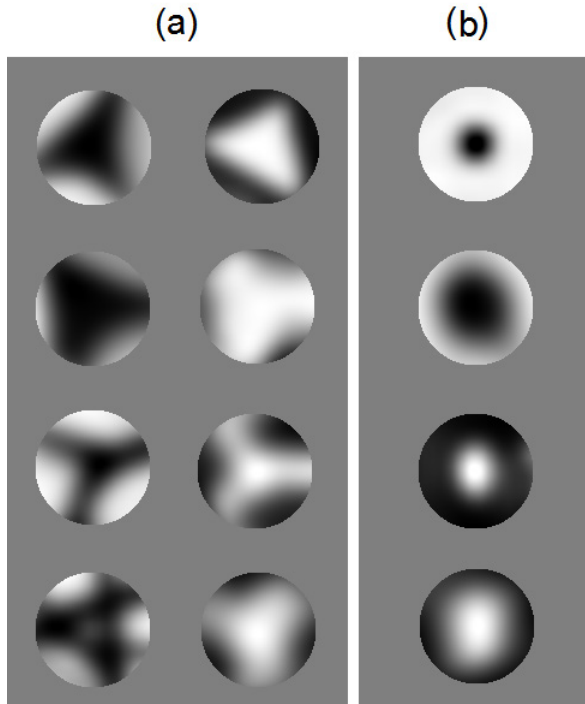
Fig. 3. Examples of patterns with symmetry types R_{120} (a) and $R_{\yen}$ (b).

Typical profiles of sections of regions with translational symmetry are shown in Fig. 4. Fig. 4 (a) shows examples of profiles for the T_{even} and T_{odd} symmetry types. Elements of this kind, corresponding to the sections of lines and edges, are very common and are a must-have element in almost all visual alphabets. Of interest are the profiles of sections with the T_{asym} symmetry type, examples of which are shown in Fig. 4 (b). They are neither edge nor lines. Despite the fact that such patterns are quite common (the number of regions with this type of symmetry is 23267, while for T_{even} it is only 9378), they have not yet been used as elements of visual alphabets.
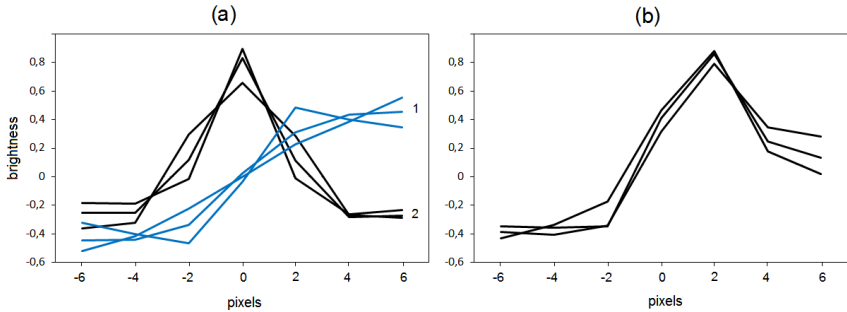
Fig. 4. Examples of profiles of patterns having symmetry types T_{even}, T_{odd} (a) and T_{asym} (b).

It should be noted that some of the elements frequently used in visual alphabets are missing from the obtained patterns. These are arc segments, T-shaped line intersections, and line and edge intersections. In the original lists of fragments, these patterns are found, but there are too few of them, and they do not form clusters with sufficient density. To eliminate this disadvantage, one can either increase the number of sample images, or supplement the set of sample images with images of artificial objects, in which such patterns are more common than in photographs of natural scenes.

A common feature for almost all the obtained patterns is the presence of mirror symmetry, at least approximate. This also applies to patterns of the T_{even}, T_{odd} and T_{asym} types, which are mirror-symmetrical about any axis that coincides in orientation with the brightness gradient. The only exceptions are the patterns in Fig. 1 (b) and Fig. 2 (b). Since the human visual system is adapted to images of natural scenes, it can be expected that the presence of mirror symmetry will be detected by it faster than other types of symmetries. The results of psychophysical experiments confirm this assumption [8]. For example, in [13], the recognition of symmetry in patterns of four types, consisting of points, squares, diagonal and rectangular segments, was investigated. For all types of patterns, the latency in recognizing figures with mirror symmetry was lower than for figures with only rotational symmetry.

Conclusion

In this work we have investigated compact symmetric patterns typical for images of natural scenes. The selection of image fragments with high

degrees of symmetry made it possible to obtain the set of samples of a relatively small size, which could be processed in a reasonable time by the DB-SCAN cluster analysis algorithm. Fifteen types of patterns were obtained, five of which were not previously used as elements of visual alphabets. Most of the patterns obtained have mirror symmetry, which is consistent with the results of research in the field of visual physiology, according to which mirror symmetry is recognized faster than other types of symmetry. The obtained experimental data can be used to form elements of universal visual alphabets, as well as visual stimuli for experiments in the field of visual physiology.

References

1. *Lillholm M., Griffin L*. Novel Image Feature Alphabets for Object Recognition // 19th International Conference on Pattern Recognition. Tampa, USA. 2008, Dec. 8–11. V. 1. P. 1–4.

2. *Aslan S., Akg*ül *C., Sankur B., Tunali T*. SymPaD: Symbolic Patch Descriptor // Proceedings of the 10th International Conference on Computer Vision Theory and Applications. Berlin, Germany. 2015, March 11–14. V. 2. P. 266–271.

3. *Lee A. B., Pedersen K. S., Mumford D*. The Nonlinear Statistics of High-Contrast Patches in Natural Images // International Journal of Computer Vision. 2003. V.54. P. 83–103.

4. *Agarwal S., Roth D*. Learning a sparse representation for object detection // Proceedings of the Seventh European Conference on Computer Vision. Copenhagen, Denmark. 2002, May 28–31. V.4. P. 113–130.

5. *Leibe B., Schiele B*. Interleaved object categorization and segmentation // Proceedings of the British Machine Conference / Ed. Harvey R., Bangham A. Durham, BMVA Press, 2003. P. 78.1?78.10.

6. *Harris C., Stephens M*. A combined corner and edge detector // Proceedings of The Fourth Alvey Vision Conference. Manchester, UK. 1988, Aug. 31 — Sept. 2. P. 147–151.

7. *Jurie F., Triggs B*. Creating efficient codebooks for visual recognition // 10th IEEE International Conference on Computer Vision, Beijing, China. 2005, Oct.17–21. V.1. P. 604–610.

8. *Wagemans J*. Detection of visual symmetries // Spatial Vision. 1995. V.9. № 1. P. 9–32.

9. *Griffin L*. Symmetries of 2-D Images: Cases without Periodic Transla-

tions // Journal of Mathematical Imaging and Vision. 2009. V. 34. № 3. P. 259–269.

10. *Keys R.* Cubic convolution interpolation for digital image processing // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1981. V. 29. № 6. P. 1153–1160.

11. *Ester M., Kriege H. P., Sander J., Xu X.* A density-based algorithm for discovering clusters in large spatial databases with noise // Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. Portland, Oregon. 1996. P. 226–231.

12. *van Hateren J. H., Ruderman D. L.* Independent component analysis of natural image sequences yields spatiotemporal filters similar to simple cells in primary visual cortex // Proceedings ofthe Royal Society of London. 1998. V. 265, P. 2315–2320.

13. *Royer E.* Detection of symmetry // Journal of Experimental Psychology: Human Perception and Performance. 1981. V. 7. № 6. P. 1186–1210.

Chapter 18.
Digital Image Procession during Video Endoscopy of Steam Turbine Blades

Liliana S. Rodikova, Valery V. Korotaev, Artyom K. Akhmerov, Aleksandr N. Timofeev, Aleksandr S. Vasilev, Gennady A. Shut, Victoria A. Ryzhova

ITMO University
Saint-Petersburg, Russia

The steam turbine is the main component of large power plants. One of the most important elements of a steam turbine is the rotor blade. The blades of the last stages are the most vulnerable to external damaging factors. They are most often subjected to serious damage due to droplet impact loads and erosion processes. It is very important to diagnose these damages quickly and immediately replace rotor blades, if a critical wear is reached [1-5].

In order to prevent damage to the turbine, timely detect and replace damaged elements, it is necessary to inspect all blades regularly. The time interval between repair and diagnostics sessions of steam turbines ranges from 6 to 2 years. Frequency and duration depends on the type of repair. For example, the duration of a major overhaul, depending on the capacity, is from 35 to 90 days. This inspection is carried out when the turbine is completely stopped [5-8]. The turbine is disassembled in order to have access to each blade. This takes a lot of time and causes large financial losses.

The ability to check the condition of the blades, to determine the residual life of the blade apparatus without opening the turbine will help to make a decision on the disassembly and repair of the turbine only when necessary. In general, examination of the condition of the turbine blades with endoscopic systems takes about two weeks [9-12], instead of several months, as in the case of a shutdown and disassembly of a turbine. Inspection with the use of video endoscopes will increase the time between repairs and thereby increase the economic efficiency of high-power installations.

Endoscopy systems are the most promising for solving this issue. The presence of the turbine barring will allow monitoring and diagnosing damage

to the rotor blades using a special video probe without opening the turbine [12-17].

A method was proposed that allows periodic monitoring of steam turbine blade wear using a video endoscope probe (video probe) on a closed cylinder. It is assumed that the video probe is inserted through a special technological hole into a closed cylinder. Further, in barring mode, all blades can be inspected within one rotor stage. The video probe includes a video camera and two light sources, captures an image of each blade (Fig. 1). The camera then transfers these frames to the synchronization unit. In the industrial computer, blade condition is analyzed using specially developed software. Based on the data obtained, the turbine operator can estimate the wear rate of the blade apparatus.
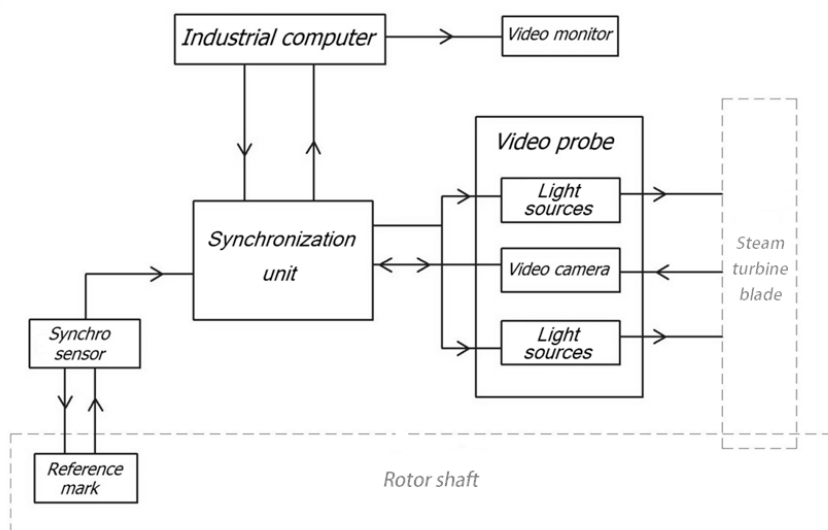
Fig. 1. Block diagram of the video endoscopy system.

The problem of rotor blade shape control at the closed cylinder of a steam turbine has been solved. This problem is complex due to the fact that the blade is a rather large object that moves at a high angular velocity at a short distance from the video probe.

The wear of the rotor blades is proposed to be estimated by the program method. The program measures the chord of the rotor blade in the captured images. Then, the calculated value of the chord of the blade is compared with

the initial (base) value. After clearing out the difference between these values, it becomes possible to determine the state of the leading and trailing edges and the presence of cracks on them.

The video endoscopy system includes a video probe, a sync sensor with a reference mark, a synchronization unit and a industrial computer with a video monitor. The video probe is the most important element of the endoscopy system as it captures the frames of the blades. It includes a light source and a video camera. To determine the index number of the inspected blade, a synchronization unit is required. It fixes a reference mark, which is attached to the turbine shaft. The images taken by the video probe are related to the synchronization unit. Then the information is sending to a industrial computer with a video monitor.

The proposed system provides the operator with information about the state of each blade. After that, the operator makes a decision on whether to continue turbine operation or on the need for repairs.

An algorithm is proposed for inspecting the turbine blade apparatus in the barring mode when the video endoscope probe is inserted into the turbine. During inspection the rotor rotates at a constant speed with a frequency of approximately 1,1 revolution per minute.

The proposed system operation algorithm is as follows. The video probe camera captures a frame of each blade, the image is transmitted to the information processing unit. In the information processing unit, the value of the chord of the rotor blade is calculated as the distance between the inlet and outlet edges in the section under consideration. The resulting value is compared with the base value stored in the computer. And then, based on the results of the comparison, a conclusion is made about to replace the blade.

The blade edge detection algorithm is based on digital image processing techniques. As an input data, we have the original grayscale image obtained by the video probe. The first step of the algorithm is to binarize the image. The next step is to find connected components in the image. Since the original image captures several blades at once, that is, the resulting image contains two connected components corresponding to two blades. An important condition for the analysis is that only one blade is analyzed at a time. For analysis, you need to determine which of the connected components in the image is the central one. This requires a calibration, that provides video probe with synchronization signals that control the imaging process.

After performing the calibration, it is known that the central pixel belongs

to the needed connected component. Further, the algorithm cuts off all non-informative connected components. The final step of the algorithm is to find the blade edges on the image and calculate the chord.

The first step to calculate the chord b_{px}, is to find the coordinates of the extreme point of the right edge (x_1, y_1). The extreme point of the right edge has the maximum vertical coordinate of the pixels included in the set. Vertical and the corresponding horizontal coordinate are the extreme point. The coordinates of the extreme point of the left edge (x_2, y_2) are found in a similar way. Next, the chord value b_{px} is calculated by finding the distance between the two extreme points. After that, the value of the chord of the blade is calculated according to the expression:

$$b_{px} = \sqrt{\left(x_2 - x_1\right)^2 + \left(y_2 - y_1\right)^2}$$

To estimate the total error in measuring the chord of the rotor blades, it is necessary to find out the main sources of primary errors of the system for measuring the geometric parameters of objects. And also evaluate the contribution of each component separately.

The influencing factors are: lens focal length error; error of distance from the axis of the rotor blade to the video probe; noise of the matrix receiver of the device; thermal deformation error; instability of the barring device.

The video probe camera captures the shots of the blades in the barring mode. Due to the fact that the object is in motion, a dynamic component is added to the main error. This is the error in the influence of the speed blur of the image.

The spatial arrangement of the video probe relative to the rotor blades is shown in Fig. 2. The value measured by the software is the segment b_i'. It is in image space. Based on the results of measuring the segment b_i', the value of the conjugate segment b_i (chord of the measured blade) is calculated. This segment is located in the plane of objects [17-23]. Segment B_i' is the size of the projection of the chord on the sensitive area of the DIS. The program calculates this segment for a given threshold value of irradiance. This size of the projection B_i' is used to calculate the required size of the chord b_i.

During diagnostics, the video probe moves along the rotor blade. When the probe moves, the chord size and the angle of the blade twist α_i change simultaneously. In order to compensate for these changes, it is necessary to orient the optical axis of the video probe at an angle γ_i to the OZ axis:

$$\gamma_i = \mathrm{arctg}(\frac{2\pi R_i}{b_i N \sin \alpha_i} + \frac{\pi R_i}{z_0 N} - \frac{D}{2z_0} - \mathrm{ctg}\alpha_i) \tag{1}$$
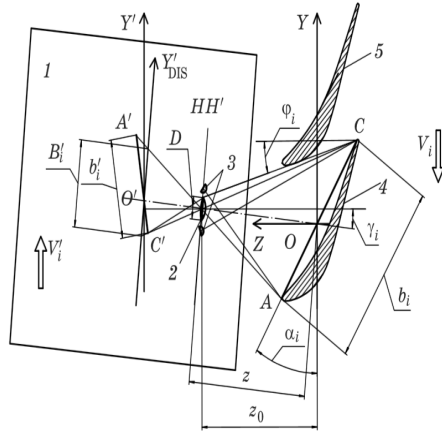
Fig. 2. Ray path in the optical scheme of the video probe when monitoring the rotor blade chord value: *1* – video probe, *2* –lens, *3* – illumination sources, *4* – cross section *i* of the rotor blade *n*, *5* – cross section *i* of the next *(n+1)*, rotor blade, A and C – inlet and outlet edges rotor blade.

The optimal position is when the optical axis of the video probe is perpendicular to the blade surface. Thus, for the distances from the shaft axis to the rotor blade, greater than the optimal distance, the equality of the angles $\gamma_i = \alpha_i$, must be fulfilled. In this case, the calculation of the size of the chord b_i will be carried out by the expression:

$$b_i = \frac{(z_0 - f')}{f'} B'_i \sqrt{1 + \mathrm{tg}^{2\,3}{}_i} \tag{2}$$

The resulting equation includes the sources of primary errors of the circuit parameters, the influence of which must be investigated. Thus, the calculation formula (2) contains the main parameters $\Delta B'$, Δz_0, Δf, $\Delta \alpha_i$. These parameters influence the result of calculating the value of the chord b_i. To assess the influence of each of these components, it is convenient to use the method of partial derivatives of function (2).

The MATLAB environment was used to create a computer model that allows to numerically estimate the total error $\delta b_{\Sigma i}$ and to tell how much it is affected by these factors.

Fig.3 shows the results of calculating the error components as a function of the distance to the rotor blades, which are the most subject to the influence of erosion.

The calculation was carried out with the following typical parameters of the blade system: the distance $R_i=1750\ mm$ from the axis of the shaft to the cross section of the blade, the chord $b_i=200\ mm$ being monitored, the angle $\alpha_i=40°$ at which the cross section of the chord is oriented, and the rotational rate $v=1.1rpm$ of the rotor.

The calculation was carried out with the following video probe parameters: focal length of the lens 3.5 mm, pixel size of the DIS 2.25×2.25 μm, frame rate 30 Hz, exposure time 0.8 of the frame time ($t_{exp}=0.8\ t_{fr}$), and $SNR=300$ (signal-noise ratio).

The relative value $\delta_i = \delta b_i/\delta b_{\Sigma i}$ was used to estimate the contribution of each error component. The calculation was carried out to estimate the impact of fabrication inaccuracy of the items and elements of the probe: relative error of the focal length, $\delta f_0=1\ \%$; indeterminacy error of the distance from the video probe to the rotor blade, $\delta z_0=0.1mm$; error caused by variation of the slope angle of the chord over the cross sections of the rotor blade, $\delta \alpha_i=1°$; indeterminacy error of the radius of the cross-section of the rotor blade, $\delta R_i=1mm$; and indeterminacy error caused by instability of the rotational rate of the rotor, $\delta v=1\ \%$.

To estimate the relative contribution and the influence of each factor on the total error, a computer model was created. The result is presented in the form of a graph in Fig. 3. The figure also shows the graphs of the dependence of all components of the error on the distance to the rotor blades z_o [24].

The calculated results based on the computer model shows that the following partial components have the strongest effect on the total error, in descending order: the indeterminacy error of the focal length of the lens of the video probe, δb_i^f; the error caused by variation of the slope angles of the chord over the cross-sections of the rotor blade, δb_i^α; the error caused by temperature distortion of the body of the probe, $\delta b_i^{\Delta T}$, and the error caused by indeterminacy of the distance from the video probe to the rotor blade, $\delta b_i^{z_o}$.

Partial components that have the insignificant effect on the total error are: the

error of determining the boundaries of the image of the blade on the DIS, δb^{DIS}_i; the error caused by the rotation-rate instability of the rotor, δb^v_i; the error caused by instability of the exposure time of the video probe, δb^{exp}_i and the error caused by the indeterminacy of the radius of the cross section of the rotor blade, δb^R_i.
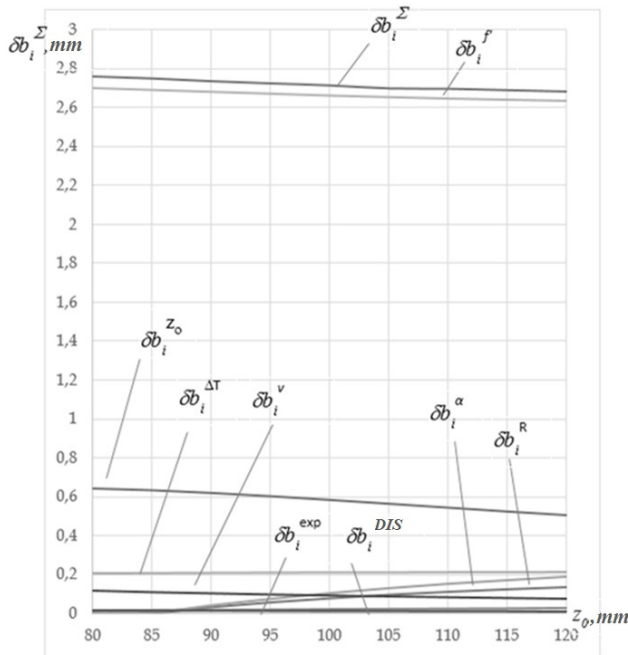
Fig. 3. Blade chord measurement error dependence on the distance to the blade.

The graph in Fig. 3 shows that the lens focal length error mostly determines the value of the overall error. The next in terms of the degree of influence are the error in setting the distance to the video unit, the error in the temperature deformations of the case and the error in the change in the chord tilt angle.

The two most influencing errors can be reduced to the magnitude of the DIS noise error when calibrating the system. This is possible because both of these errors are systematic. After the system is calibrated, the value of the total error drops significantly. This can be clearly seen in Fig.4.

Analysis of the degree of influence of each of the components of the blade chord control error after calibration showed that the total error decreases along

with increasing object distance. At a distance of 80 mm to the object axis, the total error is 0.29 mm. Thus, after calibration, the total error is reduced by a factor of 10.
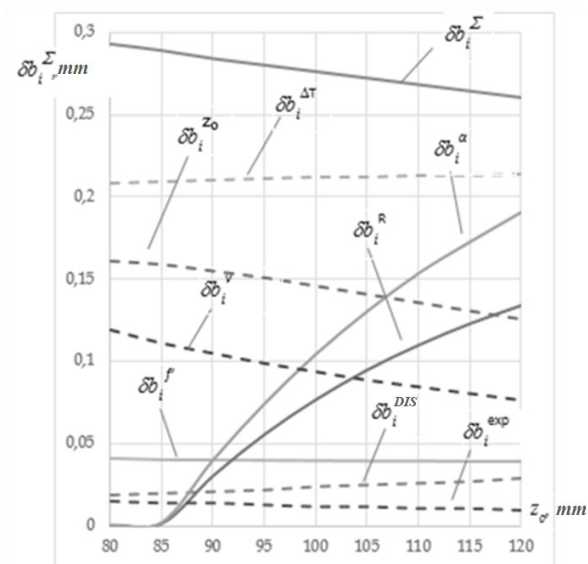
Fig. 4. Graphs of dependences of control errors after calibration.

In case when imaging is done while the turbine is moving in barring mode, dynamic errors are added to the static measurement errors. The dynamic errors are the high-speed blur of the image and the instability of the angular frequency of rotation of the turbine rotor. To study the influence of such errors and possible options for reducing their influence, an experiment was carried out on a special stand (Fig. 5).

This stand contains a shaft model with rotor blade models attached to it, representing a stage of a steam turbine. There is also a reference mark attached to the shaft. With each revolution of the shaft, the sync sensor fixes its position to determine the number of the blade on the captured frames. Rotational stage in the basement of the experimental stand simulates barring mode. The model also contains a personal industrial computer with a video monitor. A video camera with a pulse light source is used as a video probe in the stand.

An experiment was carried out, during which several frames were taken in a static mode and in a dynamic mode. First, shooting was carried out in static

mode, then in motion. As a result of the experiment, a series of frames was obtained in different operating modes.
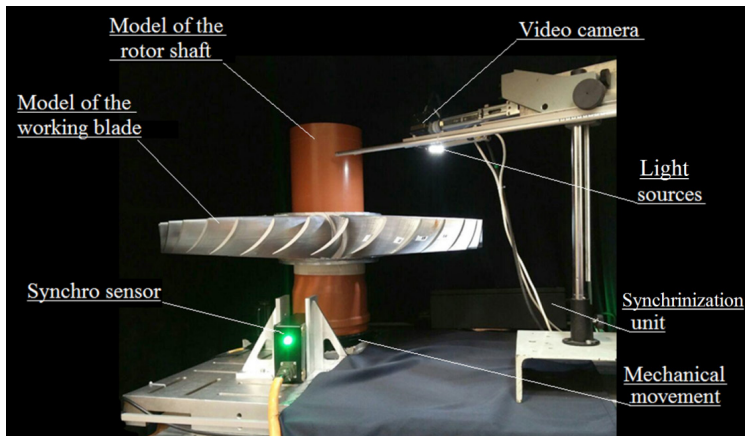
Fig. 5. Experimental setup for the dynamic error estimation.

Analysis is done by processing the received frames in the Matlab environment. Two graphs of the intensity distribution at a certain interval for the captured frames were obtained. From these graphs, one can judge the distribution of irradiance in the section of the rotor blade. After processing, both graphics of the video signals were aligned on the same coordinate system using Microsoft Excel and analyzed. The resulting graphs of the dependence of the irradiance on the pixel number of the selected section are shown in Figure 6.
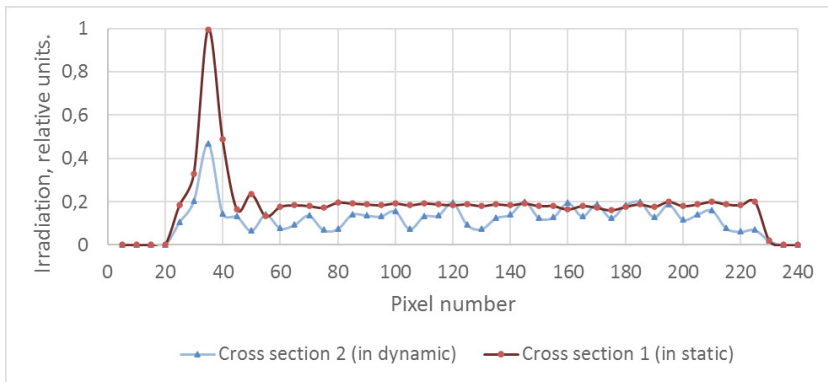
Fig. 6. Graphs of irradiance distribution on a blade.

The analysis of the plotted graphs showed that when taking shots of the blades in motion, the video signal level drops significantly. This is especially evident from the peaks on the charts. Also, when shooting in motion, the graph turned out to be smoother at the edges, which also allows us to conclude that it is necessary to increase the backlight level. The peaks of the obtained graphs do not coincide, which is another evidence of image blur, since the frame was captured in dynamics.

Thus, in the course of the study, it was found that the high-speed blur of the image significantly affects the error in measuring the wear of the blades. To reduce this influence we propose exposure time reduction and the backlight power increasement.

The estimation of the total error in determining the chord value of the tenth rotor blade b_{10} during its movement (in dynamics) and when stopped (in static) was carried out according to the software implementation of algorithm on the computer.

The study results showed that for a section of the tenth rotor blade with the design chord size $b_{10} = 103\ mm$ (for the section radius $R_i = 350\ mm$, the chord tilt angle $\alpha_i = 32°$, the power of the LED is equal, the measured value of this quantity is *103.09 mm* with an standard deviation *0.86 mm* in dynamics. During stopped the measured value of this quantity is 103.09 mm with an standard deviation 0.86mm.

Conclusion

An architecture of a video endoscopy system was developed for wear monitoring of the blades of low-pressure cylinders of steam turbines. The control is proposed to be carried out in the prohibition mode on a closed low pressure cylinder.

Algorithms were formulated for calculating the length of the chord of the rotor blade from the captured images using image processing. These algorithms makes it possible to measure the chord of the steam turbine rotor blades.

An image processing algorithm is described. It is proposed to estimate the blade wear according to the size of the chord of the rotor blade. The algorithm calculates the distance between two extreme points in each section of the blade. The calculated value is compared to the initial base (manufacturing) value.

A method is proposed for calculating the error in controlling the chord size of the rotor blades in the barring mode at a closed cylinder.

The total error of the chord size measurement of the rotor blades in the barring mode and in case of stopped turbine has been calculated. The main components of the total error have been identified. The ways to reduce the impact of the most influencing components of the final error are proposed.

Theoretical and experimental studies of the error components have shown that the following error components cause the accuracy reduction: error due to the uncertainty of the focal length of the camera lens, the error caused by the temperature deformation of the probe body, the error caused by the uncertainty in the distance from the video camera to the rotor blade and the error caused by changing the angle of inclination of the chord along the section of the rotor blades.

Experimental studies of the error of the video endoscopy system were carried out on a special experimental stand. The influence of the dynamic components of the video endoscopy error is analyzed.

Experimental studies have generally confirmed the results of computer simulations.

Proposed system and method allow to reduce the time for periodic repairing works of steam turbines of various types and make it possible to increase the economic efficiency of enterprises in the region and in Russia.

Acknowledgements

References

1. *Danilin A. I.*, *Adamov S. I.*, Chernyavskii A. Zh., *Serpokrylov M. I.* Diagnostics and monitoring of the operating condition of steam-turbine blades// Power Technology and Engineering.— 2008.— № 3.—C. 65–69.
2. Remote visual inspection for nde in power plants, Peter Samsonov, Olympus Corporation; IFD, Centerpointe Drive; Suite 430 LaPalina, CA 90623 670–8500.
3. *Pears N., Liu Y. and Bunting P.* 3D imaging, anal y sisand applications. London, UK: Springer Verlag, 2012. 500 p. Chapter 2, pp. 35–94.
4. *Gorevoy A. V.*, *Machikhin A. S.*, *Perfilov A. M.* Determination of the error

of non-contact measurement of the area of defects on surfaces of complex shapes during video endoscopic control // Scientific and technical bulletin of information technologies, mechanics and optics 2014. No. 4C. 140–148 (in Russian).

5. Patent 2655479 C1 Russian Federation, IPC G06T7 / 00. Triangulation method for measuring the area of surface areas of internal cavities of objects of known shape / *Gorevoy A. V.*, Machikhin A. S ., *Pozhar V. E.*, *Kaloshin V. A. Kologov A. V,* Appl. 28.09.2012, publ. 17.02.2017 (in Russian).

6. *Schlobohm J., Pösch A., Kästner M., Reithmeier E.* On the development of a low-cost rigid borescopic fringe projection system // Proc. SPIE9450, Photonics, Devices, andSystems VI, 6 January 2015.

7. *Korczewski Z.* Endoscopic diagnostics of marine engines // DIAGNOSTYKA.— 2007.— № 7.—С. 19–25.

8. *Korczewski Z.* Contemporary diagnostic methods for ship engines: a report on scientific research activity of Polish Naval Academy in this field // POLISH MARITIME RESEARCH2 .— 2008.— № 15.—С. 46–58.

9. *Agarwal A. K., Singh A. P., Agarwal A.* Evolution, trends and applications of endoscopy in internal combustion engines // Journal of Energy and Environmental Sustainability.— 2016.— № 1.—С. 56–66.

10. Patent 2778645 B1 European patent specification, Int Cl.: G02B23/24. Turbomachine component monitoring system and method/ Ward Jr., John David Greenville, Batzinger, Thomas James Schenectady, Janawitz, Jamison William Atlanta, Thompson, Christopher Edward Greenville., Date of filing: 11.03.2014, Date of publication 22.06.2016.

11. Patent Number WO 96/12873 F01D5/00, 21/00, A system for repairing damaged gas turbine engine airfoils/ Owens James, Nelkin Road Date of filing: 15.9.1995, Date of publication 02.05.1996.

12. *Lei Y., Cheng Y., Miller S. F.* A Flexible Endoscopic Machining Tool/ Energy Procedia.— 2012.— № 16.—С. 1033–1040.

13. *Zhongda Y.* Borescope Inspection Management for Engine// ICEESE.— 2017.—С. 1–4.

14. Patent JP 4358494 B2 2009, G02B23 / 24. Endoscopic system / Inventor Katsutoshi Tsuzuki, Hatagaya-chome, Shibuya-ku, Tokyo., declared 02.10.2002, publ. 04/11/2009.

15. Patent 2610973 C2 Russian Federation, IPC G02B23 / 24. Endoscopic system and method for examining gas turbines of a turbine / Müller V., Tiis M., Herluk K., Appl. 28.09.2012, publ. 17.02.2017 (in Russian).

16. Patent No.: US10310250 B2 CPC G0223/2484. Method and apparatus for inspection of moving parts in rotary system/ Andrew Paul Ramsbottom, Gareth Willerton Sykes, date 02.02.2017, Date of Patent 04.06.2019.

17. *Puzyrev E. I.*, *Shut G. A.* Condition monitoring device inside turbine units and parts of steam turbines // RF Patent No. 2624380 (in Russian).

18. *Rodikova L. S.*, *Korotaev V. V.*, *Shut G. A.*, *Timofeev A. N.* Investigation of the video endoscopy system for the rotor blades of steam turbines // XIV International Conference "Applied Optics-2020" (December 15–18, 2020): Proceedings of the conference — 2020 (in Russian).

19. *Gryazin G. N.* Foundations and systems of applied television / G. N. Gryazin. Gryazin; edited by N. K. Maltseva. — SPb.: Polytechnic, 2011–273 p. (in Russian).

20. *Ishanin G. G.* "Sources of optical radiation", Moscow, 2008 (in Russian).

21. *Latiev SM* Designing of precise (optical) devices. SPb: Lan, 2015. 560s. (in Russian).

22. *Rodikova L. S.* Investigation of the influence of the parameters of an opto-electronic video endoscopy system for measuring wear of steam turbine blades on a closed cylinder // Proceedings of the X Congress of Young Scientists 2020 (in Russian).

23. *Shut G., Korotaev V., Puzyrev E., Ryzhova V. A., Timofeev A. N., Akhmerov A. K., Rodikova L.* Videoendoscopy of working blades of steam turbines and control of their geometrical parameters // Journal of Optical Technology — 2020, Vol. 87, No. 11, pp. 677–683.

24. *Rodikova L. S.* Investigation of changes in the video signal during the transition to dynamic measurement of wear of steam turbine blades during barring // Proceedings of the IX Congress of Young Scientists 2020 (in Russian).

Chapter 19.
Review of the State of Elaboration of Thermal Imaging Technology: Matrices and Infrared Cameras

Pavel A. Safonov, Valery V. Korotaev

ITMO University, Saint-Petersburg, Russia

The object of this work is the market for thermal imaging equipment. The research was carried out in several stages. The key factors that have led to a change in the structure of the market in recent years are highlighted, the situation is considered and causal relationships in the global and Russian markets are identified, the subsequent comparative analysis allows us to assess how the state of affairs with thermal imaging equipment in Russia correlates with global trends. A review of receivers of optical radiation of the infrared range and optical materials is made.

As for the market overview and main market trends. Since its inception, the global market for thermal imaging technology has grown primarily due to military applications. Today the military sector provides the market with a certain growth, but the paradigm of its development has changed — the main growth of the market provides a civilian direction. A steady trend has been noted in recent years: thermal imaging equipment is gradually moving from an expensive, highly specialized market segment to a mass consumer segment (from attachments for telephones to monitoring and control devices), which gives the greatest profit.

If we talk about the geography of producing countries, it is quite obvious that these are primarily economically developed countries — the USA, France, Great Britain, Israel, Germany. This is easily explained — the development of this topic requires high technologies and was initially stimulated by defense orders. Today the leaders are such companies as Raytheon, Sofradir/Ulis, SEEK, British Aerospace, Dahua Technology and a number of others [1].

Nevertheless, in recent years, a quite understandable trend has emerged on the global market for thermal imaging equipment: the emergence of new manufacturers from South Korea, China, Japan and other countries of Southeast Asia, which intensify the competition, which, as a result, makes thermal imaging equipment more affordable. The analysis made it possible to single

out the following factors that allow companies to be among the leaders of the modern global market: implementation of innovative technologies that minimize weight and size characteristics; economic efficiency of products achieved by mass production; expanding the scope of devices; focus on the civil sector; attractive affordable price niche for the consumer.

The global market for thermal imaging systems was estimated at US $5.19 billion in 2020 and is expected to reach US $7.92 billion by 2026, with an average growth rate of 7.78 % — 9.84 % over the forecast period (2021–2026) [2]. Now these numbers may be significantly higher, experts say. Confirmation of the above was the COVID-19 pandemic, which was an additional stimulus in the wide development of thermal imaging technology, and in the civilian sector, which in turn caused the growth of the global market.

Concerning the Russian market. The Russian market has a certain specificity — a significant part of domestic instrument manufacturers use imported components. An example is thermal imaging sights for hunting. The three most popular are Russian-made products. However, the matrices for these scopes are produced by France, which increases the cost of the product. In addition, production may be in jeopardy in the event of tougher anti-Russian sanctions. Another example: in the Russian Federation, until a certain time, sensors in the LWIR range were not mass-produced, often using manufacturers from Southeast Asia.

As a result, there are products of good quality, developed in Russia, with domestic software, but assembled on imported element base. Today, the appearance of new participants has become characteristic of the domestic market of thermal imaging equipment, as well as the global one. But the reasons are of a different nature. If in the early 2000s the share of thermal imagers for measurements made by FLIR was 90 %, then the policy of sanctions against our country and price pressure led to the closure of the Russian branch of the company, and the niche was occupied by competitors from China, Germany, and Japan. At the same time, we note that in recent years, thermal imagers have been developed and produced, consisting entirely of domestic components. Back in 2016, Russia became the fourth country in the world after the United States, France and China to create its own thermal imaging matrix. A number of enterprises of JSC Shvabe and JSC Roselectronika, the Russian Academy of Sciences, and private enterprises are developing photodetectors for various purposes in Russia. As of early 2020, the following key figures can be identified on the Russian market (market share is given in brackets): Hikvision (28 %), Dahua (21 %), FLIR Systems (18 %), Axis Communications (16 %), Shvabe (11 %).

Experts are confident that multiple growth of the market volume is possible with an increase in the share of participation of Russian companies. Today

in Russia they are already developing and producing thermal imagers consisting entirely of domestic components. In particular, this is the responsibility of the Shvabe holding of the Russian Technologies State Corporation. Among the major manufacturers are Astrohn DB JSC, JSC Cyclone Central Research Institute of the Roselectronika holding, Vologda Optical and Mechanical Plant JSC of the Shvabe holding, which continue to improve their products.

So, speaking about the classification according to the scope of application, traditionally all thermal imagers are divided into three large subgroups: for civil, industrial or military applications.

When it comes to civilian applications, it must be emphasized that household appliances are an example of the most common use of thermal imagers in everyday life. The civic sector is in increasing demand from year to year, as a result, it receives more funding and development goes in many directions. For example, a thermal imager that measures temperature has become one of the symbols of last year due to the COVID-19 pandemic.

All this may mean that a thermal imager for correctly and quickly determining the temperature of a person in a stream of people, including using artificial intelligence, is a new reality and for a very long time. According to Mordor Intelligence, a US analytical company, the demand for thermal imaging devices will only increase.

Speaking about industrial devices, it should be noted that they solve a wide range of complex problems. And here two-band scanning systems made from domestic components are of certain interest—for example, the AS-TROHN-4K video surveillance complex or the development of the SON-730 JSC "PO UOMZ". Such systems, sometimes called hybrid systems, have two cameras operating in the visible and infrared ranges, and then the thermal and television images are completely combined in one frame. The software is also a domestic development.

Taking into account the transparency windows of the atmosphere and the distance to the observed object, the most effective spectral range from the point of view of thermal contrast is a thermal imaging channel with an operating range of 3–5 µm. In addition, a video channel with an operating range of 0.38–0.78 µm was used. This approach allows obtaining the most complete information by combining the characteristic features of objects in the corresponding spectral ranges of each channel.

Speaking about the demand for thermal imagers in various areas of economic activity in Russia, most experts agree on the following distribution: security and safety systems are the leaders, followed by medicine, energy and industry.

The driver for improving thermal imaging systems is the development of the element base — new optical systems, photodetectors (detectors), signal processing circuits, display systems, etc.

The work considered IR detectors as one of the most important components of a thermal imager. The whole variety of detectors can be divided into two classes: photon detectors and thermal radiation detectors (uncooled), primarily microbolometers.

Photonic detectors are used in cases where a particularly high contrast sensitivity and long detection range are required. A significant drawback is the need for a cooling system.

Microbolometers are less sensitive than photon detectors. But to date, significant results have been achieved in the creation of uncooled thermal detectors of optical radiation in the IR range, which, firstly, are comparable in their characteristics to photon detectors (in terms of speed and sensitivity), and secondly, are much cheaper. So, in 2018, the specialists of Astrohn DB JSC managed to develop the design of a pixel of a microbolometric matrix, which has equal absorption coefficients in two spectral ranges — 3 5 and 8 14 µm.

Therefore, the number of uncooled bolometers produced today is several times higher than the market for infrared matrices of other types, especially in the civil sector, which does not feel the need for high detective ability and is ready to sacrifice it for the sake of a low price.

Today, the main format of matrix photodetectors is de facto 640×512 elements with a step of 15 µm. Experts believe that in terms of "price / quality" this format will retain this position until approximately 2025. At the same time, a number of leading developers are introducing the megapixel format of 1280×1024 elements into the commercial market, although their prices are still quite high, which makes it difficult massive transition to this format.

When considering groups of thermal detectors, divided according to their principle of operation, and materials for their manufacture, we note that various silicon-based materials have shown promising qualities for microbolometric arrays, especially using the latest manufacturing technologies. Looking at the wide range of materials used for sputtering in microbolometers, we can see that the methods of making materials for microbolometers will continue to improve, leading to lower costs and higher efficiency and productivity. Extensive research on these materials indicates that in the near future the market will be much more focused on commercial sectors using a variety of materials and growing methods, resulting in a wide variety of high quality devices. Currently, the most commonly used bolometer material is still vanadium oxide, followed by amorphous silicon with several new silicon derivatives showing

promising potential; other microbolometers are rarely used. The reason for such a frequent choice of a vanadium oxide microbolometer is the optimal combination of a high temperature coefficient of resistance, a high signal-to-noise ratio, and manufacturability.

From the analysis of various sources devoted to this topic, it becomes obvious that the growth of consumers of uncooled bolometers is manifested precisely at the expense of the civil sector, which does not feel the need for high detective ability and is ready to sacrifice it for the sake of a low price. In other words, it is often about detection and, possibly, recognition, but not about identification.

A comparative analysis of the parameters of uncooled microbolometric matrices from various manufacturers led to the conclusion that Russian products [3] are close in terms of their performance to the world level. As an example, the table contains data from a number of serial products of leading enterprises. The ASTROHN-64017–1 thermal imaging detector of the Astrohn company was taken as a sample of Russian products. This photodetector operates in the spectral range of 7–14 μm, has a typical for a number of similar products, the step of photosensitive elements equal to 17 μm. In addition, the detector has a good resolution; the vanadium oxide film has a number of advantages, which we mentioned above. The indisputable advantages are resistance to external influences and built-in thermionic stabilization, as well as an economic component.

Brand (Country)	Sofradir/ULIS (France)	Raytheon (USA)	North Guang-wei Technol-ogy-GWIC (China)	FLIR Systems (USA)	Astrohn DB JSC (Russia)
Detector	Pico 1024E	AAQ-27	GWIR0302X1A	Tau640	Astrohn 64017–1
Technology	a-Si	VOx	VOx	VOx	VOx
Multiplexer matrix size	1024 × 768	2048 × 1024	640 × 512	640 × 512	640 × 480
Range of spectral sensitivity, μm	8..14	7..13	8..14	7,5..13	7..14
Pixel size, μm	17	17	17	17	17
NETD, mK	60	<50	40	<50	40

Photonic matrices are also produced in Russia — these are NPO Orion, the MZ Sapphire JSC (for a spectral range of 3–5 μm) and Astrohn DB JSC together with IPP SB RAS [4] (8–12 μm).

Manufacturer	NPO Orion	MZ Sapphire	JSC DB Astrohn and IPP SB RAS
Detector	ALPHA –9022, ALPHA –1032, ALPHA –1964	FUR-145C, FD-121, FRO-139, FD-346	Cooled MFPs of the third generation
Technology	Si	Si, Ge, InSb и CdHgTe	HgCdTe
Range of spectral sensitivity, μm	3–5	3–5	8–12

Now let us dwell on the situation with the development of matrix photo-detectors of the infrared range in Russia.

A number of enterprises concentrated in Shvabe JSC and Roselectronika JSC, in the Russian Academy of Sciences, as well as private enterprises, are engaged in the development of photodetectors for various purposes in Russia. The main suppliers are NPO Orion JSC and MZ Sapphire JSC, which are part of Shvabe JSC, as well as the private enterprise Astrohn DB JSC. NPO Orion JSC develops and manufactures cooled and uncooled photodetectors. MZ Sapphire JSC manufactures cooled and uncooled MPAs based on Si, Ge, InSb, CdHgTe (MCT: cadmium-mercury-selurium). JSC «Research Institute Polyus» develops uncooled MFPs based on InGaAs and the production of photosensitive semiconductor structures.

The enterprises of Roselectronika JSC specialize in the development and production of silicon-based matrices of the visible range; cooled MFPs based on quantum wells; a Schottky barrier made of platinum silicide and cooled FPUs based on impurity silicon (NPP Pulsar JSC, CRI Electron JSC, CRI Cyclone JSC, NPP Vostok JSC). The Institute of Semiconductor Physics of the Siberian Branch of the Russian Academy of Sciences develops semicon-ductor materials science and MPA based on CdHgTe, InAs, microbolometers and quantum wells. Astrohn DB JSC (Lytkarino) develops and manufactures thermal imaging devices for civil purposes based on uncooled MPDUs of its own production, as well as cooled MPDUs based on CdHgTe, together with IPP SB RAS.

Astrohnc DB JSC has started serial production of a matrix photodetector module in the long-wavelength range based on a cooled CMT / Si matrix (produced by the Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences) and its own micro-refrigeration system As-trohn-MKS500. heat recovery, ultrapure gaseous helium is used as a working fluid.

A few words about materials for optical elements, since the sensitivity of a thermal imager largely depends on the lens. The materials used today for

the lenses of infrared cameras are germanium, zinc selenide and chalcogenide glasses (an example is the IKS series of glasses).

Germanium transmits radiation in the wavelength range 1.8–23 µm. This makes it possible to effectively use it for the production of optical components of thermal imaging systems in the range of 3–5 µm. IR lenses, due to the weak light flux emanating from emitting objects, must have a large light transmission and at the same time a large relative aperture (the ratio of the diameter of the entrance pupil of the lens to its focal length). In addition, germanium is a non-hygroscopic material, which makes it a good infrared transmitter. The ability of germanium to provide improved carrier mobility and high non-linear coefficients will drive the growth of the infrared camera market in the near future.

Germanium has a high temperature coefficient of refractive index, which leads to a change in the focal length of lenses depending on temperature. Therefore, a passive athermal design is used, which allows mechanically compensating for this defocusing and using the devices autonomously. Such lenses are mass-produced at the Astrohn Design Bureau. As an example, the following products of this company can be cited: ASTROHN-30AT08, AS-TROHN-50AT08, ASTROHN-75AT10, ASTROHN-100AT14.

Zinc selenide for infrared cameras is expected to show significant momentum in the form of more than 10 % compound annual growth rate from 2016 to 2023. The material's ability to offer thermal shock resistance is expected to increase its share in medical devices.

It is noteworthy that since the beginning of April 2017, Astrohn DB JSC has launched installations for the production of single crystals of germanium (Ge) and silicon (Si) by the Czochralski method, and advanced technologies for producing single crystals of germanium with increased uniformity of single crystals have been introduced.

Thus, the germanium segment of optics for infrared cameras, possessing a high level of optical and mechanical properties, has been dominating the market for a long time, and this is likely to continue in the next 3 years.

At the same time, the rate of growth of the volume of production of zinc selenide lenses will remain at 10 %. The material's ability to offer thermal shock resistance is expected to increase its share in medical devices.

As for lenses made of chalcogenide glasses, the Shvabe holding, as analysts from Rostekhnologii expect, will try to develop growth and plans to occupy up to 100 % of the Russian market. The main advantage of this material is its manufacturability. If germanium and selenide glasses need to be processed mechanically, then chalcogenide glasses can be made by pressing.

This makes them much cheaper than the analogs mentioned above.

To summarize, we can say that in general, the entire industry of thermal imagers will experience an upsurge. And this trend can move from a short-term phase to a long-term one.

The results obtained in this work make it possible to get acquainted with the current state of thermal imaging equipment and the most relevant directions of its development, and will also be useful at the initial stages of the development of thermal imaging equipment, when it is required to determine the concept of a product, develop a structural diagram or select an element base.

Acknowledgements

References

1. *Safonov P. A.*, *Korotaev V. V.* Current status of thermal vision technology. Almanac of Scientific Works of Young Scientists at ITMO University in 2021. 2021. 4: 87–92 (in Russian).
2. Mordor Intelligence. Industry Reports: Thermal imaging systems market — growth, trends, COVID-19 impact, and forecasts (2022–2027). URL: https://www.mordorintelligence.com/industry-reports/thermal-imaging-systems-market (circulation date: 28.10.2021).
3. Catalog of OKB "ASTROHN" products. 2019.
4. *Kulchitsky N. A.*, *Naumov A. V.*, *Startsev V. V.* Infrared Focal Plane Array Detectors: "Post Pandemic" Development Trends. Part II // Photonics. 2020. 14(4): 320–330.

Chapter 20.
Determining the attractiveness of facial expressions under static and dynamic conditions

Evgeny L. Serov, Olga A. Korolkova

Moscow State University of Psychology & Education,
Moscow, Russia;
e-mail: bossspyhunter@hotmail.com

For a long time, the problem of the relationship between statics and dynamics in facial perception has been studied in line with how a person understands and identifies facial expressions. However, in reality, this problem is much broader. It includes ideas about such a phenomenon as the attractiveness of the displayed face. This is indicated by studies by R. B. Post in which it was found that dynamic images in various forms of their presentation (inversion, reverse) are more attractive than static images (Post, Haberman, Lica, Whitney, 2012). The authors of the above-mentioned study refer to the fact that in many respects it is the naturalness of expression that is present in a dynamic stimulus that makes it so attractive (Post, Haberman, Lica, Whitney, 2012). It also partly explains the phenomenon that people often think of themselves as ugly when viewing their photos, and when watching videos they already think of themselves more highly.

It is a well-known fact that differences in assessment of facial attractiveness can occur due to emotional expressions of the face. In particular, the presence of a positive emotion such as joy on a face makes its bearer more attractive compared to when a face displays grief or disgust (Sutherland et al., 2017). However, this phenomenon has long been studied in static images. In contrast to dynamic (video stimuli), they are less consistent with the natural conditions during which people perceive each other's faces. In addition, dynamic images can cause the effect of mimicry. The essence of this effect is that observers involuntarily copy the emotional state demonstrated by the posers, whereas this phenomenon is much rarer when viewing photographic images (Weyers et al., 2006). Thus, there is a possibility that evaluation of facial attractiveness under static and dynamic conditions can be different, even when a face demonstrates the same expression.

Thus, the relationship between statics and dynamics in the study of facial expressions has remained an unsolved problem. It is still not clear under what conditions the dynamics (movement) really contributes to the precise definition of expression, and under what conditions, on the contrary, it prevents emotion recognition (Barabanschikov, Korolkova, 2020). It is also unclear how dynamics enhances the attractiveness of facial expression images.

The main goal of our research is to examine what facial movement is actually associated with when viewing images of facial expressions. We conducted two studies to test the following hypotheses: video images of facial expressions are more attractive than photographic images, either in normal orientation as well as inverted; the attractiveness of face images and videos depends on the emotional expression of the poser.

Study 1

Materials and methods

In study 1, we were interested in the effect of the upright or inverted orientation of static and dynamic faces on the perception of the attractiveness of posers. Photo and video images of famous people (actors, politicians, TV presenters) were used as stimuli. The study involved 30 people (average age 28 years). Stimulus images were consistently shown to the participants; the task was to evaluate the attractiveness of the posers using a seven-point Likert scale.

The study was carried out in two stages. The first stage was as follows. First, a photographic image of facial expression is shown in the center of the screen with the task to evaluate the attractiveness using seven-point scale located under the displayed image, in which "1" means that the image is not attractive at all, and "7" that it is very attractive. The subjects move the cursor over the desired number on this scale and click on it with the left mouse button. After the assessment has been made, the subjects move on to the next image. Further, the subjects are presented with a two-second video image of facial expression, and they need to evaluate its attractiveness on a seven-point scale according to the same principle as with a photographic image. Photo and video images are interleaved 10 times. In total, the subjects were presented with 10 photographic images and 10 video images. At the second stage, the procedure, the task and the way of answering did not change. The only difference is that the subjects are shown photographic and video images upside down.

The stimuli were shown through Google Forms. To take part in the experiment the following equipment was required: a monitor of 17 inches with a resolution of at least 1280×720 at a distance of 60 cm, a microphone and

audio output devices (headphones and speakers), and a good internet connection. The presentation of stimuli was set up via Skype under the guidance of the experimenter. Incorrectly filled out forms were not taken into account in the processing and analysis of results. The experimenter remained in contact via Skype to observe the progress of filling the stimulus material by the subjects.

For each subject, the average score was calculated in each of four conditions: photographic images; inverted photographic images; video images; inverted video images. Friedman test was used to test the statistical hypothesis that there was no difference in the responses between the four conditions. If this hypothesis was rejected, pairwise comparisons of the responses in each of the conditions were carried out using the Wilcoxon test, adjusted for multiple Benjamini–Hochberg comparisons, in order to identify which form of images (dynamic or static), as well as which orientation (upright or inverted) appeared to the subjects the most attractive.

Results and discussion

In both normal and inverted orientations, dynamic images were more attractive than static ones. Friedman test showed that the attractiveness of the four types of images differs (χ^2 (3) = 47.548; $p < 0.001$); pairwise comparisons of average attractiveness also showed significant differences between all conditions ($p < 0.0397$).

We have found that dynamic images are more attractive than static images of facial expressions in both cases: when they were presented in normal and inverted form. Thus, the results of our study are partially consistent with the results of the study by R. B. Post. According to it, dynamic images are more consistent with the natural conditions under which facial expressions are perceived (Post, Haberman, Lica, Whitney, 2012). Also, dynamic stimuli can hide some components of the face that the observer may find unattractive and which are very clearly visible when viewing static images (Barabanschikov, 2016).

At the same time, the subjects rated the upright photo and video images the most attractive. The reason for this may be that inversion distorts the way people see faces in real life. Hence, a negative reaction arises on the part of the subjects. Similar results can be found in the studies of V. A. Barabanschikov and M. M. Marinova, in which the subjects also negatively assessed inverted images of facial expressions (Barabanschikov, Marinova, 2020). Moreover, in addition to a negative reaction, the subjects also incorrectly identify facial

expression, that is, they form an erroneous idea about it (Barabanschikov, Marinova, 2020). All this confirms once again that the perception of facial expression is a very dynamic process, which depends on how information about it will be presented.

However, we assume that the results of this study are preliminary, due to several factors. First, we did not use a set of validated facial expression images such as FACS. The situation can probably change with the use of such images. In this case, the images should be with the faces of real people, and not computer morphing, since the latter can neutralize such a phenomenon as attractiveness.

Secondly, the research was carried out in a remote format. Despite the fact that all the principles of scientific research were maximally observed and the subjects tried to fill out the stimulus material as honestly as possible, in laboratory conditions the results may be different.

Thirdly, it is necessary to further develop the problem of the attractiveness of facial expressions, since the results of existing studies do not allow making an unambiguous conclusion on this phenomenon. Perhaps this is due to the fact that its essence is difficult to explain at a theoretical level. Therefore, psychologists do not have any guidelines for the formation of studies of this phenomenon at the level of practice.

The experimental results confirmed our first hypothesis. Dynamic images of facial expressions are indeed more attractive than static images in both cases: in the conditions of normal and inverted demonstration. Thus, we were able to confirm the results of experiments by R.B. Post. Changes to images, such as inversion, distort their representation to the observer in comparison with natural conditions and lead to a deterioration in interaction with them.

Study 2

Materials and methods

In study 2, we investigated the influence of emotional expressions, both static and dynamic, on the perception of the attractiveness of posers. Thirty-nine people with normal or corrected vision, representatives of various professions, participated voluntarily in the Study 2. The average age of the subjects was 34 years. Images from the VEPEL database were used as stimuli (Barabanschikov et al., 2016). This database includes photo and video stimuli depicting the face of one male poser with different expressions. The video stimuli are dynamic transitions between basic expressions, e.g., between joy

and surprise (21 transitions in total). Photo images are still frames from these video clips. For each image, there are normative ratings of emotion perception. In our study, a seven-point Likert scale was used to assess the attractiveness of these stimuli.

Due to the COVID-19 pandemic, the study was also conducted in an online format in two stages. First, subjects are shown a short video clip on their monitors. Below the video there is a scale from 1 to 7 used to assess its attractiveness. Respondents give their answer with the mouse. Then they click on the "Next" button to move to the next video clip. This procedure is repeated 42 times (each video image is presented twice). After a 15-minute break the subjects are sequentially shown photographic images of facial expressions, which they had previously seen in the video clips. The way to evaluate them is similar to the first stage. As in the case of video clips, the subjects had to evaluate 42 photo images, among which two images were repeated twice each. The stimulus material was shown through the Google Forms. The experimenter was constantly in touch via Skype to monitor the progress of the study.

The data were analyzed in the statistical environment R (version 3.6.3). The following hypotheses were tested: 1) mean attractiveness scores of static images differ depending on the facial expression; 2) mean attractiveness scores of video images differ depending on the expressed transition between emotions; and 3) mean attractiveness scores of dynamic transitions differ from mean scores of static images extracted from these transitions. To test hypotheses 1 and 2, we used Friedman test with posterior pairwise comparisons; to test hypothesis 3, we used Wilcoxon test. Benjamini–Hochberg correction for multiple comparisons was applied.

Results and discussion

The analysis showed that the perception of the attractiveness of static images differs depending on the emotion expressed by the poser. Expressions of joy (mean value of the attractiveness scores: 4.93), joy/surprise (4.58), and calm (4.48) were perceived as the most attractive, with no significant difference in their scores ($p > 0.144$). Expression scores for surprise (3.83) and sadness (3.62) also did not differ ($p = 0.673$), but both were significantly lower than those for joy ($p < 0.001$). There were also no differences ($p > 0.09$) between fear (2.86), anger (2.68), and disgust (2.38), all of which were rated significantly lower than other expressions ($p < 0.002$).

When assessing the attractiveness of dynamic expressions, a number of differences were also revealed depending on the emotional transition ex-

pressed. Several groups were identified between which significant differences were observed, while within each group no differences between individual transitions were identified. The dynamic transition "calm–joy" (5.32) was perceived as the most appealing, whose scores differed significantly ($p < 0.04$) from all other video clips except for the transitions "joy–sadness" (4.48) and "joy–surprise" (4.55). Transitions between joy and expressions of anger (4.31), fear (4.03), or disgust (4.16) were rated significantly less attractive, but did not differ between each other ($p > 0.830$). The next group included transitions of "calm–surprise" (3.63), "sadness–calm" (3.15) and "surprise–sadness" (3.32), which also did not differ between each other ($p > 0.22$). There were no differences ($p > 0.11$) in the ratings of the "anger–surprise" (2.71), "calm–fear" (2.77), "sadness–anger" (2.61), "surprise–disgust" (3.02), and "surprise–fear" (2.77) transitions. The least appealing were perceived transitions "fear–anger" (2.18), "fear–disgust" (2.26), "fear–sadness" (2.23), "sadness–disgust" (2.35), "disgust–rage" (2.4), "disgust–calmness" (2.4) and "anger–calmness" (2.31), whose scores also did not differ ($p > 0.289$).

A comparison of perceived attractiveness depending on the type of stimulus showed that the ratings of static images and the dynamic transitions from which they were extracted differed in a number of cases. In particular, the video fragment of the "anger–joy" transition was rated as more attractive than the extracted images, most of which represented anger ($p = 0.004$). In contrast, the dynamic transitions "anger–surprise", "calm–fear", "fear–sadness", and "surprise–disgust" are rated as less attractive than the extracted individual photo images ($p < 0.011$). In the other transition series, no significant differences were found between the photo and video images ($p > 0.067$).

Thus, we were able to show that depending on the expressed emotion in a static photo image or in a video fragment, the evaluations of attractiveness can vary significantly. Static expressions of joy, calmness, joy/surprise are the most attractive; expressions of anger, disgust and fear are the least attractive. Among the dynamic transient expressions, the most attractive are the video clips that include joy as one of the expressions, and the least attractive are the video clips that contain expressions of anger, fear, and disgust.

The results of this study suggest that there are differences in determining the attractiveness of facial expressions during the viewing of their photo and video images. We hypothesize that this may be due to the fact that dynamic expression is closer to the natural communication situation than static expression, which usually forms the "first impression" of the communication partner, including his/her attractiveness.

Conclusions

In the present study, the analysis of the phenomenon of face attractiveness under static and dynamic conditions was carried out using various stimuli made with the use of photography and video recording. The influence of the type of stimulus (static photo or dynamic face), the characteristics of its exposure (upright or inverted) and emotional expression on the perception of the attractiveness of the communicant was revealed.

References

1. *Barabanschikov V.A.* Dynamics of perception of facial expressions. Moscow: Cogito-Center Publ., 2016. 378 p.
2. *Barabanschikov V.A., Korolkova O.A*. Perception of "Live" Facial Expressions // Experimental Psychology (Russia). 2020. Vol. 13. No. 3. P. 55–73.
3. *Barabanschikov V.A., Marinova M.M*. Perception of video images of the chimeric face // Poznanie i perezhivanie [Cognition and Experience]. 2020. Vol. 1. No 1. P. 112–134.
4. *Barabanschikov V.A., Zhegallo A.V., Korolkova O.A*. Perceptual categorization of facial expressions. Moscow: Cogito-Center Publ., 2016. 360 p.
5. *Post R.B., Haberman J., Lica I., Whitney D.* The frozen face effect: why static photographs may not do you justice // Frontiers in Psychology. 2012. Vol. 3. Article 22. P. 1–7.
6. *Sutherland C.A.M., Young A.W., Rhodes G.* Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties // British Journal of Psychology. 2017. Vol. 108. No 2. P. 397–415.
7. *Weyers P., M*ühlberger A., Hefele C., Pauli P. Electromyographic responses to static and dynamic avatar emotional facial expressions // Psychophysiology. 2006. Vol. 43. No 5. P. 450–453.

Chapter 21.
i-Pavlovian Technology of "Cognitive Rehabilitation"

Yu. E. Shelepin, S. V. Murav'yova, V. S. Lebedev, P. P. Vasiliev, E. Yu. Shelepin, S. Alekseenko

Pavlov Institute of Physiology Russian Academy of Sciences,
St. Petersburg, Russia

Psychophysiological investigations of the "visual brain" made it possible to develop general principles for the functional recovery of patients suffering from a deterioration in cognitive and emotional abilities. In the face of the fact that the "construction of the visual brain" is the basis of mind organization [Glezer, 1995], the proposed principles of the rehabilitation of mental functions it is possible to regulate the biological and social behavior of a person, his sensory-cognitive and sensory-affective processes, internal conative (volitional) processes, decisions making, planning and performing actions. All these processes and mechanisms are closely interconnected, during behavior of a person in a real or in virtual wold.

The rehabilitation system developed by us represents a new direction related to sensory physiology and artificial intelligence. It is based on the fundamental principles of Pavlovian cognitive learning (imitative, instructive and constructive). Accordingly, it also includes sensory, volitional and affective processes. At the same time, constructive activity forms a restructuring of neural networks in a person for his recovery. In the system of rehabilitation offered by us, the properties of visual system channels which ensure the transmission of conscious and unconscious information are used. However the transmission of visual, like any other sensory information, alone is not enough to restructure human activity. Visual processes launch algorithms for declarative and imperative action planning in opponent dorsal and ventral neural networks, which then launch decision-making and action planning processes based on the internal conative (volitional) processes.

The general principle of constructing a mental health recovery system has similarities with the organization of standard fitness environments that provide real help on their own without involving complex control with visual cues. At the same time, the patient's independent desire for creative purposeful activity in sports, work and other spheres of life activity is possible and effective only if the conative and affective processes are preserved, which,

unfortunately, are weak at patients and in addition may be suppressed by traditional psychopharmacology.

The purpose of writing this article is to describe the psychophysiological foundations of the new technology. It has been called "i-Pavlovian Cognitive training System" [Shelepin et al., 2016, 2018; Shelepin, 2017; Muravyova, Shelepin, 2018; Shelepin et al., 2020, 2021]. This paper is extreme simplification of our intelligent system that provides diagnostics, therapeutic and rehabilitation effects and is designed for personalized medicine. Figure 1 schematically shows the basic principles of the rehabilitation system developed by us and indicates the main functional connections of the system.
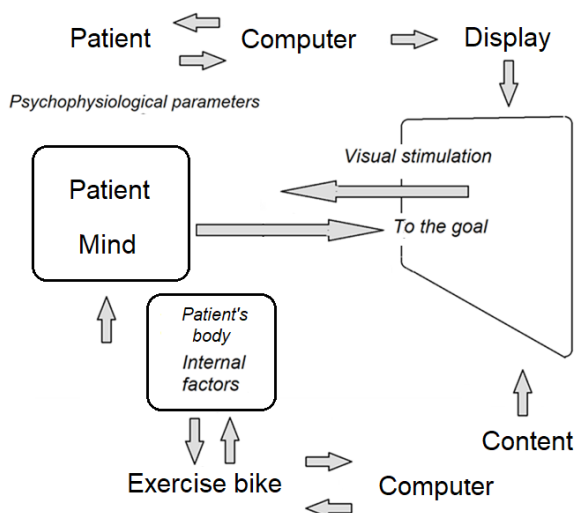
Fig. 1. General scheme of the main links of the "i-Pavlovian cognitive training System".

On fig. 1 shows sensory, predominantly visual stimulation, as well as the use of a visual signal as a control system. At the same time, the measurement of psychophysiological parameters gives the video as the main regulator of the patient's behavior, through the channels of unconscious perception [Shelepin, 2017; Shelepin et al., 2016, 2018; Muravyova, Shelepin, 2018; Shelepin et al., 2020, 2021; Murav'eva et al., 2021]. A necessary link in this system is computer-controlled rhythmic physical activity, in particular, an exercise bike. This is an important link in the orientation of the patient in achieving the goal, by involving patients in activities in a virtual or augment-

ed environment. Changes in the content of the dynamic visual image of the virtual environment is the second important part of rehabilitation, associated with the patient's muscular efforts, serving as both feedback and physiological reinforcement for the formation of a conditioned reflex to achieve the goal in the traditional sense [Pavlov, 1923].

In the proposed system, virtual or augmented reality technologies are a convenient environment that simulates the conditions of purposeful cycling.

In recent decades, the virtual environment has been actively used for the treatment and rehabilitation of patients in neurological clinics around the world. Simulators have been created that restore the orientation and movement of a person in the environment [Wilson, Foreman, 1993; Foreman et al., 2004; Chawla, Foreman, 2011; Korallo et al., 2012; Sandamas, Foreman, 2015]. They were first proposed by Nigel Foreman and co-workers and are focused on stimulating the visual dorsal neural network for better orientation in the surrounding space. These studies compared the accuracy of operators' estimation of spatial information when working in a virtual and real environment. It was shown that the acquired skills of orientation in the space of a virtual environment are transferred to orientation in an equivalent real environment. This transfer was studied not only in healthy subjects, but also in children, elderly subjects, and adults with physical disabilities who are poorly navigated in a complexly organized space. With the help of preliminary training in a virtual environment, the effectiveness of restoring the purposeful activity of patients has been shown. This preliminary preparation facilitated further behavior in the complex environment of the real world [Wilson, Foreman 1993; Pyles, Stanziano, 2001; Sandamas, Foreman, 2015].

Augmented reality is also used for rehabilitation therapy of patients after strokes in order to improve their condition and quality of life. So, for such patients, using visual copying of the patient's complex movements in the form of a "dancing man" as a feedback signal for training the quality of the patient's movements, E. Pelah developed a system that improves gait and visual control of posture and walking [Durgin, Pelah, 1999; Pelah, Thurrell, 2001; Pelah et al., 2002; Durgin et al., 2005; Nanapragasam, Pelah, 2009; Casey, Pelah, 2010; Allen et al., 2012; Pelah et al., 2015].

Another system for recovery after strokes was proposed by V. I. Skvortsova with colleagues [Skvortsova et al., 2010]. Skvortsova's technique is designed to start rehabilitation therapy after a stroke, even for use in the acute period with severe movement disorders. The authors use visual biofeedback and a virtual environment, for which goggles and a helmet are placed on the patient's head. In addition, motion sensors are installed on the head, torso and pelvic area of the patient. Some problems in patients arise due to the use of 3D

glasses or a virtual reality helmet, which in many patients cause side effects: eye fatigue, dizziness, nausea, as a result of which the patient's condition may deteriorate sharply.

In our approach to rehabilitation, in addition to external factors of visual stimulation, internal factors, such as the goal and the volitional processes that support it, are of great importance. Maintaining the desire for a goal is the development of a conditioned reflex in the brain, and in modern terms, the development of temporary connections. The formation of a temporary connection between neurons in the brain, as the basis for learning, was first proposed by I. P. Pavlov's graduate student Boris Babkin [Babkin, 1904], and the principle of reinforcement is necessary for the stable development of a conditioned reflex [Babkin, 1904; Pavlov, 1923; Hebb, 1949].

It is generally accepted that the traditional improvement in cognitive functions is achieved by recommending to patients with a slight mental deficiency to memorize poetry, solve crossword puzzles, play chess, etc. [Klimova et al., 2021; Jahn et al., 2021]. Such training trains mainly the ventral visual neural networks.

One of the modern methods to reliably assess the progress of learning to develop visual search skills are methods based on measurements of eye movement characteristics. In all areas of human activity, the ability to quickly solve visual search problems reduces the number of eye movements when searching for a target. The mechanism for controlling the development of visual skill is to train the operator to recognize the target as a blurry (low spatial frequency) image that is perceived by the receptive fields of peripheral vision to control the gaze. This gaze control involves targeting the foveal area of the retina with high spatial resolution [Skuratova et al., 2021]. It is this principle of controlling behavior with the help of preliminary low-frequency analysis that is an important part of training the visual system to achieve the goal.

For training, we set the patient the task of making a decision in a virtual environment that ensures the achievement of the goal under conditions of uncertainty. Features of the training consisted in the development of both spatial orientation and the ability to detect and recognize objects. It was the joint performance of the two tasks of spatial orientation and recognition of detected objects that provided the patient with an understanding of the content of the task facing him.

The patient's condition was monitored by measuring important psychophysiological parameters before and after training. We measured the spatio-temporal frequency range of test images and patients visual system spatial and temporal contrast sensitivity. We investigate possibility of conscious and unconscious recognition the images of objects of animate and inanimate nature,

we had recorded EEG, cognitive evoked potentials, eye movements, facial exprssion and recently provide infrared spectroscopy of the brain (NIRS). For our technique, a hardware-software recovery complex was created, which includes a computer with a large panoramic display installed in front of the patient. The images were shown on the display, the screen size of which allowed the patient to be immersed in a virtual environment. Patients could travel virtually, play games and provide interactive communication. "Hidden" control signals synthesized the dynamic images on the screen and psychophysiological data processing.

An important part of this complex is an exercise bike or a treadmill to involve patients in active activities. As a Pavlovian reinforcement, we used the result of patient-induced exercise. The muscular work of the patient improves blood circulation, which is extremely important for the brain. It is also important to note that during exercise, biologically active substances are released — endorphins and neurotrophic factors. It's all connected to the brain's pleasure systems. See reviews [Polyakov, Yachmeneva, 1981; Berridge, Kringelbach, 2015]. At the same time, movement works as a kind of Pavlovian reinforcement in the process of developing a conditioned reflex to achieve the goal, stimulating the pleasure centers of the "emotional" human brain, and normalizes neurotrophic and endocrine regulation.

Since the patient must perform actual movement on a treadmill or exercise bike with an adjustable load in accordance with the context of the game or video, the change in the load of the machine is monitored. Throughout the rehabilitation session, the patient's physical, cognitive and emotional state was monitored. The load mode is set and changed depending on the physical condition of the person. The stimulating computer is synchronized with the control computer, which provides (according to the state of visual perception and the speed of achieving the goal), a system for monitoring physiological parameters characterizing the general physical and emotional state of the patient.

At the same time, the choice of the goal is determined by the task of restoring and quantitative data on the state of the patient's brain. This provides a restructuring of the work of neural networks involved in visual perception, decision making and prediction of motor activity, as well as stimulation and adaptive capabilities of the brain associated with cognitive functions.

The developed and used "Virtual Environment Program" is an interactive video sequence that simulates, for example, cycling with different landscapes. The virtual movement is synchronized with different loads on the pedals of the exercise bike depending on the terrain. The program allows you to project a schematized image of a patient on the monitor screen, which is placed in a virtual environment with the ability to move in response to the movements of

a real object.

The results of the rehabilitation training showed that the combined effect of the virtual environment with the motor load forms a "target reflex" in the patient, i.e. there is a formation of activity aimed at achieving the goal with reinforcement in the form of conjugated muscle activity. Using Pavlov's terminology, there is a "reinforcement of the development of the target reflex" [Babkin, 1904; Pavlov, 1923]. This principle is extremely important in the effective rehabilitation of patients, since as a result of achieving the goal, the work of the whole organism is rebuilt, not only muscle activity improves, but also the state of all body systems. A person becomes satisfied with the procedures offered to him (by this method), and this improves the "quality of life" of the patient. It is important to note that in our method, by increasing the patient's motor activity, it is possible to restore (improve) the mental cognitive, conative and affective functions of the patient's brain. Therefore, the proposed method is aimed at creating neurotechnologies for the rehabilitation of patients with neurocognitive disorders. This is achieved by restructuring the work of the visual system and turning on in the brain those temporal and spatial-frequency channels of the visual analyzer that, under given threshold conditions, provide a natural combination of conscious and unconscious learning and description of the environment, orientation in this environment, recognition of individual objects in it, as well as the perception of this environment as a single, integral scene and an understanding of how in this scene it is necessary to move among "props and actors". In order to understand the scene and the role that the "props" and each of the characters play in it, it is necessary to extract the meaning of the characters and objects from memory, evaluate the positive or negative associations they cause in the mind during perception.

In other words, the answer to the first group of questions "*is this object my goal?" and "is this object good or bad?"* the observer must make the decision about the object and mentally express its attitude. The second group of questions: "*What needs to be done to achieve the goal?*" need the decision to start planning the actions that ensure the getting the goal. This new formulations of old questions: "what? and where?" are useful for constructing algorithms for the cognitive rehabilitation system that we propose.

The new formulation of questions, rephrased by us, more accurately reflect the functions of the visual dorsal and visual ventral neural networks. It should be taken into account that information about the scene and objects in this scene enters both of these neural networks, although with different spatial and temporal resolutions. Diseases predominantly caused by disruption of one of these networks have different clinical symptoms. In the dorsal system, the observed scene is consistent with the observer's body scheme and description

of the surrounding real or virtual environment. In another neural network, in the ventral one, the observed scene is consistent with the probability of danger or benefit for the subject of the observer. In the ventral network, the observed scene is consistent with the likelihood of danger, harm, or benefit to the subject of the observer. Thus, in the dorsal visual network of the brain there is a description of "where in this scene the observed object is" relative to the object of observation, and in the ventral visual network of the brain—"what this object or scene means to the observing subject." These processes occur both at the conscious and unconscious levels. Therefore, the mechanisms of conscious and unconscious extraction of information from memory are both different and similar in different aspects.

If we consider the human brain at the level of modern information technologies, then we can compare dorsal and ventral neural networks with two types of programming: imperative, as a description of how to do it, and declarative, namely, what to do. The traditional use of imperative and declarative programming languages is now supported in a hybrid approach to solving complex problems.

We associate the processing of visual information as a model of two types of psychological processes in two anatomical and physiological channels of visual information processing in the brain with the functional meaning of these two types of programming. Making this connection, we get the opportunity to consider neural networks as a dynamic structure that implements the algorithm of conscious and unconscious perception, which contributes to the survival of the organism in difficult conditions of evolutionary-historical selection.

For practice, taking into account conscious and unconscious perception is extremely important. Unconscious perception is an input to the brain that programs behavior through the interface of natural vision. By transforming brain programming and re-activating visual neural networks, it becomes a natural clinical tool for improving the patient's condition in all diseases of the so-called "cognitive sphere". Our technology can be adapted to any disease that disrupts the process of thinking, behavior and emotional sphere of a person. Therefore, our training system can be used for any mental health condition.

Patient control is based on a very subjective psychological assessment of cognitive impact. However, it is possible to quantify using the scaling method. Usually such assessments are carried out according to the school 6-point system (from 0 to the highest mark of 5). In the case of our rehabilitation methodology, patients usually give the highest score, although this may be due to side effects, entertainment from a boring hospital stay, improvement in general condition due to improved muscle function and as result good blood

circulation. Therefore, objective measurement methods are the main ones for us, such as changes in evoked potentials and restructuring of the local blood supply to the brain according to fMRI and IR spectroscopy of the brain.

Let's see how the results of monitoring the patient's condition demonstrate the effectiveness of the "i-Pavlovian Cognitive Training System". We measured cognitive visual evoked potentials to images of animate and inanimate objects [Moiseenko et al., 2015, 2016; Muravyova et al., 2016, 2017, 2018; Muravyova, Shelepin, 2018]. In healthy people, evoked potentials for images of animate and inanimate objects differ [Moiseenko et al., 2015, 2016]. These differences are manifested in the early and late components of cognitive evoked potentials and depend on the duration of the stimulus presentation. Visual cognitive potentials for these stimuli in patients with mental disorders do not have significant differences [Muravyova et al., 2017; Muravyova, Shelepin, 2018].

On average, in 25 patients, evoked potentials when following the instructions for distinguishing semantic features in images in the control group and patients with schizophrenia in the central frontal lead (Fz), central lead (Cz), central parietal lead (Pz), and occipital leads (averaging in two occipital leads (O1 and O2) are shown in Fig. 3. The most pronounced changes are characteristic of the P300 components [Muravyeva, Shelepin, 2018]. On the Figure 3 one can see very similar to the fig. 2 results after training patients with depression. According to the averaged data of the entire group of patients with depression, it can be seen that the difference is restored precisely in the late components of evoked potentials. Our results presented in fig. 2 and in fig. 3 are in good agreement with the data of other researchers. For recent work, see [Siddiqui et al., 2021], which demonstrates the importance of taking into account late waves (P300) of cognitive evoked potentials, reflecting semantic deficits in patients with mental illness. Our research clearly shows that this deficit is reversible. Cognitive training for completely different tasks effectively restores differences in evoked potentials. This restoration is shown by the example of the restoration of differences in the late components of cognitive evoked potentials on images of animate and inanimate objects [Muravyova et al., 2020], which is inherent in the differences that are observed in healthy subjects [Moiseenko et al., 2015, 2016]. These are objective indicators of the rehabilitation we carry out, which complement the subjective sensations of improving well-being.
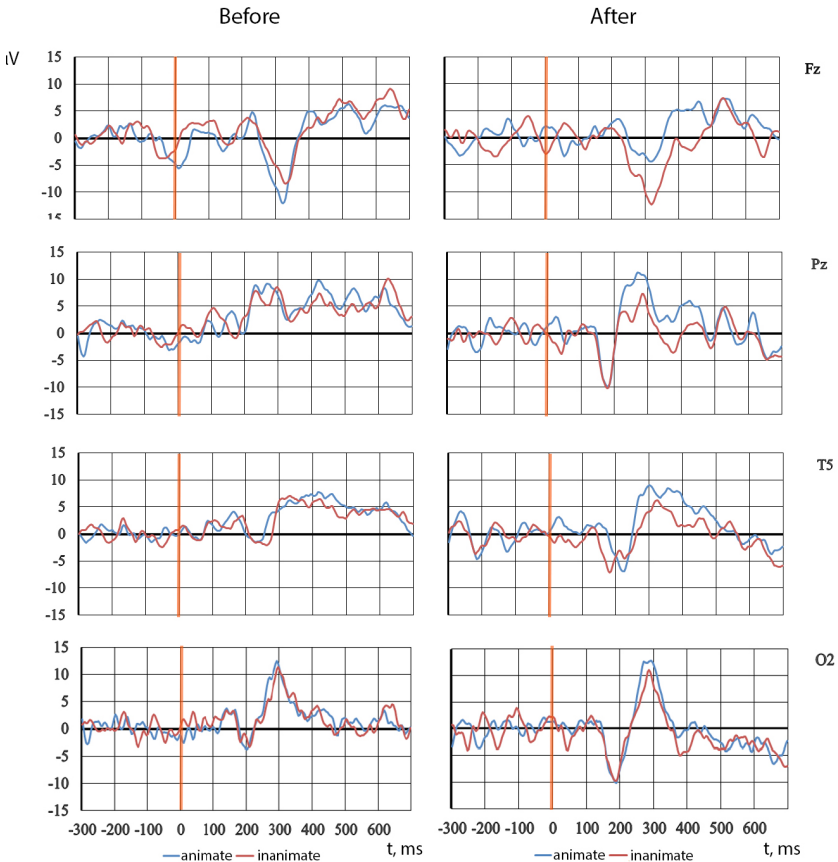
Fig. 2. Cognitive evoked potentials on presentation of contour images lasting 600 ms in one patient with paranoid schizophrenia. The evoked potentials were recorded from electrodes located on the scalp: Fz—in the center of the forehead, Pz—in the center of the crown, T5—in the temporal and O2 in the occipital regions. The column on the left is "before", on the right is "after" a two-week training course. Blue curves of evoked potentials for images of animate objects, and red curves of inanimate nature [Muravyova, Shelepin, 2018].

Fig. 3. Averaged cognitive evoked potentials for presentation of 600 ms contour images of a group of nine patients with depression, measured before (left) and after (right) training. The evoked potentials were recorded from electrodes located on the scalp: Fz-in the center of the forehead, Pz — in the center of the crown, T5 — in the temporal and O2 in the occipital regions. The column on the left is "before", on the right is "after" a two-week training course. Blue curves of evoked potentials for images of animate objects, and red curves for inanimate nature [Muravyova, Shelepin, 2018].

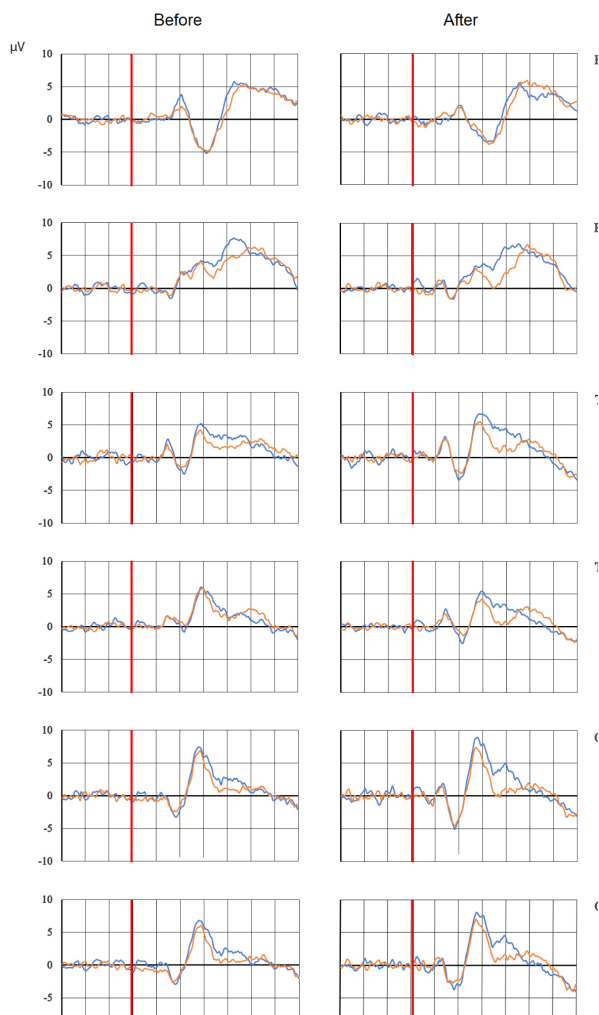Since the work of the ventral and dorsal neural networks is "intertwined", it is difficult to determine the localization of the neural network that is the source of the changes of interest to us by the localization of the electrodes from which the evoked potentials are recorded. More interesting are the temporal characteristics of the main waves of the components of evoked potentials. In particular, late component the P300 show a significant difference, for example, when images of objects of animate and inanimate nature are presented. The P300 component of the cognitive evoked potentials to the images of animate objects is usually significantly bigger. As the process of the recognition usually occurs 100–150 ms earlier, the P300 may be a reflection of a feedback signal, or demonstrate a decision-making process, even a recheck of the decision that has been already done.

The human physical activity increases the level of hormones, endorphins and the level of brain-derived neurotrophic factor (BDNF), which is very important for the growth of neurons and the differentiation of new synapses in the central nervous system in all areas of the brain associated with learning, memory and thinking. This factor is important for long-term memory [Bekinschtein et al., 2008]. BDNF is expressed in many parts of the human body, from the brain to skeletal muscles [Denham et al., 2014]. Physical exercise of patients increases its level in the blood, and it can measure in saliva [Mandel et al., 2009; Delezie and Handshin, 2018]. These measurements are important for both detecting brain disorders and predicting brain recovery.

A comparative analysis of the amplitudes of the components of the cognitive visual evoked potentials of patients before and after i-Pavlovian training is presented (Fig. 3). Cognitive visual evoked potentials in images of animate and inanimate objects in patients with mental disorders before training are similar. Training makes them different on components with a delay of 250–400 ms. This result demonstrated the training effect. Conscious or unconscious discrimination between   animate and inanimate objects is demonstrated using the example of restoration of the ventral visual neural network. The network provided a distinction of the importance of the observed object for the observer. The subjects distinguish the semantic meaning of the images. Significant differences appear in relation to the amplitudes of the components after a two-week training course of the i-Pavlovian training. Stimulation of conscious and unconscious visual channels can be carried out in various ways.

Conclusion

A method for the rehabilitation of patients has been developed that combines the use of the context of the virtual environment as a goal-forming factor synchronized with the motor load (moving treadmill or exercise

bike). That is the movement of the patient in a dynamic virtual environment, that plays the crucial role of conditioned reflex reinforcement for the general health rehabilitation. Situations created in a virtual environment require the patient to concentrate as much as possible in order to make a decision about how to get out of them in order to achieve the goal. Under these conditions, patients with neurocognitive disorders are rehabilitated by restructuring neural decision-making networks due to band stimulation of temporal and spatial-frequency channels of the visual system and selective activation of the adaptive capabilities of cognitive functions. Regular electrophysiological and psychophysical monitoring of the patient's condition demonstrates the effectiveness of "cognitive training" using the "i-Pavlovian Cognitive Training System".

The proposed training match the functions of the visual system and assessing the probability of the right optimal path choosing in the environment and assessing the probability of danger or success in choosing a target. These main functions of goal-directed activity are provided by dorsal and ventral visual neural networks. Their work proceeds on a conscious and unconscious level and is described by two well-known algorithms that provide an imperative and a declarative type of activity. The control and restructuring of these algorithms is possible in the process of training, in conditions of conscious and unconscious perception. In the process of training, we switch, through the channels of the visual system that transmit conscious and unconscious information, permit us to change the patient's attention from problems associated with his disease to other goals that we form unconsciously for him in a virtual environment [Shelepin et al., 2016; Shelepin, 2017; Shelepin et al., 2020, 2021].

Our rehabilitation technology, which we called "i-Pavlovian Cognitive Training System", is an example of effective convergence of a human and artificial intelligence aimed at restoring cognitive, emotional (affective) and volitional (conative) processes and is intended for the treatment and rehabilitation of neurodegenerative diseases of various genesis and increasing stress resistance. These technologies are intended for the diagnosis and rehabilitation of patients with common cognitive impairments, such as cerebrovascular disease ICD-10 I60-I69; mental disorders F00-F99, including mild cognitive impairment F06.7 (according to the International Statistical Classification of Diseases and Related Health Problems, 10th edition ICD-10).

References

1. *Allen L. E., Slater M. E., Proffitt R. V., Quarton E., Pelah A.* A new perimeter using the preferential looking response to assess peripheral visual fields in young and developmentally delayed children // J AAPOS. 2012. 16(3): 261–5. doi: 10.1016/j.jaapos.2012.01.006.

2. *Babkin B. P.* The experience of a systematic study of complex reflective [mental] phenomena in a dog. [Dissertation defended at the Imperial Military Medical Academy], St. Petersburg, 1904 (In Russian).

3. *Bekinschtein P., Cammarota M., Katche C., Slipczuk L., Rossato J. I., Goldin A., Izquierdo I., Medina J. H.* BDNF is essential to promote persistence of long-term memory storage // Proceedings of the National Academy of Sciences of the United States of America. February 2008. 105 (7): 2711–16. Bibcode:2008PNAS.105.2711B. doi:10.1073/pnas.0711863105. PMC2268201. PMID18263738.

4. *Berridge K. C., Kringelbach M. L.* Pleasure systems in the brain // Neuron. 2015 May 6; 86(3): 646–664. PMCID: PMC4425246; NIHMSID: NIHMS664561; PMID: 25950633; doi: 10.1016/j.neuron.2015.02.018.

5. *Casey R., Pelah A., Cameron J., Lasenby J.* Symposium On Applied Perception in Graphics and Visualization. 2010: 160. DOI: 10.1145/1836248.1836282.

6. *Chawla S., Foreman N.* Forms of Interaction in Virtual Space: Applications to Psychotherapy and Counselling. DOI: 10.1007/978–3–642–21657–2_33, Source DBLP, Conference: Universal Access in Human-Computer Interaction. 6th International Conference, UAHCI 366, 011, Orlando, FL, USA, 2011, Proceedings, Part IV.

7. *Delezie J., Handschin C.* Endocrine Crosstalk Between Skeletal Muscle and the Brain // Frontiers in Neurology. 2018. 9: 698. doi:10.3389/fneur.2018.00698. PMC6117390. PMID30197620.

8. *Denham J., Marques F. Z., O'Brien B. J., Charchar F. J.* Exercise: putting action into our epigenome // Sports Medicine. February 2014, 44 (2): 189–209. doi:10.1007/s40279–013–0114–1. PMID24163284. S2CID30210091.

9. *Durgin F. H., Pelah A.* Visuomotor adaptation without vision? // Experimental Brain Research. 1999, 127: 12–18. PMID10424410 DOI: 10.1007/s002210050769.

10. *Durgin F. H., Pelah A., Fox L. F., Lewis J., Kane R., Walley K. A.* // J Exp Psychol Hum Percept Perform. 2005. 31(3):398–419. doi: 10.1037/0096–1523.31.3.398.

11. *Foreman N., Sandamas G., Newson D.* Distance underestimation in virtual space is sensitive to gender but not activity-passivity or mode of inter-

action // CyberPsychology & Behavior. 2004, 7(4):451–7, DOI:10.1089/cpb.2004.7.451.

12. *Glezer V. D.* (1995). Vision and mind: Modeling mental functions. Lawrence Erlbaum Associates, Inc.

13. *Hebb D.* The Organization of Behaviour. 1949. John Wiley & Sons. ISBN978–0–471–36727–7.

14. *Jahn F. S.*, *Skovbye M.,* Obenhausen K., *Jespersen A. E.*, *Miskowiak K. W.* Cognitive training with fully immersive virtual reality in patients with neurological and psychiatric disorders: A systematic review of randomized controlled trials // Psychiatry Res. 2021 Jun; 300:113928. doi: 10.1016/j.psychres.2021.113928. Epub 2021 Apr 6. PMID: 33857847.

15. *Klimova B., Pikhart M., Cierniak-Emerych A., Dziuba S., Firlej K.* A Comparative Psycholinguistic Study on the Subjective Feelings of Well-Being Outcomes of Foreign Language Learning in Older Adults From the Czech Republic and Poland // Frontiers in Psychology. 2021, 12.

16. *Korallo L., Foreman N., Boyd S., Moar M., Coulson M.* Can multiple «spatial» virtual timelines convey the relatedness of chronological knowledge across parallel domains? // Computers & Education. February 2012, 58(2): 856–862, DOI: 10.1016/j.compedu.2011.10.011.

17. *Mandel A. L., Ozdener H., Utermohlen V.* Identification of pro- and mature brain-derived neurotrophic factor in human saliva // Archives of Oral Biology. July 2009, 54 (7): 689–95. doi:10.1016/j.archoralbio.2009.04.005. PMC2716651. PMID19467646.

18. *Moiseenko G. A.*, Shelepin Yu. E., *Kharauzov A. K.*, *Pronin S. V.*, *Chikhman V. N.*, *Vakhrameeva O. A.* Classification and recognition of images of animate and inanimate objects // J. Opt. Technol. 2015, 82(10), 685–693.

19. *Moiseenko G. A.*, *Vershinina E. A.*, *Pronin S. V.*, *Chihman V. N.*, Shelepin Yu. E., *Mikhailova E. S.* Latency of Evoked Potentials in the Tasks Involving Classification of Images after Wavelet Filtration // Human physiology. 2016, 42(6): 37–48.

20. *Murav'eva S. V., Kozub K. E., and Pronin S. V.* Optical and electrophysiological techniques for functional assessment of vision system neuronal networks J. Opt. Technol. 2021, 88(12): 710–715.

21. *Muravyova S., Moiseenko G., Pronina M., Shelepin E., Shelepin Y.* Dysfunction of parvo–systems and its stimulation in patients with schizophrenia with early stage of the disease // Perception. 2016, V. 45, Supplement, 39th European Conference on Visual Perception, 2016. P. 14.

22. *Muravyova S., Moiseenko G. A., Pronina M., Pronin S. V., Shelepin E. Y., Shelepin Y. E.* The impact of spatial-frequency filtering to recognize of stimuli in healthy subyects and patients with schizophrenia (psychophysical and electrophysiological methods of investigation) // Materials of the IEEE International Symposium «Video and Audio Signal Processing in the Context of Neurotechnologies», June 30— July 2, 2016. St. Petersburg.

23. *Muravyova S. V., Shelepin Yu. E.* Restoration of violations of purpose-ful activity in patients with neuropsychiatric pathology by immersion in an interactive virtual environment // In the book "Neurotechnology", Y. E. Shelepin, V. N. Chihman, Editors. 2018, Chapter 15, P. 367–384 (In Russian).

24. *Murav'eva S. V., Pronina M. V., Moiseenko G. A., Pnevskaya A. N., Polyakov Yu. I., Kropotov Yu. D., Pronin S. V., Shelepin E. Yu., Shelepin Yu. E.* Analysis of Visual Cognitive Impairments in Schizophrenia at the Early Stages of the Disease and Their Correction by Interactive Virtual Environment // Human Physiology. 2017, 43, P. 625–636 [Fiziol. Chel. 2017. 43(6), 24–36].

25. *Murav'eva S. V., Moiseenko G. A., Chomsky A. N., Sharibin E. A., Kropotov Yu. D., Shelepin Yu. E.* Stimulation of the Visual System Using a Cognitive Task in a Virtual Environment in Patients with Schizophrenia and Depression // Human Physiology. 2020, 46(5), P. 483–491 [Fiziol. Chel. 2020. 46(5), 27–36].

26. *Nanapragasam A., Pelah A., Cameron J., Lasenby J.* Visualizations for locomotor learning with real time feedback in VR // Proceedings, 2009: Symposium On Applied Perception in Graphics and Visualization. 139. DOI: 10.1145/1620993.1621033.

27. *Pavlov I. P.* Twenty years of experience in the objective study of the high-er nervous activity (behavior) of animals. M. — Pg.: Gizidat, 1923. Ch. "Reflex of the goal" [M.: Nauka, 1973. P. 214–218].

28. *Pelah A, Thurrell A. E.* Reduction of perceived visual speed during lo-comotion: Evidence for quadrupedal perceptual pathways in human? // Journal of Vision. 2001, 1:307a. DOI: 10.1167/1.3.307.

29. *Pelah A, Thurrell A.E, Berry M.* Reduction of perceived visual speed during walking: Evidence against the involvement of attention-al or vestibular mechanisms // Journal of Vision. 2002, 2: 630a. DOI: 10.1167/2.7.630.

30. *Pelah A., Barbur J., Thurrell A, Hock H. S.* The coupling of vision with locomotion in cortical blindness // Vision Research. 2015, 110: 286–94. PMID24832646 DOI: 10.1016/j.visres.2014.04.015.

31. *Polyakov E. L., Yachmeneva E. Yu.* Electrical stimulation of the reward-ing systems of the brain: Bibliography. decree, L., 1981. — 213 p. (In Rus-sian).

32. *Pyles N., Stanziano S. M.* Interactive fitness equipment, Applicant(s): [US] EPIX INC; WO0112269 (A1) — 2001–02–22; Classification A63B21/00; A63B24/00; Priority number(s): US19990374324 19990813. Also published as: WO0112269 (A9), AU7139700 (A), US2002055418 (A1).

33. *Sandamas G., Foreman N.* Active Versus Passive Acquisition of Spatial Knowledge While Controlling a Vehicle in a Virtual Urban Space in Driv-ers and Non-Drivers // SAGE Open, July-September 2015: 1–9, DOI:

10.1177/2158244015595443.

34. *Shelepin Yu. E., Muravyova S. V., Shelepin E. Yu., Yakimova E. G., Shelepin K. Yu*. A method of patient rehabilitation using a virtual environment // patent RU (11) 2 654 767 (13) C1 Starting date of the patent validity period: 02.12.2016. (In Russian).

35. *Siddiqui S. V., Nizamie S. H., Siddiqui M. A., Jahan M., Garg S., Tikka S. K., Shreekantiah U*. Evaluation of N-400 Evoked Response Potential in schizophrenia: An endophenotype or a disease marker? // Psychiatry Research. June 2021, Volume 300. 113907 PMID: 33839423, https://doi.org/10.1016/j.psychres.2021.113.

36. *Shelepin Yu. E* Introduction to Neuroiconic, SPb. 2017.

37. *Shelepin E. Y., Muravyova S. V., Yakimova E. G., Shelepin Y. E.* Neurotechnologies for managing purposeful activities // In the book "Neurotechnology", Y. E. Shelepin, V. N. Chihman, Editors. 2018, Chapter 14, P. 349–367 (In Russian).

38. *Shelepin Yu. E., Kharauzov A. K., Zhukova O. V., Pronin S. V., Kupriyanov M. S., Tsvetkov O. V.* Masking and detection of hidden signals in dynamic images // Journal of Optical Technology. October 2020, Vol. 87, No. 10, P. 624–632.

39. *Shelepin Y. E., Harauzov A. K., Vakhrameeva O. A., Zhukova O. V., Pronin S. V., Tsvetkov O. V., Skuratova K. A., Shelepin E. Y.* Unconscious visual signals and involuntary human reactions // Integrative Physiology. 2021, Vol. 2, No 4.

40. *Skuratova K. A., Shelepin E. Yu., Yarovaya N. P.* (2021) Optical search and visual skill // Journal Optical Technology, Volume, 88, № 12. Pp. 700–705.

41. *Skvortsova V. I., Ivanova G. E., Skvortsov D. V.* A method for the rehabilitation of patients in the acute stage of stroke using biofeedback and virtual reality. RF patent, RU2432 971 C1. Priority 02.04.2010. (In Russian).

42. *Wilson P., Foreman N.* Transfer of information from virtual to real space: Implications for people with physical disability // Eurographics Technical Report Series. 1993. ISSN.1017–4656, pp.21–25.

Chapter 22.
Eye Tracking as a Screening Tool for Affective and Stress-related Disorders

Ksenia Skuratova, Evgeny Shelepin, Diana Naumova, Anna Kosyakova

Pavlov Institute of Physiology RAS
St. Petersburg State University
Russia

More than half a billion people each year suffer from disorders associated, to varying degrees, with the influence of exogenous and endogenous stressors. In particular, such disorders include obsessive-compulsive disorder, anxiety disorder, and depression.

All of these disorders are characterized by attention bias to certain stimuli. For example, depression in general is characterized by decreased regulatory control of stimulus-driven affective processing. It is assumed that such a decrease in control causes the formation of rigid negative biases, which are manifested in the process of implicit information processing (for example, a negative assessment of oneself, the environment and the future, the focus of attention and memory on negative stimuli). This attention bias has been demonstrated in several eye-tracking studies, the results of which were pooled in meta-analyzes (Armstrong & Olatunji, 2012; Suslow et al., 2020). The first is more concerned with studies of the oculomotor activity of people with subclinical manifestations of depression or dysphoria. Armstrong and Olatunji found that, relative to the control group, depressed people are characterized by decreased initial orientation and decreased maintenance to positive stimuli. At the same time, they are characterized by increased gaze maintenance on dysphoric stimuli. In a meta-analysis published in 2020, turn their attention to eye movement studies in people with clinically significant depression and identify moderate group differences in maintaining attention to dysphoric stimuli and sad faces between depressed and healthy people. During free viewing, patients with acute depression pay more attention to sad faces and dysphoric pictures than healthy people. Also, there were moderate differences between these groups in maintaining attention on positive images

and happy faces. Depressed people spend less time looking at positive images than healthy people. Group differences in initial orientation were not found.

The previously mentioned meta-analysis by Armstrong and Olantunji also reported that anxious people show increased vigilance for a threatening stimulus during free viewing and visual search, and showed difficulty disengaging from threat in visual search tasks, but not during free viewing, compared to a control group. Another meta-analysis combining eye-tracking studies of people with social anxiety disorder reports that anxious people are characterized by facilitated detection and orienting of initial attention toward threat and greater maintained attention on threatening stimulus. These findings are particularly relevant for face imaging, as socially anxious people dwelled significantly longer on threat faces compared with nonanxious participants and also have difficulty disengaging attention from angry expressions.

For obsessive-compulsive disorder, it has been found that people with more severe OCD symptoms do not target (orient) specific stimuli faster than people with less symptomatics, but once a threat is identified, these people pay more attention to the stimuli associated with OCD. The OCD group made significantly more fixations upon both OCD-related and aversive stimuli compared to controls, and spent longer overall examining both of these stimuli-types.

To date, there is insufficient research into specific attention deficit disorder in people with OCD. However, it has already been found that people with OCD have a hard time ignoring irrelevant stimuli and exhibiting impaired selective attention abilities. The researchers hypothesize that the attention deficit observed in patients with OCD was mainly a by-product of their impaired executive functions, pointing to a decline in organizational skills, incoming information and optimization of cognitive resources (Snyder et al., 2015). According to these findings, depression and motor impairment, which is often seen in subjects with OCD, cannot explain the widespread executive impairment commonly seen in these patients. Consistent with this idea, it can be assumed that disruptions in cognitive and behavioral inhibitory processes underlie many of the neurocognitive symptoms found in people with OCD. It is likely that two main types of inhibitory processes are impaired in people with OCD, namely the cognitive inhibition processes, which control internal thinking and prevent obsessive thoughts and mental rituals, and the behavioral inhibition processes, which control motor actions, such as behavior testing.

Previous studies did not use a single stimulus material to compare patients with various diseases. The goal of our research is to do such a comparison so

that the information gathered may be used to develop an objective differential diagnosis method based on machine learning.

It was hypothesized that oculomotor activity in people with subclinical manifestations of anxiety disorder, obsessive-compulsive disorder and depression has specific characteristics that distinguish these people from each other and from healthy people.

For example, we hypothesize that subjects with generalized anxiety disorder will exhibit signs of hypervigilance (a tendency to pay too much attention to potentially threatening signals), resulting in an increase in the total scan path length as well as the total number of fixations on threatening images. Moreover, according to the general model of defensive behavior, hypervigilance manifests itself in potentially dangerous situations, even if a real threat has not yet been detected.

Since cognitive theories of depression suggest that a shift in attention to negative emotional stimuli plays an important role in the development and maintenance of depression (Clark, Beck, 1999), we assume that attention to dysphoric images is positively associated with depressive rumination (constant repetition of obsessive negative thoughts) and a measure of daily adaptation based on Beck's depressive cognitive triad. Increased attention to dysphoric stimuli can contribute to the sustained processing of negative information, which, in turn, can manifest itself in the form of an increase in depressive thoughts and a catastrophic attitude towards oneself, the present and the future (Beck's triad).

Research method

Eye tracking is a prominent method used to study emotional processing and the observer's visual and cognitive processes. All participants were instructed to look at the computer screen freely which had 4 images on it and pick the one that they liked the most. In total, participants had to view 20 experimental stimuli (each stimuli had 4 images: positive, threatening, dysphoric and neutral) and 5 control stimuli, on which there were only neutral images. Their oculomotor activity was recorded throughout the procedure. The position of each image was randomly selected.

Oculomotor activity was recorded using the Neurobureau software hardware complex.

Synchronous registration of GSR will allow assessing the emotional state of the subject. The choice of GSR was due to the fact that it is a highly sen-

sitive method for assessing suicidal risk (Thorell et al., 2013). Additional research methods will include self-report inventory, clinical interviews, and psychiatric history collection. The prerequisites for the subjects are normal or corrected to normal vision, as well as the absence of neurological diseases.

Because the correct interpretation of everyday scenes and facial expressions is important in communication, interpersonal relationships and in contemplation, images with facial expressions, everyday life scenes and human figures were used as stimulus to understand the individual tendency to ruminate, and other cognitive vulnerability factors such as irrational beliefs, maladaptive schemas, or cognitive biases.

The development of stimulus material for diagnostic tools is an important step in such research, so we will dwell on it in more detail.

The image search was limited to images labeled as being accessible for reuse with modification, ensuring that the images may be freely modified and distributed. A group of experts consisting of psychologists selected 200 images (50 neutral, 50 positive, 50 dysphoric, 50 threatening). Each image was placed into one of three categories— human figures, human facial expressions and scenes.

We recruited 30 participants who were blind to the purpose of the present study and exhibited considerable variability in terms of age, socioeconomic background and level of education. The gender distribution was balanced. We instructed them to look through images and assign each of them to one of 5 categories (positive, threatening, dysphoric, neutral or other). To ensure that participants were sufficiently attentive and motivated, we asked them to take a break if they felt overwhelmed during the test and then resume doing it once they had recovered.

The threatening images were rated significantly more threatening than dysphoric images. The dysphoric images were rated significantly sadder than threatening images. In general, the positive and neutral images had significantly higher levels of recognition than the other categories.

This procedure allowed us to select the most valid stimuli and create our diagnostic method.

Also at the development stage is a method of visual search for emotionally colored words among distractors belonging to the groups discussed above (neutral, positive, dysphoric threatening). Emotionally colored words for this study, as well as in the case of images, were selected as a result of double peer

review. This technique will be necessary to more accurately identify people with OCD, as they have problems with optimizing the allocation of attention resources without properly suppressing attention to distractors. In other words, people with OCD cannot selectively fixate on objects that are relevant to the task they are supposed to accomplish.

Preliminary results

In this work, we present the results collected on a control sample whose participants do not have mental disorders. The table below shows the obtained indicators of oculomotor activity.

	Neutral image (control stimulus)	Neutral image	Positive image	Dysphoric image	Threatening image
The number of fixations before the first fixation on the stimulus	4,187	4,244	4,261	5,363	4,346
Time of first fixation on the stimulus (sec)	1,334	1,312	1,372	1,604	1,347
Viewing time (sec)	1,475	1,666	1,699	1,004	0,945
First fixation duration (sec)	0,288	0,271	0,276	0,280	0,317
The number of returns	1,532	1,556	1,348	0,763	0,676
Average duration of fixations (sec)	0,278	0,270	0,287	0,257	0,325
The total number of fixations	5,306	6,170	5,919	3,907	2,907
Average saccade amplitude (degrees)	2,518	2,615	2,346	2,112	2,585
The total length of the scan-path (degrees)	6,831	9,230	8,822	4,179	4,376

From the data obtained, it can be seen that, first of all, subjects without mental disorders in anamnesis pay attention to neutral and positive images, and direct their gaze to dysphoric images last.

The total viewing time and the total number of fixations to neutral and positive images is more than one and a half times longer than the viewing time of dysphoric and threatening images. In this case, on control stimuli containing only neutral images, the viewing time is distributed evenly between all four images.

Healthy subjects are almost twice as likely to return to viewing neutral and positive images as compared to dysphoric and threatening ones, which suggests that it is more important for them to analyze the neutral and positive images.

A longer first fixation, as well as a longer average duration of fixations on threatening images, indicates that the subject is trying to obtain more information for image analysis in one fixation and is presumably associated with evolutionary mechanisms for processing threatening stimuli.

The total length of the scanning path (the sum of the amplitudes of all saccades) in neutral and positive images is more than twice that in dysphoric and threatening images, which means that healthy subjects examine neutral and positive images in more detail.

Conclusion

In general, based on the results of the pilot study, we confirmed our hypothesis that mentally healthy people focus their attention on positive and neutral images. Further research will be aimed at collecting a clinical sample, analyzing GSR indicators, as well as standardizing the data obtained for the development of screening instruments and testing its validity.

References

1. *Armstrong T., & Olatunji B.O.* (2012). Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. Clinical Psychology Review, 32(8), 704–723. doi:10.1016/j.cpr.2012.09.004.
2. *Suslow T., Hußlack A., Kersting A., & Bodenschatz, C.M.* (2020). Attentional biases to emotional information in clinical depression: A systematic and meta-analytic review of eye tracking findings. Journal of Affective Disorders, 274, 632–642. doi:10.1016/j.jad.2020.05.140.
3. *Thorell L.H., Wolfersdorf M., Straub R., Steyer J., Hodgkinson S., Kaschka W.P., & Jandl, M.* (2013). Electrodermal hyporeactivity as a trait marker for suicidal propensity in uni- and bipolar depression. Journal of Psychiatric Research, 47(12), 1925–1931. doi:10.1016/j.jpsychires.2013.08.017.
4. *Snyder H.R., Kaiser R.H., Warren S.L., & Heller W.* (2015). Obsessive-Compulsive Disorder Is Associated With Broad Impairments in Ex-

ecutive Function: A Meta-Analysis. Clinical Psychological Science, 3(2), 301–330. https://doi.org/10.1177/2167702614534210.

5. *Clark D. A., & Beck A. T.* (1999). Scientific foundations of cognitive theory and therapy of depression. New York: John Wiley & Sons.

Chapter 23.
Neurophilosophy of Inner Silence

N.A. Solovyev, G.A. Moiseenko

St. Petersburg State Institute of Psychology and Social Work;
Pavlov Institute of Physiology Russian Academy of Sciences

Corresponding author: Nikita Solovyev
E-mail: solovyovnikita@mail.ru

General Information about the state of inner silence

Altered states of consciousness attract the attention of researchers from the point of view of philosophy and psychology as well as from the point of view of neurophysiology. Controlled states of consciousness, which a person can independently control, are of considerable interest. This interest is induced by both the possibility of investigation of the nature of consciousness itself and by various practical applications. Controlled altered states of consciousness should include, first, various meditative states known in the Eastern spiritual tradition, and second, states of inner silence related to the tradition of hesychasm, a mystical trend in Orthodox Christianity. As far as the Eastern tradition is concerned, the Western term "meditation" completely mislables the essence of this practice. Indeed, the literal translation of the word "meditation" from the Latin language means "reflection", and the essence of the Eastern meditative practice is related precisely with stopping the thought process and transitioning into a state of inner silence: "The first duty of the yogin is to think-that is, not to let himself think." (Eliade, 2019, p. 48). In the tradition of hesychasm, the state of inner silence and withdrawal from everything intelligible is also of paramount importance: "And if you <> zealously strive to join the contemplation of mystical visions, then get away from your activities and feelings, and mind, and from everything sensually perceived, and from everything intelligible, and from everything that exists, and from everything that does not exist, in order to the best of our ability to strive for supernatural unity with the One who surpasses any essence and any knowledge " (Dionysius the Areopagite, 1991).

The question of stopping the thought process and removing from every-thing intelligible may, at first glance, seem too abstract and insignificant, but the answer to it allows describing the structure of human consciousness and better understanding Descartes' famous formula "Ego cogito ergo sum", which despite all criticism and different interpretations, remains the corner-stone of modern philosophy. The problem is that this formula can be under-stood in two ways. In the natural sciences it is usually interpreted that the process of thinking is an indispensable condition for human existence. At the same time, in European philosophy emphasizes the importance of the exis-tence of an extra-physical Self observing the content of consciousness. In this regard, M. Heidegger wrote the following: "The representing I is far more *essentially* and necessarily *co*-represented in every "I represent" namely as something toward which, back to whicn, and *before* which every represented thing is placed" (Heidegger, 1984, p. 107). Thus, the state of inner silence (lacking contents of consciousness) can help in understanding and interpret-ing the famous formula of Descartes.

Modern neurophysiological studies of altered states of consciousness

Recently, a large number of studies of neural correlates of various medita-tive states, associated mainly with various Buddhist practices, was published (for review see Lee et al., 2018). The main conclusion of various studies is that advanced mediatation practitioners (average experience of about 30 years) demostrate an increase in the intensity of high-frequency oscillations of beta and gamma waves regardless of methods of meditation (with and without concentration on the object). This conclusion is confirmed by research data presented in (Lutz et al., 2004; Fell et al., 2010; Lippelt et al., 2014; Tang et al., 2015). Quite detailed studies of the EEG activity of the brain during meditation in the Buddhist tradition are presented in works (Lutz et al., 2004; Ricard et al., 2014). Voluntaries, whose meditative experience was 10,000–50,000 hours for 15–40 years, demonstrated an increase in brain gamma ac-tivity during meditation. In this studies, consistent with the conclusion in (Lee et al., 2018), no such changes were observed in the control group. In addition, the presence of phase synchronization in the gamma range during meditation is noted: " long-term Buddhist practitioners self-induce sustained electroen-cephalographic high-amplitude gamma-band oscillations and phase-synchro-ny during meditation" (Lutz et al., 2004). The authors emphasized that they

investigated the state during the objectless meditation of love and compassion when the attention is not focused on specific objects: "First-person descriptions of objectless meditations, however, differ radically from those of concentration meditation. Objectless meditation does not directly attend to a specific object but rather cultivates a state of being. Objectless meditation does so in such a way that, according to reports given after meditation, the intentional or objectoriented aspect of experience appears to dissipate in meditation. This dissipation of focus on a particular object is achieved by letting the very essence of the meditation that is practiced (on compassion in this case) become the sole content of the experience, without focusing on particular objects" (Lutz et al., 2004, p. 16372). Moreover, it was found that neuronal activity during objectless meditation may differ from activity during meditation with concentration on an object, although this was not noted in the review (Lee et al., 2018). It is important to note that different parts of the brain are activated in various meditation practices, although an increase in the power of the gamma rhythm is observed in all cases.

Formulation of the problem

The complexity of the description of consciousness outlined above is multiplied by the complexity of the interpretation of various spiritual teachings belonging to different traditions, cultures and civilizations. It makes it difficult not only to study the phenomenon of consciousness, but also to interpret the results of objective studies of the neural correlates of various meditative states. «Meditation practice is difficult to access because of its countless forms of appearances originating from the complexity of cultures it has to serve. This makes a suitable categorization for scientific use almost impossible. However, empirical data suggest that different forms of meditation show similar steps of development in terms of their neurophysiological correlates» (Fell et al., 2010, p. 218). In addition, a fundamental difficulty in the study of meditative states and the state of silence is that an outside observer cannot penetrate into the inner world of the subject. An outside objective observer can investigate only the material component of the process, and the contents of *consciousness* can be observed and interpreted only by the person. Thus, the results of objective observation of the state of the subject's brain can be compared only with the description of the subject's subjective experience. Moreover, the inner experience is very difficult to express in words and, secondly, it is always interpreted by the subject on the base of his/her own worl-

dview and metaphysical attitudes.

The described difficulties in the study of altered states of consciousness (and of consciousness in general) require the development of fundamentally new approaches to the problem. In fact, the researcher should become both a subjective observer of his/her inner world from the inside and a objective observer comparing the results of these subjective observations with the results of objective measurements of the states of his/her own brain. It is a very rarely used approach to the problem of psychophysical parallelism.

Based on the foregoing, here we attempted to combine the researcher and the subject in one person. In principle, this approach can be classified as a kind of modification of the introspection method, widely known in psychology, but in neurophysiology it is not generally accepted. Despite the unusualness of this approach, this work will present results of experiments conducted with one of the authors, N. A. Solovyev (67 years old), an Orthodox Christian having experience of a state of mind similar to the state of "inner silence", known in the tradition of hesychasm.

Objective EEG studies of the state of the subject's brain were carried out at the Pavlov Institute of Physiology Russian Academy of Sciences under the guidance of Professor Yu. E.Shelepin, Ds, MD and at the Clinic of Bioacoustic Correction under the guidance of K. V. Konstantinov, PhD. We will present the evidence that studies of the electroencephalographic activity of the subject's brain in a state of silence showed a significant increase in the power of the gamma rhythm (the maximum increase in the average power was 2.72) compared to the baseline calm state. This allows us to say that the state of the subject was, from the point of view of neurophysiology, similar to the states of the subjects in the published earlier studies of the Buddhist tradition subjects. In other words, in this case, we could focus precisely on the neurophilosophical side of research, describing the state of inner silence in the first person in accordance with the European tradition, taking into account the results of our previous works (Solovyev 2019; Solovyev 2021a, 2021b).

Experimental methods and results

In the literature cited above the main attention was paid to comparison of the EEG activity in experienced practitioners with control subjects not familiar with meditation. In our experiments, the subject confirmed that he was capable to reach meditative states indicated by an increase in the power of the

gamma-band. It made possible to focus on obtaining information about this state from the first-person, as well as to slightly modify these states by changing conditions of experiments, while measuring the electrical activity of the brain. The experimental methods were as follows.

A series of measurements of the EEG activity of the brain was carried out in the sequence of states: 1. A calm state, which was taken as a reference point; 2. A state of inner silence; 3. Counting from different initial numbers 0, 1, 2, 3, etc. to 100 and back in sevens while maintaining inner focus on the counting process and the absence of extraneous thoughts; 4. Concentration on the left temple and maintaining the state of inner silence. All EEG measurements were performed when the subject's eyes were open. The duration of each phase in the series varied from 3 to 5 minutes. In some experiments, two series of experiments were carried out consecutively 10–15 minutes apart without removing the electrodes from the subject's head.

Registration and analyses of the EEG was carried out using an encephalograph manufactured by Mitsar-EEG-201 with a sampling rate of 500 Hz and the WinEEG software. The electrodes were positioned in accordance with the international 10–20 system. Connected electrodes located on the earlobes were used as a reference. Power spectra and topograms were analyzed within the following bands of EEG rhythms: delta- (1.5–4 Hz), theta- (4–7 Hz), alpha- (8–13 Hz), beta1- (14–20 Hz), beta2 — (20–29 Hz) and gamma (30–115 Hz). The average power of EEG rhythms was calculated for the selected ranges of EEG rhythms in all areas of the brain separately for each state (state of calm wakefulness and state of inner silence, counting to 100 and back by sevens and concentration on the left temple). The average power of EEG rhythms was calculated for the ranges for all electrodes and for individual electrodes, and the ratios of the average power of EEG rhythms in percent between different states were compared. The mean values, standard deviations, minimum and maximum values of the averaged power of EEG rhythms for all electrodes were calculated for each state separately. In the mode of inner silence and in the basic calm state, the number of blinks was compared. Within these two states, the percentage of blinks, the mean and the standard deviation were calculated. The data were analyzed using Statistica 10 and Excel 10 software.

As it was said above, one of the results of our research was confirmation of the fact that in a state of inner silence, the average power of high-frequency rhythms significantly increases. Fig. 1 shows an increase in the average

power of the beta2-rhythm (20–29 Hz), and the gamma rhythm (30–115 Hz) during the transition from a calm state to a state of inner silence. In total, 10 series of measurements were carried out. The maximum increase in the average power of the gamma-rhythm in one of the series was 2.72 times. As noted above, in a number of studies we carried out double series of experiments 10–15 minutes apart (totalling 4 double series, consisting of 8 single series). At the same time, in the second series of experiments, we always observed a smaller percentage increase in the power in the beta2 and gamma bands. In one case, in the second series of experiments, we observed even a 0.72-fold decrease in beta2 and gamma activity. Apparently, this was due to the fact that in the second series of experiments the subject did not have time to return to a calm state, and the initial calm state in the second series of experiments was always accompanied by more powerful beta2 and gamma activities. The average power of the alpha rhythm, as can be seen from Fig. 1, in our experiments remained at approximately the same level.
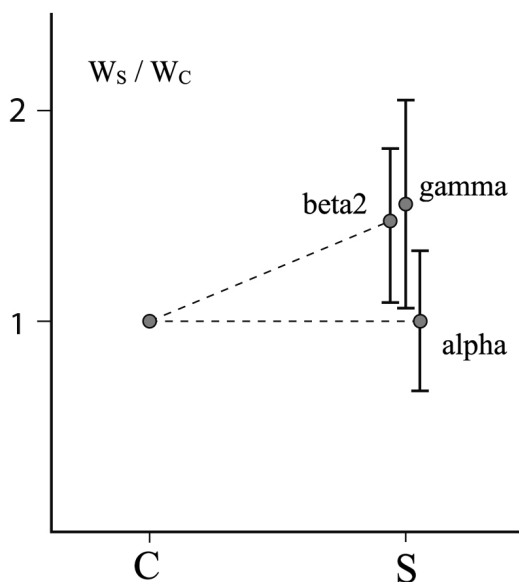

Fig. 1. The ratio of the average power of gamma, beta2 and alpha rhythms in the state of inner silence (W_S) to the average power of the basic calm state (W_C) in the same bands. The points on the graph represent the data averaged over all experiments with the indication of the average statistical deviation C—calm state, S—inner silence.

In our experiments, the maximal increase in the power of the gamma-rhythm was observed in the temporal regions. Fig. 2 shows topograms of the EEG activity of the subject's brain in two states: the calm state — *a*, the state of inner silence — *b*.

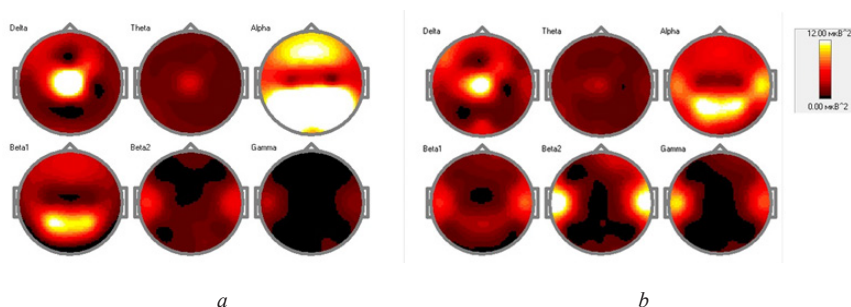*a*   *b*

Fig. 2. Topograms of the distribution of electroencephalographic activity of the subject's brain in a calm state (a) and a state of inner silence (b). Lighter areas correspond to higher power in the corresponding band.

In different series of experiments, a change in the distribution of the power of gamma activity between the right and left temporal regions was observed. However, we did not find a correlation of these changes with the internal state. In the mode of inner silence, compared with the basic calm state, there was a significant decrease in the frequency of blinking by an average of 4.8 ± 2.6 times. Strictly speaking, reducing the frequency of blinking and stopping eye movements is necessary to enter a state of silence.

In addition to the state of inner silence, our experiments studied the brain activity in the counting mode up to 100 and back in sevens starting with different initial numbers 0, 1, 2, 3, and so on, in order to avoid memorizing the sequence of numbers when counting. This is a quite difficult intellectual task. In the counting mode, the subject tried to preserve as much as possible the state of inner silence and concentration only on counting without the appearance of extraneous thoughts. During the transition from the state of inner silence to counting by sevens, a slight decrease in the average power of the gamma waves in the temporal regions (electrodes T3 and T4) was observed with a simultaneous significant increase in the average power in the beta1 range (14–20 Hz) in the occipital and parietal regions on the O1-O2 electrodes and on the electrodes P3-Pz-P4 (Fig. 3).
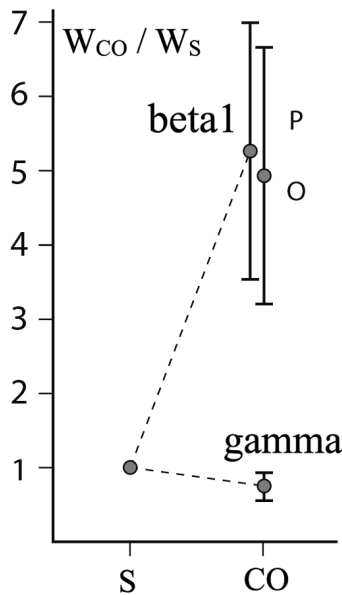
Fig. 3. The ratio of the average power in the counting mode by sevens up to 100 and back (W_{CO}) to the average power in the silence mode (W_S). Gamma band on electrodes T3 and T4, beta1 band on electrodes O1-O2 (O) and P3-Pz-P4 (P).

The next aim of our experiments was to study the effect of the subject's concentration on the power of gamma activity in the brain. After measurements in the state of silence, the subject concentrated his attention on the left temple, trying to increase its activity and while remaining in a state of inner silence. Objective studies have shown an increase in the power of the gamma rhythm in the left temple (T3 electrode) and a decrease in the power in the right temple (T4 electrode). An important feature here is that, as we said above, in different series of experiments the maximum activity in the gamma band was observed in different temples, and sometimes the activity was approximately the same. When concentration was on the left temple, the power of the gamma rhythm in it increased substantially, if in the preceeding state of silence, the power of the gamma rhythm in the left temple was lower than in the right temple. If, before concentration in a state of silence, the power in the left was higher than in the right, then the power in the left temple increased insignificantly, but the intensity in the right temple dropped quite dramatically.

The most prominent variable describing this effect was the ratio of the power of the gamma rhythm in the left temple to the power in the right temple. When concentrating on the left temple, this ratio increased on average by 6.0 ± 2.4 times relative to the state of inner silence.

Philosophy and physiology of inner silence

We begin our discussion of the results and their philosophical interpretation with a quote from Ricard and co-authors which essentially summarizes 15 years of research on the neural correlates of the state of non-objective Buddhist compassion meditation: "Meditation explores the nature of the mind, providing a way to study consciousness and subjective mental states from the firstperson perspective of the meditator. $\Leftrightarrow$ we have studied the brain's electrical activity using electroencephalography (EEG) during compassion meditation in which the meditators described the well-defined sense of self as becoming less fixed and permanent. We found that these long-term Buddhist practitioners were able, at will, to sustain a particular EEG pattern. Specifically, it is called high-amplitude gamma-band oscillations and phase synchrony at between 25 and 42 hertz. $\Leftrightarrow$ Changes in electrical activity may reflect an increased awareness in expert meditators of their surroundings and their internal mental processes, although additional research is needed to better understand the functioning of gamma oscillations." (Ricard, 2014, p. 45). The presented conclusions are consistent with our results considering that in our experiments we studied not the meditation of compassion, when the subject concentrated on a well-defined inner feeling, but the state of complete inner silence, i.e., without content of consciousness in the subject's inner mental space. Putting aside a comparison of either the psychological or the spiritual and theological aspects of compassion meditation and the state of inner silence, we will focus on discussion of the latter.

By the state of inner silence, we mean a state in which a person has no content of consciousness in his inner mental space, including thoughts framed in words, images, desires, intentions, unformed by words, etc. At the same time, however, awareness of oneself is preserved, i.e., my I is perceived as the presence of oneself in being. The self is perceived not as a certain center located in a certain place, for example, in the head, but rather as a self dissolved in space, in the volume of the body and even around it. Perhaps it was this state that Heidegger meant when he used the hard-to-translate concept of Dasein, which he, however, attributed to the state of horror. Dasein is pure

presence when Nothing is manifested because of the absence of the contents of consciousness: "We have already initially given what, for our purposes, is the only essential answer to our question, if we take care that the question about No-thing has actually been posed. For this demands that we carry out the conversion of man into his being there [des Menschen in sein Da-sein], which every instance of dread occasions in us, in order to apprehend No-thing, which is obvious in it as it manifests itself" (Heidegger, 1976, p. 106). Of course, the state of horror is quite difficult to control and even more so to study. The state of silence appears much more evident in yogic practices. In the classic work of M. Eliade "Yoga: Immortality and Freedom" this state is described as follows: "It woud be wrong to regard this mode of being of the Spirit as a simple "trance" in which consciousness was emptied of all content. Nondifferentiated enstasis is not "absolute emptiness." The "state" and the "knowledge" simultaneously expressed by this term refer to a total absence of objects in consciousness, not to a consciousness absolutely empty. For, on the contrary, at such a moment consciousness is saturated with a direct and total intuition of being" (Eliade, 2019, p. 93).

Of course, such a description simplifies the situation, since the state of silence is interrupted from time to time by the spontaneous appearance of thoughts that the silent person can destroy, simply by realizing them and drawing attention back to the state of silence. Ricard and co-authors devide this process into four stages: "Researchers identified four phases of a cognitive cycle: an episode of mind wandering, a moment of becoming aware of the distraction, a phase of reorienting attention and a resumption of focused attention." (Ricard et al., 2014, p. 42).

It is interesting to note the rather obvious fact that in a state of silence our Self cannot control anything simply because there is nothing to control: there are no contents of consciousness that could be controlled. Indeed, the possibility of control, which is associated with the realization of free will, arises when contents appear in the mental space. When a thought, image or desire appear, a person can reject them or allow their further evolution. In other words, free will is exercised in some negative way. Here we can trace quite obvious analogies with the well-known experiments of Libet, where the subject was aware of a spontaneously arising desire to perform a certain action and could subsequently prevent it. Libet associates free choice with a conscious choice in which consciousness participates (Libet, 1999). In fact, the conscious contents of consciousness appear from the depths of

the unconscious and only after reaching a certain threshold do they pass into the category of the conscious. On the basis of objective experiments, it was shown (Shelepin et al., 2020) that unconscious contents of consciousness have a significant impact on human behavior, and conscious behavior is what came through filters of actions that are born in the ocean of the unconscious, causing inhibition or cancellation of unconscious impulses.

From a natural-scientific point of view, the spontaneous appearance of thoughts can be viewed as an intensification of certain initial fluctuations in a nonequilibrium system, which is the brain. It is shown (Danilov et al., 2020) that a single neuron can be considered as a quantum system. Then, neural network is like a multimode laser capable of generating and amplifying various modes. In other words, we can view the brain as a macroscopic quantum system. Indeed, the brain operating in the mode of inner meditation, as is well known even from everyday life, turns out to be quite well isolated from the environment. External stimuli violate this isolation and interfere with the brain work. This is very similar to the decoherence of a closed quantum system by the environment. One may assume that the brain in the mode of thinking operates with a superposition state, and an external stimulus can interrupt the unitary evolution of a closed system. However, in addition to decoherence by the environment, a process associated with the collapse of the brain wave function by the person's own extraphysical Self is possible. This idea, based on the Copenhagen interpretation of quantum mechanics, was formalized by von Neumann, who, in fact, introduced into physics the concept of an extra-physical abstract self capable of collapsing the wave function of a complex macrosystems, including the human brain (Neumann von, 1964). It should be understood that the Copenhagen interpretation of quantum mechanics, which declares the ability of the human self to collapse the wave function of a closed particle-device-brain system, is quite controversial because the decoherence of a particle is usually carried out by the environment. However, this interpretation and von Neumann's ideas are quite suitable for describtion of the impact of the extra-physical self on our own brain. This approach allows solving the problem of free will. In fact, considering the issue of brain control from the outside of the physical self, we come to the Schrödinger's cat paradox, in which the collapse of the wave function of the closed system of one's own brain is carried out by the observer's Self. "Schrödinger cats can be observed by me in my body" (Grib, 2013). These Schrödinger cats that the Self observes in my body are nothing more than thoughts in a superposition state. It turns

out that thinking works in two modes: in the mode of unitary evolution of mental superposition and in the mode of translating it into a discrete form. In a state of silence, this is observed quite clearly. The four phases in meditation (see Introduction) explain this process in more detail. These phases consists in the spontaneous generation of thought due to quantum fluctuations in some neuron, in the amplification and unitary evolution of this thought, in awareness of the thought due to the self of the silent person, and, finally, in the collapse of the mental superposition upon its awareness. The collapse, or reduction of the wave function, upon awareness, manifests one alternative from the mental superposition which results in stopping the thought process and transferring the person back to a state of silence. With ordinary thoughts, after the reduction of the brain wave function, the process of unitary evolution continues until the next collapse. This is how a chain of thoughts arises, which has a kind of reference points associated with the translation of the superposition into a discrete form. It is interesting that this is exactly how Roger Penrose describes the operation of a quantum computer and points out the need for a regular collapse of its wave function in the process of calculations: "Thus, for the most part, it is the action of **U** [unitary evolution] that governs the evolution of the device, with the preservation of such superpositions being an essential part of its action. The **R**-procedure [reduction of the wave function] would become relevant mainly only at the end of the operation, when the system is 'measured' in order to ascertain the result of the computation. In fact (although this is not always recognized) the action of **R** must also be invoked from time to time in a more minor way during the course of the computation, in order to ascertain whether or not the computation has yet terminated." (Penrose, 1994, p. 355–356).

Summarizing this section, let us dwell on those details of our experiments that distinguish them from the already traditional studies of meditative states. Changes in gamma activity at switching to counting mode and during concentration on a specific area of the head (brain). As noted above, the transition to the counting mode by sevens up to 100 and back is accompanied by a slight decrease in gamma activity in the temporal regions most pronounced in the silent mode and by the significant increase in the power in the occipital and parietal regions in the beta1 band on the O1-O2 electrodes and on the P3-electrodes. Pz-P4 (Fig. 3). Note that in the counting mode, an attempt was made by the subject to preserve the silence mode as much as possible, although this may seem like a rather contradictory condition. The data obtained

on changes in the brain activity in the counting mode seem to indicate that in this mode the brain regions responsible for figurative thinking are switched on, since counting is usually accompanied by a visual representation of a numerical sequence. At the same time, a slight decrease in gamma activity (typical for the silent mode) indicates a partial exit from it. The second distinguishing feature of our experiments is the demonstration of the ability to control the spatial distribution of gamma activity in the silence mode. As mentioned above, focusing on the left temple led to an increase in gamma activity in it and a decrease in gamma activity in the right temple. This result is quite interesting but we do not have a clear interpretation of it, except obviously that the Self can control the activity of the brain. It is also possible that this effect demonstrates that figurative and discursive thinking are not very strongly associated with different hemispheres of the brain (at least with activity in the gamma band), since when the gamma activity in the hemispheres changed in the mental space, neither images nor internal monologue characteristic of discursive thinking. From the point of view of subjective experience, mental concentration on different parts of the head and body is accompanied by a feeling of redistribution of a certain substance, for which the most appropriate term may be the word "energy". In the deepest state of inner silence, "energy" is distributed fairly evenly throughout the body and there is a feeling that this "energy" can even go beyond body's limits.

Conclusion

In the present study, the state of inner silence was investigated. From the point of view of neurophysiology this state is identical to the meditative states of the Buddhist tradition actively studied at the present time. The results of objective (according to EEG data) control of the current brain activity and its change under subjective introspective control were obtained. It has been shown that in the state of silence, like in Buddhist meditation, the power of the gamma rhythm increases relative to the basic calm state. The subject's ability to control the spatial distribution of gamma activity when concentrating on various parts of the head was demonstrated. The spatial distribution of activity in the gamma and beta1 ranges during the transition to the performance of intellectual tasks with partial preservation of the mode of inner silence was also documented.

The presented work has shown that the study of controlled altered states of consciousness is a fairly suitable method for studying the phenomenon of consciousness, since in this case the subject is able to give the most con-

scious account of his inner world. However, one needs to understand that the interpretation of such experiments very much depends on the philosophical concept of consciousness. This is due to the fact that the description of subjective experience in altered states of consciousness is an extremely difficult philosophical problem. Indeed, the description of the state of silence, in which there are no contents of consciousness but only awareness of self as a kind of supreme contemplative and governing entity, cannot be carried out in terms of traditional natural sciences. This is because the absence of contents of consciousness presupposes the absence of any form whatsoever. The description of forms is what natural sciences do. That is why in Western philosophy there are difficulties with the interpretation of Descartes' statement "Ego cogito ergo sum". On the one hand, the Descartes-Husserl-Heidegger approach and the Christian tradition assert the extra-physical Self as the center of the personality. On the other hand, the natural-scientific interpretation of Descartes's formula emphasizes the thinking process, and the 'I' as the center of the personality is ignored. A similar situation takes place in the Eastern tradition: " knowledge of one's Self—not in the profane sense of the term, but in its ascetic and spiritual sense—that is the end pursued by the majority of Indian speculative systems" (Eliade, 2019, p. 14). However, in the Buddhist tradition, the concept of the substantial Self is absent, and Buddha " denied the possibility of discussing any absolute principle, as he denied the possibility of having an even approximate experience of the true Self, so long as man was not "awakened."" (Eliade, 2019, p. 163).

From the point of view of philosophical understanding of the results of the presented here experiments and the problem of consciousness in general, it is also fundamentally important that the concept of an extra-physical Self arises in the Copenhagen interpretation of quantum mechanics and von Neumann's formalism, where the Self is the final instance that carries out the collapse of the brain wave function. It means that creation a complete picture of the world reqires a synthesis of natural-scientific rationalism with philosophical concepts and metaphysical models.

Academy of Sciences, K. V. Konstantinov for conducting independent experiments at the Clinic of Bioacoustic Correction, V. L. Tsibulsky for reading the manuscript and useful comments.

References

1. *Danilov O. B.*, *Rozanov N. N.*, *Solovyev N. A.*, *Soms L. N.* Multimode lasers as analogs of complex biological systems. Optics and Spectroscopy, 2016, 120(4): 682–690. [*Данилов О. Б.*, *Розанов Н. Н.*, *Соловьев Н. А.*, *Сомс Л. Н.* Многомодовые лазеры как аналоги сложных биологических систем. Оптика и спектроскопия. 2016. 120 (4): 682–690].

2. Dionysius the Areopagite. On the Mysterious Theology. Historical and Philosophical Yearbook-90. Moscow: Nauka, 1991, 227–232. [Дионисий Ареопагит. О таинственном Богословии. Историко-философский ежегодник-90. М.: Наука, 1991, 227–232].

3. *Eliade M.* Yoga: Immortality and Freedom. New York: Routledge, 1958.

4. *Fell J., Axmacher N., and Haupt S.* From alpha to gamma: electrophysiological correlates of meditation-related states of consciousness. Med. Hypotheses. 2010. 75: 218–224.

5. *Grib A. A.* On the interpretation of quantum physics. Advances in Physical Sciences, 2013, 183(12): 1337–1352. [*Гриб А. А.* К вопросу об интерпретации квантовой физики. УФН. 2013, 183(12): 1337–1352].

6. *Heidegger M.* European Nihilism. In: Nitzsche, V. 4: Nihilism. D. F. Krell (ed.). San Francisco: Harper, 1984.

7. *Heidegger M.* Was ist Metaphysik? In: Wegmarken (1919–1961). Frankfurt am Main: Vittorio Klosterman, 1976.

8. *Lee D. J.*, *Kulubya E., Goldin P., Goodarzi A., and Girgis F.* Review of the Neural Oscillations Underlying Meditation. Front. Neurosci. 2018. 12: 178. doi: 10.3389/fnins.2018.00178.

9. *Libet B. W.* Do We Have Free Will? Journal of Consciousness Studies. 1999. 6 (8–9): 47–57.

10. *Lippelt D. P., Hommel B.,* and *Colzato L. S.* Focused attention, open monitoring and loving kindness meditation: effects on attention, conflict monitoring, and creativity (a review). Front. Psychol. 2014. 5: 1083.

11. *Lutz A., Greischar L. L., Rawlings N. B., Ricard M.,* and *Davidson R. J.* Long-term meditators self-induced highamplitude gamma synchrony during mental practice. Proceedings of the National Academy of Sciences of USA. 2004. 101: 16369–16373.

12. *Neumann von J.* Mathematical Foundations of Quantum Mechanics. Princeton, NJ: Princeton University Press, 1996.

13. *Penrose R*. Shadows of the Mind A Search for the Missing Science of Consciousness. Oxford University Press, 1994.

14. *Ricard M., Lutz A., Davidson R. J.* Mind of the meditator. Scientific American. 2014 (Nov.). 311 (5): 39–45.

15. *Shelepin Yu. E.,   Kharauzov A. K.,   Zhukova O. V.,   Pronin S. V., Kupriyanov M. S., Tsvetkov O. V.* Masking and detection of hidden signals in dynamic images. J. Optical Technology. 2020, 87(10): 89–102. [*Шелепин Ю. Е.*, *Хараузов А. К.*, *Жукова О. В.*, *Пронин С. В.*, *Куприянов М. С.*, *Цветков О. В.* Маскировка и обнаружение скрытых сигналов в динамических изображениях. Оптический журнал. 2020. 87 (10): 89–102].

16. *Solovyev N. A.* Quantum neurophilosophy and rehabilitation of the Cartesian model of consciousness. Journa of Higher Nervous Activity, 2019, 69(1): 120–129. [*Соловьев Н. А.* Квантовая нейрофилософия и реабилитация картезианской модели сознания. Журн. высш. нервн. деят. 2019, 69(1): 120–129].

17. *Solovyev N. A.* Trinity metaphysics and quantum revolution. Sankt Petersburg: Demetra, 2021a. [*Соловьев Н. А.* Троичная метафизика и квантовый переворот. СПб.: Деметра, 2021a].

18. *Solovyev N. A.* Trinity Metaphysics. Voprosy Filosofii. 2021b. 2: 107–116. [*Соловьев Н. А.* Троичная метафизика. Вопросы философии. 2021b. 2: 107–116].

19. *Tang Y. Y.*, *Holzel B. K.*, and *Posner M. I.* The neuroscience of mindfulness meditation. Nat. Rev. Neurosci. 2015. 16: 213–225.

Chapter 24.
Understanding the mechanism of invariant object vision from monkey
inferior temporal cortex

Manabu Tanifuji

RIKEN, Tokyo, Japan

Neural mechanisms of invariant object recognition can be addressed as a classification problem whether or not images of an object are separable from other object images in the neuronal space (DiCarlo and Cox, 2007). Here, the neuronal space means a multidimensional space where axes give responses of individual neurons. For example, faces with the same view of different individuals are more similar than the different views of the same individual in the pixel space. Thus, it is hard to classify individual faces in the V1 neuronal space where each neuron encodes a spatially localized feature. Recent studies have shown evidence that, through a cascade of networks from V1 to the inferior temporal (IT) cortex, the neuronal space to represent objects is transformed into the one where a linear boundary can separate images of an object from those of other objects (Hung, et al., 2005; Chang, et al., 2017; Nam et al., 2021). For example, Nam and colleagues recorded columnar responses to faces from IT cortex (Fig. 1A). They showed that different face views of an individual were well separated from those of other individuals' faces in the neuronal space defined by columnar responses (Fig. 1B, the trace in red). The question is what are the features that make images of an object separable from other object images. Using face columns in IT cortex as a representative case, we addressed this question (Nam, et al., 2021).

How to address the feature encoding of face columns in IT cortex

Previous studies suggested that IT neurons encode visual features with intermediate complexity (Desimone, et al., 1984; Takaka, et al., 1991; Kobatake, et al.,1994). These pioneering studies also clarified several basic characteristics of IT neurons. First, IT neurons have large receptive fields and certain ranges of size invariance (Ito, et al., 1995). The receptive field may

not cover entire visual field, but it is as large as at least the foveal region (Op De Beeck and Vogels, 2000). Second, IT neurons responding to similar visual features are clustered together in columns, and thus we can use responses of columns or multi-unit activities to address visual features encoded by IT neurons (Fujita, et al., 1992; Sato, et al., 2009). Third, recordings from monkeys well trained with a set of visual features revealed that learning causes emergence of neurons responding to the learned stimuli or changes neuronal response property (Kobatake, et al., 1998).
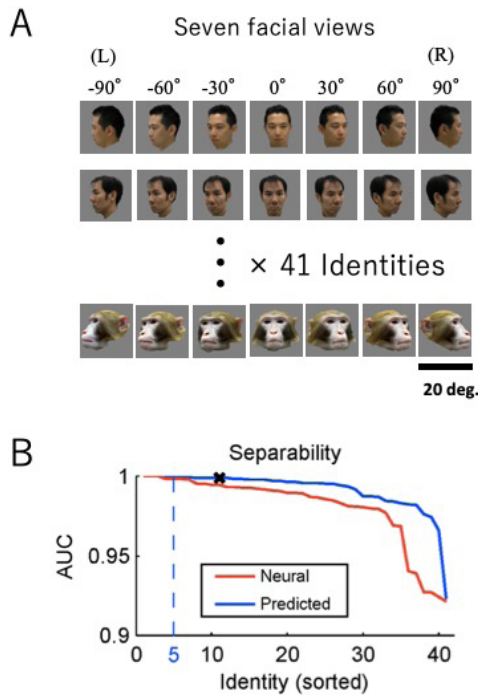

Fig. 1. Face images of an individual were well separated from other faces in the neuronal space and in the visual feature space. (A) The face stimuli consisting of 7 views for 41 individuals. (B) The AUC analysis to evaluate how well images of an individual are separable form others. We searched for the projection vector w that maximally separated the two groups of faces in each space, using simple linear regression. Then, the separability achieved by w was evaluated by AUC (the area under the receiver operating characteristic curve). Vertical and horizontal axes represent AUC and target individuals, respectively. Red and blue represent separability in the neuronal space and the visual feature space, respectively.

Although these pioneering studies provided insights into feature encoding of IT neurons, the way they address the visual features was not sufficient enough to understand invariant object representation because of two reasons. First, these studies explored the visual feature by simplifying a responsive object. In this way, the feature space searching for the visual feature of neurons was limited and thus the approach was more or less subjective. Second, even if we identify the simplest feature that activate the neuron as the feature encoded by the neuron, there is no mathematical model to link responses to arbitrary objects and the feature, which makes it difficult to relate feature encoding and classification of objects in the IT neuronal space. Our challenge was to overcome these two problems.

Toward this goal, first, we searched for the visual features from a set of image fragments that consisted of a massive number of image fragments (n=560,000) instead of stimulus simplification (Nam, et al., 2021). These image fragments were derived from natural images that comprise only small part of all possible images. We think that this constraint is reasonable since our visual system is essentially dedicated to recognition of natural objects. In fact, previous studies searched for features as a part of natural images (Desimone, et al., 1984; Takaka, et al., 1991; Kobatake, et al.,1994; Tsunoda, et al., 2001). By applying this constraint, we made our search space as large as possible within a finite number of elements in our fragment set.

Second, we hypothesized that IT neurons calculate object responses based on Euclidean distances between the visual feature and the objects. This is a template matching method except for calculating distances in the V1 neuronal space but not in the pixel space, meaning that our hypothesis approximated image processing from V1 to IT cortex (Fig. 2). Previous studies revealed that visual features of IT neurons were smaller than the object stimuli (Takaka, et al., 1991; Kobatake, et al.,1994). Because of large receptive fields, IT neurons seemed to respond to a part of objects that matched to the object stimuli (Ito, et al., 1995). Similarly, in our model, we calculate Euclidean distances at various location within each of objects, and took the minimum distance to calculate responses. With this hypothetical model, we can estimate responses to arbitrary objects from identified visual features, and argue classification of objects in the neural space in term of visual feature encoding.
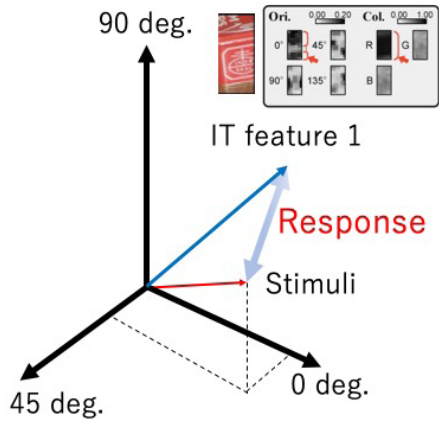
Fig. 2. How responses to visual stimuli were determined by the model. Here, objects and the visual feature are represented in three-dimensional space, but actual dimension is (4 (orientations) + 3 (colors)) x m x n, where m and n are the vertical and horizontal size of the feature encoded by the neuron. Euclidean distance between the visual feature (blue arrow) and each of object stimuli (red arrow) at the V1 neural space. The distance is then fed to radial basis function to make short and long distances tend to 1 and 0, respectively.

Under these considerations, we first calculated responses to object images for each of the fragments in our fragment set. Then, these predicted responses were compared with neuronal columnar responses to these objects. For a given column, if we can find a fragment with its predicted responses to be highly correlated with neural responses to objects, we took the fragment as a good estimate of visual feature encoded by the column. In fact, the best fragment revealed correlation between predicted and columnar neural responses as high as 0.5 for the best visual features (see below).

Neural recording and identified features

We recorded object responses from 88 columns that were reliable and face-selective. Among them, 39 columns exhibited significant response variations across seven face views (Fig. 3). In addition to the seven face views (n = 287) (Fig. 1A), we also recorded responses for view-uncontrolled faces (n = 532), and non-face objects (n = 690). The view-uncontrolled faces and non-face objects, were used to search visual features of these columns. Altogether, the number of visual stimuli were 1,509.
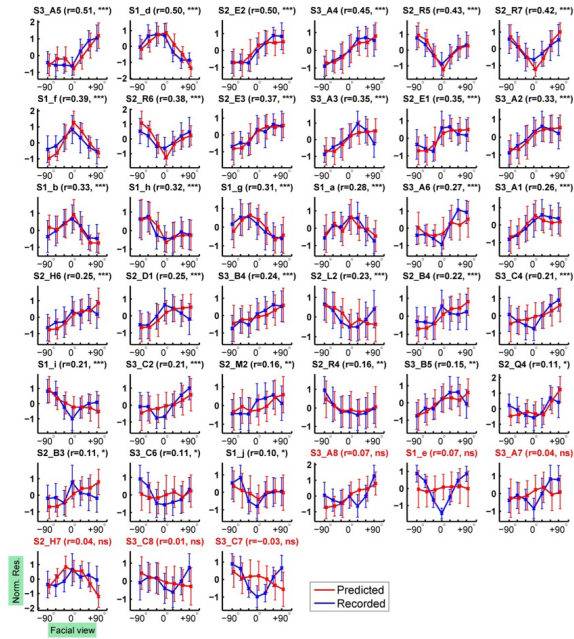
Fig. 3. Neuronal and predicted view tuning curves from 39 columns. Mean and standard deviation across identities are plotted after taking a z-score. The sites are sorted by the correlation between neuronal and predicted view tuning curves (). Among 39 sites, significant correlation (p < 0.05) was found from 33 sites (= 84.6 %, the site IDs are in black). *: p < 0.05, **: p < 0.01, and ***: p < 0.001.

Fig. 4A depicts the visual features for a representative column (site S1_a) which is represented in the pixel space (left) and also in the V1 neuronal space (right). The feature was characterized by a combination of reddish color and horizontal orientation components (Fig. 4A, arrows). This feature matched the faces (stimuli a, b, c, and d, in Fig. 4B) more readily than the non-face objects (stimulus e). Thus, the feature predicted higher responses to faces (8.91 ± 4.95 sp/s; n = 532) than to the non-face objects (0.64 ± 5.60 sp/s; n = 690). These predicted responses were significantly correlated with the neural responses. The correlation across all stimuli including faces and non-face objects and the correlation across faces, = 0.58 and 0.41, respectively (Fig. 4C). The predicted responses also reproduced view tuning and identity tunings at each view of the column (Fig. 4D). The correlation between the view tuning

curves and identity tuning curves were 0.281 (p = 1.3×10^{-6}, n = 287) and
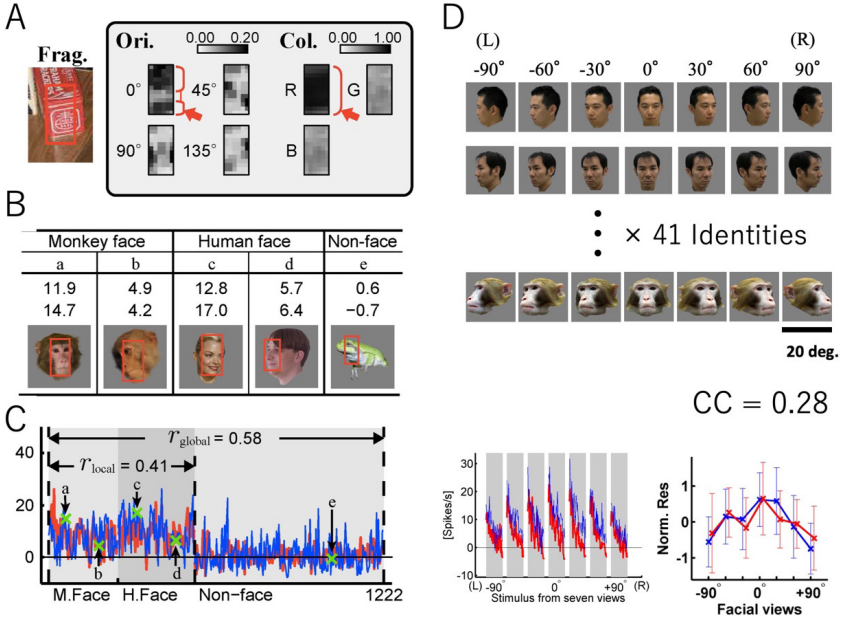0.902 (p < 10^{-6}, n = 41), respectively.

Fig. 4. The feature encoded by a representative column (S1_a). (A) The feature
given in the pixel space (left) and in the V1 neuronal space with 4 orientations and
three colors (right). (B) Neural (above) and predicted (below) responses to the five
representative stimuli. Please note that the predicted responses ranging from 0 to 1
were converted to spike rates using the mean value of the face responses and of the
non-face object responses. (C) The tuning curve of neural (blue) and predicted (red)
responses for view-uncontrolled faces (n = 532), and non-face objects (n = 690). The
objects are grouped in categories in the horizontal axis. From left to right, the cate-
gories are monkey faces (M.face), human faces (H.face), and non-face objects (Non-
face). The stimuli in (B) are denoted by green "x". (D) Neural (blue) and predicted
(red) responses to seven views of 41 identities. In the left panel, the order of identi-
ties is kept the same in individual views. Tendency that responses decreasing from
left to right is the same for seven views, meaning that identity tunings in different
views are similar to each other both in neural (blue) and predicted (red) responses. In
the right panel, responses were averaged for identities to emphasize view tuning of
the column.

For 88 face selective columns, we applied a standard cross-validation test where we divided the objects (n = 1,222) into equal-sized training and test sets (n = 611). Visual features identified for the training set well predicted responses for the test set: the average correlation coefficient between the predicted and neural responses was 0.51 ± 0.09 for all stimuli including faces and non-face objects (global correlation), and was 0.37 ± 0.12 for the stimuli across faces (local correlation). Furthermore, although we identified the features only with view-uncontrolled faces and non-face objects, we found that the identified visual features well predict view tuning of the view sensitive columns (36 among 39 columns) (Fig. 3).

The identified features explained invariant representation of individuals in the neuronal space spanned by a set of face selective columns

As shown at the beginning, face views of one individual were well separated from those of other individuals' faces in the neuronal space defined by columnar responses to these faces (Fig. 1B, the trace in red). The critical question is whether the identified features replicate the view invariant representation of faces. Fig. 5 gives the features of 39 face selective columns. Some of the face selective columns share the same features so that independent features are 32 among 39. Furthermore, the predicted view tuning curves of three columns were not significantly correlated with recorded view tuning curves of these columns (Fig. 3, bottom three panels). Therefore, we addressed the above question using independent and reliable 29 features among 32, and found that 29-dimensional space defined by the identified visual features gave view invariant representation of faces (Fig. 1B, the trace in blue). Although there were variations across individuals, performance was above 90 % and performance was 100 % for five individuals (Left to vertical broken line).

The necessary characteristics of the features to make invariant face representation possible

For a future perspective, it would be more informative to give general characteristics of the visual features instead of describing how individual features contributed to invariant representation of faces.

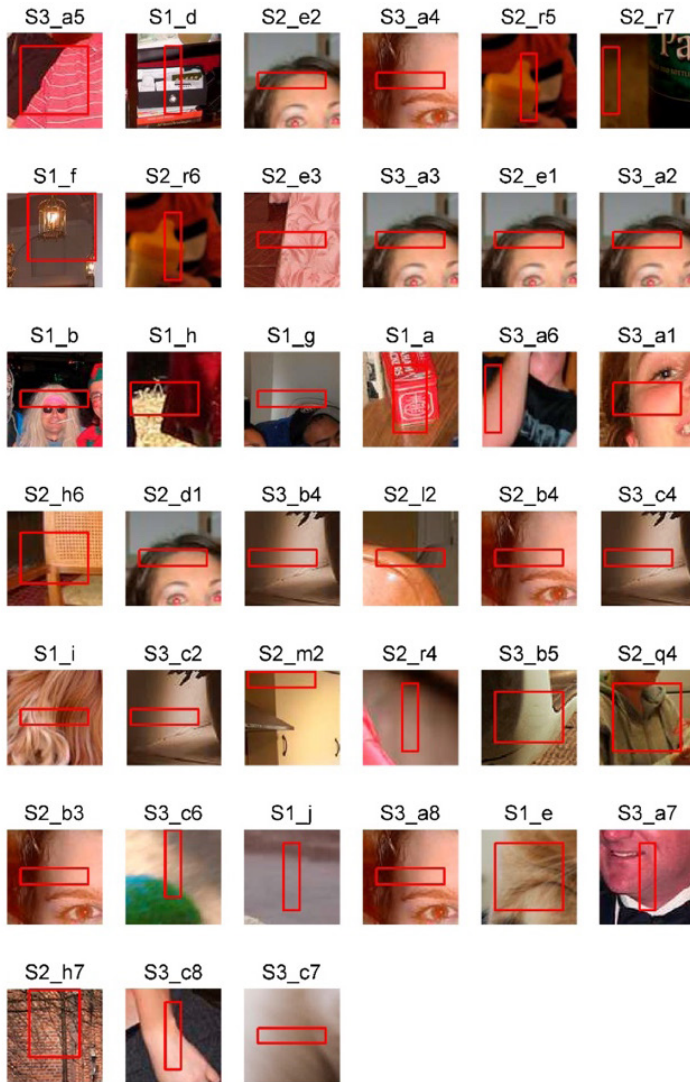**Fig. 5.** The features identified for 39 face columns represented in the pixel space. The rectangular region demarcated in red in each panel indicates the feature. The features are arranged in the same order as the tuning curves shown in Fig. 3, so one can relate features to their view tuning curves. We observed the same feature identified form multiple sites. suggesting that there is a canonical feature set in high-level vision.

First, face neurons are defined as larger responses to faces than non-face objects on average. But, this does not mean that face neurons equally respond to all of faces and do not respond to any of non-face objects. As shown in Fig. 4C, neural responses showed variations across faces and across face and non-face objects. We chose the visual features from the fragment set with the highest correlation among faces and with the highest correlation among faces and non-face objects. Therefore, the estimate of visual features of IT columns also showed response variations across faces and across face and non-face objects. This was not only for the column in Fig. 4. As shown in Fig. 3, for example, there were response variation in views both in neuronal and predicted responses.

Second, we calculated Euclidean distances in the V1 neuronal space but not in the pixel space. It does not simply mean that we projected pixel space images to local orientations and colors space before calculating Euclidean distance, but we also took local invariance of these local features into account such as complex cells. Extracting local orientation and color and local invariance for fragments and objects may be essential for invariant representation of faces at the end. Since we have not yet systematically compared the performance of feature selection with the Euclidean distance calculated in the pixel space and in the V1 neuronal space, this factor remains for future investigations.

In our study, we used neural responses to find the visual feature necessary to make invariant face recognition possible. However, the recording form IT neurons is not easy and takes time. One possible direction would be to search for natural image fragments by taking the above points into account, and to test whether faces are invariantly represented in the space where the axes defined by the fragments. If this is the case, we can achieve invariant object representation with a shallow network that consists of early processing of local orientation with small invariance and late processing of calculating Euclidean distance with large positional invariance.

References

1. *Op De Beeck H., Vogels R.* Spatial sensitivity of macaque inferior temporal neurons // J Comp Neurol. 2000, 426: 505–518.
2. *Chang L., Tsao D. Y.* The code for facial identity in the primate brain // Cell. 2017, 1013–1028.
3. *Desimone R., Albright T. D., Gross C. G., Bruce C.* Stimulus-selective

properties of inferior temporal neurons in the macaque // J Neurosci. 1984, 4: 2051–2062.

4. *DiCarlo J.J., Cox D. D*. Untangling invariant object recognition // Trends Cogn Sci. 2007, 11: 333–341.

5. *Fujita I., Tanaka K., Ito M., Cheng K*. Columns for visual features of objects in monkey inferotemporal cortex // Nature. 1992, 360: 343–346.

6. *Hung C. P., Kreiman G., Poggio T., DiCarlo J. J*. Fast readout of object identity from macaque inferior temporal cortex // Science. 2005, 310: 863–866.

7. *Ito M., Tamura H., Fujita I., Tanaka K*. Size and position invariance of neuronal responses in monkey inferotemporal cortex // J Neurophysiol. 1995, 73: 218–226.

8. *Kobatake E., Tanaka K*. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex // J Neurophysiol. 1994, 71: 856–867.

9. *Kobatake E., Wang, G., Tanaka K*. Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys // J Neurophysiol. 1998, 80: 324–330.

10. *Krizhevsky A., Sutskever I., Hinton G*. ImageNet classification with deep convolutional neural networks // In: Advances in Neural Information Processing Systems. 2012. p. 1097–1105.

11. *Nam Y., Sato T., Uchida G., Malakhova E., Ullman S., Tanifuji M*. View-tuned and view-invariant face encoding in IT cortex is explained by selected natural image fragments // Scientific Report. 2021, 11:7827.

12. *Op De Beeck H., Vogels R*. Spatial sensitivity of macaque inferior temporal neurons // J Comp Neurol. 2000, 426: 505–518.

13. *Sato T., Uchida G., Tanifuji M*. Cortical columnar organization is reconsidered in inferior temporal cortex // Cereb Cortex. 2009, 19: 1870–1888.

14. *Tanaka K., Saito H-A., Fukuda Y., Moriya M*. Coding visual images of objects in the inferotemporal cortex of the macaque monkey // J Physiol. 1991, 66: 170–189.

15. *Tsunoda K., Yamane Y., Nishizaki M., Tanifuji M*. Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns // Nat Neurosci. 2001, 4: 832–838.

Chapter 25.
The role of second-order visual features in the classification of objects

Yavna D. V., Babenko V. V., Gorbenkova O. A., Plavelsky I. V., Voronaya V. D.

Southern Federal University

Rostov-on-Don, Russia

Second-order visual mechanisms are the neuronal systems that enable us (and other mammals) to detect spatial changes in contrast, orientation, and spatial frequency in the visual scene. The existence of these mechanisms was experimentally shown at the 70s of the last century (Henning, Hertz, & Broadbent, 1975), but more detailed study began in the late 80s (Chubb & Sperling, 1988; Fogel & Sagi, 1989; Sutter, Beck, & Graham, 1989 et al.). At that time a theoretical scheme (Figure 1) that is partially relevant till nowadays was proposed. It explains the principle of the second-order visual mechanisms operation: rectified or transformed in other nonlinear way outputs of simple striate cortex neurons form the input of second-order stage neurons that have a much larger receptive field. This allows them to be tuned to lower spatial frequencies; the bandpass peaks of second stage filters are 8–16 times lower (Sutter, Sperling, & Chubb, 1995) than those of first order filters. The scale and orientation invariance of second-order visual mechanisms is achieved due to their multichannel organization (see Shelepin et al., 2008 for review). Each channel has a band-pass tuning for a certain range of orientations and spatial frequencies.
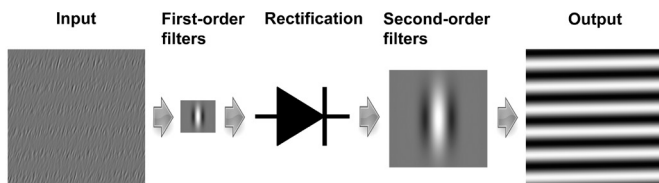

Fig. 1. Theoretical scheme of the second-order visual mechanisms' model.

So, first, simple striate neurons transmit luminance gradients of a certain spatial frequency, and then second-order neurons filter the rectified result at a

reduced frequency. This scheme with minimal differences, in fact, represents the principle of operation of the amplitude modulation (AM) detector, which is well known to radio fans. It detects envelope changes using a bandpass filter tuned on a carrier frequency, rectification process, and a low-pass filter. The main differences in the scheme of second-order visual mechanisms are only that neurons always perform as band-pass filters, and not low-pass ones. In addition, an AM detector and the second-order visual mechanism deal with signals that differ in spatial dimensions: spatial luminance changes in an image are two-dimensional, while instantaneous values of a radio signal are single-dimensional (that means they are changing in time).

This scheme explains the principles of detecting contrast modulations in visual domain, from which studies of second-order visual mechanisms began. Images detection by observers could not be explained in terms of the multiple detectors theory (Graham, 2011) that was popular that days and based on the classical (D. H. Hubel and T. N. Wiesel) neurophysiological studies of the striate neurons properties, and is well enough for explaining the principles of luminance modulations detection.

However, the classic filter-rectify-filter scheme does not explain some aspects of the perception of modulations of various local image features. A mechanism based on this particular scheme would detect contrast modulations best of all. It would also detect the modulations of orientation and spatial frequency, but the proposed filter system does not have the options to distinguish them. At the same time, the human observer distinguishes them easily and preattentively. Currently, there is a sufficient amount of experimental data that allows us to consider the mechanisms for detecting different modulations as relatively independent (Babenko & Ermakov, 2015; Ellemberg, Allen, & Hess, 2006; Kingdom, Prins, & Hayes, 2003; Schofield & Cruickshank, 2005, etc. .). Corresponding corrections to the classical model have been proposed by some authors (Babenko & Yavna, 2018; Kingdom et al., 2003). They do not affect the general principle of its structure and concern only the receptive fields organization of the second-order filters, which implement the specificity of the mechanisms to spatially modulated image feature. The specificity of second-order mechanisms has also been successfully modeled by training convolutional neural networks (Yavna, Babenko, & Ikonopistseva, 2019).

Traditionally, second-order visual mechanisms were considered mainly in the context of detecting and distinguishing textures. Recently, understanding the role of these mechanisms has become significantly wider: a number of studies have shown the mechanisms' participation in the processes of restoring the three-dimensional shape of objects, in figurative-background segmentation and in the visual attention control (Frey, König, & Einhäuser, 2007; Johnson & Zarei, 2010;

Khastkhodaei, Jurjut, Katzner, & Busse, 2016). The purpose of this work is to assess the role of second-order mechanisms in the processes of object classification.

During the study we tested the hypothesis that the information extracted by second-order visual mechanisms from images may be sufficient to classify the objects they contain. Testing the hypothesis helps to answer the question of whether second-order features are informative in the classification task as they are, or whether they are used by the visual system only for preliminary scene segmentation and setting the focus of attention.

We used the method of computer modeling: a neural network, the input of which was formed by before-created and trained models of second-order visual mechanisms, was trained to classify the objects of 20 classes. We are aware that the modeling results cannot be directly extrapolated to a living brain. However, they can serve as important arguments in favor of one or another hypothesis about information processes that are carried out by some brain mechanisms or structures.

Models of second-order mechanisms were presented as fully convolutional autoencoders trained to demodulate spatial contrast, orientation, and spatial frequency modulations on artificial textures. The encoding part of the model included 3 convolutional layers, each of them contained 32 neurons with sizes of 3×3, 3×3, and 5×5, respectively. After each convolutional layer, there was a 2×2 pooling layer. The output layer contained 8 neurons with size 5×5. The input layer of decoding part corresponded to the encoder output layer. The decoding part also contained 3 convolutional layers of 32 neurons with sizes 5×5, 5×5 and 3×3, after each of which 2×2 upsampling was performed; the output layer was convolutional with one 3×3 neuron. A linear ReLU activation function was used for all convolutional layers. The width and height of the encoder input layer have not been specified, which allows the autoencoder to be used to deal with images of any size. The autoencoder scheme is shown in Figure 2.
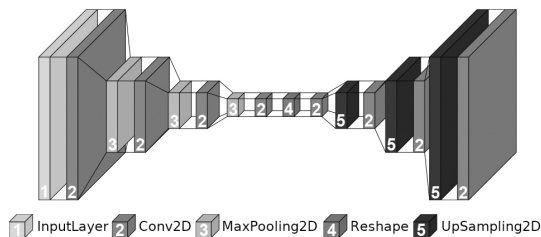

Fig. 2. Schema of a fully convolutional autoencoder that implements a model of a second-order visual mechanism specific to the modulated characteristic of an image. Plotted with the visualkeras library (Gavrikov, 2020).

Training, validation, and testing of the second-order visual mechanisms models were carried out on artificial modulated textures created by randomly superimposing Gabor micropatterns. The algorithm for producing such textures is described in detail (Prins & Kingdom, 2003); it is based on making changes (that are envelope) to the parameters of synthesized micropatterns (that are carrier) depending on their location. The textures randomly varied in the modulated local characteristic (contrast, orientation, or spatial frequency), tilt of the carrier and envelope axes, carrier and envelope wavelengths, envelope phase shift, average brightness, and overall contrast. Continuous trigonometric functions were used as modulators. The pixel dimensions of the textures were 256×256, the pixels were grayscale.

A total of 15000 images were generated, 5000 for each modulated feature. 3500 images from 5000 ones were taken for training, while the rest were used for validation and testing. Thus, the number of images in the training dataset was 10500. During training, textures were fed to the input, and maps of the modulation function values were fed to the output. These maps were saved during the texture synthesis and encoded in such a way that the minimum was 0 and the maximum was 255. The training process was aimed to fitting the model parameters that provide the greatest possible similarity between the expected values of the modulation functions and the results of the encoding-decoding operations. Examples of training textures and the results expected from each modulation-specific model are shown in Figure 3.
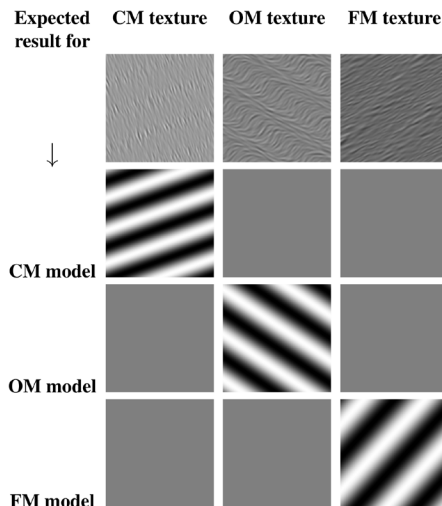

Fig. 3. Training textures and expected results.

The MobileNetV2 network (Sandler, Howard, Zhu, Zhmoginov, & Chen, 2019) was used as a classifier. The input layer accepted images scaled to a pixel size of 224×224, which is due to the presence of a fully connected part in the network. Models of second-order visual mechanisms were located in parallel after the input of the network, and their weights were frozen during training. The concatenation layer was followed by the MobileNetV2 layers. The output was a fully connected layer of 20 softmax-activated neurons encoding the belonging to category (Figure 4). The dropout layer located before the output one had a very low rate in the final version of the model and had no significant effect on training. In the control series of the study the classifier was trained without models of second-order visual mechanisms at the input.
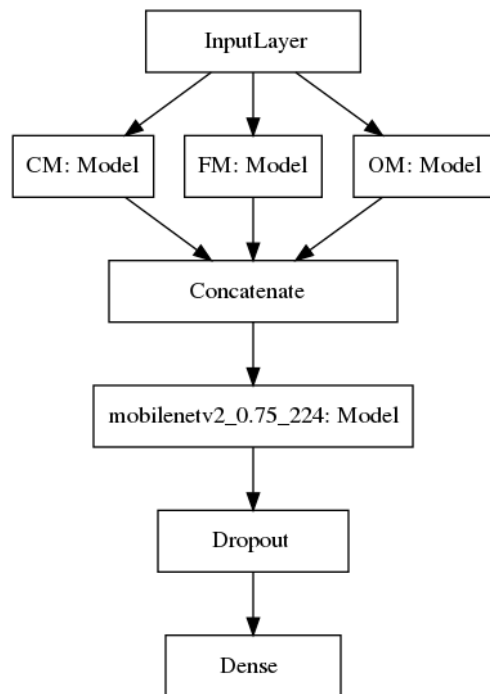

Fig. 4. Block-scheme of a classifier based on MobileNetV2 with models of second-order visual mechanisms at the input.

The LSUN image collection (Yu et al., 2016) was used to train the classifier network. The training was carried out on one Nvidia Tesla M40 12 Gb accelerator, which made it practically impossible to use the entire collection

containing 51 181 482 images. In addition, LSUN contains a significant number of misclassified objects: according to the authors' assessments, they are about 10 % (Yu et al., 2016, p. 6). A random sample was taken; each of the 20 categories of the collection was represented by 11000 training images, 2200 validation images and 2200 test images. Training the classifier became possible only after we removed wrong classified objects from the sample; the final sample included 181717 images for training, 34601 for validation, and 36360 for testing (82 % of the original random sample).

The training was carried out in 200 epochs, and the weights providing the best achieved performance on the validating image set were kept after the end of the epoch. The network, which receives only second-order features as input, did not produce universal representations of class features, and demonstrated strong overfitting: despite the accuracy close to 100 % on the training set, it classified test objects at a random level of 5 % (Figure 5). At the same time, the level of training appropriate for MobileNetV2level was achieved in the control series: the accuracy reached 100 % on the training sample and 84 % on the test sample. The results obtained do not support the hypothesis of the sufficiency of second-order features for the classification of objects.
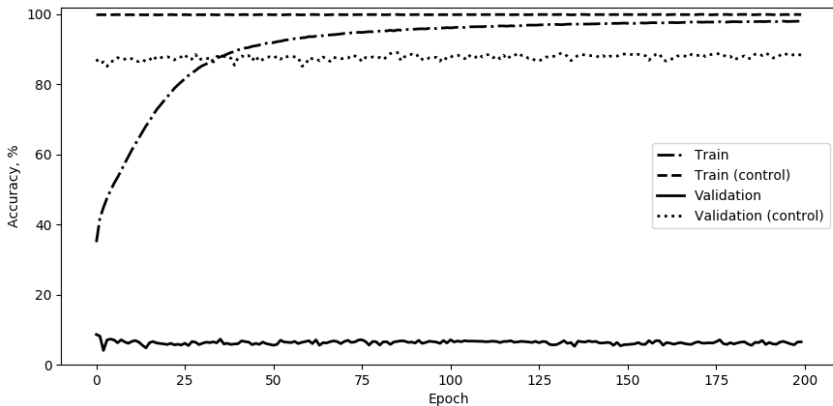
Fig. 5. Dynamics of training of classifiers in main and control series.

However, it is too early to assert that second-order visual mechanisms do not participate in the classification in general and the classification of objects in particular. Previously, we did a similar study with scene images. The network with second order mechanisms models at input has identified a number of classes with high accuracy. The images were selected from the SUN collection (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010), containing 397 classes of

scenes. The selection was not completely random—each class had to contain *objects* of the same basic category: highways, gazebos, telephone booths, etc. The training sample contained 7432 images, and the validation sample contained 1100 images. We formed 20 classes in total.

Classifier models consisted of convolutional and fully connected parts. The input layer accepted images scaled to a pixel size of 256×256. As the input there were three parallel branches, each of which began with a pretrained model of second-order mechanisms specific to contrast, orientation, or spatial frequency modulations. The weights of the models were frozen. Each of the branches included one convolutional layer of $32\ 7 \times 7$ neurons and a 2×2 pooling layer. Single branches were connected into a broad one, containing sequentially convolutional layers of 128, 64 and $16\ 7 \times 7$ neurons, each of which was followed by a 2×2 pooling. The convolutional part was connected to three fully connected layers of 80, 40 and 20 neurons. The image belonging to the class was encoded by activating one of 20 neurons in the output layer. In the models of second-order mechanisms, as in all inner layers of classifiers, the relu activation function was used; the output layer used the softmax function. There were no pre-trained models in the network trained for control purposes (Figure 6).
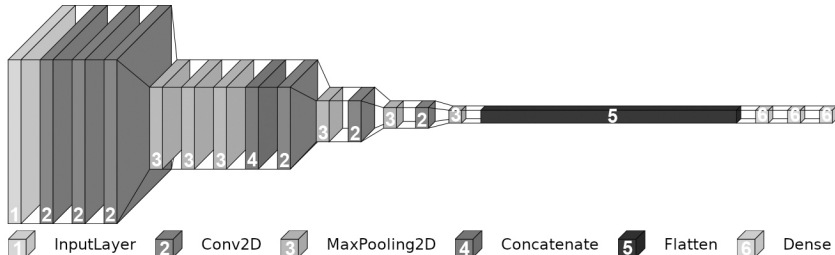
Fig. 6. Scheme of the network used to classify scenes.

The models were trained over 300 epochs, with the best results kept after each epoch. It turned out that by about the 20th epoch the models began to confidently recognize several classes of validating images, reaching 100 % accuracy on the training set, which is associated with the effect of retraining. After training we found classes that were accurately identified by networks:- models with second-order mechanisms at input—bottle storage, phone booth, tower, vineyard;—models without second-order mechanisms at input—amphitheater, barrel storage, cavern, coral reef, gazebo, iceberg, phone booth, staircase, tower, vineyard Thus, the performance of the control model was higher. However, it is import-

ant to note that second-order information is sufficient for correct identification of some scenes that contain objects. The obtained result may be an evidence of the fact that visual mechanisms of the second order play only supporting part in the identification of objects, although they may be sufficient for identification of some classes of scenes.

References

1. *Babenko, V. V., & Ermakov, P. N.* (2015). Specificity of brain reactions to second-order visual stimuli. *Visual Neuroscience*, *32*, E011. doi: 10.1017/S0952523815000085.

2. *Babenko, V. V., & Yavna, D. V.* (2018). Competition for attention among spatial modulations of brightness gradients [in Russian]. *Russian Psychological Journal*, *15*(3), 160–189. doi: 10.21702/rpj.2018.3.8.

3. *Chubb, C., & Sperling, G.* (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *J. Opt. Soc. Am. A*, *5*(11), 1986–2007.

4. *Ellemberg, D., Allen, H. A., & Hess, R. F.* (2006). Second-order spatial frequency and orientation channels in human vision. *Vision Res*, *46*(17), 2798–2803. doi: 10.1016/j.visres.2006.01.028.

5. *Fogel, I., & Sagi, D.* (1989). Gabor filters as texture discriminator. *Biol. Cybern.*, *61*(2), 103–113.

6. *Frey, H.-P., König, P., & Einhäuser, W.* (2007). The role of first- and second-order stimulus features for human overt attention. *Perception & Psychophysics*, *69*(2), 153–161. doi: 10.3758/bf03193738.

7. *Gavrikov, P.* (2020). *Visualkeras*. GitHub. Retrieved from https://github.com/paulgavrikov/visualkeras.

8. *Graham, N. V.* (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. *Vision Research*, *51*(13), 1397–1430. doi: 10.1016/j.visres.2011.02.007.

9. *Henning, G. B., Hertz, B. G., & Broadbent, D. E.* (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Res.*, *15*(8–9), 887–897.

10. *Johnson, A., & Zarei, A.* (2010). Second-order saliency predicts observer eye movements when viewing natural images. *Journal of Vision*, *10*(7), 526–526. doi: 10.1167/10.7.526.

11. *Khastkhodaei, Z., Jurjut, O., Katzner, S., & Busse, L.* (2016). Mice can use second-order, contrast-modulated stimuli to guide visual perception. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(16), 4457–4469. doi: 10.1523/JNEUROSCI.4595–15.2016.

12. *Kingdom, F. A. A., Prins, N., & Hayes, A.* (2003). Mechanism independence for texture-modulation detection is consistent with a filter-rectify-filter mechanism. *Visual Neuroscience*, *20*(1), 65–76. doi: 10.1017/s0952523803201073.

13. *Prins, N., & Kingdom, F. A. A.* (2003). Detection and discrimination of texture modulations defined by orientation, spatial frequency, and contrast. *Journal of the Optical Society of America A*, *20*(3), 401. doi: 10.1364/JOSAA.20.000401.

14. *Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.‑C.* (2019). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *ArXiv:1801.04381 [Cs]*. Retrieved from http://arxiv.org/abs/1801.04381.

15. *Schofield, A., & Cruickshank, A.* (2005). Transfer of tilt after-effects between second-order cues. *Spatial Vision*, *18*(4), 379–397. doi: 10.1163/1568568054389624.

16. *Shelepin, Yu. E., Chikhman, V. N., Vakhrameeva, O. A., Pronin, S. V., Foreman, N., & Pasmore, P.* (2008). Invariance of visual perception (in Russian). *Experimental Psychology (Russia)*, *1*(1). Retrieved from https://www.elibrary.ru/item.asp?id=13019577.

17. *Sutter, A., Beck, J., & Graham, N. V.* (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. *Percept. Psychophys.*, *46*(4), 312–332.

18. *Sutter, A., Sperling, G., & Chubb, C.* (1995). Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision Research*, *35*(7), 915–924. doi: 10.1016/0042–6989(94)00196-s.

19. *Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A.* (2010). SUN database: Large-scale scene recognition from abbey to zoo. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3485–3492. doi: 10.1109/CVPR.2010.5539970.

20. *Yavna, D. V., Babenko, V. V., & Ikonopistseva, K. A.* (2019). Neural network models of second order visual filters. *Neural Networks and Neurotechnologies*, 198–203. St. Petersburg, Russia: BBM.

21. *Yu, F., Seff, A., Zhang, Y., Song, S., Funkhouser, T., & Xiao, J.* (2016). LSUN: Construction of a Large-scale Image Dataset using Deep Learning with Humans in the Loop. *ArXiv:1506.03365 [Cs]*. Retrieved from http://arxiv.org/abs/1506.03365

Chapter 26.
Interval Analysis of the Duration of Fixations in the Model Card Game
«I Believe / I Don't Believe It»

Zhegallo A. V., Basul I. A.

The Institute of Experimental Psychology, Moscow State University of Psychology and Education, Moscow, Russia.

Corresponding author: Zhegallo A. V., zhegs@mail.ru

When examining a static face, the observer is completely devoid of information about the dynamics of facial expressions. As a result, eye movements when examining a static face are characterized by significantly greater variability than when examining a dynamic face (Barabanshchikov, Zhegallo, 2018). Against the background of high individual variability when examining static facial expressions, it is possible to establish differences in examining a normally oriented and an inverted face, but not differences in examining various "basic" emotional expressions (Barabanshchikov, 2012). In recent years, we have gradually refused the research of the perception of a static face, moving on to the research of the perception of natural dynamic facial expressions in conditions close to natural.

The first step here was to compare the indicators of oculomotor activity in the perception of a static and dynamic face (Barabanshchikov, Zhegallo, 2018). Based on the material of the research, the predictors of the reliability / unreliability of the reported information were further identified, which were binary markers of non-verbal activity manifestations identified by experts (Barabanshchikov, et al., 2018). In addition, predictors of the reliability / unreliability of the reported information were identified in the form of integral patterns of facial expressions, identified as main components when analyzing the dynamics of a set of reference points on the surface of the face (Barabanshchikov, Zhegallo, 2021). The possibilities of this experiment are limited by the situation of vicarious communication with the communicant; the participants of the research were shown the fragments of the video recording of the living face of the sitter in artificially constructed situations, suggesting that the sitter must provide knowingly reliable or inaccurate in-

formation, as well as in a professional interview situation.

As part of the research, a special experimental procedure for a pair experiment with communication mediated by video system was developed (Zhegallo, Basyul, 2019a). The experiment was carried out in the form of the «I Believe / I Don't Believe It», model card game. Analysis of the experimental data obtained showed that during the game, the subjects actively use video communication to interact with their opponents (Zhegallo, Basyul, 2019b). In our last research (Zhegalo, Basyul, 2021), the duration of fixations was analyzed at certain stages of the model game (making a move, observing the opponent's assessment of the move, observing the opponent's move, making a decision "to believe" / "to not believe" the reported information). It was found that the average duration of fixations in experimental situations (ES) associated with the decision «to lie» is shorter than in experimental situations associated with the decision «to tell the truth.» Also, the average duration of fixations in ES related to making a decision "not to believe" is shorter than in ES related to making a decision "to believe". Additional analysis of the dynamics of facial expressions made it possible to identify a complex facial expression pattern associated with the decision «to tell the truth» / «to lie».

In this work, we additionally conduct a detailed analysis of the duration of fixations at distinct phases of the model game. Each phase is divided into 1 second analysis intervals. The analysis includes, without any averaging, all fixations related to the given analysis interval. In this case, the result of the analysis should be considered not individual significant differences in the duration of fixations at specific intervals of analysis, but the general pattern of the presence / absence of differences. Obtaining a regular pattern that allows meaningful interpretation of the results is an argument in favor of the fact that the results obtained are not an artifact due to multiple comparisons.

Experimental procedure

We analyzed the data obtained during the experiment described in (Zhegallo, Basul, 2019b). The participants in the experiment took part in the model card game «I Believe / I Don't Believe It». During the experiment, eye movements, facial expressions of the participants, and the protocol of the experiment (participants' moves) were recorded. Both players have their game cards and a button for adding cards at the bottom of the screen. Adding cards is carried out at the beginning of the game and after the «consumption» of cards

in the course of the game. A video image of the opponent's face is displayed in the window in the center left of the screen.

During the game, the participants alternately perform moves, during which a game card and a proposed card are selected. The proposed card (which matches or does not match the game card) is shown to the opponent. Thus, the player either «tells the truth» (the offered card is the same as the game card), or «lies» (the offered card does not match the game card). The opponent must decide whether he believes that the proposed card is the same as the game card. If the correct answer is "I believe", both players receive 1 point. If the correct answer is "I don't believe", the guessing player gets 2 points. If the answer is incorrect, the player who guessed the card gets 2 points.

The recearch used the original hardware and software complex (Zhegallo, Basyul, 2019a, b). Registration of eye movements was performed using SMI REDm eye trackers, frequency 120 Hz. The playing field was demonstrated on a ViewSonic VA903m LCD monitor with a screen resolution of 1280×1024. The distance from the subjects' eyes to the surface of the screen was 60 cm. The subjects were in motionless, non-rotating KE-1 chairs (Neurobotics), their head resting on the headrest.

The analysis of the characteristics of eye movements was performed on a sample of 13 games, which corresponded to 26 individual records of eye movements. Records of 3 couples, in which one of the participants only used the "telling the truth" strategy, were excluded from the analysis). The duration of the game was from 10 to 21 minutes, m = 15.5 minutes, iqr = 13.6 minutes — 16.5 minutes. Fixations were detected using the IDT (Dispersion thresholdid entification) algorithm from the ETRAN package (Zhegallo, Marmalyuk, 2015) for the R environment (R CoreTeam, 2020), the minimum duration was 6 data samples (50 ms), the maximum variance was 40 pixels (1 ° at a distance to the screen of 60 cm.).

The individual proportion of moves in which the participants reported the truth ranged from 0.22 to 0.78; m = 0.51, iqr = 0.44–0.67. The individual proportion of moves in which the participants gave the answer «I believe» ranged from 0.28 to 0.83; m = 0.44, iqr = 0.39–0.50. The number of points earned ranged from 22 to 47, m = 35, iqr = 32–38. There is no significant correlation between the proportion of truthful messages and the proportion of "I believe" responses. There is no significant correlation between the proportion of truthful messages and the number of points earned. The proportion of "I believe" answers is negatively correlated with the number of points earned; r = –0.39, p = 0.05.

Analysis of eye movements

The analysis of the eye movements was carried out on a total sample of 1864 experimental situations (ES). For each ES, the average duration of fixations and the quality of the recording were calculated as the ratio of the total duration of fixations to the duration of ES. The distribution for the duration of the fixations is different from the normal one.

From the player's point of view, the following 4 phases can be distinguished during the game.

1 — The player makes a move deciding «to tell the truth» or «to lie».

2 — The player observes how the opponent evaluates his move, making a choice of «to believe» or «not to believe».

3 — The player observes how his opponent makes his move, the opponent chooses «to tell the truth» or «to lie».

4 — The player makes a choice «to believe» or «not to believe»

1 '— The player makes the next move.

The data sample contained 466 ESs related to each phase of the game. The analysis of eye movements was performed separately for each phase of the game. The duration of the move was from 4.50 sec to 208.57 sec; Me = 15.45 sec, 95 % interval = 7.46 sec, 38.10 sec. The duration of the decision to believe / not to believe ranged from 0.05 sec to 38.35 sec, Me = 3.92 sec, 95 % interval 1.22 sec — 16.72 sec.

Phase 1 (the player makes a move)

For analysis, out of 466 ESs, 374 ESs were selected. Selection criteria: execution time no more than 50 sec, recording quality above 0.7. Of these, in 175 ES the subjects «lied», in 199 — «told the truth.» The difference in the average duration of fixations for ES, where the player «lies» and «tells the truth» is significant; Mann — Whitney test, p = 0.05. Intersample shift according to Hodges — Lehmann $\Delta = -10$ ms; Me (false) = 200.0 ms, Me (true) = 212.5 ms. Assuming that the reduction in the duration of fixations is associated with the decision «to lie» / «to tell the truth», we additionally localized the time interval in which the reduction in the duration of fixations is most pronounced. For this, among the fixations related to the previously identified ES, the selection of those fixations was carried out, the start time of which is within a given interval of 1 sec. The results are presented in Tab. 1.

Table 1

Median duration of fixations for 1-second time intervals for the 1st phase of the game. N0 is the number of fixations, related to the ES in which the player «lied». N1 is the number of fixations related to the ES in which the player «told the truth». p—level of significance of differences (Mann-Whitney test). Me(false)—the median duration of fixations related to ES, in which the player «lied» (ms). Me(true)—the median duration of fixations related to ES, in which the player «reported the truth» (ms). Δ—inter-sample shift according to Hodges—Lehmann (ms)

Interval start time (sec)	N0	N1	p	Me(false)	Me(true)	Δ
0	823	898	0.006	183	200	-17
1	664	707	0.00025	175	192	-25
2	669	738	0.016	192	204	-17
3	638	792	insignif	192	200	
4	632	734	insignif	200	208	
5	610	685	insignif	200	200	
6	605	681	insignif	208	217	
7	582	634	0.029	200	217	-17
8	560	619	0.046	200	208	-17
9	530	590	insignif	208	217	
10	479	561	insignif	208	208	
11	418	520	insignif	208	208	
12	381	453	insignif	208	208	
13	338	403	0.069	208	217	-17
14	300	351	insignif	208	208	
15	234	301	insignif	208	208	
16	233	254	insignif	217	229	
17	198	203	insignif	221	208	
18	158	163	insignif	208	242	
19	149	149	insignif	200	217	

As can be seen from the table, the main decision-making interval, which conditions the differences in the duration of fixations, is 0–3 seconds. The visualization of the fixation positions, related to this interval, showed that they

are localized in the right area of the screen (interface elements, cards, conceived and shown by the opponent, information about the number of points received). The second additional interval (with lower significance levels) corresponds to times of 7–9 sec. The visualization of the fixation positions related to this interval showed that they are localized in the upper area of the screen (informational message), the right area (conceived and shown cards), the area of the opponent's face display. There is also a tendency towards the presence of the 3rd additional interval of 13–14 sec.

Phase 2 (the player observes how the opponent evaluates his move)

For analysis, out of 466 ESs, 303 ESs were selected. Selection criteria: duration of execution is more than 1 sec, recording quality is higher than 0.7. Of these: 88 ES—«a lie», «I do not believe»; 53 ES «a lie», «I believe»; 77 ES "a truth", "I don't believe", 85 ES "a truth", "I believe". The difference in the average duration of fixations for the ES, where the player previously «lied» and «told the truth» at the level of tendencies; Mann—Whitney test, $p = 0.08$. Intersample shift according to Hodges—Lehmann $\Delta = -19$ ms, Me (false) = 277.7 ms, Me (true) = 296.2 ms. An additional analysis of the duration of fixations at 1-second intervals (Table 2) showed that the noted tendency towards a reduction in the duration of fixations takes place only for the first two intervals.

Table 2

Median duration of fixations for 1-second time intervals for the 2nd phase of the game

Interval start time (sec)	N0	N1	p	Me(false)	Me(true)	Δ
0	544	602	0.088	217	233	-17
1	410	460	0.095	200	217	-17
2	297	329	insignif	217	217	
3	221	242	insignif	217	217	
4	171	166	insignif	217	229	
5	119	115	insignif	242	250	
6	97	87	insignif	233	283	
7	79	63	insignif	200	208	
8	51	54	insignif	225	204	
9	41	37	insignif	250	250	

Phase 3 (the player watches his opponent make his move)

For analysis, out of 466 ESs, 309 ESs were selected. Selection criteria: execution time no more than 50 sec, recording quality above 0.7. Of these, in 147 ES, the subjects «lied», in 162 — «told the truth.» The difference in the average duration of fixations for ES, where the opponent is "lying" and "telling the truth" is significant; Mann — Whitney test, $p = 0.008$. Hodges-Lehmann inter-sample shift $\Delta = -27$ ms; Me (false) = 269.9 ms, Me (true) = 306.4 ms. Additional analysis of the duration of fixations at 1-second intervals (Table 3) showed the presence of the following intervals, for which there is a change in the duration of fixations, presumably due to non-verbal information coming from the opponent: 1–4 sec., 10–12 sec., 15–19 sec.

Table 3

Median duration of fixations for 1-second time intervals for the 3rd phase of the game

Interval start time (sec)	N0	N1	p	Me(false)	Me(true)	Δ
0	544	605	insignif	192	208	
1	524	559	0.039	192	200	-17
2	543	554	0.005	183	208	-17
3	527	535	0.049	200	208	-17
4	508	562	insignif	200	217	
5	480	501	insignif	200	208	
6	462	482	insignif	208	221	
7	413	444	insignif	217	217	
8	404	396	insignif	208	233	
9	367	384	insignif	217	225	
10	366	305	0.006	208	233	-33
11	274	257	0.018	229	250	-33
12	287	257	insignif	208	217	
13	226	230	insignif	233	225	
14	215	219	insignif	233	225	
15	193	165	0.00047	217	275	-50
16	160	137	0.001	217	292	-58
17	123	115	0.035	208	267	-42
18	123	95	0.02	233	283	-58
19	93	68	insignif	217	254	

Phase 4 (the player makes the choice «to believe» or «not to believe»)

For analysis, out of 466 ESs, 345 ESs were selected. Selection criteria: duration of execution is more than 1 sec, recording quality is higher than 0.7. Of these: 97 ES—«a lie», «I do not believe»; 66 ES «a lie», «I believe»; 88 ES "a truth", "I do not believe", 95 ES "a truth", "I believe". The difference in the average duration of fixations for ES, where the opponent is "lying" and "telling the truth" is significant; Mann—Whitney test, $p = 0.005$. Inter-sample shift according to Hodges—Lehmann $\Delta = -23$ ms; Me (false) = 257.4 ms, Me (true) = 282.7 ms. The difference in the average duration of fixations for ES, where the player "believes" and "does not believe" is significant; Mann—Whitney test, $p = 0.01$. Intersample shift according to Hodges—Lehmann $\Delta = -21$ ms; Me (false) = 263.6 ms, Me (true) = 276.2 ms. An analysis of 4 combinations of the player's and opponent's strategy showed the significance of the differences; Kruskull-Wallace test $p = 0.004$; Me (a lie, I don't believe) = 251.86 ms; Me (a lie, I believe) = 269.52 ms; Me (a truth, I don't believe it) = 278.09 ms; Me (a truth, I believe) = 287.90 ms.

An additional analysis of the duration of fixations at 1-second intervals (Table 4) for ES in which the opponent is "lying" / "telling the truth" has been carried out. For 1-second intervals of 0–1 sec and 7–8 sec, there is a significant reduction in the duration of fixations for ES, in which the opponent is "lying".

Table 4

Median duration of fixations for 1-second time intervals for the 4th phase of the game (for ES, in which the opponent «tells the truth» / «lies»).

Interval start time (sec)	N0	N1	p	Me(false)	Me(true)	Δ
0	674	722	0.01	204	217	-17
1	515	540	insignif	200	208	
2	390	420	insignif	200	208	
3	267	294	insignif	225	208	
4	215	203	insignif	208	208	
5	145	134	insignif	242	212	
6	123	103	insignif	208	233	
7	99	74	0.008	183	254	-58
8	76	65	insignif	188	208	
9	60	43	insignif	238	217	
10	56	37	insignif	217	250	

An additional analysis of the duration of fixations at 1-second intervals (Table 5) for ES in which the player "believes" / "does not believe" the information reported has been carried out.

Table 5

Median duration of fixations for 1-second time intervals for the 4th phase of the game (for ES, in which the player «believes» / «does not believe» the information reported).

Interval start time (sec)	N0	N1	p	Me(don't belive)	Me(belive)	Δ
0	762	634	0.003	200	217	-17
1	557	498	insignif	208	200	
2	435	375	insignif	200	208	
3	314	247	0.0045	200	233	-33
4	217	201	insignif	200	217	
5	132	147	insignif	225	242	
6	102	124	insignif	233	200	
7	75	98	insignif	200	212	
8	77	64	insignif	200	212	
9	45	58	insignif	200	242	
10	43	50	insignif	225	225	

The discussion of the results

In this research, we used the interval analysis of the duration of fixation as a technique that allows us to localize the time intervals corresponding to the decision-making by the participants of the experiment. How informative was this technique? Let's start with the 1st phase of the game. When making his move, the player is maximally involved in the game process and focuses on depicting the game situation. In the data presented, this is clearly expressed in a sufficiently high number of visual fixations up to the last analysis interval. It can be assumed that the main algorithm of the player's actions is to first make a decision about the strategy to be used, and then translate it into an appropriate choice of cards. This algorithm of actions can be associated with the first decision-making interval: 0–3 seconds. Late decision-making intervals can be associated with the fact that in some cases the participants in the game start actions without a ready-made strategy. Verification of this

assumption requires an increase in the sample size, including the amount of material related to each of the participants. In this case, it becomes possible to construct individual profiles of the duration of fixations, as well as to identify ES, in which a late decision was made.

It should be noted that in our recent additional series of experiments, we suggested that the participants play two games sequentially, while in the second game, as a rule, the speed of players making moves significantly increased, so that it becomes possible to increase the amount of collected experimental data (the number of analyzed ESs) without a radical increase in the duration of the experiment. At the same time, it can be expected that as the volume of the game increases, the number of cases of late decision making will decrease. It should also be noted that for the SMI REDm eye trackers we used, the recording frequency is 120 Hz, which gives an accuracy of 8.3 ms in determining the duration of fixations. The use of a high-speed eye-tracker (for example, Tobii Pro Spectrum) would provide greater accuracy in determining the duration of fixations and thereby increase the efficiency of localizing decision-making intervals with a relatively small amount of experimental data.

Summarizing the results related to the 1st phase of the game, we believe that the experimental data obtained reliably indicate the presence of the main decision-making interval (0–3 sec), during which the player makes the decision «to tell the truth» / «to lie». At the level of the parameters of eye movements, this decision is expressed by a reduction in the duration of fixations when making a decision «to lie» by 17–25 ms. It can be expected that with an increase in the individual volume of the model game, it will become possible to build individual profiles of the participants and further predict the strategy used by a particular player based on the indicators of eye movements.

In the 2nd phase of the game, the player passively observes the actions of the opponent, who decides whether to believe the information provided. At this stage, the result does not depend in any way on the actions of the player. The characteristic duration of phase 2 (Me = 3.92 sec) is significantly shorter than that of phase 1 (Me = 15.45 sec). As a result, already by the 10th 1-second interval of analysis, the number of fixations included in the sample is drastically reduced. Differences in the duration of fixations for ES, in which the player previously «lied» / «told the truth» at the level of tendencies, are present only in the interval of 0–2 sec. This result has no independent practical significance. In theoretical terms, the observed effect can be interpreted as a repeated weakly expressed confirmation of a previously made decision at the moment the «Play» button is pressed.

The 3rd phase of the game is a mirror image of the 1st phase. Here the opponent makes a decision «to lie» or «to tell the truth.» The player at this moment can observe the expression on the opponent's face, trying to highlight the signs of the message of reliable / unreliable information. However, the player does not always follow this strategy, as a result of which the number of ESs suitable for analysis turns out to be less than in phase 1. Thus, the number of fixations included in the analysis at each of the 1-second intervals is reduced. Nevertheless, the sample size remains sufficient for analysis up to the last 20th interval. Within 20 seconds, 3 intervals are distinguished, at which significant differences are observed in the duration of fixations between the ES, in which the opponent "lies" / "tells the truth"— 1–4 seconds, 10–12 seconds, 15–19 seconds. This result can be interpreted in two ways.

Option 1: there are high-sensitivity periods in the process of the player's perception. During these periods, the player is more sensitive to the received non-verbal information. Differences in the localizations of fixations at different intervals of analysis could be an indirect confirmation of this assumption. An earlier analysis (Zhegallo, Basyul, 2021) shows that the differences in the duration of fixations are maximum at the optimal number of fixations in the video display area, with $2 < nf < 11$; n (false) = 56, n (true) = 60; p = 0.003; Δ = –50 ms; Me (false) = 253.4 ms, Me (true) = 312.9 ms. Thus, it can be assumed that there are complex internal determinants that determine the processing of incoming non-verbal information. It should be additionally emphasized that according to the results of the experiment, the incoming non-verbal information does not lead to an increase in the accuracy of assessing the reliability of the information being reported.

Option 2: in the dynamics of the opponent's facial expressions, there are periods when the desired non-verbal signs are more actively manifested. This assumption can be verified by performing additional interval analysis of the dynamics of the opponent's face. On the material of marking the key points related to the video recordings of the opponent's facial expressions in phase 3, a principal component analysis is performed and then the principal component loads are calculated for 20 consecutive 1-second intervals. The necessary analysis is planned to be carried out at the next stage of research.

The 4th phase of the game is, up to a reversal of roles, equivalent to the 2nd phase. The player who is completely passive in the 2nd phase of the game must now make a decision "to believe" / "not to believe" of the information being reported. In this phase, there are two periods for which there are differences in the duration of fixations associated with the opponent's decision and

two periods associated with the player's "I believe" / "I do not believe" rating. In this case, the first period in both cases corresponds to the interval 0–1 sec. The reduction in the duration of fixations associated with obtaining unreliable information is less pronounced than during phase 3. This result is fully explained by the shorter duration of phase 4 compared to phase 3. In principle, an additional analysis of the opponent's facial expressions is possible here during the first 1- second interval.

The results in terms of the reduction in the duration of fixations associated with the assessment «I do not believe» are of interest in connection with the possible transfer to a wide class of problems associated with decision-making under conditions of uncertainty. It can be assumed that in psychophysical experiments related to decision-making under conditions of uncertainty, low confidence in decision-making will be associated with shorter visual fixations. According to the discrete decision-making model (Shendyapin, Skotnikova, 2015), the decision is made on the basis of a series of repeated measurements. Assuming that in our case visual fixation corresponds to a separate measurement, it can be assumed that the initial low confidence in the decision made will lead to a reduction in the duration of visual fixations, which will make it possible to obtain a larger number of repeated measurements in a limited time. This hypothesis can be tested experimentally in a psychophysical experiment with confidence fixation and parallel registration of eye movements.

The performed recearch shows the possible perspective of applying the methods of registration of eye movements and video recording of facial expressions for a practical assessment of the reliability / unreliability of the information reported. This assumption can be verified in experiments with parallel registration of standard polygraphic indicators (GSR, respiration, photoplethysmogram), eye movements and facial expressions. Facial expressions are analyzed at the level of integral expressive movements, distinguished from the dynamics of a set of previously detected key points (facial landmarks) by the method of principal components (Zhegallo, Basyul, 2021). Verification of the hypothesis involves comparing standard polygraphic indicators, the duration of fixations at discrete 1-second intervals of analysis, the severity of the main components in the dynamics of key points (with discreteness of one video frame). For the experiment, it is preferable to use high-speed eye trackers, which provide high accuracy in measuring the duration of visual fixations. When performing video recording of facial expressions, it is advisable to use cameras of Full HD cameras, 60 fps. The video processing

we are currently using with the OpenFace software provides the coordinates of 68 key points. Switching to the use of Google MediaPipe allows you to increase the number of selected key points to 468 with partial coverage of the forehead area.

Summary

This research continues the analysis of eye movements in a model experiment — a computer version of the card game «I Believe / I Don't Believe It», mediated by video communication. For analysis, the model game is divided into 4 phases. Phase 1 — the player taking a move. Phase 2 — the player observing the opponent's decision-making. Phase 3 — the player observing the opponent's taking a move. Phase 4 — the player making a decision. To perform a detailed analysis of eye movements, each phase was additionally divided into 1-second analysis intervals. Samples of visual fixations were compared for each analysis interval.

For phase 1, intervals are discovered that correspond to the player making a decision on the strategy of the game «to tell the truth / to lie». The duration of fixations in the case of the decision «to lie» is significantly shorter than in the case of the decision «to tell the truth.» For phase 3, the intervals are discovered, during which the duration of fixations is determined by non-verbal information about the decision made by the opponent. The results obtained can be further used to develop a methodology for assessing the reliability / unreliability of the reported information on eye movements and the dynamics of facial expressions.

Funding

References

1. *Barabanschikov V. A.* Ekspressii lica i ih vospriyatie [Facial expressions and perception]. — Moscow, IPRAS, 2012.
2. *Barabanschikov V. A.*, *Zhegallo A. V.* Okulomotornaya aktivnost' pri vospriyatii dinamicheskih i staticheskih vyrazhenij lica [Okulomotor activity at the perception of dynamic and static expressions of the face]. Eksperimental'naâ psihologiâ = Experimental Psychology (Russia), 2018. Vol. 11, no. 1, pp. 5–34. doi:10.17759/exppsy.2018110101. (In Russ., abstr. in Engl.)

3. *Barabanschikov V. A.*, *Zhegallo A. V.*, *Khoze E. G.*, *Solomonova A. V.* Neverbal'nye prediktory ocenok dostovernosti/nedostovernosti soobshchaemoj informacii [Nonverbal predictors in the estimates of truthful and deceptive statements]. Eksperimental'naâ psihologiâ = Experimental Psychology (Russia), 2018. Vol. 11, no. 4, pp. 94–106. doi:10.17759/exppsy.2018110408. (In Russ., abstr. in Engl.)

4. *Barabanschikov V. A.*, *Zhegallo A. V.* Dinamika klyuchevyh tochek lica kak indikator dostovernosti soobshchaemoj informacii. [Dynamics of key facial points as an indicator of the credibility of reported information]. Eksperimental'naâ psihologiâ = Experimental Psychology (Russia), 2021. Vol. 14. no. pp. 101–112. doi:10.17759/exppsy.2021140207.

5. *Barabanschikov V. A.*, *Zhegallo A. V.*, *Korolkova O. A.* Perceptivnaya kategorizaciya vyrazhenij lica [Perceptual categorization of facial expressions]. Moscow, Cogito-Center, 2016.

6. *Zhegallo A. V.*, *Basul I. A.* (2019a) Organizaciya issledovanij mezhlichnostnogo obshcheniya, oposredovannogo videokommunikaciej. [Organization of research on interpersonal communication mediated by video communication] // Lico cheloveka: poznanie, obshchenie, deyatel'nost'. [Human face: cognition, communication, activity]. In *Editor K. I.* Ananieva, V. A. Barabanschikov, A. A. Demidov. Moscow, Cogito-Center — Moscow Institute of Psychoanalysis, 2019. pp. 77–89.

7. *Zhegallo A. V.*, *Basul I. A.* (2019b) Parnyj eksperiment dlya vyyavleniya indikatorov pravdy i lzhi s vysokoj ekologicheskoj validnost'yu [Paired experiment to identify indicators of truth and lies with high environmental validity] Eksperimental'naâ psihologiâ = Experimental Psychology (Russia), 2019. Vol. 12. no. 4. pp. 151–159. doi:10.17759/exppsy.2019120412.

8. *Zhegallo A. V.*, *Basul I. A.* Pokazateli okulomotornoj aktivnosti kak indikatory dostovernosti / nedostovernosti soobshchaemoj informacii [Oculomotor activity as indicator of the reliability / unreliability of the provided information] // Cognition and experience.(RUS), 2021 Vol. 2 Issue 4. PP. 37–47.

9. *Shendyapin V. M.*, *Skotnikova I. G.* Modelirovanie prinyatiya resheniya i uverennosti v sensornyh zadachah [Modeling decision making and confidence in sensory tasks] Moscow, IPRAS, 2015.

10. *Zhegallo A. V.*, *Marmalyuk P. A.* ETRAN — R Extension package for Eye Tracking Results Analysis // Perception. 2015. Vol. 44. Issue 8–9. pp. 1129–1135.

Chapter 27.
Neurotechnology for Investigation of Non-verbal Communication between
Russians and Chinese, Native Speakers of Different Languages

**Zhukova O. V., Chu N. N., Pronin S. V., Lee P-L., Hsu H-T.,
Shelepin E. Yu., Vasiliev P. P.,
Lebedev V. S., Moiseenko G. A., Morozov S. A., Shelepin Yu. E.**

Pavlov Institute of Physiology RAS
St. Petersburg, Russia
CWLab International
Los Angeles, USA
National Central University
Taiwan

The purpose of our research is to study the main mechanisms of non-verbal communication between representatives of historically different cultures, speakers of different languages with fundamentally different writing systems. Thus, to explore the elements of the individual and social "archetype" of human behavior, hidden in the centuries-old layers of socio-cultural influence and fleeting individual experience. The influence of social and individual experience is often explored as different types of sociocultural learning (Tomasello, Kruger, Ratner, 1993). There are three types of training — imitative, instructive, "constructive".

Imitation learning refers to the reproduction of behavioral strategies in accordance with genetic inclinations and childhood impressions that lay the archetypes of behavior, and with fashion that affects a person at puberty and adolescence, with the context of the situation, with an understanding of intentions, and the nature of the behavior of the environment — "teacher" ubiquitous artificial selection. Fashion is the most striking example of the development of rather rigid patterns of visual images, due to which the fastest recognition occurs when determining friend or foe. The importance of following and highlighting the features of patterns corresponding to fashion occurs earlier than all other features, which forces many to follow it (Glezer, 1993; Demidov, 1987).

Instructional learning occurs when students internalize teacher instructions and use them to self-regulate their own cognitive functions during subsequent behavior. These two types of learning play a key role in cultural transmission from generation to generation because adults pass on cultural knowledge, values and beliefs to their children.

Constructive learning, referring to the process during which peers collaborate to create something new, from games and music to new technologies. This higher, creative, constructive form of learning and the formation of behavior in the process of joint activity was chosen by us from many other ways to achieve this goal, since it is in this type of learning that archetypes of behavior developed in childhood can be preserved and manifested, which, in particular, are also necessary for joint creativity with strangers, and at certain stages on the basis of only non-verbal interaction.

Our project is aimed specifically at researching the archetypes of non-verbal interaction. Probably, some of the archetypes are transmitted from generation to generation genetically, while others are formed as a result of different forms of learning, and their relationship and restructuring is determined by context, life situation and fashion. Learning shapes the brain in many ways. Therefore, parents from different cultures teach children differently, which leads to the development of different cognitive strategies and the corresponding functional organization of the brain. For example, during play, Chinese toddlers spend more time in direct physical contact with their mother than Canadian toddlers, and they need more time before they approach strangers to play together (Chen et al., 1998). During early childhood education, children usually receive information related to their own culture. For example, in the United States, mothers are more likely to actively talk and play with their children, while in East Asia, mothers are more likely to calm and lull their children (Minami, McCabe, 1995). As a result of this early interaction, basic specific facial expressions, gestures and postures are formed.

To recognize the features of facial expressions, gestures and posture, it is necessary to simultaneously describe global and local features. This is important both for humans and for artificial recognition systems. The search for local and global signs occurs constantly and in the process of non-verbal communication. To isolate a face, a low-frequency description and predominantly horizontal components in a two-dimensional spatial-frequency spectrum are sufficient. The emotional state is hidden in the diagonal components of the two-dimensional spectrum of the face (Shelepin et al., 2014; Logunova, Pronin, Shelepin, 2014).

A clear balance of choice in the face of local and global features can be carried out using the control of eye movements. In particular, eye movement techniques have been applied to distinguish between static images of West European and East Asian faces. East Asian and Western cultures are built on a combination of a holistic global cognitive strategy and an analytical local cognitive strategy for processing information and highlighting significant objects. The chapter considers whether it is possible to find cultural differences in the processing of global and local information in face recognition tasks? Miellet explored this issue by developing a gaze-dependent technique they called "Expanding Spotlight" (Miellet et al., 2013). Adult East Asian and Western European subjects were presented with a set of facial images before the experiment to remember the series of faces. After a 30-second pause after the training stage, images began to show this series of faces, but half diluted with new ones. The subjects were asked to judge whether the person was represented during the training phase. At this stage, each face was masked with a Gaussian aperture centered on the participant's fixation point so that information about the face corresponding to their identity was only available within the Gaussian aperture. The Gaussian aperture was limited to 2 °, but dynamically expanded for each new fix (1 ° every 25ms). It was assumed that if the expanding spotlight hypothesis was correct, observers would maintain fixation at a given location until they had enough foveal and parafoveal information at that location to solve the problem. Then Miellet et al. processed information about facial images in the spatial frequency domain (for example, in the high and low frequency range) in each cultural group. The findings showed that Western participants used a local sample of high-frequency spatial information covering all features critical to effective face recognition (eyes and mouth), while participants from East Asia achieved a similar result by using global, spatial low-frequency information when observing those the same persons. Behavioral and physiological studies have shown that global or contextual information is transmitted over a low frequency channel, while local or detailed information is transmitted over a high frequency spatial frequency channel during visual perception. These data are consistent with the results of our research (Shelepin, 1985; Shelepin, 2017).

The results of Miellet et al. show that preference for holistic or analytical processing during facial perception is associated with different neurocognitive strategies in East Asians and Westerners. The authors distinguish an informative triangle of the face: "eyes—upper lip". According to the published data, during face recognition, at large angular values, it is possible to obtain T-shaped "attention maps", in which the greatest number of fixations

and their average duration are observed in the areas of the eyes, nose (bridge of the nose), upper lip and corners of the lips (Barabanshikov, 2012, 2018). In general, there are five possible types of face examination: — "linear horizontal" — the areas of interest of this pattern fall on the left and right eyes of a person (4.5 %); — "linear vertical" — areas of interest are grouped around the nose and mouth (7.5 %); — "triangular" — areas of interest for this pattern include the areas of the eyes and mouth, sometimes with fixation of the nose area (70 %); diagonal "- areas of interest fall on only one of the eyes (right or left), as well as the area of the nose and mouth (4.5 %); "Topical" — grouping of fixations (11.5 %). But it is the triangular pattern that is the most common (Ananyeva et al., 2010). It should be noted that the choice of the type of pattern depends on psychological factors. For example, a longer viewing of the right side of the face suggests that the subject is guided by important and significant signs in the face of the interlocutor.

Based on the work carried out to study the direction of gaze in facial images, it can be assumed that the gaze of the interlocutors at each other will most often be directed within the famous triangle "eye — upper lip". What are the observers looking for in these areas of the face? The answer to this question is contained in the famous story "Childhood" by Lev Nikolaevich Tolstoy. All human emotions are contained in the corners of the eyes and the corners of the lips. The observer sees these significant signs of expression of emotions at the limit of the resolving power of the eye. (Campbell, Shelepin, 1990). Therefore, the search for the most significant features occurs precisely within this triangle. It is there that the observer directs the fovea, the region of the retina with the highest resolution, which ensures the search for local features.

Switching attention from global to local signs occurs both voluntarily and spontaneously. This switching occurs quickly enough. In the studies that we conducted back in 1992–1997 (Shelepin et al., 1997), we answered the question: is it possible to "objectively" study the moment of activation of the local and global mechanism of visual information processing due to voluntary (on command) switching the subject's attention from local to global features in Kinchl's figures (Kinchla, 1974, 1977, 1979). We observed significant ($p < 0.05$) differences between two different categories of responses in the occipital region of subjects with latencies of 70 and 110 ms after the onset of the stimulus. An important result was that, based on the evoked potentials, we found that the switching of selective attention occurs in the early stages of visual information processing, 70–100 ms. Another study also recorded event-related evoked potentials in Chinese subjects, reflecting brain responses

Chapter 27. Neurotechnology for Investigation of Non-verbal Communication between

to Asian and European faces expressing pain or a neutral state (Sheng, Han, 2012). The differences appeared in the component with 170–200 ms latency. These stimuli were probably more complex than in studies with Kinchl figures, which led to longer latencies of EP components reflecting the difference.

To localize the mechanisms of recognition of local and global facial features in the brain, fMRI studies were carried out (Sheng, Han, 2012). Brain activity was analyzed according to fMRI data in the process of identifying causality and direction judgment tasks in Chinese and American subjects (Han et al., 2010). Causal judgments about physical events and human behavior are highly influenced by parenting. Thus, when making judgments about the causality of social types of behavior, subjects from East Asia were more sensitive to contextual analysis in comparison with Americans of European descent, who were more inclined to identify causal relationships (Choi et al., 1999; Morris, Peng, 1994). In Western cultures, an analytical style of cognitive processes is developed, while in East Asian culture, a holistic type of cognition is developed (Nisbett, 2003). This reminds us of the assumptions put forward by us about the possibility of dividing all types of behavior into an imperative and declarative algorithm of the brain, associated mainly with the dorsal and ventral visual neural network (Shelepin et al., 2021).

To elucidate the neurophysiological mechanisms associated with cultural differences, Han and colleagues scanned two independent groups of Chinese and European-American participants and focused on the medial prefrontal cortex and left parietal region associated with causal attribution of a physical event (Han et al., 2014). The search for the causality of events is a unique human trait that is associated with the work of the medial prefrontal cortex (Penn, Povinelli, 2007).

In this regard, it has been suggested that activation in the left parietal region, whose activity is sensitive to contextual information during causal judgments, should be more pronounced in Chinese than in American participants. Indeed, it was found that while medial prefrontal activity involved in judgments of causality was comparable across the two cultural groups, left parietal activity associated with judgments of causality was more pronounced in Chinese than in American participants. Thus, these findings revealed universal mechanisms for identifying causal relationships with the work of the medial prefrontal cortex. And culturally-specific descriptions are associated with the work of the left parietal cortex, which, apparently, is specifically engaged in processing contextual information during the perception of cause-and-effect relationships.

Cultural differences in left parietal activity are interesting in that the left parietal cortex was activated in Chinese participants, regardless of whether they were asked to pay attention to causal judgments or to contextual information (that is, in both dispositional and contextual causal judgments), whereas this area of the brain failed to show activation in American participants in any of the types of causal judgments. It appears that the Chinese and American participants in this study exhibited two extremes of neural strategies when participating in context processing during causal attribution. Differential interactions of the left parietal cortex in causal attribution of physical events provide a neurobiological explanation for the discovery that Chinese, compared to Americans, are more likely to attribute the causes of the same events to contextual factors (Peng, Knowles, 2003). Is it possible that people from other cultures can demonstrate the activity of the left parietal cortex in terms of determining cause-and-effect relationships, taking into account the context of the situation, i. e. based on the strategy typical for the subjects — immigrants from East Asia.

To test this hypothesis, Han and colleagues scanned a group of German participants while they identified causal relationships in the test and found that German subjects in some working conditions showed increased activity in the left parietal cortex during causal relationships in context (Han, Ma, 2015). However, no activation in the same area of the brain was observed when the German participants made a dispositional causal judgment about a physical process. Thus, the left parietal activity underlying the processing of contextual information during judgments of causality is highly sensitive to the cultural experience of people.

Studies of visualization of brain activity (fMRI) made it possible to compare the involvement of different neural networks when subjects used different strategies for constructing a visual picture of the world. Recall that the Europeans use mainly analytical, and the Chinese mainly use global analysis. Note that both types of situation analysis and tests are common to all people, but they have only a few peculiarities. This result is important for the planning and implementation of our project of searching for common archetypes of behavior: is it possible to see the same brain activity in Russians and Chinese in the process of performing a common joint activity that is understandable to both samples in an adult "sandbox"? Therefore, we began the solution of our project with the development of a game (developer Pronin S. V., Researcher of the Pavlov Institute of Physiology RAS), acceptable for working at a long distance via the Internet, a game in which it is possible to study various

algorithms of behavior, despite the differences related to both with the genetic characteristics of our subjects, and the peculiarities of culture, due to fundamental differences in the design of written speech.

Developed two programs — First and Second — for two players. The first of them is the "leading" ("master") one: all parameters are set in it, the experiment is controlled from it, and the results are displayed in it. The second program, Second, is "follower" ("subject"): it receives commands from the first program. Of the configurable parameters, it contains only the IP address of the computer on which First is running and its port number. In the First program, when you click on the "Open playing field" button, a playing field is opened, divided into rectangular cells. If there is a connection with the Second program, then the same field will also open in it. The parameters of this field (number of cells, line width, etc.) are set in the "Playing field" section. The meaning of the game is as follows. One of the cells contains a hidden object — a green rectangle. One of the participants (follower) does not see this object, the other (leader) sees. The task of the first, "follower" player is to go through the cells to find the one in which the object is hidden. If any additional information is missing, it will iterate over the cells in random order. But if, in addition to the playing field, he demonstrates on the screen an image of the face of the second participant (the leader), who sees how close he is to the goal, then, observing his emotional reaction, the "follower" can correct his choice and find the object on average faster than "Blindly" (Figure 1).

To demonstrate the face of the leading player, any program was used that allows the image from the camera of another computer to be displayed on the screen of one computer. When searching for a hidden object, the "follower" player moves the mouse cursor over the cells of the playing field. If the cursor is within this cell for longer than the specified time (this is the "Delay, ms:" parameter in the "Playing field" section of the settings), this cell is considered selected. If there is no hidden object, the cell is highlighted with a frame (by default — red), if there is an object, the cell is filled with green and the game ends.

On the computer of the "leading" participant, both the hidden object (the cell with it is marked in green) and the cells selected by the "follower" player are visible. If you select the option "Show cursor of the second player" in the settings section, then the "leading" participant will also see the movement of the cursor on the monitor of the "follower" player. By default, the playing field on both computers is made "transparent". This allows each of the partic-

ipants in the experiment to see the face of the other participant on the screen during the game.
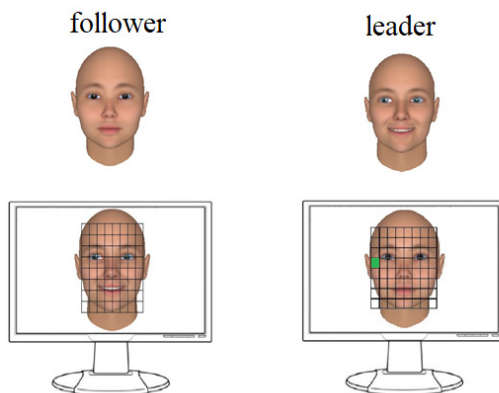

Fig. 1. Scheme and functional roles of players during the game.

Additionally, the program has thought out the possibility of organizing a game, during which the hidden object is not shown on the screen of the "leading" player. Accordingly, this player, since he does not see the target, cannot tell the location of the hidden target by facial expressions to the other player. After the game is over, the program waits for further action from the lead participant who runs the First program. He can either continue the game or exit the game mode ("Exit" button). If the second participant leaves the game, the First program will also exit the game mode. In both cases, the Results section of the First program will show the results of the games played by the following parameters: the duration of each game and the number of cells "passed" during each game, as well as the average values for the entire cycle of games played. If the hidden object was not shown on the screen of the "leading" player, then there will be a corresponding mark in front of such a game. An application for a patent for an invention is being prepared.

In the course of testing the program (when playing games in human-computer mode), the optimal size of the playing field is 8 by 8 cells, which corresponds to the size of the chessboard. This size of the playing field was chosen as the standard for all series of games.

Several modes of player interaction have been developed. The first mode—each player has his own role. One player plays the role of "leader"—

a hidden object appears on his monitor. The player's task is to use non-verbal means of communication to prompt the second player where the target is. The second player plays the role of a "follower"—his task is to find a hidden object as quickly as possible, focusing on the non-verbal prompts of the first player. The second mode is similar to the first. But the hidden object does not appear on the lead player's screen. Accordingly, his role is to passively observe the actions of the second player. The role of the "follower" player remains the same. But he acts "blindly"—in the absence of facial expression. The appearance of the game with no hidden object was orchestrated in a pseudo-random sequence. This condition was introduced to assess the impact of non-verbal means of communication on the effectiveness of performing a joint game task.

A hardware-software complex has been developed for conducting psychophysiological research in the paradigm of hyperscanning. The scheme of the complex is shown in figure 2.
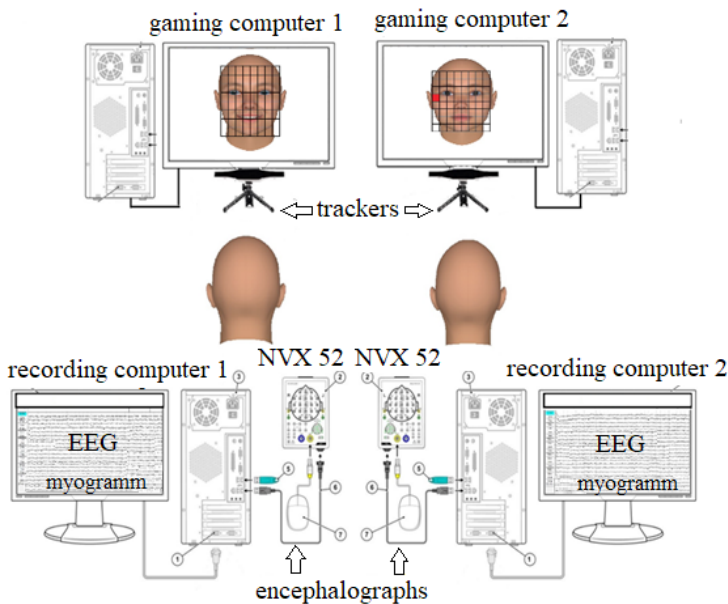
Fig. 2. Diagram of the experimental complex for research in the paradigm of hyperscanning.

The complex includes:

1. Two recording computers with monitors.

2. Two gaming computers with monitors.

3. Two trackers (Gazepoint GP3, Canada)

4. Two encephalographs (NVX52, Medical Computer Systems)

5. Accessories (electrodes for EEG and myogram, cables, mice, etc.).

Synchronization of all computers is organized via com ports. The program recorded the marks of both the beginning and the end of each game. In the course of the research, sociological, psychophysical (game statistics) and psychophysiological research methods were used — the electroencephalogram, eye movements and myogram of each player were recorded. To register the myogram, two electrodes were used, which were symmetrically attached to the zygomaticus major muscle (Figure 3).
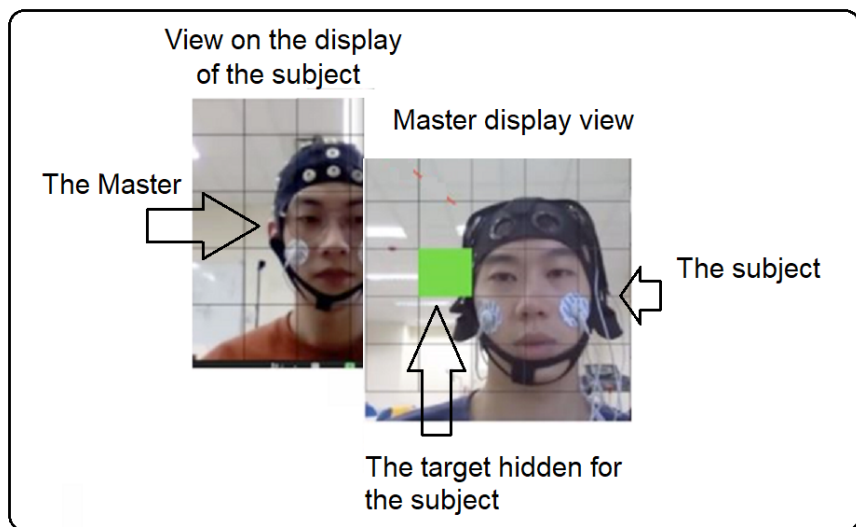
Fig. 3. Scheme of attachment of electrodes for recording myogram.

Psychophysical Analysis Results

In 2021, preliminary studies were prepared and conducted, in which 6 pairs of players from the same culture and 4 pairs of players from different

(Chinese and Russian) cultures took part. Each pair of subjects sequentially went through two series of studies, 48 games in each, where conditions alternated with the presence and absence of a "hidden object" on the monitor of the "leading" player. The series differed only in those who played the role of "leader" and who performed the role of "follower". Changing roles between players occurred once during all 96 games. As an example, Figure 4 shows the results of one pair of players—graphs of the dependences of the duration of games on the number of "passed" cells in different series of the study.

B
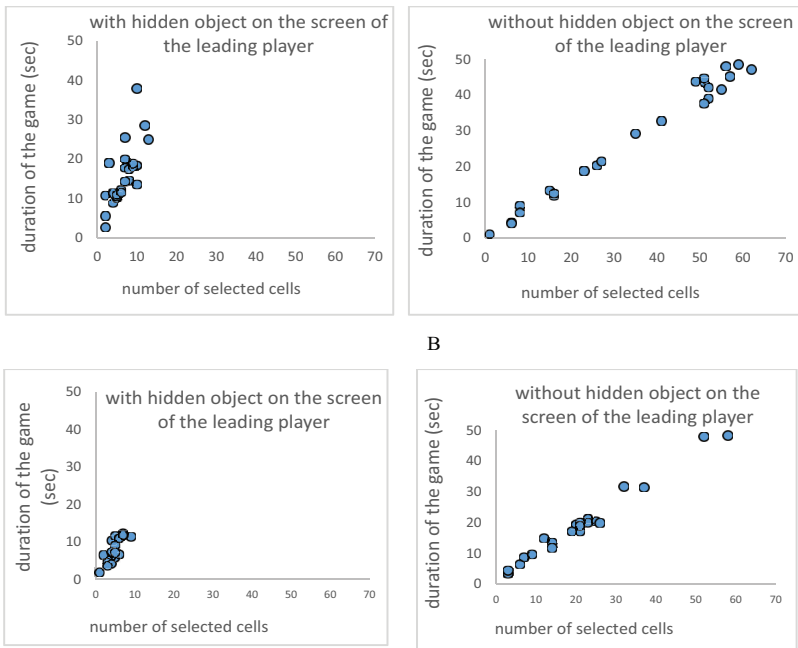
Fig. 4. As an example, graphs of the dependences of the game duration on the number of "passed" cells, one pair at a time, are shown in the first (A) and second (B) series of research in the conditions of presentation and absence of a hidden object on the screen of the "leading" player (analysis of variance, $p < 0.001$)

It can be seen that in both the first and second series of the study, the experimental paradigm with the absence of a hidden object on the monitor screen of the "leading" player, and, accordingly, the absence of the very possibility of non-verbal prompting, led to the fact that the game of finding a hidden goal took more a long time in comparison with the condition when the

hidden target was presented on the screen. Thus, the mimic hint in the game played an important role that significantly influenced the efficiency of the game. In the second half of the study, the duration of the game is shortened, which is probably related to the effect of learning the rules of the game and the adaptation of the players to each other. All revealed differences were statistically significant (analysis of variance, p <0.001).It is of interest that similar patterns were revealed in Russian-Chinese pairs (Figure 5).
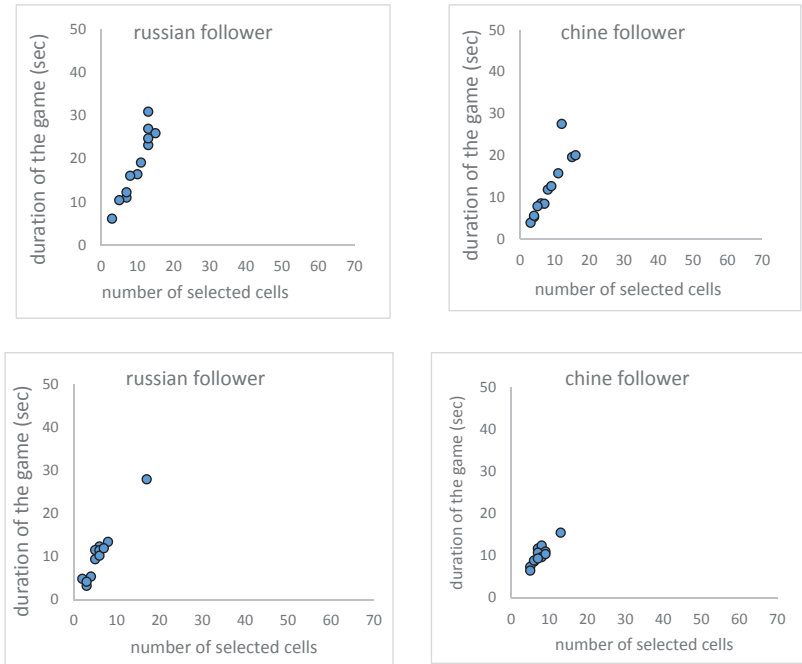
Fig. 5. As an example, graphs of the dependences of the game duration on the number of "selected" cells for one pair are given—representatives of different cultures in the first (A) and second (B) series of the study (12 games in each) in the conditions of presenting a hidden object on the screen of the "leading" player.

Also, as in pairs of players from the same culture, there is a learning effect. There were no significant differences between the players (p> 0.05).

Spectral Analysis Results

It is known that brain rhythms reflect the processes of functional modulation of the activity of large-scale neural networks in the human brain, and are

probably associated with many aspects of human activity (Buzsaki, 2006). A large number of studies demonstrate the role of brain oscillations in different frequency ranges when performing a variety of cognitive and emotional tasks (Buzsaki, 2006; Kropotov, 2010; Kostandov, Cheremushkin, 2013; Ponomarev et al., 2017). Therefore, it can be assumed that rhythms are associated with various aspects of communication. The neural mechanisms of verbal communication have been studied in sufficient detail (Rimmele et al., 2018). Speech is inherently organized rhythmically. For example, the timelines of phonemes and syllables, as well as formal prosodic aspects such as intonation and stress, fall into separate frequency bands. The power of rhythms reflects the processes of segmentation, coding, and speech production at different time scales (Rimmele et al., 2018). The role of rhythms in the process of non-verbal communication is much less clear.

Simultaneous registration of the bioelectric activity of the brain was carried out — an EEG was taken during the game of two participants. A total of 6 couples took part. EEG registration and processing was carried out using two 48-channel NVX 52 analyzers (sampling frequency — 2000 Hz) and WinEEG programs developed at the Institute of the Human Brain of the Russian Academy of Sciences. N. P. Bekhtereva. The electrodes were positioned in accordance with the international 10–20 system in leads Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2. Reference electrodes were placed on the earlobes. The work used the installation of EEG derivations relative to the combined ear referent. The resistance of the electrodes did not exceed 5 kΩ. Blink artifacts were corrected using the independent component method (Vigario, 1997). Additionally, epochs containing artifacts associated with body and head movement caused by strong emotional reactions during social interaction were excluded from the analysis. These threshold values were selected empirically by performing multiple processing with different parameters and then visualizing the results. Blinks and eye movements were recorded to detect artifacts. The influence of these artifacts was removed using the WinEEG program. Additionally, EMGs were recorded. The spectral density was estimated using the discrete Fourier transform. When calculating the spectral density, the entire analyzed interval was divided into epochs with a duration of 4 s, overlapping by 50 %. The Hann window was used to smooth the spectra.

Analyzed 4 pairs of EEG studies on 12 games in each when the players perform the role of "leader" and "follower". The results are shown in Figure 6.
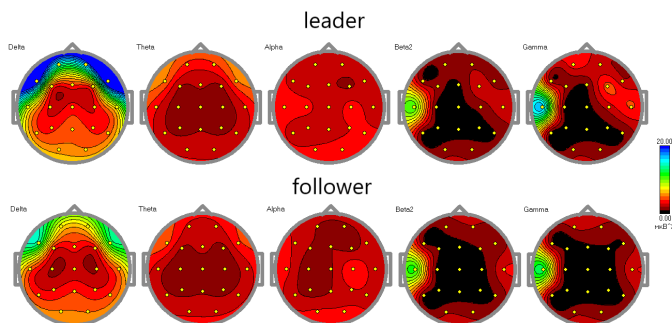
Fig. 6. Averaged EEG spectral power topograms for 4 pairs of players and 12 games — comparison of the role of Leader and Follower. The figure shows the following frequency ranges: 1.5–4 Hz (delta), 4–7.5 Hz (theta), 7.5–14 (alpha), 14–20 Hz (beta 1), 20–30 Hz (beta 2), 30–40 Hz (gamma) brain rhythms.

An increase in the power of signals is observed when the players perform the role of "leader" in the range of 1.5–4 Hz and 30 Hz EEG rhythms. Significant increase in signal power when playing the role of "leader" is observed in the range of 1.5–4 Hz in the frontal (Multivariate Tests of Significance, Fp2, F7, F4, F8), central (C3, Cz) and temporal (T3) regions of the brain (F = 1340.478; p <0.05).

Comparison of two conditions is carried out: if there is a "hidden goal" on the screen of the "leading" player or in its absence. Figures 7–8 show an example of comparing the power of EEG signals in a game with the presence or absence of a hidden object.

It can be seen that the signal strength in all frequency ranges is higher when playing with facial expressions. The maximum power of rhythms in the "leading" player, when he sees a hidden target, is expressed in the ranges of 1.5–4 Hz and 4–7.5 Hz in the frontal (leads Fp1 and Fp2), occipital (O1, Oz, O2) and temporal (T6, T5) areas of the brain; in the range of 10 Hz — in the occipital (O1, Oz, O2), temporal (T6, T5, T3) and frontal (F7 and F8) areas of the brain; in the ranges of 14–30 Hz — in the occipital (O1, Oz, O2) regions of the brain and in the range of 30–40 Hz — in the occipital (O1, Oz, O2) and frontal (Fp1) regions of the brain. Other couples have obtained similar results when playing the role of "lead" and "follower" players. Thus, when comparing the maximum values of the spectral power for 6 games when performing different roles, an increase in power in the slow and high-frequency EEG

ranges was found when playing the role of the "leading" player. When performing various roles, the conditions of the game with mimic prompts had a significant effect on the redistribution of signal powers in different frequency ranges.
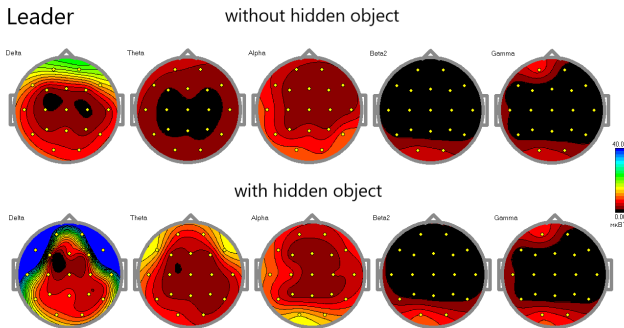

Fig. 7. Averaged EEG spectral power topograms (6 games)—comparison of conditions "without hidden object" / "with hidden object" for one "leading" player. When there was a hidden target on the screen (condition with a mimic hint), the "leading" player had to indicate to the second player the location of the green square by facial expressions and other non-verbal means. During the absence of a hidden object on the monitor screen (condition without a mimic prompt), the "leading" player simply watched the actions of another player, who "blindly" searched for the target. The figure shows the following frequency ranges: 1.5–4 Hz (delta), 4–7.5 Hz (theta), 7.5–14 (alpha), 14–20 Hz (beta 1), 20–30 Hz (beta 2), 30–40 Hz (gamma) rhythms of the brain.

Let us consider in more detail the specifics and role of the selected frequency ranges in the non-verbal interaction of players. In general, many authors associate changes in slow rhythm with an assessment of the emotional and motivational significance of stimuli (Rimmele et al., 2018). Thus, in the case of non-verbal and verbal communication, a number of studies have shown an increase in the synchronization of delta waves in the perception of images of faces (Güntekin, Basar, 2009) and speech by ear (Ding et al., 2016). Moreover, with non-verbal communication, the power of delta rhythms is higher in the perception of emotional ones in comparison with neutral persons in the parietal and occipital leads of the brain. An increase in the synchronization of delta waves in the central leads was also revealed with emotional involvement in communication (Knyazev, 2009). In this work, a more significant increase

in the power of slow rhythms in the frontal leads of the brain is observed, which, according to published data, is found when solving a wide range of cognitive tasks (Knyazev, 2009).
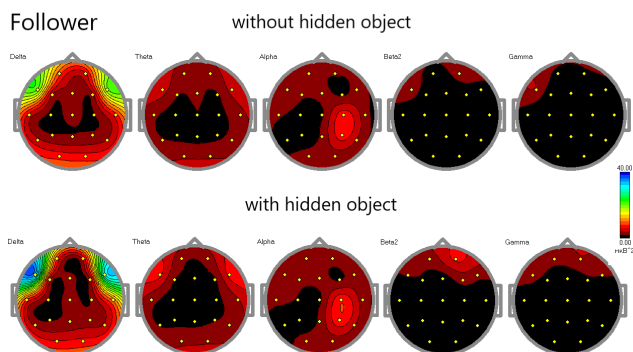

Fig. 8. Averaged topograms of the spectral power of the EEG (6 games) — comparison of the conditions "without a mimic prompt" / with a "mimic prompt" for one "follower" player. When there was a hidden target on the screen (condition with a mimic hint), the "follower" player had to find the hidden object as quickly as possible, focusing on the mimicry of his partner in the game. During the absence of a hidden target (condition without a mimic prompt), the "follower" player "blindly" searched for the target. The figure shows the following frequency ranges: 1.5–4 Hz (delta), 4–7.5 Hz (theta), 7.5–14 (alpha), 14–20 Hz (beta 1), 20–30 Hz (beta 2), 30–40 Hz (gamma) rhythms of the brain.

However, it should be noted that the players actively interacted with each other using eye movements, so it can be assumed that in this frequency range we see both the delta rhythm itself and the artifacts associated with eye movements. This assumption is supported by the fact that the saccade frequency is 1.6–5 Hz, which corresponds to the delta frequency range (1.5–4 Hz). For each EEG recording, a standard routine processing was applied to remove motion artifacts. An eye movement removal filter was used during recording. But non-verbal communication in our conditions is based on the movements of the oculomotor muscles, the manifestation of which we see in the delta frequency range.

When comparing the conditions of the presence or absence of a hidden target, the expressed power of 10 Hz signal in the occipital and parietal leads of the brain is shown. In general, most of the equivalent dipole sources of this

rhythm are observed in a relatively local area in the area of the furrow at the border of 17 and 18 fields according to Brodmann. According to published data, it is more pronounced with the eyes closed in a relaxed state and is blocked during concentration (Buzsaki, 2006). In our conditions, in addition to the occipital alpha rhythm, we observe fluctuations in the electric potential with the same frequency — 10 Hz, but recorded near the sensorimotor areas — the parietal leads of the brain. This type of oscillation is called mu-rhythm, rolandic or sensorimotor rhythm. A feature of this rhythm is that it reacts poorly to the opening of the eyes, but desynchronization of its waves occurs when the subject performs a movement, observes the movement of another, or imagines performing an action (Muthukumaraswamy, 2004). Moreover, the blockade of the rhythm can occur several hundred milliseconds before the movement itself. In this regard, it is assumed that suppression of the mu rhythm can be considered a probabilistic index of the activity of mirror neurons. It can be assumed that averaging all players will reveal differences between all conditions and modes of game interaction in a given frequency range.

Of interest is the analysis of rhythms in the range of 14–40 Hz in various modes of interaction between players. The relationship between beta and gamma rhythms and muscle contraction in the motor and premotor cortex has been shown for a long time, and the increase in power coincides with the moment the movement begins and continues throughout the motor response (Crone, 1998). Therefore, it is likely that the recorded activity in this frequency range, especially in the temporal leads, is associated with the movement of facial muscles. This assumption is supported by the fact that the activity of gamma rhythms has a very low amplitude (5–7 µV) and it is very difficult to separate it from EMG artifacts (Ponomarev, 2016).

Thus, we can conclude that recognition of facial expressions is not just an emotional reaction, but a complex cognitive process, probably the most complex visual cognitive process (Campbell, Shelepin, 1990). These conclusions remain valid even if the power of delta and gamma rhythms is a simultaneous reflection of the EEG and EMG of the activity of the facial and oculomotor muscles arising during non-verbal communication. Next year, additional processing techniques will be applied to separate and evaluate both rhythms themselves and muscle movements.

Correlation Analysis Results

The analysis of the correlation of two EEG signals (for players playing in pairs) has been carried out. For this, the Correlation program (developed by

Pronin S. V.) has been developed, which allows you to calculate the correlation level both without calculating, and with calculating the power of EEG signals. To assess the level of correlation, the Pearson correlation coefficient was used. To assess the level of correlation, all data were filtered in six frequency ranges. On average, the level of correlation between two identical leads among players is at a low level (0.1–0.3 %). This is typical for all frequency ranges and leads of the brain (Figure 9).
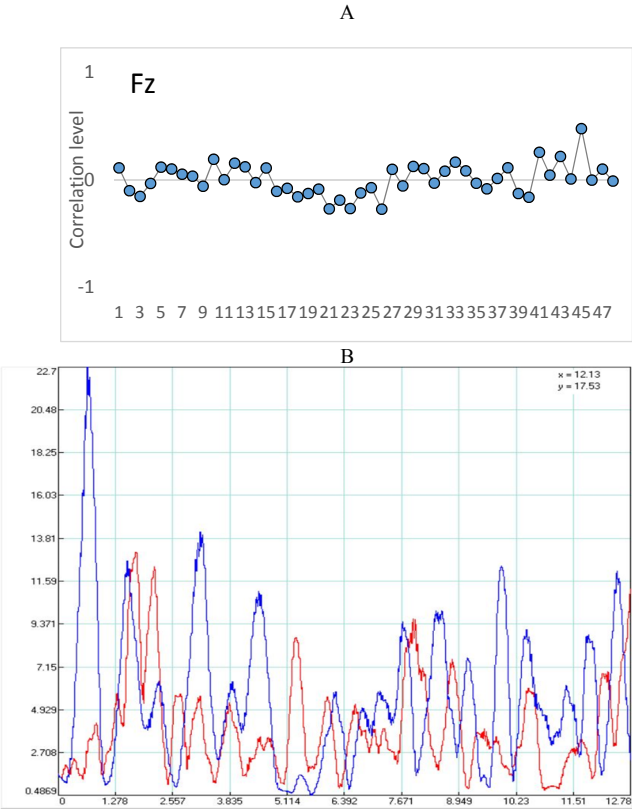

Fig. 9. An example of the level of correlation between two identical leads Fz in two players in the range from 7.5–14 Hz (A). The abscissa represents the sequence of 48 games in the first series of the study. The ordinate is the Pearson correlation coefficient. B — graphs of the powers of two EEG signals in leads Fz in games numbered 7 in one series of the study.

A low level of correlation is quite expected, since each of the players fulfilled their instructions and roles, which were significantly different from each other. It should be noted that this tendency is typical for both electrodes (EEG) and biochannels that record activity from the zygomaticus muscle (EMG). At the same time, when performing a series of games, there are quite often individual games with an average or high (0.8–0.9 %) level of correlation between players. Figure 10 shows an example of such synchronization in one pair of subjects (in one series of studies) in the frequency range 1.5–4 Hz in the frontal leads (F7).
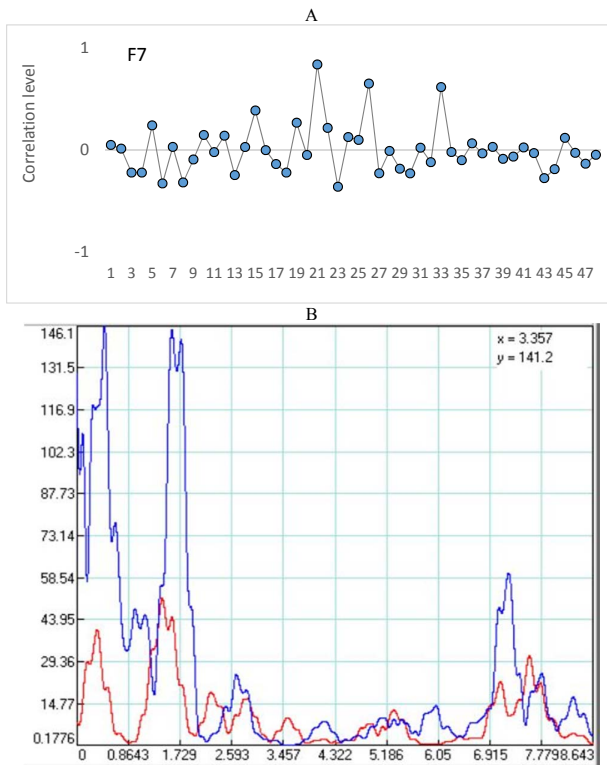
Fig. 10. An example of the level of correlation between two EEG signals in the same lead F7, filtered in a slow frequency range from 1.5 to 4 Hz (A). The abscissa represents the sequence of 48 games in the first series of the study. The ordinate is the Pearson correlation coefficient. B — graphs of the powers of two EEG signals in lead F7 in the game number 33.

Most often, these "spikes" are observed in the delta and gamma frequency ranges. As we can see, the level of correlation throughout the series of studies, consisting of 48 games, is uneven. This is true for both EEG and EMG. At one point in time, the level of correlation between players may increase significantly, in other cases, on the contrary, decrease. Probably, at high levels of correlation in the delta and gamma frequency ranges, we see synchronous movements of the facial and oculomotor muscles in players during the game. But since these non-verbal signals are extremely important for effective communication between people, at this stage it is not so important to separate the received effects from EEG and EMG, but it is more important to consider them as a whole, as a marker of non-verbal communication. Next year, the analysis of rhythm synchronization in different frequency ranges and different experimental paradigms will be continued. We assume that a meaningful difference will be found in games with and without a "hidden object". Also, the results of the analysis will be provided not only between the same leads from the "master" and "follower" player during the game, but between different electrodes (according to the principle—comparison of all with all), as shown in Figure 11. This is of interest, since one the same area of the brain, for example, the visual, can be synchronized with a number of other areas that play an important role in non-verbal communication.
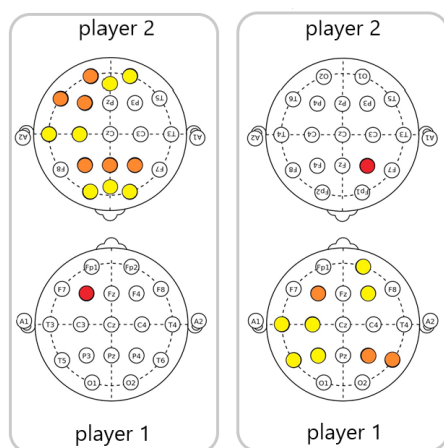

Fig. 11. The level of correlation of one lead F3 (marked with a red circle) for player 1 with all leads for player 2. And vice versa, one lead F3 for player 2 with all leads for player 1. A low level is indicated in yellow (0.3–0.4 %), and orange—the average level (0.5–0.7 %) of the correlation between the electrodes in two players.

Eye movements Analysis Results

Joint attention, that is, the ability to follow the gaze of another person and determine his fixation point, is one of the abilities of human perception used in communication and problem solving. It is now possible to study joint attention using the eye-tracking technique—a method of analyzing eye movements and the point of focus of gaze—which is often used to assess the strategy of behavior and play (Polonio, 2015). We used Gazepoint GP3 and GP3HD equipment, switched on at 60 Hz. The recording was carried out using Gazepoint Analysis software (Gazepoint, 2021). Screen Capture has been selected. The incentive design is presented as follows: after the calibration process, the screen recording (Screen Capture) began, which was stopped by the button at the end of the series of games for each participant. To analyze the data for each game, a file was exported with fixations for BPOGX, BPO-GY, SACCADE_MAG, SACCADE_DIR indicating the fixation coordinates along the X and Y axes, respectively, as well as the magnitude (saccade path) and direction. Heatmaps and gaze maps were also exported.

The eye movements of both players are investigated. Eye movement maps during one game (18 sec) in the presence of a hidden object on the monitor of the "leading" player are shown in Figure 12.
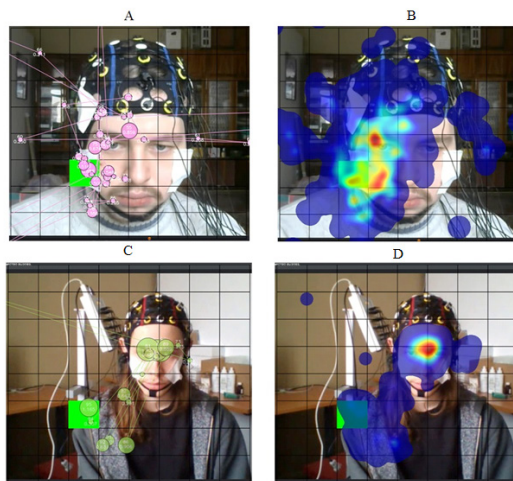

Fig. 12. Map of gaze movement and heat map obtained from the screen of the "leader" (A, B) and "follower" (C, D) players in the presence of a hidden object (condition with a mimic prompt)

In this condition of the game, the "leading" player saw the target and could tell the "follower" player its location by mimicry. The analysis of the maps shows that the "follower" player is fixed in front of the eyes of the "lead" player and the current position of the cursor, and there is a constant switching between these areas of interest. The "lead" player's attention is more distributed, but there is also an area of interest between the current cursor position and the Player's eyes. The maps of eye movements during one game (23 seconds) in the absence of a hidden object on the monitor screen of the "leading" player are shown in Figure 13. The screenshot was taken from the moment the game ended, when the required field was found.
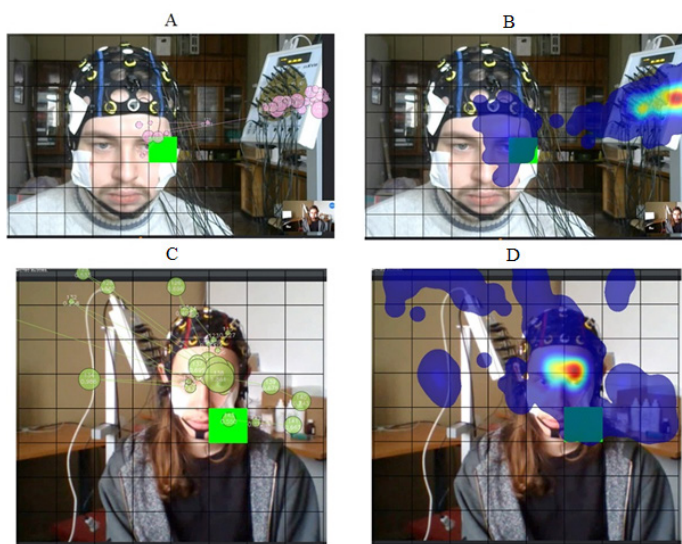

Fig. 13. Map of gaze movement and heat map obtained from the screen of the "master" (A, B) and "follower" (C, D) players in the absence of a hidden object (condition without mimic hints)

In the task, in the absence of a hidden object on the monitor screen, the attention of the "leading" player switches between the player's current face and distracting elements on the screen. On the screen of the "follower" player, there is also only one fixation on the target field (due to the lack of a hint on the screen until the end of the game), but a large number of fixations are perfect on the face of the lead player. The results obtained are consistent with previously published data (Barabanshchikov, 2012). In our conditions, the

linear patterns of face examination are more pronounced, which is possibly associated with the individual preferences of this pair of players or with the experimental paradigm itself. To assess group preferences, it is necessary to average the parameters of eye movement on a sample of subjects.

The fixation values were aligned with the game's grid (Figure 14). Below is a map of the sequence of cell selection in games. The combined map shows that the fixations occurred mainly in the areas where the cell was selected. In the course of further analysis, it is planned to test the hypothesis that under different conditions of the game (in the presence and absence of a mimic hint), the number of movements from the selected cell to the area of the eyes of the "leading" player will be different.
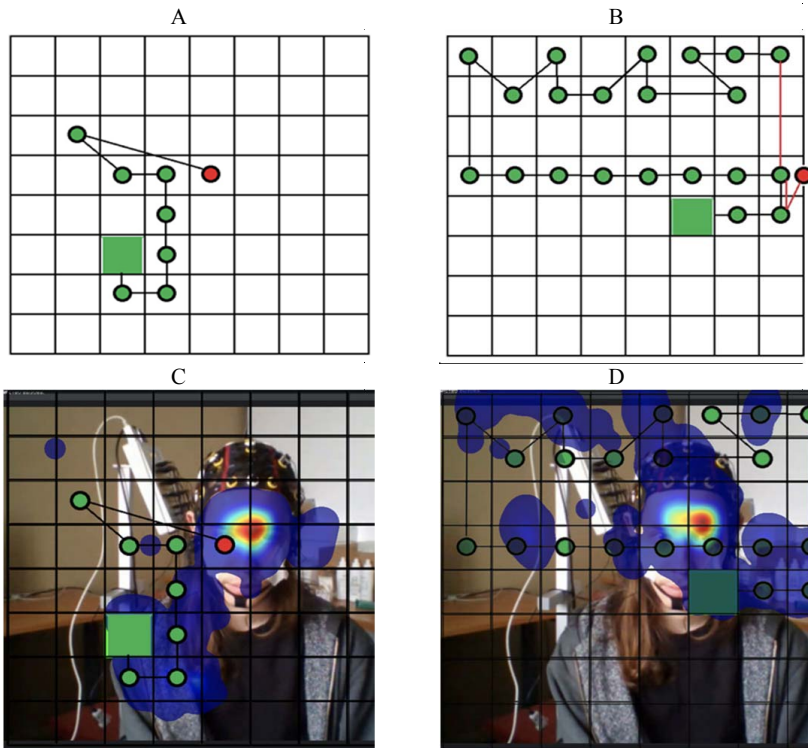

Fig. 14. Maps of the sequence of cell selection by games in the presence (A) and absence (B) of a hidden object. Below are the combined heat maps and cursor movements in the presence (C) and absence (D) of a hidden object.

Thus, in the first year of work on the project, we, on the basis of the principle of constructive interaction and training of representatives of such cultures, created a new algorithm for the study of neurophysiological mechanisms for achieving a common goal. In its simplest form, this is a game in which non-verbal communication plays a key role, and the speech communication channel is turned off. We have found that there is interaction between partners and it works the same for representatives of both cultures. The algorithm of constructive interaction between partners created by us, in view of its apparent efficiency, will be applied in 2022 as well. To develop the technical capabilities of the implementation of our project in 2022, already in the reporting year, we created a new hardware and software complex based on the convergence of the method of infrared spectroscopy of the brain, electroencephalography, electromyography of facial muscles and eye movement control, which allows subjects who are actively moving and working in a team, separate out the responses of the brain, the activity of mimic, chewing, speech muscles and eye movements. The convergence of new and traditional methods of neurophysiological research opens up a promising direction in the study of the interaction of subjects in the process of constructive, purposeful activity.

In 2022, we will continue to study the neurophysiological mechanisms of constructive interaction in the context of only non-verbal communication between research participants in St. Petersburg and Taipei based on non-verbal communication between representatives of historically different cultures, with fundamentally different languages and scripts. The software will be developed. From colleagues from the Taiwan University of Technology, we receive a new type of active electrodes developed by them, which are not yet available on the world market of equipment and are specially designed for recording EEG in difficult hyperscanning conditions. In a separate series of studies, fMRI will be used primarily for studying the deep structures of the brain and comparing them with literature data. The algorithms and complexes of psychophysiological research technologies developed by us will allow us to gain new knowledge about the work of the brain in the process of constructive interaction by subjects in St. Petersburg and Taipei. First, we will establish the dynamics of the activity of large-scale neural networks in the context of non-verbal communication in order to jointly achieve a goal in the process of a joint game. Secondly, we will show that the open (Han et al., 2010) significantly more costly brain activity according to fMRI data in Chinese subjects, compared to American subjects, is not a manifestation of differences in genetics, culture or upbringing, but in total only in acquain-

tance with the tests. We make this assumption based on our previous research (Zhukova, Vasiliev, 2020). Thirdly, correlations will be established between the activity of various neural networks in the process of non-verbal communication in order to achieve a goal. Fourthly, we will establish the possibility of predicting the most probable further behavior of the subjects in the course of joint purposeful activity. Fifthly, the obtained results will allow to deduce the archetypes of non-verbal means of communication.

The results obtained within the framework of this project will be directed to the development of technologies in the field of personalized medicine. And they will also be used in other successfully carried out works on the creation of intelligent devices for personality restoration in case of depression of various origins, successful socialization of patients with paranoid form of schizophrenia, with some forms of neurodegenerations disease. They will contribute to the creation of intelligent devices to optimize interaction for business development of Russian and Chinese partners through the development of non-verbal communication means.

References

1. *Ananyeva K.I., Drummers V.A., Kharitonov A.N.* Isostatic patterns of eye movements in the perception of a human face // Experimental psychology in Russia. M. Publishing house of the Institute of Psychology of the Russian Academy of Sciences. 2010.S. 195–199).

2. *Barabanshikov V.A.* Expressions of the face and their perception. M.: Publishing house "Institute of Psychology RAS".— 2012.

3. *Barabanshikov V.A.*, *Korolkova O.A.*, *Lobodinskaya E.A.* Recognition of emotions in the conditions of a stepped stroboscopic exposure of facial expressions // Experimental Psychology, 2018. V. 11. No. 4. P. 50–69. doi: 10.17759 / exppsy.2018110405.

4. *Buzsaki G.* Rhythms of the Brain. Oxford. New York: Oxford University Press, 2006. 448 p.

5. *Campbell F. W.* The transmission of spatial information through the visual system // Neurosciences. 1974. Boston: MIT press. P. 95–103.

6. *Campbell F., Shelepin Yu.* Possibilities of foveola in distinguishing objects // Sensory systems. 1990. T. 4. No. 2. P. 181–185.

7. *Chen X., Hastings P. D.*, *Rubin K. H.*, *Chen H., Cen G., Stewart S. L.* Childrearing attitudes and behavioral inhibition in Chinese and Canadian tod-

dlers: a crosscultural study // Developmental Psychology. 1998. Vol. 34. P. 677–686.

8. *Choi I.*, *Nisbett R. E.*, *Norenzayan A.* Causal attribution across cultures: variation and universality // Psychological Bulletin. 1999. Vol. 125. P. 47–63.

9. *Crone N. E.*, *Miglioretti D. L.*, *Gordon B.*, *Lesser R. P.* Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event– related synchronization in the gamma band // Brain. 1998. V. 121. P. 2301–2315.

10. *Demidov V. E.* As we see, what we see. M.: Znanie, 1987. 238 p.

11. *Ding N., Melloni L., Zhang H., Tian X., Poeppel D.* Cortical trackingof hierarchical linguistic structures in connected speech // Nat. Neurosci. 2016. V.19. P. 158–164.

12. *Glezer V. D.* Vision and mind. L.: Nauka, 1993. 284 p.

13. *Güntekin B., Basar E.* Facial affect manifested by multiple oscillations // Int J Psychophysiol. 2009. V.71 (1). P. 31–36.

14. *Han S., Mao L., Qin J., Friederici A. D., Ge J.* Functional roles and cultural modulations of the medial prefrontal and parietal activity associated with causal attribution // Neuropsychologia. 2010. Vol. 49 (1). P. 83–91, DOI: 10.1016/j.neuropsychologia. 2010. 11. 003.

15. *Han S., Ma Y.* Cultural differences in human brain activity: a quantitative meta-analysis // NeuroImage. 2014. Vol. 99. P. 293–300.

16. *Han S., Ma Y.* A culture-behavior-brain loop model of human development // Trends in Cognitive Sciences. 2015. Vol. 19. P. 666–676.

17. *Kinchla R. A.* Detecting target elements in multi element arrays: A confusabi1ity model // Perception & Psychophysics. 1974. Vol. 15. P. 149–158.

18. *Kinchla R. A.* The role of structural redundancy in the perception of visual targets // Perception & Psychophysics. 1977. Vol. 22. P. 19–30.

19. *Kinchla R. A.*, *Wolfe J. M.* The order of visual processing: "Top-down," "bottom-up," or "middle-out" // Perception & Psychophysics. 1979. Vol. 25. P. 225–231.

20. *Knyazev G. G., Slobodskoj-Plusnin J.Y., Bocharov A. V.* Event-related Delta and Theta Synchronization During Explicit and Implicit Emotion Processing // Neuroscience. 2009. Dec 29;164(4):1588–600. doi: 10.1016/j.neuroscience.2009.09.057. Epub 2009 Sep 29.

21. *Kostandov E. A.*, *Cheremushkin E. A.* Changes in low- and high-frequency fluctuations in the alpha-range of the EEG in the intervals between significant visual stimuli // Physiol. person. 2013.Vol. 39. No. 4. P. 5–12.

22. *Kropotov Yu. D.* Quantitative EEG, cognitive evoked potentials of the human brain and neurotherapy. Donetsk.: Publisher Zaslavkiy Yu.A., 2010.

506 p.

23. *Logunova E.V, Pronin S.V, Shelepin Yu. E.* Modeling the operation of spatial-frequency filters in the perception of complex dynamic scenes // Optical journal. 2014. T. 81. No. 11.

24. *Miellet S., Vizioli L., He L., Zhou X., Caldara R.* Mapping face recognition information use across cultures // Frontiers in Psychology. 2013. Vol. 4. P. 1–12.

25. *Minami M., McCabe A.* Rice balls and bear hunts: Japanese and North American family narrative patterns // Journal of Child Language. 1995. Vol. 22. P. 423–45.

26. *Morris M., Peng K.* Culture and cause: American and Chinese attributions for social and physical events // Journal of Personality and Social Psychology. 1994. Vol. 67. P. 949–971.

27. *Muthukumaraswamy S. D.*, *Johnson B. W.*, *McNair N. A.* Murhythm modulation during observation of an object-directed grasp // Cognit. Brain Res. 2004. Vol. 19. P. 195–201.

28. *Nisbett R. E.* The geography of thought: how Asians and Westerners think differently, and why // New York: Free Press. 2003.

29. *Peng K., Knowles E. D.* Culture, education, and the attribution of physical causality // Personality and Social Psychology Bulletin. 2003. Vol. 29. P. 1272–84.

30. *Penn D. C.*, *Povinelli D. J.* Causal cognition in human and nonhuman animals: a comparative, critical review // Annual Review of Psychology. 2007. Vol. 58. P. 97–118.

31. *Polonio L., Di Guida S., Coricelli G.* Strategic sophistication and attention in games: An eye-tracking study // Games and Economic Behavior. 2015. Vol. 94. P. 80–96.

32. *Ponomarev V. A.* Latent sources of the electroencephalogram and event-associated potentials and their significance // Author's Abstract of Doctoral Dissertation, St. Petersburg, N. P. Bekhterova Institute of the Human Brain. 2016.

33. *Ponomarev V. A.*, *Pronina M. V.*, *Kropotov Yu.D*. Dynamics of the spectral density of the electroencephalogram in theta, alpha and beta ranges in the visual go / nogo test // Human Physiology. 2017.V.43. No. 4. P. 13–24.

34. *Rimmele J., Gross J., Molholm S., Keitel A.*, eds. Brain Oscillations in Human Communication. Lausanne.: Frontiers Media. 2018. doi: 10.3389/978–2–88945–458–7.

35. *Shelepin Yu. E., Kolesnikova L. N., Levkovich Yu.I.* Visocontrastometry, Leningrad.: Nauka, 1985. 105 p.

36. *Shelepin Y. E.*, *Danilova M. V.*, *Harauzov A. K.*, *Kropotov Y. D.*, *Sevostianov A. V.* Attention and preattentive vision // Abstracts of

XXXIII International. Congress of physiological sciences. St. Petersburg. 1997.

37. *Shelepin Yu. E., Borachuk O. V., Pronin S. V., Kharauzov A. K., Vasiliev P. P., Fokin V. A.* Face and non-verbal means of communication // Petersburg psychological journal. 2014. No. 9. Pp. 1–43.

38. *Shelepin Yu. E.* Introduction to neuroiconics: Monograph.—SPb.: Troitsky Most, 2017. 352 p.

39. *Shelepin Yu. E., Kharauzov A. K., Vakhrameeva O. A., Zhukova O. V., Pronin S. V., Tsvetkov O. V., Skuratova K. A., Shelepin E. Yu.* Unconscious visual signals and involuntary reactions of a person // Integrative physiology. 2021. T. 2. No. 4.

40. *Sheng F., Han S.* Manipulations of cognitive strategies and intergroup relationships reduce the racial bias in empathic neural responses // NeuroImage. 2012. Vol. 61 (4). P. 786–797. DOI:10.1016/j.neuroimage.2012.04.028.

41. *Tomasello M., Kruger A. C., Ratner H. H.* Cultural learning // Behavioral and Brain Sciences. 1993. Vol.16. P. 495–511.

42. *Vigario R. N.* Extraction of ocular artifacts from EEG using independent component analysis // EEG and Clin. Neurophysiol. 1997. V. 103. P. 395

43. *Zhukova O. V., Vasiliev P. P.* Neural network restructuring and changes in operators' strategies in the process of face image recognition // Journal of Optical Technology. 2020.Vol. 87. N. 10. P. 581–589.